

S. J. Mack
Y. Tsai
A. Sanchez-Mazas
H. A. Erlich,
Editors

13th International Histocompatibility Workshop Anthropology/Human Genetic Diversity Joint Report

Chapter 3: Anthropology/human genetic diversity population reports

Editors' addresses

Steven J. Mack^{1,2}, Yingssu Tsai³, Alicia Sanchez-Mazas⁴,
Henry A. Erlich^{1,2}

¹Children's Hospital Oakland Research
Institute, Oakland, CA; ²Roche Molecular
Systems, Alameda, CA; ³University of
California, Berkeley, CA; ⁴Laboratory of
Anthropology, Genetics, and Peopling History,
Department of Anthropology and Ecology,
University of Geneva, Switzerland

Acknowledgements

This work was supported by the Swiss SNF
grant #3100-49771.96 to A.S.-M.

Each participating laboratory has prepared a report describing the origins and historical, ethnographic and demographic information for each population sample, along with relevant results of previous genetic studies and summaries of the IHWG Biostatistic Core's analyses.

These reports are organized by the geographic region in the sections that follow. In some cases, population samples from different regions are described in the same report, and these instances are noted in the table for each region. When there is a discrepancy between the name of the population described in a report and the name used for the population sample in the analyses, the population sample name is provided in parentheses.

The allele, genotype, haplotype data and analyses for each sampled population can be found in appendices C, Tables 1 and 2 respectively, posted at the IHWG website under the Anthropology component (<http://www.ihwg.org/components/diversr.htm>), and in the eBook edition of HLA 2004 to be published by PubMed (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books>).

Sub Saharan Africa

Population: Kenyan from Kenya

Authors: Ma Luo¹, Joanne Embree¹, Suzie Ramdahin¹, Jeckoniah Ndinya-Achola², Simon Njenga², Job B. Bwayo², Kristine Jacobson³, Luvina Kwan³, Marlis Schroeder³, Sha Pan¹, Marc Jevan Narayansingh¹, Shehzad Iqbal¹, Robert C. Brunham^{1,4} and Francis A. Plummer^{1,2,5}

Authors' addresses: ¹Department Of Medical Microbiology, University Of Manitoba, 730 William Avenue, Winnipeg, Manitoba Canada. ²Department Of Medical Microbiology, University Of Nairobi, Nairobi, Kenya. ³HLA Lab, Cancer Care Manitoba, Winnipeg, Canada. ⁴University Of British Columbia

HLA 2004:
Immunobiology of the Human MHC.
Proceedings of the 13th International
Histocompatibility Workshop and Congress

Centre Of Disease Control, Vancouver, Bc, Canada. ⁵National Microbiology Laboratory, Winnipeg, Manitoba, Canada.

As the “cradle of humanity”, Kenya’s diverse population is the result of early settlements and migrations. The oldest human remains, predating tool-use, have been found in the Lake Victoria Highlands and Rift Valley regions of Kenya. Three major historical migrations into Kenya, along with the fusion of immigrants with early Kenyan hunters and gatherers, resulted in three major linguistic groups – Bantus, Nilotes and Cushites – and the formation of more than 70 ethnic groups. The Bantu people migrated into east Africa through several routes; those who settled on the hilltops behind the coastline formed the Mijikenda. Those who settled around Lake Victoria and Mt. Elgon became the ancestors of the Abaluhya, Abagusii and Abakuria. In the central regions, they became the Taveta, the Dawida and the Akamba people, while in the Kenya highlands they became the Agikuyu, Aembu, Chuka, Tharaka, and the Ameru. The Cushitic migrations brought Somali people to the northeastern part of Kenya, in the area located between the Juba and the Tana rivers. The eastern Cushites are represented by Yaaku, who live near Doldol, northwest of Mount Kenya. Migration also brought Orma, Boranna and Rendille into northern Kenya from southern Ethiopia. The Nilotic migrations brought three groups of Nilotic tribes into Kenya. The riverlake Nilotes are made up of the Luo, who settled around Lake Victoria. The plains Nilotes are made up of the Teso, Turkana, Maasai, Njemps and Samburu. The highland Nilotes are made up of the Omotik, Kalenjin-Pokot, Marakwet, Elgeyo, Tugen, Nandi and Terik. Since 600 AD, Arabs, Europeans and Asians have settled in the coastal area. Today these groups make up the smallest percentage of the population and are scattered all over Kenya (1–6). The origins and great diversity of the Kenyan population are of great interest for anthropological studies.

We studied the HLA allelic diversity of a group of unrelated

Kenyan women who were enrolled in a study of mother-to-child HIV transmission as a control population at the onset of the current HIV epidemic, when prevalence rates were approximately 2%. The women were enrolled in the study when they gave birth in the hospital of Nairobi. They originate from different areas of Kenya and represent diverse ethnic origins, therefore, this population is a representative of the larger Kenyan population. A high-resolution sequence-based method using a taxonomy-based sequence analysis (TBSA) to assign alleles (7–8) was used to type HLA class I (A, B and C) and class II (DPA, DPB1, DQA1, and DQB1) loci. The quality of the typing was controlled using the IHWG QC samples. The population genetic analysis was carried out by the IHWG Biostatistic Core.

Of seven HLA loci examined, the overall distributions of genotypes at the HLA-A, B, C, DPA1, DPB1, and DQB1 loci do not deviate significantly from expected Hardy-Weinberg proportions, whereas the distributions of genotypes at the HLA-DQA1 locus deviate significantly from what would be expected (p -value=0.0145). An excess of homozygous genotypes observed for the DPB1 locus (p -value=0.0122) could be due to population admixture, as most of the DPB1 homozygotes observed are originally from regions inhabited by different tribes.

Ewens-Watterson analysis of Homozygosity indicated that while the values of homozygosity at HLA-A, B, C, DQB1 and DPA1 loci are lower than the expected neutral values, they do not deviate significantly from the expected neutral values. A slightly higher than expected homozygosity value was observed at DPB1 locus. Ewens-Watterson analysis showed that the heterozygosity at DQA1 locus is significantly higher (p value=0.0250) than the expected neutral value, and is consistent with the action of balancing selection in this population at this locus. These data suggest that, in the face of an extremely diverse ethnic composition, the Kenyan population is overall under balancing selection.

References

1. Wills AJ. The Story Of Africa From The Earliest Times. London, University Of London Press. 1968–69.
2. Kenya, An Official Handbook. Published For The Government Of The Republic Of Kenya. East African Pub. House 1973.
3. Kenya. An Official Handbook. [Nairobi, Kenya: Ministry Of Information And Broadcasting, 1989.
4. Sharman M 1972 Africa Through The Ages; An Illustrated History. [2nd Edition, London] Evens Brothers. 1972.
5. Were GS. Western Kenya Historical Texts: Bbaluyia, Teso, And Elgon Kalenjin. Nairobi, East African Literature Bureau. 1967.
6. Luo M, Embree J, Ramdahin S, Jeckoniah Ndinya-Achola, Simon Njenga, Job B. Bwayo, Pan S, Mao X, Cheang M, Stuart T, Brunham RC, And Plummer F. 2002 HLA-A and HLA-B in Kenya, Africa: Allele Frequencies and Identification of HLA-B*1567 and HLA-B*4426. *Tissue Antigens* **59**, 370–380.
7. Luo M, Blanchard J, Pan Y, Brunham K, Brunham RC. 1999 High-Resolution Sequence Typing Of HLA-DQA1 and DQB1 Exon 2 DNA with Taxonomy-Based Sequence Analysis (TBSA) Allele Assignment. *Tissue Antigens* **54**, 69–82.

Population: Luo and Nandi from Kenya

Authors: A. M. Moormann¹, K. Cao^{2,3}, C. Masaberg^{2,3},
O. P. Sumba⁴, D. Koech⁵, J. Ng^{2,3}, R. J. Hartzman², J. Kazura¹,
M. A. Fernández-Viña^{2,3}

Authors' addresses: ¹The Center for Global Health and Diseases at Case Western Reserve University, Cleveland Ohio, USA, ²C.W. Bill Young/DoD Marrow Donor Program, Naval Medical Research Center, Bethesda MD, USA,

³The Department of Oncology, Georgetown University, Washington DC, USA,

⁴The Center for Vector Biology and Control Research, KERMI, Kisumu, Kenya, ⁵Kenya Medical Research Institute, Nairobi, Kenya

Kenya is home to many different populations that speak more than 60 languages. Two homogeneous populations of Luo (N=265) and Nandi (N=241) from Kenya were studied. The Luo were from Kanyawegi, (0.11 degrees south latitude, 34.63 degrees east longitude) which is located in the lowland Kisumu district of the Nyanza Province in western Kenya. The Luo (aka Kenyan Lowlander) population studied here was predominantly (97%) of Luo ethnicity and their primary language was Luo. The Nandi, a subgroup of Kalenjin, were from Kipsamoite, (0.33 degrees north latitude, 35.01 degrees east longitude) which is located in the highland Nandi district in the Rift Valley Province of western Kenya. The Nandi (aka Kenyan Highlander) population was entirely of Nandi ethnicity and their primary language was Nandi. The Luo and Nandi languages belong to the major linguistic Nilo Saharan and Eastern Sudanic-Nilotic groups.

Alleles at HLA class I (HLA-A, C, B) loci were typed using PCR-SSOP methods (1). Genotype ambiguities were resolved using group-specific amplification (1). All genotypic ambiguities were tested and resolved for the WHO alleles recognized up to January 2002, with exception of alleles differing outside exons 2 and 3.

The distribution of genotypes at the HLA class I loci in these populations did not show significant deviation from Hardy-Weinberg expectations. High frequencies of heterozygous genotypes were observed at all HLA class I loci. Significantly low values of homozygosity were detected using the Ewens-Watterson homozygosity test at the HLA-A and B loci in the Nandi and at the HLA-A and C loci in the Luo.

We observed 28 and 30 HLA-A alleles, 39 and 45 HLA-B alleles, and 21 and 22 HLA-C alleles in the Nandi and Luo, respectively. Both populations show similarities in the allele distribution at HLA-A and B loci, and there is no single predominant allele at either locus. The nine most frequent HLA-A alleles in both populations have frequencies ranging from 0.1183 to 0.0498. Among these alleles, A*0201, A*0101, A*6802,

A*2301, A*6601, and A*2902 are frequent in both groups. The A*0202, A*0205, A*3402, A*3001, and A*3002 alleles are present with relatively high frequencies in one population (e.g. A*0202, A*0205, and A*3402 in the Nandi, and A*3001 and A*3002 in the Luo) but have low frequencies in the other. For HLA-B the most common alleles are B*5801, B*5802, B*1503, B*5301, B*4201, B*4501, and B*1801 with frequencies ranging from 0.1245 to 0.0434. The B*1402, B*2703, B*5703, and B*8101 alleles have high frequencies in one group but low frequencies in the other (B*2703 and B*5703 are high in the Nandi, and B*1402 and B*8101 are high in the Luo). For HLA-C, the most frequent alleles in both populations are Cw*0602, Cw*0701, Cw*0401, Cw*1701, Cw*0202, and Cw*0704, with frequencies ranging from 0.2167 to 0.0509. Cw*0702 and Cw*0802 are present in the Luo with higher frequencies, whereas Cw*1402 and Cw*1602 have higher frequencies in the Nandi.

We identified 35 and 29 HLA-A-C-B haplotypes in the Nandi and Luo, respectively. The most common haplotypes observed in the Nandi population were A*3001-Cw*1701-B*4201, A*0202-Cw*0602-B*5802, A*6601-Cw*0602-B*5802, A*2902-Cw*1701-B*4201, and A*0202-Cw*0202-B*1503. In the Luo population, the A*6601-Cw*0602-B*5802, A*0201-Cw*0202-B*1503, and A*2902-Cw*1701-B*4201 haplotypes had the highest frequencies. Significant overall linkage disequilibrium exists between all pairs of loci typed in the two populations. HLA-C-B loci exhibited stronger associations than HLA-A-B loci or HLA-A-C loci. Many pairs of HLA-C and B alleles had absolute ($D' = 1$) or almost complete associations, that is, each of these HLA-B alleles has predominant or absolute associations with only one HLA-C allele. Complete or almost complete associations were observed for Cw*0602-B*1302, Cw*0304-B*1510, Cw*0202-B*2703, Cw*0404-B*4012, Cw*1701-B*4101, Cw*1701-B*4201, and Cw*0602-B*5802 pairs. Certain HLA-B alleles show strong associations with two specific HLA-C alleles in both populations: B*1503 associates with Cw*0202 and Cw*0401, B*4501 associates with Cw*0602 and Cw*1601, B*5301 associates with Cw*0401 and Cw*0602, and B*8101 associates with Cw*0804 and Cw*1801. Some HLA-B alleles display associations with multiple HLA-C alleles: B*0801 associates with Cw*0701/0702/0704, B*1801 associates with Cw*0501/0701/0702/0704, and B*5801 associates with Cw*0302/0701. Some HLA-B alleles associate with certain HLA-C alleles in only one population: B*0801-Cw*0701 and B*1801-Cw*0501 in the Nandi, and B*0801-Cw*0702 and B*1801-Cw*0701 in the Luo.

The estimated genetic distances based on the distributions

of class I alleles showed that the Luo and Nandi are close to each other. They are also close to some other African populations such as Ugandans at HLA-A, C loci (data not shown), Cameroonians at HLA-A, B loci (2) and Zulus from South Africa at the HLA-A locus (3). The Luo display close genetic distance to Zulus at HLA-B (4) and to Zambians at HLA-C (data not shown). Interestingly, the two Kenyan populations are also close to African-Americans at HLA-A and C (5). The

differences in the contribution of alleles to genetic distances are likely to be related to the evolutionary histories of the HLA loci in different continents.

Acknowledgements

The research was funded by the NIH AI43906 “Immunologic Memory to pre-erythrocytic *Plasmodium falciparum*” and is supported by funding from the Office of Naval Research N00014-00-1-0898 to the C.W. Bill Young Marrow Donor Recruitment and Research Program.

References

1. K Cao, M Chopek, MA Fernández-Viña. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). *Reviews in Immunogenetics* **1**, 177–208, 1999.
2. Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, *2612, *3006 and HLA-B*1403, *4016, *4703. Ellis JM, Mack SJ, Leke RF, Quakyi I, Johnson AH, Hurley CK. *Tissue Antigens* **56**(4), 291–302, 2000.
3. D Middleton, F Williams, A Meenagh, AS Daar, C Gorodezky, M Hammond, E Nascimento, I Briceno, MP Perez. Analysis of the distribution of HLA-A alleles in populations from five continents. *Hum Immunol* **61**(10), 1048–1052, 2000.
4. F Williams, A Meenagh, C Darke, A Acosta, AS Daar, C Gorodezky, M Hammond, E Nascimento, D Middleton. Analysis of the distribution of HLA-B alleles in populations from five continents. **62**(6), 645–650, 2001.
5. K Cao, JA Hollenbach, XJ Shi, WX Shi, M Chopek, MA Fernández-Viña. Analysis of the frequencies of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* **62**, 1009–1030, 2001.

Population: Malian Admixed from Mali

Authors: K. Cao^{1,2}, K. E. Lyke³, C. Masaberg^{1,2}, O. K. Doumbo⁴, J. Ng^{1,2}, R. J. Hartzman¹, C. V. Plowe³, M. B. Sztein³, M. A. Fernández-Viña^{1,2}

Authors' addresses: ¹C.W. Bill Young/DoD Marrow Donor Program, Naval Medical Research Center, Bethesda MD, USA, ²Department of Oncology, Georgetown University, Washington DC, USA, ³The Center for Vaccine Development, University of Maryland Baltimore, Maryland, USA, and ⁴Malaria Research and Training Center, Faculty of Medicine, Pharmacy and Dentistry, University of Bamako, Mali, West Africa

To study the diversity and associations of the HLA class I (A, C, B) alleles, 138 subjects from Mali were tested. The population studied is relatively homogeneous. In this sample, 80.7% of the subjects self reported as Dogon, 5.3% as Peuhl (also known as Fulani), 3.5% as Bambara, and 10.5% as other ethnic groups. The Dogon speak various dialects of the Dogon language, part of the broad Niger/Atlantic/Volta Congo group. Peuhl, Bambara and French are also spoken in the study area.

Allele-level typing of the HLA class I (A, C, and B) loci was performed using PCR and hybridization with sequence-specific oligonucleotide probes (SSOP) (1). Allelic resolutions for genotypic ambiguities were achieved using group-specific amplification and SSOP hybridization (1).

Distribution of alleles at HLA-A, C, and B loci

Twenty-one HLA-A alleles, 31 HLA-B alleles and 19 HLA-C alleles were identified in this population. Compared with other African populations studied, fewer alleles were observed at all three HLA class I loci. In addition, the Malians showed a few predominant alleles at each of HLA-A, C and B loci, in contrast with other African populations. The most frequent alleles for the three class I loci include: A*2301, A*3001, B*5301, B*4201, B*3501, Cw*1601, Cw*0401, and Cw*1701, and their allele frequencies (a.f.) are 0.2283, 0.1413, 0.1594, 0.1377, 0.1268, 0.2830, 0.2132, and 0.1434, respectively. These alleles are present with lower frequencies in other populations including Kenyans (data not shown), Ugandans (data not shown), Zambians (data not shown), Zulus from South African (2–3), Cameroonians (4), the Mandenka tribe from Senegal (5) and with some exceptions, African Americans (6). A*2301 is seen at a similar frequency in Cameroonians (a.f.=0.187); and A*3001, B*4201, and Cw*1701 exhibit similar frequencies in Zambians (data not shown). Together with the other common alleles (having allele frequencies >0.5), the eight predominant alleles mentioned above make up 68.48% of the population coverage for HLA-A, 76.44% of the population coverage for HLA-B, and 81.02% the population coverage for HLA-C. The alleles with intermediate frequencies include A*3303, A*0201, A*0202, A*6802, B*5201, B*1503, B*7801, B*4501, B*0702, Cw*0701, Cw*0602, and Cw*0202. The B*7801 allele was one of the most frequent al-

leles observed around the African continent, yet was absent in Kenyans, Ugandans, Zambians, Zulus and Cameroonians.

Hardy-Weinberg equilibrium and Ewens-Watterson homozygote test of neutrality

The distribution of heterozygous genotypes at each of the HLA class I loci did not deviate from the proportions expected under Hardy-Weinberg equilibrium. Relatively high frequencies of heterozygous genotypes were observed at all class I loci, with frequencies of 0.913, 0.949 and 0.876 for the HLA-A, C and B loci, respectively. The diversity measured by frequencies of heterozygous genotypes at these loci in Malians was not as remarkable as that in other African populations. Malians displayed lower than expected values of homozygosity under the Ewens-Watterson model at the HLA-A and B loci, but were not significantly different from the values expected under neutrality for these loci. The observed homozygosity at the HLA-C locus was virtually identical to that expected under neutrality.

Haplotypic associations

One-hundred-nineteen HLA-A-C-B haplotypes, 106 HLA-A-B haplotypes, 89 HLA-A-C haplotypes, and 57 HLA-C-B haplotypes were observed in Malians. The most common three-locus haplotypes are A*3001-Cw*1701-B*4201, A*2301-Cw*1601-B*5201, A*0201-Cw*1601-B*4501, and A*2301-Cw*0401-B*3501 with frequencies ranging from 0.1044 to 0.0342. As seen in other populations, the associations between the HLA-B and HLA-C loci were stronger than that between the HLA-A and HLA-B loci or between the HLA-A and HLA-C loci. Many pairs of C-B alleles had absolute ($D' = 1$) or almost complete associations. The most frequent HLA-C-B haplotypes are Cw*1701-B*4201, Cw*0401-B*5301, Cw*1601-B*5201, Cw*0401-B*3501, and Cw*1601-B*7801, with frequencies ranging from 0.1395 to 0.0698. Interestingly, the three predominant HLA-C alleles display associations with multiple HLA-B alleles. Similarly, two of the three most frequent HLA-B alleles associate with multiple HLA-C alleles. The frequent B*5301 allele was observed in haplotypes with Cw*0401,

Cw*0404, Cw*0602, Cw*0701, Cw*0702, Cw*1402, and Cw*1601. The Cw*1601 allele was associated with B*0702, B*1516, B*3501, B*4501, B*5101, B*5201, B*5301, and B*7801; haplotypes with Cw*0401 included the B*0702, B*1503, B*1509, B*1510, B*3501, B*3517, B*4403, B*4501, B*5001, and B*5301 alleles. An absolute association between B*4201 and Cw*1701 was observed. The cumulative frequencies of the three predominant HLA-C alleles (Cw*1601, Cw*0401, and Cw*1701, combined a.f. = 0.6396) was higher than the combined frequencies of the three predominant HLA-B alleles (B*5301, B*4201 and B*3501, combined a.f. = 0.4239). This observation could suggest the action of directional selection on the HLA-C locus, and hitchhiking of HLA-B alleles in C-B haplotypes under such selection could result in an overall decrease in diversity at the HLA-B locus. However, neither directional selection at the HLA-C locus, nor reduced diversity at the HLA-B locus are indicated by the Ewens-Watterson homozygosity test.

Genetic distance at HLA class I loci

Genetic distances were estimated based on the distributions of alleles at the HLA class I loci. These were compared with other populations including Kenyans (data not shown), South African Zulus (2–3), Cameroonians (4), the Mandenka tribe from Senegal (5), and African Americans (6). The comparison of the genetic distances between these African populations showed that Malians displayed greater distances with all the other African populations. Malians appear to show the closest relationships with Cameroonians (A and B loci), Zambians (C and B loci), Zulus (A locus), and the Mandenka tribe (C locus) which correlate with either common broad linguistic groups and/or their geographic location.

Acknowledgements

The research was funded by Contract no. N01-AI-85346 from the National Institutes of Health (NIH), and is supported by funding from the Office of Naval Research N00014-00-1-0898 to the C.W. Bill Young Marrow Donor Recruitment and Research Program.

References

1. K Cao, M Chopek, MA Fernández-Viña. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). *Reviews in Immunogenetics* 1, 177–208, 1999.
2. D Middleton, F Williams, A Meenagh, AS Daar, C Gorodezky, M Hammond, E Nascimento, I Briceno, MP Perez. Analysis of the distribution of HLA-A alleles in populations from five continents. *Hum Immunol* 61(10), 1048–1052, 2000.
3. F Williams, A Meenagh, C Darke, A Acosta, AS Daar, C Gorodezky, M Hammond, E Nascimento, D Middleton. Analysis of the distribution of HLA-B alleles in populations from five continents. 62(6), 645–650, 2001.

4. JM Ellis, SJ Mack, RF Leke, I Quakyi, AH Johnson, CK Hurley. Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, *2612, *3006 and HLA-B*1403, *4016, *4703. *Tissue Antigens* **56**(4), 291–302, 2000.
5. A Sanchez-Mazas, QG Steiner, C Grundschober, JM Tiery. The molecular determination of HLA-Cw alleles in the Mandenka (West Africa) reveals a close genetic relationship between Africans and Europeans. *Tissue Antigens* **56**(4), 303–312, 2000.
6. K Cao, JA Hollenbach, XJ Shi, WX Shi, M Chopek, MA Fernández-Viña. Analysis of the frequencies of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* **62**, 1009–1030, 2001.

Population: Mandenka from Senegal

Authors: Alicia Sanchez-Mazas¹, Barbara Kervaire², Jean-Marie Tiercy²

Authors' addresses: ¹Laboratory of Anthropology, Genetics and Peopling History, Department of Anthropology and Ecology, University of Geneva, Switzerland, and ²Transplantation Immunology Unit, University Hospital, Geneva, Switzerland

The Mandenka (plural Mandenkalu) population is an endogenous community of about 3,000 individuals living in a hilly area of the Niokholo District, near the town Kédougou, in Eastern Senegal, West Africa (12 degrees 35 minutes north latitude, 12 degrees 09 minutes west longitude). Their language belongs to the Mande family, the first to differentiate from a proto-Niger-Congo phylum emerging about 8,000 years ago in an area surrounding present Nigeria and Cameroon.

The Mandenkalu invaded the Niokholo region, at that time a territory of the Bedik population, during the Mali Empire (14th Century). At the end of the 18th Century, a third population, the Bande Fulani, also settled in that area. After a period of trouble at the end of the 19th Century due to a war with the Fouta Djallon Fulani, the Mandenkalu, Bedik and Bande Fulani established peacefully in close villages. Despite their cultural differences, they maintain a certain amount of gene flow (1).

The Mandenkalu were previously analyzed for numerous polymorphisms, including serological markers (1, 2), HLA-C (3) and HLA class II (4) molecular polymorphisms, mtDNA (3), random nuclear probes (6), the Y chromosome (7), and α - and β -globin (8, 9). Following HLA-ABC serological typings carried out in 1991 (1), these 3 class I loci were further typed by PCR-SSOP, using the IHWG RLS protocol, for A and B, and local primers and probes, for HLA-C (3). HLA-DRB1, DRB3/B4/B5 and DQB1 were previously typed by PCR-SSOP (4), using local primers and probes, and HLA-DPB1 by commercial reagents (10). Of the initial sample of 154 individuals commonly typed at high-resolution for the 7 HLA loci mentioned above, 85 could be considered as “unrelated” (i.e. not including close relatives). The results are summarized in part in reference #11 listed below (13th Workshop book).

The results briefly described in this report concern the set of 93–94 unrelated individuals typed for A and B with the IHWG RLS reagents, and analyzed by the IHWG Biostatistic Core (12).

Locus A exhibits high frequencies for alleles *2301 (0.177) and *3002 (0.124), leading to a relatively L-shaped frequency distribution not rejected by the Ewens-Watterson (EW) selective neutrality test. Locus B exhibits a more even distribution, with only one frequency exceeding 10%, the frequency of *3501 (0.122). For this locus, the EW test rejects the hypothesis of neutrality at the 5% level ($0.025 < P < 0.050$). There is no detectable deviation from Hardy-Weinberg equilibrium at HLA-A and B loci. Finally, two haplotypes are observed at an estimated frequency around 5%: A*0202-B*5801 (0.053), and A*3002-B*3501 (0.048).

Compared to other sub-Saharan African populations tested in this Workshop, the Mandenkalu share the same most frequent HLA-A alleles with the Shona from Zimbabwe (A*2301, A*3002), at close frequencies. The neighboring Malian population also exhibits A*2301, but at a higher frequency than in Mandenkalu, as well as the subtype A*3001, rather than A*3002. At the B locus, the Malian population exhibits a very similar frequency of B*3501, but also a high frequency of two alleles (B*5301, B*4201) that are uncommon in Mandenka.

These preliminary results indicate that, in general, the Mandenkalu exhibit some common features with other sub-Saharan Africans, but also that their genetic profile is clearly distinct even from that observed in the Malians, their closest geographic neighbors. Overall, HLA-B distributions observed among sub-Saharan African populations seem more difficult to interpret in relation to historical or geographical relationships than HLA-A distributions.

Acknowledgements

We warmly thank the Swiss National Foundation for Scientific Research (grants #3100-49771.96 to A.S.-M. and #3100-057557.99 to J.-M.T). We are also particularly grateful to the Mandenkalu for their welcome and cooperation.

References

1. Blanc M, et al. Inter-ethnic genetic differentiation: Gm polymorphism in Eastern Senegal. *Am J Hum Genetics* **46**, 383–392 (1990).
2. Dard P, et al. Polymorphisme des HLA loci de classe I, HLA-A, -B, et -C, dans la population Mandenka du Sénégal Oriental. *C. R. Acad Sci Paris* **314**, série III, 573–578 (1992).
3. Sanchez-Mazas A, Steiner QG, Grundschober C, Tiercy J-M. The molecular determination of HLA-Cw alleles in the Mandenka (West Africa) reveals a close genetic relationship between Africans and Europeans. *Tissue Antigens* **56**, 303–312 (2000).
4. Tiercy J-M, et al. HLA-DR polymorphism in a Senegalese Mandenka population: DNA oligotyping and population genetics of DRB1 specificities. *Amer J Hum Genet* **51**, 592–602 (1992).
5. Graven L, et al. Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. *Mol Biol Evol* **12**, 334–345 (1995).
6. Poloni ES, Excoffier L, Mountain JL, Langaney A, Cavalli-Sforza LL. Nuclear DNA polymorphism in a Mandenka population from Eastern Senegal: comparison with eight other human populations. *Ann Hum Genet* **59**, 43–61 (1995).
7. Poloni ES, et al. Human genetic affinities for Y-chromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics. *Amer J Hum Genet* **61**, 1015–1035 (1997).
8. Martinson JJ, et al. High diversity of alpha-globin haplotypes in a Senegalese population, including many previously unreported variants. *Amer J Hum Genet* **57**, 1186–1198. (1995).
9. Currat M, et al. Molecular analysis of the beta-globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the beta(S) Senegal mutation. *Am J Hum Genet* **70**, 207–223 (2002).
10. Sanchez-Mazas A, Tiercy J-M, Kervaire B, Langaney A. HLA-DPB1 in anthropology: A unique polymorphism. Abstract 245. *Human Immunology* **61**, Suppl. 1, S127 (2000).
11. Sanchez-Mazas A, Tiercy J-M. The HLA molecular diversity of an African population from Eastern Senegal. In: Hansen JA, Dupont Bo (eds.) *HLA 2004, Immunobiology of the Human MHC*. Seattle, IHWC Press (2003).
12. Results of data analysis performed on the “CHESAN-Mandenka population” by the IHWG Biostatistics Core.

Population: Hutu and Tutsi from Kigali, Rwanda

Authors: J Tang, S Allen, E Karita, and RA Kaslow

Authors' address: Departments of Medicine and Epidemiology, University of Alabama at Birmingham, Birmingham, AL, USA; National AIDS Control Program, Kigali, Rwanda

Kigali is the capital city of Rwanda, the mostly agricultural “Country of a Thousand Hills” located in central Africa. The local residents speak either French (the official language) or the local Bantu language known as Kinyarwanda. The total population in Kigali is over 240,000.

There are two main tribes in Rwanda. The original inhabitants are the shorter Bantu people known as the Hutu. The tall and thin Tutsi came from Ethiopia during the Middle Ages. The Tutsi ruled the Hutu for many centuries before Germany colonized Rwanda in the late 19th century. Kigali (1 degree 57 minutes south latitude, 30 degrees 4 minutes east longitude) became a regional center from 1919 to 1962 as part of the Belgian colony. When Rwanda gained independence in 1962, Kigali was chosen to be the nation's capital. Many of the Tutsi fled to Uganda at that time.

In 1986, Rwandan women seeking antenatal care at the Central Hospital of Kigali were screened for antibodies to HIV-1 (1). A random sample of 1458 (460 HIV seropositive and 998 HIV seronegative) women were enrolled into prospective studies in 1988 (2). The annual mortality averaged 4–5% among the HIV-infected, compared with 0.1–0.2%

among HIV-uninfected women, between 1988 and 1994 (3). The study was interrupted by genocide and civil war from April through August 1994. Participants returning after September 1994 were re-enrolled for intensive 3-monthly follow-up including medical history, physical examination, laboratory studies, and sampling of peripheral blood mononuclear cells (PBMCs).

Details of the 280 Kigali participants (202 HIV-1 seropositive and 78 seronegative) retained for HLA typing have been described elsewhere (4). Two lines of evidence indicated that loss of other participants due to civil war and lack of data or banked biological specimens did not seem to create a genetic bias. First, genotyping data from several microsatellites (5) and additional loci encoding CC (β) chemokine receptor 2 (CCR2) (5), CCR5 (5), and peptide transporters TAP1 (6) and TAP2 (7) closely fit into Hardy-Weinberg equilibrium (HWE). Second, at the 2-digit-specificity level, the distribution of HLA class I and class II alleles did not deviate from Hardy-Weinberg equilibrium (HWE) (8).

The high-resolution HLA-DRB1 and DQB1 alleles described here were defined by sequencing-based typing (SBT) and PCR with sequence-specific primers, respectively (8). Global exact tests did not reveal clear genetic heterogeneity between HIV-1 seronegative and seropositive groups ($P=0.14$ for DRB1 locus and 0.27 for DQB1). The top 12 DRB1 alleles in this cohort were *1503 (freq.=0.170), *1302 (0.145), *1101 (0.093), *0102 (0.079), *0701 (0.057), *1303 (0.057), *0301 (0.055), *1102 (0.048), *1301 (0.042), *0302 (0.041),

*1001 (0.036), and *0804 (0.034). Eight common DQB1 alleles, i.e., *0602 (Freq.=0.213), *0501 (0.195), *0301 (0.186), *0201 (0.155), *0604 (0.107), *0609 (0.057), and *0402 (0.034), and *0603 (0.018) accounted for 96% of the total. These common alleles produced 17 DRB1-DQB1 haplotypes with a frequency exceeding 0.01.

The observed DRB1 homozygosity (35.0%) in the Kigali cohort differed ($P=0.016$) from the frequency (23.4%) expected based on HW proportions. Allele dropout in SBT could not explain this difference since DRB1-DQB1 haplotypes were assigned unambiguously. Meanwhile, the observed and expected frequencies matched closely for the individual heterozygous genotypes involving common alleles at the DRB1 locus. The distribution of DQB1 genotypes supported HWE as well as the Ewens-Watterson neutrality model (Appendices 6.1 and 7.1).

Among the 17 major DQB1 genotypes (allele pairs) present

in the Kigali cohort, the *0301/*0501 combination was much less frequent than expected assuming HW proportions ($P=0.023$). This was initially considered as a chance phenomenon resulting from multiple comparisons. However, similar decrease was also seen in HIV-1 seropositive Caucasians in another cohort ($P=0.058$) (J. Tang and R. A. Kaslow, unpublished data). Thus, the finding might have broader implications in the context of HIV-1 infection or population genetics in general.

Overall, the Kigali cohort resembled several other African populations in terms of HLA-DRB1 and DQB1 polymorphisms alone (8). However, HLA haplotypes spanning the alpha (HLA-A), beta (HLA-B and HLA-C), and delta blocks are highly diversified among native Africans. Accordingly, genetic relatedness cannot be simply measured by genotypes defined at a single HLA locus or several loci within the same evolutionarily conserved block.

References

- Allen S, Van de Perre P, Serufilira A, Lepage P, Carael M, DeClercq A, Tice J, Black D, Nsengumuremyi F, Ziegler J. Human immunodeficiency virus and malaria in a representative sample of childbearing women in Kigali, Rwanda. *J Infect Dis* **164**, 67–71, 1991.
- Allen S, Lindan C, Serufilira A, Van de Perre P, Rundle AC, Nsengumuremyi F, Carael M, Schwalbe J, Hulley S. Human immunodeficiency virus infection in Urban Rwanda. Demographic and behavioral correlates in a representative sample of childbearing women. *JAMA* **266**, 1657–1663, 1991.
- Lindan CP, Allen S, Serufilira A, Lifson AR, de Perre PV, Chen-Rundle A, Batungwanayo J, Nsengumuremyi F, Bogaerts J, Hulley S. Predictors of mortality among HIV-infected women in Kigali, Rwanda. *Ann Intern Med* **116**, 320–328, 1992.
- Tang J, Costello C, Keet IPM, Rivers C, LeBlanc S, Karita E, Allen S, Kaslow RA. HLA class I homozygosity accelerates disease progression in human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses* **15**, 317–324, 1999.
- Tang J, Myracle AD, Allen SA, Karita E, Musonda R, Fultz PA, Kaslow RA. Novel alleles at the lymphotoxin alpha (LTa) locus mark extended HLA haplotypes in native Africans. *Hum Immunol* **62**, 269–278, 2001.
- Tang J, Freedman DO, Allen S, Karita E, Musonda R, Braga C, Margolick J, Kaslow RA. TAP1 polymorphisms in several human ethnic groups: characteristics, evolution, and genotyping strategies. *Hum Immunol* **62**, 256–268, 2001.
- Tang J, Freedman DO, Allen SA, Karita E, Musonda R, Braga C, Jamieson BD, Louie L, Kaslow RA. Genotyping TAP2 variants in North American Caucasians, Brazilians, and Africans. *Genes Immun* **2**, 32–40, 2001.
- Tang J, Naik E, Costello C, Karita E, Rivers C, Allen S, Kaslow RA. Characteristics of HLA class I and class II polymorphisms in Rwandan women. *Exp Clin Immunogenet* **17**, 185–198, 2000.

Population: Shona from Harare, Zimbabwe

Authors: Louie, L.¹, Mather, K.², Meyer, D.², Hollenbach, J.², Jackman, R.¹, Schultz, K.¹, Yiu, C.¹, Sohsmann, M.¹, Nguyen, V.¹, Katzenstein, D.³, Mason, P.⁴, Thomson, G.²

Authors' addresses: ¹Children's Hospital Oakland Research Institute, Oakland, CA, USA; ²University of California, Berkeley, CA, USA; ³Stanford Univer-

sity Medical School, Stanford, CA, USA; ⁴Biomedical Research and Training Institute and University of Zimbabwe Medical School, Harare, Zimbabwe

Zimbabwe is a landlocked country located at the southern end of the Great Rift Valley in South Central Africa. The country covers 150,803 square miles of land (1) ranging in topography from volcanic mountains and plateaus to low lying

plains and rivers. It is bordered on the north by Zambia, on the northeast and east by Mozambique, on the west and southwest by Botswana, and on the south by South Africa. Formerly known as Southern Rhodesia, the population of Zimbabwe in 2000 was estimated at 12 million (2).

Humans began to populate the area 30–40,000 years ago, according to rock paintings that exist across the country. The Mashona (people of the Shona tribe) are a heterogeneous group of several assimilated tribal entities that have inhabited Zimbabwe since migrating northwest from the coastal regions beginning in 500 AD (3). The earliest evidence of organized civilization was Great Zimbabwe, the stone city at the center of international gold trade from 1200–1500 AD with East Africa, the Persian Gulf, India and China. External trade was established with Muslim and Portuguese people in the late 1400s with the Mutapa State, a group that migrated northeast from Great Zimbabwe. However the Shona maintained control of Mutapa State and the slave trade had little impact on Zimbabwean plateau communities (4). Because domestic slavery was also minimal, the Mashona experienced little admixture and no bottlenecks. Five major Bantu subgroups (north, central, south, east, and west) comprise the present-day Mashona. The language spoken is Shona. Prior to colonialization in the late 1800's, polygamy was commonly practiced.

Harare is the capital city of Zimbabwe, at the heart of the central plateau at an elevation of 5,000 feet. In Harare, the Mashona comprise ~80% of the local population with the next most prevalent group present at a frequency of 11%.

Subjects were recruited from the Harare area as part of a study of tuberculosis and HIV. Because individuals were collected for a case-control study of disease, a subset was chosen in the following way to approximate a random population sample. All 229 subjects submitted to the Anthropology Component for analysis were ethnically Shona. Subjects were asked by questionnaire to identify their tribal affiliation and mother tongue as well as that of their parents. Only those whose parents were both Shona were eligible for inclusion. For the HLA Anthropology study, subjects assigned to each of four categories (HIV+/TB+, HIV-/TB+, HIV-/TB+, HIV-/TB-) were randomly selected from the disease study participants based on the known prevalence of HIV-1 infection, TB disease, and TB/HIV co-morbidity in the community (Table 1)(5, 6). The total sample contained individuals in each category in proportion to their frequency in the population (Table 2).

From blood collected into anticoagulant acid citrate dextrose, DNA was extracted using a standard salting out

Table 1. HIV and TB prevalence in Zimbabwe

	HIV+	HIV–	Totals
TB+	0.5%	0.13%	0.63%
TB–	24.56%	74.81%	99.37%
Totals	25.06%	74.94%	100%

Table 2. Subjects used in reconstructed random population

	HIV+	HIV–	Totals
TB+	1	0	1
TB–	56	172	228
Totals	57	172	229

technique (7). Allele-level typing of class I (A, B, C) and class II (DRB1, DQB1, DQA1, DPB1) was done using a combination of methods, including reverse line strip hybridization (Roche Molecular Systems, Inc, Alameda CA, (8)), dot blot hybridization with sequence specific oligonucleotide probes (7), sequence-specific PCR (9), and single strand conformational polymorphism to confirm homozygotes.

We find that all loci but HLA-DQA1 are consistent with expected Hardy-Weinberg equilibrium proportions. In the Shona, applying Ewens-Watterson F-statistic analysis, we find significantly lower homozygosity than expected under neutrality at C, DRB1, and DQA1 and lower (but not significant) homozygosity at A, B, and DQB1. In contrast, homozygosity at DPB1 is higher than that expected under neutrality, although not significantly so.

Within class I, linkage disequilibrium (LD) is high ($D' > 0.85$) only between B and C (separated by 85 kb). Excluding DPB1, all class II loci (DRB1, DQA1, DQB1, encompassing 75 kb) have high pairwise LD ($D' > 0.85$). Three pairs of loci had D' values of approximately 0.6 (A-B, A-C, B-DRB1). These loci are separated by much greater distances: 1.17–1.56 Mb and three of them (A, B, DRB1) are highly polymorphic. Relatively high LD across long distances may be explained by their high polymorphism: LD increases with polymorphism. DPB1 is not in LD with any of the other loci. The most common A:B:DRB1 haplotype is *6802-*0702-*1503 followed by *3002-*1801-*1101 (11 and 7 copies observed, respectively, for a combined frequency of 0.04).

African populations are expected to show high variation and low linkage disequilibrium due to historically large effective population sizes (10, 11). Comparisons to other populations in this volume will shed light on these predictions. At several loci we find evidence for balancing selection suggesting that the allelic repertoire in the Shona may have been shaped by past infectious diseases.

References

1. Zimbabwe. In: The Columbia Encyclopedia. sixth ed: Columbia University Press; 2000.
2. Xinhua News Agency. Zimbabwe's population officially put at 12.65 million. In: newafrica.com, Green Growth Ltd; 2001.
3. Radonjic B. Africa South of the Sahara Chronology: Great Zimbabwe. In: Koeller DA; 1996.
4. Olufunwa B, Mazrui A, Yefru Y, Ajayi JFA, Donkoh W, Odhiambo A, et al. Central Africa & the Coming of the Portuguese. In: The Story of Africa. London: BBC World Service.
5. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *Jama* **282**(7), 677–686, 1999.
6. UNAIDS. The Report on the Global HIV/AIDS Epidemic: "The Barcelona Report". XIV International Conference on AIDS, Barcelona, 7–12 July 2002. In: <http://www.unaids.org/barcelona/presskit/report.html>; 2002.
7. Twelfth international histocompatibility workshop technical handbook. In: Charron D, Fauchet R, editors.; 1996; Paris: HLA et Medecine; 1996.
8. Erlich H, Bugawan T, Begovich AB, Scharf S, Griffith R, Saiki R, et al. HLA-DR, DQ and DP typing using PCR amplification and immobilized probes. *European Journal of Immunogenetics* **18**(1–2), 1991.
9. Olerup O, Aldener A, Fogdell A. HLA-DQB1 and -DQA1 typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. *Tissue Antigens* **41**(3), 119–134, 1993.
10. Tang J, Naik E, Costello C, Karita E, Rivers C, Allen S, et al. Characteristics of HLA class I and class II polymorphisms in Rwandan women. *Exp Clin Immunogenet* **17**, 185–198, 2000.
11. Louie L, Francisco AM, W. K, Katongole-Mbidde E. HLA class II among the Baganda in Uganda. In: 12th International Histocompatibility Conference; 1996; Paris, France: Human Immunol; 1996. p. 141.

Populations: Ugandans from Kampala, Uganda and Zambians from Lusaka, Zambia

Authors: K. Cao¹, C. Masaberg¹, J. Yu², D. L. Mann³, M. A. Fernández-Viña¹

Authors' addresses: ¹previous: American Red Cross National Histocompatibility Laboratory; current: C.W. Bill Young/DoD Marrow Donor Program, Naval Medical Research Center, Bethesda MD, USA and the Department of Oncology, Georgetown University, Washington DC, USA, ²American Red Cross National Histocompatibility Laboratory, Baltimore, USA, and ³Division of Immunogenetics, Department of Pathology, University of Maryland School of Medicine, Baltimore, USA

To study the diversity and associations of HLA class I (A, C, and B) alleles, 163 and 45 subjects from Uganda and Zambia, respectively, were tested. Allelic typings of HLA class I (A, C, and B) loci were performed using PCR and hybridization methods with sequence-specific oligonucleotide probes (SSOP) as described (1). Allelic resolutions for genotype ambiguities were achieved utilizing group-specific amplification and SSOP hybridization methods (1).

Samples were collected from random individuals living in the capital cities Kampala, Uganda and Lusaka, Zambia. Both countries are located in Central East Africa and present a variety of populations and languages. Even though the collected samples may have limited application for anthropological studies, the results reported here illustrate the genetic diversity of Africa, and may suggest directional selection.

Distribution of alleles at HLA-A, C, B loci

34 and 20 HLA-A alleles, 50 and 30 HLA-B alleles, and 24 and 12 HLA-C alleles were observed in Ugandans and Zambians, respectively.

In Ugandans, the high frequency HLA-A alleles are A*0201, A*0101 and A*2301, with allele frequencies (a. f.) ranging from 0.184 to 0.0675. In Zambians, the high frequency HLA-A alleles are A*3002, A*3001, A*0201 and A*2301, with allele frequencies ranging from 0.2325 to 0.0814. More alleles are observed for HLA-B than for HLA-A and HLA-C; however, no highly predominant alleles are seen in either population for HLA-B. Ugandans exhibit many intermediate and low frequency HLA-B alleles; B*0702, B*0801 and B*1503 were the only alleles with an allele frequency above 0.05. In Zambians, the high frequency HLA-B alleles are B*4201, B*5301, B*1503, B*5101 and B*5703, with allele frequencies ranging from 0.1477 to 0.0568. The HLA-C alleles Cw*0202, Cw*0303, Cw*0304, Cw*0401, Cw*0602, Cw*0701, Cw*0702, Cw*0704, Cw*0802, Cw*1601, Cw*1701 and Cw*1801 were present in both Ugandans and Zambians; however, these alleles displayed contrasting frequencies. For example, Cw*0602 has a relatively low frequency in Zambians (a.f.=0.033) compared to the frequency in Ugandans (0.117). Cw*1601, Cw*1701 and Cw*1801 have high frequencies in Zambians (a.f.=0.10, 0.156, 0.10 respectively) and low frequencies in Ugandans (a.f.=0.049, 0.028 and 0.009 respectively). Novel alleles similar to A*2601, B*0702 were observed in two Ugandans. A novel allele similar to Cw*0304 was observed in both Ugandans and Zambians.

Hardy-Weinberg equilibrium and Ewens-Watterson homozygote test of neutrality

The distributions of heterozygous genotypes at each of the HLA class I loci did not deviate from the distributions expected under Hardy-Weinberg equilibrium in Ugandans. At

all loci, Ugandans displayed values of homozygosity lower than expected according to neutrality expectations from the Ewens-Watterson homozygote test of neutrality. In Zambians, homozygosity values at the HLA-B and HLA-C loci were below the neutrality expectations, while the value of homozygosity at HLA-A was identical to the neutral value predicted by the Ewens-Watterson model. The high value of homozygosity was caused by the high frequency of the A*3002 allele. A*3002 has a low frequency in other populations, and the increase in its frequency in Zambians may result from directional selection. The neutral distribution of HLA-A genotypes observed in the Zambian population may be due to the balance between directional selection and overdominant selection, which seems to act on many African HLA loci.

Haplotypic associations

These populations presented relatively few haplotypes overall; only six HLA-A-B haplotypes and nine HLA-C-B haplotypes were observed in both Ugandans and Zambians. All A-B haplotypes in Ugandans have low haplotype frequencies (h.f.) (h.f. <0.028), with the exception of A*0101-B*0801 (h.f. = 0.037). In Zambians, the A*3001-B*4201, A*3002-B*5301, A*3002-B*5703 and A*2902-B*4201 haplotypes display higher frequencies, ranging from 0.116 to 0.035. The

Cw*0702-B*0702, Cw*0701-B*0801, Cw*0202-B*1503, and Cw*0401-B*5301 haplotypes have high frequencies ranging from 0.056 to 0.04. HLA-B and C loci exhibited stronger associations than HLA-A and B loci or HLA-A and C loci. The most frequent alleles have multiple associations with alleles of the other loci; e.g., in Zambians A*3002 was associated with B*5301, B*5703 and B*8101.

The most common three-locus haplotypes in Ugandans were A*0101-Cw*0701-B*0801 and A*7401-Cw*0202-B*1503. A*3002-Cw*0401-B*5301, A*3002-Cw*1801-B*5703, and A*2902-Cw*1701-B*4201 had high frequencies in Zambians.

Genetic distance at HLA class I loci

Genetic distances were estimated according to the distributions of alleles at HLA class I loci and compared with other populations including Malians (data not shown), Kenyans (data not shown), Zulus from South Africa (2–3), Cameroonians (4), the Mandenka tribe from Senegal (5), and African Americans (6). Zambians had the closest distance with the Kenyan Luos at HLA-A and C loci, and with Zulus at HLA-B locus. The Ugandans appear closer to the Kenyans. Both Ugandans and Zambians displayed close genetic distance with African Americans.

References

1. K Cao, M Chopek, MA Fernández-Viña. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). *Reviews in Immunogenetics* **1**, 177–208, 1999.
2. D Middleton, F Williams, A Meenagh, AS Daar, C Gorodezky, M Hammond, E Nascimento, I Briceno, MP Perez. Analysis of the distribution of HLA-A alleles in populations from five continents. *Hum Immunol* **61**(10), 1048–1052, 2000.
3. F Williams, A Meenagh, C Darke, A Acosta, AS Daar, C Gorodezky, M Hammond, E Nascimento, D Middleton. Analysis of the distribution of HLA-B alleles in populations from five continents. **62**(6), 645–650, 2001.
4. JM Ellis, SJ Mack, RF Leke, I Quakyi, AH Johnson, CK Hurley. Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, *2612, *3006 and HLA-B*1403, *4016, *4703. *Tissue Antigens* **56**(4), 291–302, 2000.
5. A Sanchez-Mazas, QG Steiner, C Grundschober, JM Tiery. The molecular determination of HLA-Cw alleles in the Mandenka (West Africa) reveals a close genetic relationship between Africans and Europeans. *Tissue Antigens* **56**(4), 303–312, 2000.
6. K Cao, JA Hollenbach, XJ Shi, WX Shi, M Chopek, MA Fernández-Viña. Analysis of the frequencies of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* **62**, 1009–1030, 2001.

Population: Zulu from Natal Province, South Africa

Authors: MG Hammond¹, D Middleton² and D Anley¹

Affiliation: ¹Transplantation Unit, National Blood Transfusion Services Pinetown, South Africa, ²Northern Ireland Histocompatibility and Immunogenetics Laboratory, City Hospital, Belfast, Northern Ireland

The Zulu-speaking people belong to the Natal Nguni linguistic division of the central Bantu (Niger-Congo family). Natal was populated by a variety of tribes at the beginning of the 19th century. In 1815, Shaka took the Zulu throne and began to reorganize his army. A number of tribes fled from the area, initiating a period of intense inter-tribal warfare in the interi-

or of Southern Africa. The Ndebele of Zimbabwe and the Shangana of the Eastern Transvaal lowveld are the descendants of the people who fled from Natal during this period under the leadership of Umzilikazi and Soshangana respectively. Today there are more than 6 million Zulus in the province of Natal.

Samples from 955 random, unrelated Zulu blood donors were typed using several methods (11th Workshop DNA reagents, IHWG Reverse Line Strip reagents, PCR-DNA) at the HLA-A, B, C, DRB1, DQA1, DQB1, and DPB1 loci. Of these, high-resolution genotypes for 290 donors were analyzed by the IHWG Biostatistics Core. The allele frequencies observed in this dataset are consistent with those observed in previous studies of this population.

Overall deviations from Hardy Weinberg equilibrium proportions were observed at the DQA1 (p-value=0.0001) and DPB1 (p-value: 0.0140) loci. While the overall genotype distributions at other loci were consistent with Hardy-Weinberg expectations, an excess of homozygotes was observed at the DRB1 locus (p-value=0.0350), and a deficiency of homozygotes was observed at the HLA-A locus (p-value=0.0444).

North Africa

Populations: Moroccan from Souss, Morocco and Algerian from Oran, Algeria

Authors: H. Izaabel¹, S. Djoulah^{2,3}, H.-J. Garchon⁴, S. Caillat-Zucman⁴, G. Beaurain⁴, O. Akhayat¹, J.-F. Bach⁴, I. Khalil², S. Benhamamouch³, L. Degos², I. Deschamps⁴, J. Hors², A. Sanchez-Mazas⁵

Authors' addresses: ¹Laboratory of Cellular and Molecular Biology, University Ibnou-Zohr, Agadir, Morocco, ²INSERM U93, Hôpital St Louis, Paris, France, ³Institute of Biology, Molecular Biology and Laboratory of Genetics, Es-Senia University, Oran, Algeria, ⁴Laboratory INSERM U25 and U30, Hôpital Necker, Paris, France, ⁵Laboratory of Anthropology, Genetics and Peopling History, University of Geneva, Switzerland

The present North African population is constituted both by Berber ("Tamazight") and Arab-speaking populations, who speak languages belonging to the Semitic branch of the Afro-Asiatic family (1). The origin of Berbers is still unknown. Most sources refer to these populations as the descendants of an autochthonous Capsian civilization living in North Africa in the Mesolithic period, which further intermixed with Neolithic people (2). The region was further settled by several Mediterranean populations (e.g., Phoenicians) some 3,000 years ago. Other important events were the Roman conquest in 146 BC and the establishment of Jewish communities dur-

ing the 1st century. Finally, Arab invasions in the 7th and 11th centuries also contributed to the present North African Berber genetic pool, with the exception of populations in some very remote geographic areas.

Two different datasets are considered in the present report. The first dataset is a sample of 98 Moroccan individuals living in the Souss area near Agadir, in southwestern Morocco (30 degrees 26 minutes north latitude, 9 degrees 36 minutes west longitude), and is likely to be of Berber origin. This sample was previously tested by PCR-SSOP for DRB1, DQA1, and DQB1 polymorphisms, and compared to other North African and European populations (3, 4). The most remarkable result was to find a higher level of homozygosity (i.e. a lower level of genetic diversity) in the Souss Moroccans compared to surrounding North African and European populations, as only three DRB1-DQA1-DQB1 haplotypes (0701-0201-0201, 0301-0501-0201, and 11-0501-0301) accounted for about 45% of the total class II haplotype frequencies (20–21%, 17–18%, and 7–8%, respectively). Close genetic relationships were also found with southwestern Europeans (e.g., Spanish and Basques). Both our previous results and those obtained by the IHWG Biostatistics core (5) agree in detecting a significant deviation (5% level) from Hardy-Weinberg (HW) equilibrium at the DRB1 and DQB1 loci, but not at DQA1. On the other hand, no deviation from selective neutrality was detected using the Ewens-Watterson (EW) test.

The second dataset is a sample of 99 individuals from an Algerian population of the Oran area in northwestern Algeria (35 degrees 40 minutes north latitude, 0 degrees 40 minutes west longitude). DRB1 and DQB1 loci were previously tested by PCR-SSOP (DQA1 was partially tested as well), first in a sub-sample of 47 individuals (6) and then in a larger sample (7), and the data were submitted to the 12th IHWC (8) (12th IHWC Map number #235). As in the Souss Moroccans, the DRB1*0301 (15%), *0701 (14%), *1101 (9%), DQA1*0501 (24%), and DQB1*0201 (30-32%), *0301 (19%) alleles are the most frequent. However, the level of diversity is much higher in the Algerian population, whose genetic profile is close to that observed in Tunisians. There is an apparent deviation (5% level) from Hardy-Weinberg equilibrium at the DQA1 locus, but no deviation from selective neutrality according to the EW test.

The present brief report shows that these two North African populations share a high level of HLA genetic similarity, which can be attributed to a common history and/or gene flow through neighboring populations found in the 1,000 kilometer region between southern Morocco and western Algeria. Nonetheless, the results also underscore a relevant dif-

ference in the level of internal genetic diversity of each population, the Moroccan from Souss being much more homozygous for HLA class II than the Algerian from Oran. This suggests that the Souss population has been more isolated than the Oranese (which may also explain its significant HW deviation), in agreement with their remote location south of the Atlas. Due to very heterogeneous geographic locations in

coastal, semi-desert, desert or mountainous areas, and to different historical backgrounds and degrees of gene flow with successive immigrant populations, North African populations are apparently very diverse genetically, even among different Berber populations (9). As HLA data on North African populations are still scarce, this region would be worth investigating in future population genetic studies.

References

1. Ruhlen M. A guide to world language Vol. 1 Classification. Edward Arnold London, 1987.
2. Murdock JP. Africa. Its people and their culture history. New York: McGraw-Hill Book Company, 1959.
3. Izaabel H, Garchon H-J, Caillat-Zucman S, Beaurin G, Akhayat O, Bach J-F, Sanchez-Mazas A. HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* **51**, 106–110, 1998.
4. Hors J, et al. 1997. HLA in North African populations: 12th International Histocompatibility Workshop : NAFL report. In: Charron D (ed) Genetic diversity of HLA: functional and medical implications, vol. 1. Sèvres: EDK, p. 328–334.
5. Results of data analysis performed on the “CHESAN-Moroccan population” by the IHWC Biostatistics core.
6. Djoulah S, Sanchez-Mazas A, Khalil L, Benhamamouch S, Degos L, Deschamps I, Hors J. HLA-DRB1, DQA1 and DQB1 DNA polymorphisms in healthy Algerian and genetic relationships with other populations. *Tissue Antigens* **43**, 102–109, 1994.
7. Benhamamouch S, Boudjemaa A, Djoulah S, Le Monnier de Gouville I, Bessaoud K, Hors J, Sanchez-Mazas A. Anthropological study of west Algerian population. In: Charron D (ed) Genetic diversity of HLA: functional and medical implications, vol. 2. Sèvres: EDK, p. 181–182, 1997.
8. Bodmer J, Cambon-Thomsen A, Hors J, Piazza A, Sanchez-Mazas A. Anthropology report: introduction. In: Charron D (ed) Genetic diversity of HLA: functional and medical implications, vol. 1. Sèvres: EDK, p. 269–284, 1997.
9. Sanchez-Mazas A. The Berbers of North Africa: genetic relationships according to HLA and other polymorphisms. In: Arnaiz-Vilena A (ed) Prehistoric Iberia: Genetics, Anthropology, and Linguistics. New York: Kluwer Academic Press / Plenum Publishers, p. 65–77, 2000.

Population: Chaouya from Morocco

Authors: Angelica Canossi¹, Daniela Piancatelli¹, Khadija Oumhani², Anna Aureli¹, Marilena Di Rocco¹, Tiziana Del Beato¹, Gabriella Liberatore¹, Carlo Umberto Casciani¹, Rajae El Aouad² and Domenico Adorno¹

Authors' addresses: ¹Istituto C.N.R. Trapianti D'Organo e l'Immunocitologia, L'Aquila, Italy, ²Immunology Laboratory, Institut National D'Hygiene, Rabat, Morocco

The people of Morocco speak either Arabic or the Berber language and are, in some places, bilingual. Arabic speakers outnumber Berber speakers, and live mainly in Northern Morocco and in the South on the Atlantic Coast (1). Chaouya (CH) is the name given to Berber and Arabic-speaking ethnic groups present in different regions of North Africa, including Algeria. The Chaouya population sample reported here is an Arabic-speaking population from West Morocco (the Atlantic coast), living in the Settat area. This population derives from the Zenata (8th century, Berber speakers) and Houwara (13th century, Berbers that were bilingual) populations that came

from Africa (especially Lybia) and spread throughout North Africa. Despite their common Berber origins, some Chaouya populations of the North Africa have remained Berber-speaking, others became bilingual, and others became totally Arabic-speaking, as the Moroccan Chaouya.

In order to provide genetic and anthropological information with which to clarify the historical origins of this population, we collected blood samples from 99 healthy unrelated Chaouya volunteers in collaboration with the Institut National D'Hygiene of Rabat. Both parents and the grandfathers of all subjects belonged to the same ethnic group.

We present here a high-resolution study of Chaouya polymorphisms at the HLA-A, B and DRB1 loci. HLA class I (HLA-A and HLA-B, exons 2,3,4) and class II (HLA-DRB1, exon 2) allele polymorphisms were determined by sequence-based typing (SBT) using Big Dye terminator chemistry and fluorescently labeled primers (Applied Biosystems), respectively.

Until now, studies of HLA in Arabic-speaking Moroccans have been limited to one ethnic group from El Jadida city (2), Arabic speakers of Berber origins, and recently, one group from Casablanca (3). Our group has previously investi-

gated HLA-DRB1 polymorphisms in the Chaouya and Berber-speaking Metalsa (ME) groups (4).

This population was in Hardy-Weinberg equilibrium for all the HLA loci studied. The normalized deviate of homozygosity (F_{nd}) was negative for all loci, and was significantly diverged from the neutral value for the HLA-B locus ($F_{nd} = -2.640$, $p = 0.0214$), suggesting the operation of balancing selection (heterozygote advantage). Gene diversity was elevated (A locus: 0.92, B locus: 0.97, DRB1 locus: 0.94), as described in many African populations.

A total of 28 alleles were found at the DRB1 locus, the most frequent (present in at least 10% of the population) being DRB1*0701 (allele frequency (a.f.) = 16.16%), DRB1*15011 (a.f. = 12.62%) and the DRB1*03011 allele (a.f. = 11.61%) (4).

By comparing Chaouya HLA-DRB1 polymorphism with that of other Moroccan populations, we identified some specific variants (DRB1*0101, *03021, *04011, *1103, and *08032) as well as some uncommon alleles (DRB1*03021, *1503, *1406, and *0806) in the Chaouya. Only the DRB1*03011 allele frequency differed significantly between the Chaouya and Metalsa (CH a.f.% = 11.7; ME a.f.% = 20.2, $p = 0.0311$). These differences could be indicative of the distinct origins of the Berber and Arabic-speaking groups. As described in our genetic distance analysis (4), the Chaouya are closely related to Algerians, and then to the Metalsa, El Jadida and Souss groups, and to a lesser degree with Iberians, French and Ethiopians, confirming the historically documented common origin of these north African populations (4).

Of the 27 alleles identified at the HLA-A locus, the most frequent were A*02011 (a.f. = 17.9%), A*01011 (a.f. = 12.7%) and A*3001 (a.f. = 7.5%). Several alleles unique to Africans (A*0102, *0202, *3001, *3002, *3402, *6802, *7401, *8001) (5–9), as well as some variants previously observed at high frequencies in Spaniards (A*0101, 3001) and Orientals (A*2402, 2601) (5, 10) were observed. The

novel A*3010 allele was observed in the Chaouya and submitted to GeneBank (January 2001, AF323494-6). The frequencies of the A*3001 (7.5% vs 1.4%, $p = 0.028$) and A*3301 (5.6% vs 0.7%, $p = 0.05$) alleles differed significantly from the Metalsa, and differences in the A*29 (6.7% vs 2.0%, $p = 0.057$) and A*33 (2.1% vs 7.8%, $p = 0.047$) allele groups were observed in comparison with Casablanca.

Of the 40 observed B locus alleles, the most frequently observed were B*5001 (a.f. = 8.8%) and B*4403 (a.f. = 7.3%). Several alleles previously observed at high frequencies in Africa were recognized, (e.g., B*1503, *1510, *3508, *3910, *57031, *5301, *4102, *4201, *1403, and *5801) (5, 6, 11, 12, 13). We also observed alleles frequently present in Spaniards (B*0801), Mediterraneans (B*0801, 4403, 4901) (14) and Asians (B*5201, *4001), and defined the previously unknown B*1403 exon 4 sequence (November 2000, AF279664). Chaouya HLA-B polymorphism differed significantly from Casablanca for the B*45 (6.6% vs 0.45%, $p = 0.003$) group, and from El Jadida for the B*44 (11.0% vs 5.3%, $p = 0.08$) and the B*5(51,52) allele groups (5.8% vs 11.1%, $p = 0.05$).

The most common HLA-A-B-DRB1 haplotypes observed in the Chaouya were A*0201-B*5101-DRB1*1501 and A*0201-B*5001-DRB1*0701, both occurring at a frequency of 3.03%. In addition to these, 11 other haplotypes appeared more than once.

The Chaouya shared some of the most common haplotypes (A*0201-B*4402-DRB1*0402 and A*2402-B*0801-DRB1*0301) with the Metalsa and a mixed Moroccan population from Casablanca (3) (A2-B50 and A68-B44), demonstrating common origins of these Moroccan Berber and Arabic-speaking groups. Nevertheless, this analysis highlights some unique genetic characteristics of Chaouya within Morocco and has demonstrated the influence of intermingling with various populations (4).

References

1. Grimes BF (eds): *Ethnologue*, 13th edition. Summer Institute of Linguistic Inc., 1996. Web site: www.sil.org/ethnologue/countries/Moro.html#MEY.
2. Gomez-Casado E, Del Moral P, Martinez-Laso J, et al. HLA genes in Arabic-speaking Moroccans: close relatedness to Berbers and Iberians. *Tissue Antigens* **55**, 239–249, 2000.
3. Choukri F, Chakib A, Himmich H, Raissi H, Caillat-Zucman S. HLA class I polymorphism in a Moroccan population from Casablanca. *Eur J Immunogen* **29**(3), 205–211, 2002.
4. Oumhani K, Canossi A, Piancatelli D, et al. Sequence-based analysis of HLA-DRB1 polymorphism in Metalsa Berber and Chaouya Arabic-speaking groups from Morocco. *Hum Immunol* **63**, 129–138, 2002.
5. Geer L, Terasaki PI, Gjertson DW. Gene frequencies and world maps of their distributions. Chapter 7. In: Gjertson DW and Terasaki PI. *HLA 1998*, UCLA Tissue Typing Laboratory, 1998: 348.
6. Ellis JM, Mack SJ, Leke RFG. Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, 2612, 3006 and HLA-B*1403, *4016, *4703. *Tissue Antigens* **56**, 291–302, 2000.

7. Modiano D, Luoni G, Petrarca V, et al. HLA class I in three West African ethnic groups: genetic distances from sub-Saharan and Caucasoid populations. *Tissue Antigens* **57**, 128–137, 2001.
8. Cao K, Hollenbach J, Shi X, et al. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five mayor ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* **62**, 1009–1030, 2001.
9. De Pablo R, Garcia-Pacheco JS, Vilches C, et al. Bubi Normal In: Gjertson DW and Terasaki PI. HLA 1998, UCLA Tissue Typing Laboratory, 1998: 104–105.
10. Kashiwase K, Tanaka H, Ishikawa Y, et al. Japanese normal In: Gjertson DW and Terasaki PI. HLA 1998, UCLA Tissue Typing Laboratory, 1998: 207.
11. Arnaiz-Villena A, Benmar D, Alvarez M, et al. HLA allele and haplotype frequencies in Algerians. relatedness to Spaniard and Basques. *Hum Immunology* **43**, 259–268, 1995.
12. Andrien M, Defleur V, Vandercruys M, et al. Senegalese Normal. In: Terasaki PI and Gjertson DW. HLA 1997, UCLA Tissue Typing Laboratory, 1997: 178.
13. Gorodezky C, Alaez C, Vazquez-Garcia N, et al. The genetic structure of Mexican Mestizos of different locations: tracking back their origins through MHC genes, blood group systems, and microsatellites. *Hum Immunol* **62**, 979–991, 2001.
14. Gebuhrer L, Betuel H. Caucasian French Normal. In: Gjertson DW and Terasaki PI. HLA 1998, UCLA Tissue Typing Laboratory, 1998: 143.

Population: Metalsa from Morocco

Authors: D. Piantatelli¹, A. Canossi¹, K. Oumhani², A. Aureli¹, M. Di Rocco¹, T. Del Beato¹, G. Liberatore¹, C. U. Casciani¹, R. El Aouad², D. Adorno¹

Authors' addresses: ¹CNR Istituto Trapianti d'Organo e l'Immunocitologia, L'Aquila, Italy; ²Laboratoire d'Immunologie, Institut National D'Hygiene, Rabat, Morocco

Class I and Class II HLA polymorphism was examined in the Metalsa (ME) population, a Berber population from north Morocco. The Berber people have an ancient origin and represent the autochthonous basis of the modern Moroccan population. An oriental origin for the Berbers has been suggested (1), but Morocco has experienced so many colonization and immigration events during its history (2) that the modern Moroccan people demonstrate great heterogeneity. The people of Morocco speak either Arabic or the Berber language and are, in some places, bilingual.

Moroccan Berbers are represented by three main communities (3). The Metalsa belong to the Berber-speaking group from north Morocco (Rif mountains, Nador area), and speak the Tarifit language. Until now, few HLA studies have focused on Moroccans. HLA-DRB1 polymorphism in Metalsa Berbers and in Chaouya Arabic-speakers from Morocco was discussed by our group (3). Other studies have included a Souss Berber population from the Agadir area (4), a population sample from the El Jadida area (Arabic-speakers of Berber origins) (5), and recently class I polymorphism was examined in a mixed population of Arabic and Berber speakers from Casablanca (6).

To genetically characterize this Berber population, 100 unrelated individuals were collected in collaboration with the Immunology Laboratory of the Institut National D'Hygiene of Rabat and HLA typed. These samples were not collected in

the context of a disease study. All donors were born in and living in the same area, and both parents and grandfathers belonged to the Metalsa ethnic group. The polymorphism of the HLA-A, B, and DRB1 loci was investigated using a sequence-based typing (SBT) system. SBT of the HLA-A and -B loci was performed on a 2-kb DNA fragment. This fragment was sequenced with primers specific for exons 2, 3 and 4 (7), using dye terminator chemistry (HLA-A and HLA-B Sequence-Based Typing Kits, PE, Applied Biosystems). After a preliminary SSP-PCR screening of the major DRB1 allele groups, DRB1 SBT was carried out using fluorescently labeled primers (8).

The Metalsa population was in Hardy-Weinberg equilibrium for all the studied loci. The normalized deviate of homozygosity (Fnd) was negative for all loci and was significantly diverged from the neutral value for the A locus (Fnd = -1.1103), suggesting that balancing selection was acting. Gene diversity was elevated, as it is in African populations (A locus, 0.92; B locus, 0.95; DRB1 locus, 0.91).

Twenty-one HLA-A alleles were present. The most frequently detected alleles were A*0201 (allele frequency (a.f.)=0.1736) and A*0101 (a.f.=0.1458), which are common in many populations, especially Caucasians. The frequencies of A*3002 (a.f.=0.0903) and A*2301 (a.f.=0.0833) were typical of north and sub-Saharan Africa. Comparison with other Moroccan populations demonstrated distinct differences in Metalsa and Chaouya A*30 alleles, with a significant difference for A*3001 allele frequency (p=0.028). No differences from the El Jadida and Casablanca groups were observed.

Of the 34 HLA-B alleles observed, the high frequency of B*44 alleles is notable in comparison with other Caucasian (European and north American) populations, as well as with other Moroccan groups (El Jadida, p=0,0002; Casablanca, p=0,0220; Chaouya, p=0.0320). An unexpectedly higher

frequency of B*4403 with respect to B*4402 (B*4403, a.f.=0.1250, B*4402 a.f.=0.0882) was found, as has been seen in Spaniards (9). B*0801 (a.f.=0.0882 in ME), an allele previously observed at high frequencies in the Mediterranean area, and B*50 (B*5001, a.f.=0.0735; B*5002, a.f.=0.04412) were also frequent. A novel HLA-B allele (Genbank AF279663) was identified in one individual, and was officially designated as B*570302 (June 2000, HWS100000605).

Of the 27 alleles found at the DRB1 locus, DRB1*03011 ((a.f.)=0.2020), DRB1*0701 (a.f.=0.1212), and DRB1*1302 (a.f.=0.1111) were the most frequent (3). The frequency of DRB1*03011 differed significantly between the Metalsa and Chaouya populations, and some alleles found in the Metalsa were absent in other Moroccan populations (e.g., DRB1*1304, absent in all the other groups; DRB1*1502, absent in Souss Berbers and Chaouya; DRB1*1201 and *08041, absent in Chaouya; DRB1*1406, absent in El Jadida group and Souss Berbers; and DRB1*1503, absent in Souss Berbers).

Overall, several alleles previously observed to be common in African populations (A*3402, *6802, and *7401, B*1503,

*4102, *4202, *7801, and *5802, and DRB1*01021, *0806, and *1001) were observed in the Metalsa, along with alleles observed less commonly in African populations (A*3004, B*2702, *2703, *5001/02, and *3503, and DRB1*1503, *1406, *0806, and *1304)(10). Genetic distances calculated for DRB1, demonstrated a close relationship between the Metalsa, El Jadida and Chaouya groups, followed by Algerians and Tunisians, and the Souss Berbers to a lesser extent (3).

The predominant A-B-DRB1 haplotypes in the Metalsa (frequencies >3%) were A*0101-B*4402-DRB1*0402, A*2301-B*4403-DRB1*0405, A*2402-B*0801-DRB1*0301. This last haplotype was shared with the Chaouya group. This study demonstrates the close relatedness of the Metalsa to the other Moroccan and north African populations as well as the common genetic characteristics of Berber and Arabic-speaking groups, and highlights some unique characteristics of the Metalsa population in comparison to other ethnic groups within Morocco, demonstrating the influence of immigration over several centuries.

References

1. Kossowitch N: Anthropologie et groupes sanguins des populations du Maroc. Ed. Masson et Cie, 1953.
2. Julien CA. Histoire de l'Afrique du Nord. Paris, Payot, 1961.
3. Oumhani K, Canossi A, Piancatelli D, et al. Sequence-based analysis of the HLA-DRB1 polymorphism in Metalsa Berber and Chaouya Arabic-speaking groups from Morocco. *Hum Immunol* **63**, 129–138, 2002.
4. Izaabel H, Garchon HJ, Caillat-Zucman S, et al. HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* **51**, 106–110, 1998.
5. Gomez-Casado E, Del Moral P, Martinez-Laso J, et al. HLA genes in Arabic-speaking Moroccans: close relatedness to Berbers and Iberians. *Tissue Antigens* **55**, 239–249, 2000.
6. Choukri F, Chakib A, Himmich H, et al. HLA class I polymorphism in a Moroccan population from Casablanca. *Eur J Immunogenet* **29**, 205, 2002.
7. Johnston-Dow L, Bengston A, Scheltinga S, et al. DNA sequencing based typing of HLA-A. In: Charron D, ed. HLA-Genetic Diversity of HLA – Functional and Medical Implication. Paris, France: EDK Medical and Scientific International Publisher 250–254, 1997.
8. Fernandez-Vina M, Bignon JD. Primers and oligonucleotide probes (SSOP) used for DNA typing of HLA class II alleles. In Charron D (ed.): Genetic diversity of HLA. Functional and medical implications. Paris, EDK, 1997.
9. Gallardo D, Arostegui JJ, Rodriguez-Luaces M, et al. HLA-B44 subtyping in the Catalan population using reference strand mediated conformation analysis. Implications for the selection of unrelated bone marrow donors. *Tissue Antigens* **56**, 173–177, 2000.
10. Gjertson DW, Terasaki PI (eds.) HLA 98. Los Angeles, CA: UCLA Tissue Typing Laboratory, 1998.

Population: Moroccan from El Jadida, Morocco

Authors: Arnaiz-Villena A, Moscoso J, Martínez-Laso J, Zamora J, Lowy E, del Moral P and Gómez-Casado E.

Authors' address: Dept. Immunology and Molecular Biology, H. 12 de Octubre, Universidad Complutense, 28041-Madrid, Spain

It has been shown that most Moroccan and North African people are of ancient Berber origin and that the genetic HLA

profile of urban and country Moroccan populations is similar whether they speak only Arab (imposed by the 7th century AD), Berber, or Berber-Arab. (1–3). The Berber language is spoken by approximately 40% of the Moroccan population, particularly by the people living in the non-urban and Atlas Mountains areas (3).

The Moroccan population sample of 96 unrelated healthy individuals was collected from El Jadida with all four grandparents also coming from this region. El Jadida is an immigra-

tion center. People from the surrounding areas have come to live in the city, as a result of recent industrial development in the last part of the 20th century. Most people coming from country areas spoke Berber at home as well as Arab learned in school.

High resolution HLA class II typing (DRB1, DQB1) was performed by a reverse dot-blot technique with the Automated Innolipa system (Innogenetics, Belgium). HLA-DQA1 alleles were characterised using 11th and 12th International Histocompatibility Workshop and local reagents, as well as PCR-SSO. These and statistical methodologies are detailed in (3).

HLA-A and B allele frequency distributions and estimated HLA-A-B haplotypes in the El Jadida population were first presented in a study in which an additional 29 Mediterranean and Sub-Saharan African populations (Moroccans, Moroccan Jews, Algerians, Berbers (Souss), Egyptians, Ashkenazi Jews, Non Ashkenazi Jews, Lebanese-NS, Lebanese-KZ, Iranians, Turks, Armenians, Greeks (from Cyprus, Attica and Aegean), Cretans, Macedonians, Sardinians, Italians, Spaniards, Span-

ish-Basques, Portuguese and French; Mossi, Fulani, Rimaibe, Nubians, Oromo and Ahmara) were compared with this population in order to determine genetic relatedness (1). The main conclusions from this comparison of the HLA genetic profiles of Arab-speaking Moroccans with those of other Mediterranean populations are:

(i) Northern (Iberians) and southwestern Mediterraneans (North Africans) show a similar HLA genetic background. Iberians, including Basques, are HLA related to North African Berbers. Present day Arab-speaking Moroccan and Algerian urban and country people show an indistinguishable HLA profile (1).

(ii) A gene flow is noticed among all Mediterraneans, which probably occurred in a long-lasting circum-Mediterranean contact, which included cultural and language exchanges during glacier and inter glacier periods (4).

Acknowledgements

This work was supported by grants from the Spanish Ministry of Science PM-1999-23 and BMC-2001-1299.

References

1. Arnaiz-Villena A, Gómez-Casado E and Martínez-Laso J. Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens* **60**, 111–121, (2002).
2. Izaabel H, Garchon HJ, Caillat-Zucman S, et al. HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* **51**, 106–110, (1998).
3. Gómez-Casado E, del Moral P, Martínez-Laso J, et al. HLA genes in Arabic-speaking Moroccans: close relatedness to Berbers and Iberians. *Tissue Antigens* **55**, 239–249, (2000).
4. Arnaiz-Villena A, Martínez-Laso J, Alonso-García J. Iberia: population genetics, anthropology and linguistics. *Human Biol* **71**, 725–743, (1999).

Europe

Population: Croatian from Zagreb, Croatia

Authors: E. Cecuk-Jelicic, Z. Grubic, V. Kerhin-Brkljacic, K. Stingl, A. Kastelan

Authors' address: National Referral Organ Transplantation and Tissue Typing Centre, University Hospital Zagreb, Zagreb, Croatia

HLA class I polymorphism in the Croatian population was previously studied as part of the Anthropology Component of the 11th International Histocompatibility Workshop. In that study, HLA-B44, B18, B8, B35, and B51 were five of the most frequent antigens among Croatians (1). The aim of the 13th IHWG study was to establish molecular level HLA-A and B allele diversity and haplotype associations, as well as to

compare our results with the results from other populations so far tested.

The Croatian population sample consisted of 150 unrelated healthy individuals, all inhabitants of the Croatian capital city of Zagreb (15 degrees 53 minutes 42 seconds north latitude, 45 degrees 47 minutes 4 seconds east longitude) and surrounding areas. These sampled individuals originated from different regions of Croatia and should be considered a representative sample of the general population. DNA was obtained from peripheral blood leukocytes by a standard salting-out protocol (2). All samples were amplified using PCR and tested with the IHWG Reverse Line Strip method. The allele and genotype frequencies were estimated by the IHWG Biostatistics Core.

No deviation from Hardy-Weinberg equilibrium was detected for the HLA-A and B loci. Twenty-seven HLA-A allele-

es were found in our population, the most frequent of which were A*0201 at 26.3%, A*2402 at 16.0%, A*0301 at 11.0%. The frequency of all other HLA-A alleles identified among Croatians was less than 10%. The frequencies of HLA-A alleles were in good agreement with our previously reported serological data for this population, with the exception that A*0101 was observed at a lower frequency in this sample (9.7% vs 12.2%).

In contrast with the HLA-A locus, 44 distinct HLA-B alleles were detected in this study, but only six were present with a frequency greater than 5%; B*5101 (12.0%), *1801 (11.3%), *0702 (9.7%), *3801 (6.7%), *4402 (6.0%) and *1501 (5.7%). When compared with serological data, the unusually low frequency of the group of B*08 alleles (allele frequency (a.f.)=3.7%) was not statistically significant. In this group of alleles, B*0801 was the most frequent (1.7%), but we also observed B*0804 (1.3%) and B*0805 (0.7%). Among B*35 alleles, B*3501 predominated with a frequency of 4.3%, followed by B*3503 (3.0%), B*3502 (2.7%) and B*3520 (1.3%). Two other alleles for the same group, B*3508 and B*3521, were observed only once. Comparison

of HLA-A and -B frequencies between Croatians and other Europeans did not reveal significant differences (3). The frequency of blank alleles was 0, which indicated that this method is very precise and very useful in solving problems with alleles that could not be serologically typed.

On the basis of the polymorphism observed at the HLA-A and B loci in our population, 1188 different HLA-A-B haplotypes were possible. One hundred thirty-eight different combinations were determined from this sample of 300 chromosomes. The HLA-A*0301-B*0702 haplotype (which has been observed previously at high frequencies in Europe) was the most frequent haplotype (5.3%) while A*0201-B*5101 (4.8%) was the second most frequent haplotype. The surprisingly low frequency of the A*0101-B*0801 (1.3%) haplotype is probably due to low individual frequencies of these two alleles. In this study, we also focused on the analysis of A9-B12 haplotypes because of their relatively high frequency in our population. Interestingly, B*4402 allele was always found in combination with A*2402 (2.1%) and never with A*2301, while B*4403 was always associated with the other A9 detected subtype, A*2301 (1.3%).

References

1. Cuccia M, Astolfi P, Gyodi E, et al 1992 HLA in five populations of southern Europe. In: Tsuji K, Aizawa M, Sasazuki T (ed) HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference, Vol. 1, Oxford University Press, Oxford, 655–656.
2. Miller A, Dykes D, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**, 1215–1217, 1988.
3. Imanishi T, Akaza T, Kimura A, et al. 1992 Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T, ed. HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference, Vol. 1, Oxford University Press, Oxford, 1065–1220.

Population: Czech from Prague, Czech Republic

Authors: E. Ivaskova and N. Bendukidze

Authors' address: Institute for Clinical and Experimental Medicine (IKEM), Prague, Czech Republic

Czechs make up nine-tenths of the population of the Czech Republic, a 78,900 square km nation situated in the approximate geographic center of Europe. According to the 1998 census, the Czech Republic has a population of 10.3 million people (5 million males and 5.3 million females). Three quarters of the population live in urban areas. The West Slavic language is spoken in the historical regions of Bohemia, Mo-

ravia, and Silesia, all of which are now part of the Czech Republic, as well as in émigré communities, including perhaps close to a million speakers in North America.

According to historical evidence, the Celts were the first ethnic group to occupy the present-day Czech Republic, having arrived around the 4th century BC. The Celtic Boii tribe gave the country its Latin name: Boiohaemum (Bohemia). They were pushed out by the German tribes (Marcomanni, Quidi) before the beginning of our era. At the end of 5th and the beginning of 6th century, Slavs settled in the territory of Bohemia and Moravia during the period known as the "Migration of Peoples". The first half of the 7th century marked the first successful attempt to unite the Slavonic tribes.

The Moravian Kingdom was formed early in the ninth century, and rapidly expanded to include Bohemia, Slovakia, southern Poland, and western Hungary. This expanded kingdom, known as the Great Moravian Empire, was situated at the crossroad of the German lands in the west and Byzantium in the east. In this period Slavonic (the Cyrillic alphabet) script was created by monks Cyril and Methodius and used in Great Moravian Empire. The Franks (a Germanic people) conducted destructive raids into Moravian territory, and German priests and monks came to spread Christianity in its Roman form among the Slavs. German and Byzantine influences contended until 885 AD, when the Great Moravian Empire was drawn into the sphere of influence of the Roman Catholic Church. As a result, the Czechs and Slovaks adopted the Latin alphabet and became further differentiated from the Eastern Slavs, who continued to use the Cyrillic alphabet and adhere to Eastern Orthodoxy (1).

DNA samples from 106 unrelated individuals were collected in Prague and typed for HLA-A, B, Cw, DRB1, DQA1, DQB1 and DPB1 loci. A, B and C loci were typed using IHWG RLS, 12th Workshop SSP (2, 3, 4, 5) and SSP commercial reagents. DRB1, DQA1, DQB1 DPB1 were typed using SSOP and SSP reagents (5, 6).

The frequencies of the most common alleles in the Czech population are A*0201, 0.28; B*0702, 0.13; Cw*0602, 0.16; DRB1*0701, 0.17; DQA1*0501, 0.27; DQB1*0301, 0.21; and DPB1*0401, 0.40.

No locus-wide deviations from expected Hardy-Weinberg

equilibrium proportions were detected. Non-significant positive values of the normalized deviate of homozygosity (F_{nd}) were observed for the HLA-A ($F_{nd}=0.07$) and DPB1 ($F_{nd}=0.79$) loci, resulting from higher observed homozygosity values than were expected under neutrality, a trend in the direction of positive selection. We suppose that this increased homozygosity could result from the non-significant excess of homozygous DPB1*0401 genotypes (p -value=0.33). Significant negative values of the normalized deviate calculated for the HLA-B and DQB1 loci reflected low values of observed homozygosity. Expected homozygosity based on neutrality assumptions was much higher than observed homozygosity, consistent with the operation of balancing selection for these loci. F_{nd} values at Cw and DRB1 loci were consistent with the operation of neutral drift.

HLA-A*0301 -B*0702 - DRB1*1501 (haplotype frequency (h.f.)=0.04) and HLA-A*0101 - -B*0801 - DRB1*0301 (h.f.=0.03) were the most frequent A-B-DR haplotypes. HLA-A*0301 - -Cw*0702 - -B*0702 (h.f.=0.06) was the most frequent A-Cw-B haplotype. DRB1*0701 - DQA1*0201 - DQB1*0201 (h.f.=0.12), DRB1*0701 - DQA1*0201 - DPB1*0401 (h.f.=0.08) and DRB1*1501 - DQB1*0602 - DPB1*0401 (h.f.=0.07) were the most frequent 3 loci class II haplotypes in Czech population.

Acknowledgements

This research was supplied by a grant of the Ministry of Health of the Czech Republic NM/6522-3.

References

1. Britannica Online. <http://www.eb.com/>.
2. Bendukidze N, Meschishvili E, Ivaskova E, Cerna M, Churadze T, Guladze R, Shalamberidze K. Determination of HLA Class I Alleles in Georgian, Kurdish and Czech Populations by PCR-ARMS. *Human Immunology* **47**(1–2), 54, 1996.
3. Tilikainen A, Fischer G, Grubic Z, Gyödi E, Ivašková E. et al. Anthropological features of the East European region. Genetic diversity of HLA. Functional and Medical Implications 1997, p. 307. Editor: D.Charron.
4. Ivašková E, Bendukidze N, Kupková L, Sajdlová H, Hana I. Czech Normal. HLA 1998, Editors: Paul I. Terasaki, David M. Gjertson (Lenexa, Kansas, USA) p. 126–128L.
5. Bendukidze N, Ivaskova E, Zahlavova L, Hana I, Kupkova L, Sajdlova H, Slavcev A. HLA-C typing by PCR-SSP and haplotype analysis in Czech population. *Human Immunology* **Vol. 61 (suppl. 1)**, S102, 2000.
6. Cerna M, Fernandez-Vina M, Ivaskova E, Stastny P. Comparison of HLA class II alleles in Gypsy and Czech populations by DNA typing with oligonucleotide probes. *Tissue Antigens* **39**, 111–116, 1992.

Population: Finnish from Finland

Authors: Marja-Liisa Lokki^{1,2} and Elina Honkavaara¹

Authors' addresses: ¹FRC Blood Transfusion Service, and ²Transplantation Laboratory, University of Helsinki

Situated between Sweden, Norway and Russia, Finland spans 59 to 70 degrees north latitude but has a milder climate than expected for a northern nation. The population is unevenly distributed, with 13% living in the northern part of the coun-

try while a third lives in Finland's six largest cities. Along with Estonian, Saame, and Hungarian, Finnish belongs to the Uralic language family. More than 90% of Finland's population speaks Finnish as their mother tongue, with Swedish and Saame spoken by native minority populations.

Archeological studies in Finland have shown evidence of 4000 years of continuous habitation, and suggest that the area was first inhabited by eastern Uralic speakers. While Indo-European speakers are thought to have inhabited Finland's southern and western coastlines for the last 2000 years, the spread of these groups was slow and most of the country was settled from the southeastern-central area as late as the 16th and 17th centuries AD (only 15–20 generations ago). This phase of settlement was typified by isolated breeding units of a few hundred people (subisolates), which experienced little gene flow (isolation by density). The population had reached 500,000 by the mid-18th century, and grew rapidly over the last 250 years to the current size of 5.1 million.

Admixture has occurred between the Finnish-speaking, Swedish-speaking, and resident Saami populations. However, the Finnish Saami population is distinct with respect to “Finnish” diseases (1–3) (see below). Finnish Y-chromosomal diversity is reduced when compared with neighboring populations, indicating a period of narrow bottlenecks and a reduction in effective population size relative to other European populations (4). When compared with other European groups, HLA-A, B, and C gene and haplotype frequency analyses (5) as well as HLA-A, B, C and DR antigen frequency analyses (6) show unique deviations specific for the Finnish population as well as prominent differences between geographically remote areas within Finland. This corresponds to the previous and recent findings that local isolates in Finland may show significant deviations from the population mean (7). Overall, the Finnish population has been shaped by major bottlenecks, small numbers of isolated founders, and rapid expansion, all of which have resulted in genetic drift. This demographic history has led to the “Finnish disease heritage”, a unique catalogue of genetic diseases that occur in Finland at higher frequencies than in other European countries.

The population sample in this study is comprised by 90 unrelated Finnish individuals representing 42 sequentially collected cadaver kidney donors and 48 donors from the Finnish bone marrow donor registry. These samples were obtained from across the entire country and were previously typed with HLA-A and B low-resolution kits and with class II high-resolution kits.

The A, B and C loci were typed using IHWG RLS reagents, DRB1 with commercially available SSP or reverse format SSOP kits, and DQB1 and DPB1 with reverse format SSOP kits. Some samples at the DRB1 and DQB1 loci were typed at low (“two-digit” or “serological level”) resolution. Common northern European alleles were assigned in these cases (e.g., DRB1*15 was assigned as DRB1*1501).

While no overall deviations from expected Hardy-Weinberg equilibrium proportions were detected, an excess of homozygous genotypes was observed at the DQB1 locus ($p\text{-val}=0.0394$).

This is the first report of high-resolution HLA class I allele and haplotype distributions in Finland. As with the low-resolution information from previous studies, this study shows that the most common haplotype in Finland is A*0301-B*3501-DRB1*0101 (haplotype frequency (h.f.)=0.0779), a haplotype with an exceptionally high frequency among Caucasian populations. The three other most frequent haplotypes are A*0101-B*0801-DRB1*0301 (h.f.=0.0611), A*0301-B*0702-DRB1*1501 (h.f.=0.0556), and A*0201-B*1501-DRB1*0401 (h.f.=0.0500), all common in Caucasian populations. Several unexpected estimated haplotypes, which had not been previously identified in the HLA-A-B-DRB1-haplotype report from Northern Finland (8), such as A*2402-B*4001-DRB1*1301, A*0201-B*4001-DRB1*1301, A*3101-B*5101-DRB1*1301, A*0201-B*5101-DRB1*1301, and A*0201-B*2705-DRB1*0801 were also identified. The A*2402-B*3901-DRB1*0404 haplotype was observed only once and A*28-B*08-DRB1*03 haplotypes were not observed. In this study population, the highest linkage disequilibrium estimates were found in locus pairs C:B ($D'=0.91100$) and DRB1:DQB1 ($D'=0.94647$).

References

1. Peltonen L, Palotie A, Lange K. Use of population isolates for mapping complex traits. *Nature Reviews* 1, 182–190, 2000.
2. Kere J. Human population genetics: Lessons from Finland. *Ann Rev Genomics Hum Genet* 2, 103–128, 2001.
3. Varilo T, Paunio T, Parker A, et al. The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum Mol Gen* 12, 51–59, 2003.

4. Sajantila A, Salem A-H, Savolainen P et al. Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *Proc natl Acad Sci USA* **93**, 12035–12039, 1996.
5. Lokki M-L, Julin M. HLA-A,B,C gene and haplotype frequencies in the Finnish population. *Tissue Antigens* **20**, 239–250, 1982.
6. Sirén M-K, Sareneva H, Lokki M-L, et al. Unique HLA antigen frequencies in the Finnish population. *Tissue Antigens* **48**, 703–707, 1996.
7. Nevanlinna H. Rare hereditary diseases and markers in Finland: an introduction. In *Population Structure and Genetic Disorders*, ed. A Eriksson. London: Academic, 1980.
8. Ikäheimo I, Silvennoinen-Kaakinen S, Tiilikainen A. HLA five-locus haplotypes in Finns. *Eur J Immunogen* **23**, 321–328, 1996.

Population: Georgians from Tbilisi, Republic of Georgia

Authors: N. Bendukidze, E. Ivaskova

Authors' address: Institute for Clinical and Experimental Medicine (IKEM), Prague, Czech Republic

The Republic of Georgia – called *Sakartvelo* in Georgian and *Gruzia* in Russian – is situated at the crossroads of Europe and Asia, in the Western part of the Transcaucasia, on the southern slopes of the Caucasian Mountains, and is a multi-ethnic nation. Ethnic Georgians, who refer to themselves as *Kartvelebi* in their native tongue, form the majority of the population in most of the regions. According to the census of 1989 they comprised ~70.0 percent of the population (3,787,393) and cognate Ibero/Caucasian stock. Other major ethnic groups are Armenians (8.1%), Russians (6.3%), and Azerbaijanians (5.7%). There are also Ossetians, Greeks, Abkhaz, Ukrainians, Kurds, Jews, Assyrians, and other ethnic groups. The great majority of the population of the Republic of Georgia belongs to the Eastern Orthodox Church (1, 2). The ethnic Georgian population can be subdivided into Svans, Meshetins, Ajars, Megreles, Kakhetins, Karts and other smaller groups with regional identities (3). The population census of 1989 estimated that 26% (824,000) of all ethnic Georgians lived in Tbilisi (4).

The Georgian republic counts its history from the 12th century BC. At the end of the 2nd and the beginning of the 1st millennium BC, two major tribal unions arose: those of the Diakhi (Taokhi, Tao) and the Qolha (Colchis). Their union disintegrated in the mid-8th century B.C. In the 8th–7th centuries B.C., the Karts, Mengreles, Chans and Svans came to the fore among the Georgian tribes, and as a result of their consolidation, a two-state confederation took shape in the 6th–4th centuries BC.

Due to its productive land and strategic location astride the crossroads between Europe and Asia, Georgia (the kingdom of Kartli) has attracted the attention of traders as well as potential conquerors. The Georgians' long history has been marked by intensive interactions with other nations and cul-

tures, such as those from Assyria, Urartu, Greece, Rome, Persia, Arabia, Byzantium, Mongolia, Iran, Turkey, Russia and the Western European countries. An isolated Christian kingdom (Christianity was accepted by Georgia as the state religion in the 4th century AD) situated in the centre of the Caucasus was of special strategic significance. For that reason, it was the main target of foreign aggression. Persian, Arabic, Mongolian, Turkish and Iranian invasions were almost permanent and continued several centuries.

In 1783 AD, Georgia became a protectorate of Russia, but was again conquered by Persia in 1795. East Georgia was annexed by Russia in 1801.

Through more than 30 centuries of such interactions, Georgians have managed to preserve their unique language, culture, and identity. The Georgian alphabet was created in the 4th century BC, but the oldest surviving text in the Georgian language is dated to the 5th century AD.

Recent studies of mitochondrial DNA (mtDNA) have indicated that Georgian mtDNA lineages are characteristic of the European gene pool (5). They occupy an intermediate position between European and Near Eastern populations in gene trees and principal co-ordinate analyses, suggesting that they are either ancestral to European populations or derived via admixture from European and Near Eastern populations (6).

HLA antigen and haplotype diversity studies in the Georgian population began during 12th IHWS (5, 8) and were continued using IHWG RLS reagents and SSP commercial kits. 108 DNA samples were typed at the HLA-A, B, and C loci. This Georgian population sample is representative of all ethnic Georgians.

A*0201 (allele frequency (a.f.)=0.30952), B*5101 (a.f.=0.16204), Cw*0401 (a.f.=0.21963) were shown as the most frequent Class I alleles in Georgian population. HLA-A*0201 – HLA-Cw*0401 – HLA-B*3503 was the most frequent three Class I antigen (A-Cw-B) haplotype (haplotype frequency (h.f.)=0.03653). HLA-A*0201 – HLA-B*5101 was the most frequent two antigens (A-B) haplotype (h.f.=0.08026).

No locus-wide deviations from expected Hardy-Weinberg

equilibrium proportions were observed at the HLA-A locus. Under the Ewens-Watterson homozygosity test of neutrality, a positive value of the normalized deviate was observed for HLA-A ($F_{nd}=0.4988$, non-significant), which was in the direction of positive directional selection. Observed homozygosity was higher than expected under neutrality for this locus. This could be the result of non-significant excess of homozygous HLA-A-0201 genotypes (p -value=0.7383). Non-significant negative normalized deviations (F_{nd}) from natural expectations were calculated for the Cw locus ($F_{nd}=-0.7043$), which suggests the operation of balancing selection for the Cw locus.

Although no locus-wide deviation from expected Hardy-Weinberg equilibrium proportions were observed at HLA-B,

an excess of homozygous genotypes was detected at the HLA-B locus (p -value=0.0240). The B*5101 allele comprised 32.4%, B*3503 – 22.2% and B*3501 – 16.7% of the allelic diversity for this population, but no genotypes with expected counts greater than 5 were observed. The high value of k (42) put the HLA-B locus allele data out of the range of simulated homozygosity values, and the Ewens-Watterson homozygosity test of neutrality could not be used. The non-significant, but elevated p -value of F (0.6193) calculated under the Slatkin implementation of the Ewens-Watterson homozygosity test suggested the operation of directional selection for this locus. An increased sample size should help to estimate the natural selection process for Georgian population.

References

1. Country profile: Georgia – General Information <http://www.mfa.gov.ge/general.html>.
2. About Georgia. <http://ggdavid.tripod.com/georgia/>.
3. People. <http://www.tbilisipastimes.com/people.html>.
4. Revaz Gachechiladze. "Population Migration in Georgia and Its Socio-Economic Consequences". Discussion Paper Series – UNDP, 1997.
5. Comas D, Calafell F, Bendukidze N, Fananas L, Bertrapeit J. Georgian and Kurd mtDNA sequence analysis shows a lack of correlation between languages and female genetic lineages. *Am J Phys Anthropol* **112**(1), 2000, 5–16.
6. Nasidze I, Stoneking M. Mitochondrial DNA variation and language replacements in the Caucasus. *Proc R Soc Lond B Biol Sci* **7**, **268** (1472), 1197–1206, 2001.
7. Bendukidze N, Meschishvili E, Ivaskova E, Cerna M, Churadze T, Guladze R, Shalamberidze K. Determination of HLA Class I Alleles in Georgian, Kurdish and Czech Populations by PCR-ARMS. *Human Immunology* **47**(1-2), 54, 1996.
8. Alexeev L, Khaitov R, Boldyreva M, Trofimov D, Guskova I, Akopian A, Fedoseeva T, Unsrinov I, Pospelov L, Bendukidze N, Meschishvili E, Ivaskova E, Cerna M, Churadze T, Guladze R, Shalamberidze K. HLA in some ethnic groups of the former Soviet Union (FSU). In: Genetic diversity of HLA Functional and Medical Implication. Ed. D.Charron 1996, Vol. 1: 364–374.

Population: Roma from Bulgaria

Authors: M. Ivanova, A. Michailova, E. Naumova

Authors' address: Central Laboratory of Clinical Immunology, Medical University, Sofia, Bulgaria

It has been suggested that the Roma people (a.k.a. Romanies) first migrated to Eastern Europe from Hindustan. They arrived in Western Europe during the 14th and 15th centuries. Most Roma groups settled down into a single location and only 5 percent of European Romanies are thought to be nomadic.

There is historical evidence of several major migrations and settlements of Romanies in the Bulgarian lands. The largest number migrated westward into the present Bulgarian territory during the 13th and 14th centuries (1) and by the end of

the 14th century, they had established settlements in Bulgaria and throughout the Balkans (2). The number of Romanies in the Bulgarian lands increased under Ottoman rule in the Balkans (14th-15th centuries). More recent migrations of Roma from Wallachian lands into Bulgaria occurred between the 15th and 18th centuries.

The Roma people make up 3.4 percent of the Bulgarian population and represent one of the largest minorities. They are present throughout Bulgaria, and are distributed relatively equally in both rural and urban areas. The languages spoken by the Roma minority in Bulgaria are Romani, Bulgarian, Turkish and Vlach. About 39.2 percent of Romanies in Bulgaria are Muslims, while 60.4 percent are Orthodox Christians. Roma living in Bulgaria are heterogeneous in their identity, and three major groups (the Jerlii, the Kardarashi,

and the Rudari) can be identified. The Jerlii are the settled Roma in the country, descended from those who arrived and settled in the Bulgarian lands during the Ottoman rule. The Jerlii are split into two major groups, the Christian Dasikane Roma and the Muslim Horohane Roma. The Wallachian Roma comprise a third group of settled Roma. The Kardarashi are historically nomadic Christian groups that were forced to settle during the 1950s. This group is divided into Lovari and Kalderari, which in turn are divided into numerous groups and subgroups. The Rudari, also often called the Vlach Roma are recently settled Christians who speak an ancient Romanian dialect. They are divided into the Lingurari, Ursari, and Majmunari groups.

In the percent preliminary study 12 healthy unrelated Romanies, randomly selected from different regions of Bulgaria were included. All sampled Romani subjects were Orthodox Christians speaking Romani and Bulgarian. HLA-A, -B, -C genotyping was performed by IHWG RLS reagents. No deviation from Hardy-Weinberg equilibrium was observed for these loci. Significant positive deviation of homozygosity from neutral expectations under the Ewens-Watterson model ($F_{nd}=2.27$) was observed for the HLA-B locus. While this value could suggest directional selection, it may result from the small size of the sampled population. Thirty-three alleles: 7 HLA-A, 13 HLA-B and 11 HLA-C alleles were observed. The most frequent alleles for the three loci tested were as follows: A*0101 (allele frequency (a.f.) = 0.333), A*0201

(a.f.=0.208), A*1101 (a.f.=0.208); B*4006 (a.f.=0.273), B*5201 (a.f.=0.182); B*4101 (a.f.=0.091); Cw*1502 (a.f.=0.273), Cw*1507 (a.f.=0.136). Although the A*1101; B*4006, B*5201, Cw*1502, and Cw*1507 alleles have been detected in the general Bulgarian population (3,4), their frequencies in Romanies are much higher. Similar frequencies for these alleles were observed in Asian populations (5,6). Additional alleles thought to be characteristic of Asian populations were observed in this Bulgarian Roma population (e.g., A*0211 (a.f.=0.083), A*2601 (a.f.=0.042); B*2704 (a.f.=0.045)) (5,6). The Bulgarian Roma HLA-A-C-B haplotype distribution was quite different from the general Bulgarian population, and although some of the observed haplotypes have been described in European groups, the small size of this population sample makes it difficult to assess the significance of these haplotypes, as accurate three-locus haplotype estimation may not have been possible via the EM methods employed. As a result, these results may not represent an accurate sampling of Bulgarian Roma haplotypic diversity. However, haplotypes similar to the two most frequent Bulgarian Roma A-C-B haplotypes (A*0101-Cw*1502-B*4006 (haplotype frequency (h.f.) = 0.150) and A*0101-Cw*1507-B*4006 (h.f.=0.150)) have been observed in other Roma groups (7).

In conclusion, these results suggest low admixture between Romanies and other Bulgarians in support of the proposed historical origins of the Roma ethnic groups.

References

1. Kenrick D. From India to the Mediterranean: The Migration of the Gypsies, Paris, Gypsy Research Centre. 1993
2. Eminov A. Turkish and Other Muslim Minorities in Bulgaria, New York: Routledge, 1997.
3. Ivanova M, Spassova P, Michailova A, Naumova E. Distribution of HLA class I alleles and haplotypes in Bulgarians – contribution to understanding the origin of the population *Tissue Antigens* **57**, 2001, 208–215
4. Ivanova M, Rozemuller E, Tyufekchiev N, Michailova A, Tilanus M, Naumova E. HLA polymorphism in Bulgarians defined by high-resolution methods in comparison with other populations. *Tissue Antigens* **60**, 496–504, 2002.
5. Clayton J, Lonjou C. Allele and haplotype frequencies for HLA loci in various ethnic groups. In: Charron D, ed. *Genetic diversity of HLA. Functional and Medical Implication*. 1. Paris: EDK, 1997: 665–820.
6. Geer L, Terasaki PI, Gjertson D. Gene Frequencies and World Maps of Their Distribution. In: Terasaki PI, Gjertson D, ed. *HLA 1998*. Lenexa: American Society for Histocompatibility and Immunogenetics, 1998: 327–363.
7. Ogawa A, Tokunaga K, Lin L, et al. Diversity of HLA-B61 alleles and haplotypes in East Asians and Spanish Gypsies. *Tissue Antigens* **51**, 356–366, 1998.

Population: Slovenian from Slovenia

Authors: B. Vidan-Jeras¹, B. Jurca¹, V. Dolzan², M. Jeras¹, K. Breskvar², M. Bohinjec¹

Authors' addresses: ¹Blood Transfusio Center of Slovenia, Ljubljana, Slovenia,

²Institute of Biochemistry, Faculty of Medicine, Ljubljana, Slovenia

With 2 million inhabitants, the territory of modern Slovenia borders on Italy, Austria, Hungary, Croatia and the Adriatic Sea, and has been intensively inhabited since the Bronze Age. Although historians and linguists are convinced that the Slovenian population belongs to the group of Slavic nations, the

issue of origin of Slovenians has not been settled. The Veneti are possible Slavic ancestors of the Slovenians, but it remains unclear whether they were present in this region during the Bronze age (1, 2) or if Slovenians are descendants of Slavs who migrated across the Carpatian mountains of modern-day Ukraine in the 6th century.

Slovenia lies on a trade route between Central and South Europe, and this geographical position has made it a bridge between different cultures and populations. A region of about 20,000 km² has always been open to migrations. For example, archaeological remains of the Illyrian culture date back to the beginning of the 9th century BC, and antedate the evidence of the Celtic culture from the mid 4th century BC.

During the transition era between the 1st century BC and 1st century AC, the region was under Romans rule until the Western Roman Empire was invaded by Huns and Germanic tribes. The presence of Goths and Lombards in Slovenia is documented by archaeological findings. The intrusion of Hungarians separated Slovenians from other Slavic neighbours. The defeat of the Slovenian army in 955 was followed by the influx of German clergy, nobility and peasant colonists, most of which were from Bavaria.

In the Middle Ages, the invasion of Turks into Europe influenced the structure of the Slovenian population. Southern Slavs of the Greek Orthodox religion were prevalent among refugees pushed into the Slovenian region in the 15th and 16th centuries. Under the Habsburg monarchy (1250–1918) many Germans and Czechs came to Slovenia. The influence of Napoleon's Illyrian Provinces (1809–1813) upon the Slovenian population was negligible. Following the foundation of Yugoslavia ("The kingdom of Serbs, Croats and Slovenes") in 1918 many soldiers and policemen from Eastern and Southern regions became permanent residents of Slovenia. The influx of new citizens from other Yugoslav regions and their naturalization was actively promoted after the Second World War.

While the modern Slovenian people have a mixed and diverse heritage, they have remained a distinctive national community (4). The Slovenian language belongs to the group of western Slavic (Indo-European, Balto-Slavic) languages (5).

A population sample of 100 randomly chosen Slovenian individuals (all of whom had parents who had declared themselves as Slovenians), representing all parts of the country, was collected in the capital city. The study design was approved by the Slovenian Ethical Committee for Research in Medicine (6). An analysis of HLA-A, B and C diversity was carried out after medium resolution PCR-SSP based typing during the 12th IHW (7). The DRB1, DQB1, DQA1, DPB1, DPA1 loci (8), and the promoter regions of DQB1 (9) and DQA1 (10) were typed by PCR-SSO and PCR-SSP methods. In a study of global allele distributions, the DPB1*0401 allele was observed at a very high frequency in Slovenia (allele frequency=0.45), and was observed to decline in frequency in surrounding European populations (11, 12). In addition, strikingly different frequency distributions were observed in a comparative study of B44 subtypes and their A-B-DR haplotypes in populations from Slovenia, the Netherlands, and Switzerland. In contrast to Swiss and Dutch populations, where for example, B*4402 is predominantly linked to A2, the predominant linkage in Slovenians is with A24 and A28 (13). A study on Y-chromosomal diversity in Europe, which included the Slovenian population, revealed that the diversity is primarily influenced by geography, rather than by language (14).

An analysis of the HLA-class II data carried out by the IHWG biostatistics core showed an excess of homozygotes relative to H-W proportions for the DRB1, DQA1 and DQB1 loci, suggesting possible typing drop-outs for these loci, very likely due to the limited number of sequence specific primers and probes used for typing during the past IHW.

References

1. Bor M, Savli J, Tomazinic I eds. Veneti-our ancient ancestors, (orig. in Slovenian). Wien, Editiones Veneti, 1989.
2. Tomazinic I ed. The Slovenians, (orig. in Slovenian). Wien, Editiones Veneti, 1999.
3. Stih P, Simoniti V, eds. History of Slovenia (orig. in Slovenian). Ljubljana, Korotan, 1996.
4. Granda S. (Centre of Scientific Reserch of the Slovene Academy of Sciences and Arts) personal communication.
5. Ruhlen M. A Guide to World's Languages. Vol. 1: Classification. London, Edward Arnold, 1987.
6. Bohinjec M, Vidan-Jears B, Dolžan V, Jurca B. A HLA genetic analysis of the Slovenian population, (orig. in Slovenian). Ministry of Science of the Republic of Slovenia, Research Project #C3-0562-311/92, Ljubljana, 1992.
7. Tonks S, Marsh SGE, Bunce M, et al. HLA Class I DNA typing study. In: Charron D, ed. Genetic diversity of HLA: Functional and medical implications. EDK, Paris, 199–215, 1997.
8. Tiilikainen A, Fischer G, Grubic Z, et al. Anthropological Features of the East European Region. In: Charron D, ed. Genetic Diversity of HLA: Functional and Medical Implications. Paris, EDK, 307–313, 1997.

9. Reichstetter S, Adorno D, Alaez C, et al. DQB1 promoter polymorphism: 12th International Histocompatibility Workshop Study. In Charron D. Genetic diversity of HLA: Functional and medical implications. Paris, EDK, 176–183, 1997.
10. Dolžan V, Vidan-Jeras B, Breskvar K, Bohinjec M. Polymorphism in the promoter region of HLA-DQA1 gene in Slovene population. *Pflugers Arch* 431, Suppl. 2, 123, 1996.
11. Vidan-Jeras B, Jurca B, Dolžan V, Jeras M, Breskvar K, Bohinjec M. Slovenian Normal. In: Terasaki PI, Gjertson DW, eds. HLA 1998. Lenexa, ASHI, 181–182, 1998.
12. Geer L, Terasaki P, Gjertson DW. Gene Frequencies and World Maps of Their Distributions. In Terasaki PI, Gjertson DW, eds. HLA 1998. Lenexa, ASHI, 327, 1998.
13. Vidan-Jeras B, Breur-Vriesendorp B, Bohinjec M, Jeannet M, Roosnek E, Tiercy J-M. HLA-B44 allele frequencies and haplotypic associations in three European populations. *Eur J Immunogenet* 24, 335–345, 1997.
14. Rosser ZH, Zerjal T, Hurles ME, et al. Y-chromosomal diversity in Europe is clinical and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67, 1526–1543, 2000.

South-West Asia

Population: Arab Druse from Israel

Authors: Chaim Brautbar¹, Shoshana Israel¹, Kais Firro², Cilly Safirman¹, Keren Miller¹, Avraham Amar¹ and Campbell Witt³

Authors' addresses: ¹Tissue Typing and Immunogenetics Unit, Hadassah Hospital, Jerusalem, ²Middle East Studies, University of Haifa, Haifa, Israel, ³Royal Perth Hospital, Perth, Australia

The Druse are geographically dispersed over Syria, Lebanon, Israel and Jordan. The earliest evidence of the existence of the Druse community in Galilee in northern Israel dates to the eleventh century. Of approximately forty villages inhabited by the Druse between the 11th and the 17th centuries in Galilee, sixteen villages remain, with an additional two villages established in Carmel after the 17th century.

In 1593 the Druse gained control of the district of Sidon and five years later the district of Safad in Galilee. Druse immigrants from Lebanon and Northern Syria began to move into the Galilee. In 1633, the Druse community in Galilee went into decline, and subsequently migrated into southern Syria. However, at the end of the 17th century, Druse from the Galilee and Lebanon moved to Mount Carmel. At the end of the 18th century, a new wave of migration into Galilee and the Carmel came from northern Syria, near Aleppo (Halab). Nearly every village in Israel has families who can trace their origin to the district of Aleppo, many of them called Halabi. Those who made it to the largest village in the Carmel through the years do not speak the Palestinian or Lebanese dialects of Arabic but retain their Aleppo accent today.

Although there is no information on the kinship relation between people from the different villages, it is known that many Druse families are descendants of clans that accepted the Druse faith in the 11th century. Soon after its formation in the 11th century, the members of the Druse community ceased inter-

marrying outsiders. Endogamy within the patrilineal family and consanguinous marriage within the extended family are prevalent among the Druse community in Israel.

The population studied comprised 100 healthy unrelated Arab Druse collected in the largest Druse village in Israel. All donors were interviewed and first degree relatives were not included. The DNA samples were obtained from the National Laboratory for the Genetics of Israeli Populations at Tel Aviv University, a national repository for ethnic groups residing in Israel. HLA class I molecular typing was performed using the IHWG RLS kits. HLA-DRB1 generic and allele specific analysis was performed using the Lipa HLA reverse hybridization kit. Ambiguities were solved by the Olerup SSP DRB1 allele specific kits.

The HLA-A, B and DRB1 loci of the Arab Druse were tested for Hardy-Weinberg equilibrium. Despite samples having been collected in a random fashion, all three loci were found to depart significantly from Hardy-Weinberg equilibrium (p-values, 0.013, 0.020 and 0.062, respectively). The departure from Hardy-Weinberg equilibrium was manifest by a surplus of homozygotes. In order to exclude the possibility of alleles being undetected by the typing methods employed all homozygous samples were tested by a second method. This included 18, 18 and 16 samples at the HLA-A, B and DRB1 loci, respectively. In only one case was an additional allele identified by the second technique for HLA-A and in two cases an additional allele was identified at DRB1. However, the p-values given above are for the corrected data indicating that undetected alleles explained only a small part of the departure from equilibrium. The most likely explanation for the observed departure from equilibrium (homozygote excess) is therefore an underlying population sub-structure to the sample. This may have occurred due to the fact that the Druse in Israel consists of several large “clans”. Marriage tends to occur within the clan. Due to genetic drift, each clan may have higher frequencies of different alleles resulting in homozygosity for different alleles.

The most frequent HLA-A alleles were A*02, A*11 and A*24. These alleles are also common in Ashkenazi and Moroccan Jews (1) and Lebanese (2). In contrast HLA-A*01 was present in a relatively low frequency (4%) as compared to Ashkenazi and Moroccan Jews and Lebanese (16%, 26%, 23% respectively). The most common HLA-B allele was B*35. This is also the most frequent allele in Ashkenazi Jews (21%) and Lebanese (36%). Other HLA-B alleles that were frequent among Druse B*18, B*51, B*07, B*50 and B*44 are also common among Ashkenazi and Moroccan Jews and Lebanese. At the HLA-DRB1 locus, DRB1*1104, DRB1*1101 and DRB1*0701 were frequent in the Arab Druse as in most of other Jewish populations, Israeli Arabs (3) and Lebanese (4).

The most frequent three locus haplotypes were A*24-B*35-DRB1*1104, A*02-B*50-DRB1*0701, A*24-B*18-DRB1*1104, A*02-B*44-DRB1*1104 and A*29-B*07-DRB1*1401.

In order to understand the genetic relationship between the Arab Druse and surrounding ethnic groups, DRB1 allele frequencies have been analyzed in terms of their standard genet-

ic distance (SGD) to those previously published for Israeli Arabs, Ashkenazi Jews, Jews from Morocco, Lybia, Iran, Yemen and Ethiopia, Arab Emirates, Bedouins, Moroccans, German, Japanese and African-Americans. SGDs indicate that the Druse are most closely related to other Israeli Arabs ($SGD = 0.17$), followed by all of the Jewish groups ($0.2 < SGD < 0.3$) except for the Ethiopian Jews. Arab Emirates, Moroccans and Germans were clearly more distantly related along with the Ethiopian Jews ($0.45 < SGD < 0.6$). Japanese and African-Americans were even more distantly related ($SGD > 0.8$). Although Israeli Arabs were the closest ethnic groups to the Druse, their DRB1 allele distributions were significantly different ($p < 0.0001$), best exemplified by a higher frequency of DRB1*0301, 1001, 1302, 1303 and 1502. Thus, despite their geographical proximity, the Arab Druse population investigated in this study has been sufficiently secluded so as to develop significantly different HLA allele frequencies from the Israeli non-Druse Arabs.

References

1. Roitberg-Tambur A, Witt CS, Friedmann A, Safirman C, Sherman L, Battat S, Nelken D, Brautbar C. Comparative analysis of HLA polymorphisms at the serologic and molecular level in Moroccan and Ashkenazi Jews. *Tissue Antigens* **46**, 104–110, 1995.
2. Nuwayri-Salti N, Shaya M. Major histocompatibility class I antigens in the Lebanese population. *WHO Eastern Mediterranean Region Health Journal*. 1977, 3, 101–107.
3. Amar A, Kwon OJ, Motro U, Witt C, Bonne-Tamir B, Gabison R, Brautbar C. Molecular analysis of HLA class II polymorphism among different ethnic groups in Israel. *Human Immunology* **60**, 723–730, 1999.
4. Bias WB, Nuwayri-Salti N, Wood WH, Cissell B, Berger S, Ng-Wong C, Schmeckpeper BJ. 12 International Histocompatibility Anthropology SWAS regional report. Proceedings of the 12th International Histocompatibility Workshop and conference. D. Charron, editor. 1977, 1, 353–363.

Populations: Golla from Andhra Pradesh, and north Indians from New Delhi, India

Authors: Steven J. Mack^{1,2}, Michael H. Crawford³, Nilmani Saha⁴, Andrea J. Jani¹, Laura N. Geyer², Mohan Reddy⁵, Henry A. Erlich^{1,2}

Authors' addresses: ¹Children's Hospital Oakland Research Institute, Oakland, CA, ²Roche Molecular Systems, Alameda, CA, ³Department of Anthropology, University of Kansas, Lawrence, KS, ⁴Department of Human Genetics, University of Pittsburgh, PA, ⁵Indian Statistical Institute of Calcutta, Calcutta, India

The Indian subcontinent has been populated since the Paleolithic, and has experienced a history of repeated invasion, colonization and demographic change in that time. Early mi-

grants from Africa are thought to have followed the Indian coastline to reach southeast Asia and Austronesia. Prior to the arrival approximately 4000 years ago of Indo-European speakers from the Iranian plateau, the region is thought to have been inhabited by Tibeto-Burman, Austroasiatic and Dravidian speaking groups. Indo-European speakers likely imposed a caste system on themselves and the indigenous populations, which migrated south. Extensive contacts with and invasions by groups from east and southwest Asia and Europe have continued over the last several thousand years, culminating with British colonial control of India, which ended in 1947. As a result of these contacts, the Indian subcontinent remains a region defined by tribal and caste subdivisions, which are reinforced by geography, language and religion. In modern India, Indo-Aryan (or Indic) languages (such as

Hindustani) are spoken by 74% of the population, primarily in northern regions of the nation, while Dravidian languages are spoken by 24% of the population, primarily in the southern regions (1–3).

Situated between the northern states of Haryana and Uttar Pradesh, and close to the states of Rajasthan, Punjab, and Uttaranchal, New Delhi is the capital of India. In general, northern India has been a crossroads for trade between east Asia, the Middle East, Europe and Africa for some time, and this area has been shown to have a high degree of HLA diversity, in part due to genetic contributions from outside of India (4). In addition, the region around the capital has been an urban center for at least three thousand years, and much of the ethnic and cultural diversity of the nation can be found in New Delhi. Hindi is spoken in New Delhi along with many of the 14 additional official languages of India.

The Golla are a Telugu-speaking Dravidian caste located in the state of Andhra Pradesh in southern India. The Golla population sample reported here is comprised by six endogamous Golla subcastes from the Chittor district of southern Andhra Pradesh. Consanguinity within each subcaste has resulted in a high degree of inbreeding, with resulting increases in homozygosity. Class I serological data and analyses, and DRB1 allele frequency data and analyses for the individual Golla subcastes have been published previously (4). Because the sample size for each Golla subcaste is low (n ranges from 10 to 26), the Golla are considered here as a single population to permit additional statistical analysis.

Subsets of 70 unrelated individuals from New Delhi (28 degrees 37 minutes north latitude, 77 degrees 13 minutes east longitude) were typed at the HLA-A (66 individuals), C (56 individuals) and B (66 individuals) loci using the IHWG RLS typing system. As described previously (5), samples from 110 unrelated members of the Golla subcastes in Andhra Pradesh (17 degrees 30 minutes north latitude, 78 degrees 30 minutes east longitude) were collected in thirty villages in the Chittoor district of Andhra Pradesh, and subsets of these samples were typed at the HLA-A (88 individuals) and B (104 individuals) loci using the IHWG RLS system and at the DRB1 (109 individuals) locus using the Dynal Auto-RELI DRB1 general and subtyping system. Statistical analyses were carried out by the IHWG Biostatistics Core.

The genotypic proportions at the HLA-A, C and B loci in the New Delhi population and at the HLA-B locus in the Golla population were consistent with expected Hardy-Weinberg equilibrium proportions (HWE). However, deviations from HWE were observed at the HLA-A locus (p -value=0.0246) and at the DRB1 locus (p -value=0.0025) in the Golla. This

latter deviation is the result of an overall excess of homozygote genotypes (expected 10.02, observed 23, p -value < 0.0001). These deviations from HWE in the Golla likely result from inbreeding in each of the six Golla subcastes.

While homozygosity values (F_{nd}) were consistent with the action of neutral evolution at all typed loci, homozygosity at the HLA-C locus in the New Delhi population approached a statistically significant low value (F_{nd} =-1.10, p -value=0.0507).

Allelic diversity (h) was high ($h>0.9$) at all typed loci in both populations. The HLA-A locus showed the lowest diversity with eighteen ($h=0.91$) and 21 ($h=0.92$) HLA-A alleles in the New Delhi and Golla populations, respectively, while the HLA-B locus showed the highest diversity, with 31 ($h=0.95$) and 33 ($h=0.93$) alleles observed in the New Delhi and Golla populations, respectively. Twenty-three HLA-C alleles ($h=0.94$) were observed in the New Delhi population, and 22 DRB1 alleles were observed ($h=0.92$) in the Golla population.

Several high frequency (allele frequencies (a.f.) greater than 0.06) alleles were observed at all loci in both populations, with 8 HLA-A, 8 HLA-C and 8 HLA-B alleles comprising 78%, 71% and 69% of the HLA-A, C and B allelic diversity in the New Delhi population, and 7 HLA-A, 7 HLA-B, and 6 DRB1 alleles comprising 76%, 68%, and 67% of the Golla HLA-A and B allelic diversity.

In particular, the A*1101 and *2402 alleles were observed at very high frequencies (0.23 and 0.11 in New Delhi and 0.12, and 0.14 in the Golla) in both populations, and the *0201, *0211, *0301, and *3303 alleles were seen at frequencies greater than 0.06 in both populations. The Golla were distinguished from the New Delhi population by the presence of the A*0101 and *3101 alleles at frequencies greater than 0.1, while these alleles were present in the New Delhi population at frequencies of approximately 0.03.

At the HLA-B locus, only the B*4006 allele was observed at a very high frequency (0.13 in the Golla and 0.11 in New Delhi) in both groups, but the *0801, *3503, *5101, *5201, and *5801 alleles were observed at frequencies of at least 0.06 in both populations. The B*4403 allele was seen at a very high frequency in New Delhi (0.11), but was seen at a low frequency 0.02 in the Golla. In addition, both groups displayed considerable diversity of A2 and B15 alleles, and extensive A24 allelic diversity (6 alleles) was seen in the Golla.

Extensive two and three-locus haplotypes were inferred in both populations. Eight high-frequency (haplotype frequency (h.f.) greater than 0.03) A:B haplotypes comprised 36% of

the haplotype diversity in the New Delhi population, while six high-frequency haplotypes comprised 27% of the Golla A:B haplotype diversity. However, while several high-frequency HLA-A and B alleles were shared between these two populations, the A*1101:B*5201 and *3303:*5801 haplotypes were the only high-frequency haplotypes so shared (h.f.=0.048 and 0.032 in New Delhi, and 0.048 and 0.054 in the Golla, respectively), suggesting that north and south Indian groups have been isolated for considerable time. The highest frequency A:B haplotype in the New Delhi population was A*2402:B*3501 (h.f.=0.056), while A*3303:*5801 was the highest Golla A:B haplotype. Five high-frequency A:C:B haplotypes comprised 22% of the haplotype diversity

in the New Delhi population, and two A:B:DRB1 haplotypes were observed at high frequency in the Golla population (combined h.f.=0.08).

In general, the New Delhi and Golla populations display several alleles and haplotypes characteristic of east Asian populations. For example, the A*2402 and *1101 alleles are observed at high frequencies in almost all east Asian and Oceanian groups, as are the A*3303:B*5801 and A*3303:B*5801:DRB1*0301 haplotypes (the latter is seen in the Golla at a frequency of 0.036). These patterns of allele and haplotype sharing data suggest considerable east Asian contributions to the Indian gene pool.

References

1. Breton RJL. Atlas of the languages and ethnic communities of South Asia. New Delhi: Sage Publications India Pvt. Ltd., 1997.
2. Masica CP. The Indo-Aryan languages. Cambridge: Cambridge University Press, 1991.
3. Gonen A. The encyclopedia of the Peoples of the World. Jerusalem : Jerusalem Publishing House Ltd., 1996.
4. Mehra NK, Rajalingam R, Kanga U, McEneny L, Cullen C, Agrawal S, Middleton D, Pllack MS, Amirzagar A, and Singal DP. Genetic diversity of HLA in populations of India, Sri Lanka and Iran. Genetic Diversity of HLA: ED Dominique Charron, Paris EDK. Handbook of the 12th International Histocompatibility Workshop, 1996, 314–320.
5. Crawford MH, Reddy BM, Martinez-Laso J, Mack SJ, Erlich H. Genetic variation among the Golla pastoral caste subdivisions of Andhra Pradesh India, according to the HLA System. *Human Immunology* **62**, 1031–1041, 2001.

Population: Ashkenazi and non-Ashkenazi Jews from Israel

Authors: Ephraim Gazit, Rivka Kalt and Ron Loewenthal

Authors' address: Tissue Typing Laboratory Sheba Medical Center and Sackler Faculty of Medicine, Tel Aviv University

When Israel's independence was proclaimed in 1948, the Jewish population numbered about 600,000, and an unprecedented wave of migration tripled the population within 3–4 years. The size of the population in Israel is presently about 6,000,000, of whom 80% are Jews and the majority of the remaining 20% Muslim Arabs. Waves of migration brought Jews to Israel from all parts of the world. The main groups came from the Middle East, North Africa, and Eastern and Western Europe. Those who were born in Western and Eastern Europe and in North America are traditionally defined as Ashkenazi, and the remaining are defined as non-Ashkenazi. The exile of Jews from Spain in 1492 brought primarily non-Ashkenazi Spanish immigrants (called Sephardic Jews after the Hebrew word Sepharad for Spain) into Central and North Europe adding to the admixture of Jews worldwide. The immigration of Diaspora Jews into Israel created another melting

pot pressure and it is possible that within a generation the separation of the Jewish population into sub-groups will become obsolete.

Previous genetic studies have indicated that certain genetic diseases are more common in certain Jewish ethnic groups. For example, Tay-Sachs, Gaucher's disease, and Familial Dysautonomia are more common among Ashkenazis. Familial Mediterranean Fever is common among North African Jews and Glucose 6-phosphate dehydrogenase (G6PD) deficiency is prevalent among Iraqi Jews.

Linguistically, the Ashkenazi Jews, especially those of Eastern Europe, spoke Yiddish, a dialect composed of a combination of German, Russian and Hebrew words. Non-Ashkenazis spoke either their native language or Ladino, an old Spanish dialect spoken by Sephardic Jews since the 15th Century. In Israel, the main official language is Hebrew, which is spoken and taught in the school system.

The tremendous ethnic admixture that presently occurs in Israel has been the logic behind our present Workshop study. We randomly selected 146 unrelated volunteer donors to determine their class-I gene polymorphism without separating them into their respective ethnic groups, and used the IHWG

RLS technique to type these samples at the HLA-A, B and C loci.

At the HLA-A locus, the frequency of A*0101 and *0201 were about equal (allele frequency (a.f.)=0.171 and 0.128), and these alleles were found in about a third of the individuals tested. A relatively high frequency of A*2402, *2601 and *1101 was also observed.

At the C locus, Cw*0602 was the most common allele, followed by the Cw*0401, *0802 and *0202 alleles. At the B locus, the most common allele was B*5201, followed by the B*3801, *5001, *1404, *1801, *3501 and *4403 alleles.

The most common A-C-B haplotypes found in this population were A*0101-Cw*0602-B*5703, 1101-1202-5201, 3303-0802-1404, 0101-0401-3501, 0301-0802-1404 and 0205-0602-5001. Such haplotypes as A*2601-Cw*1202-B*3801 and 2402-0401-3502, which have been observed in previous studies of Jewish populations, were under-represented in this cohort.

Due to the extreme polymorphism of the class I alleles at the DNA level, it is difficult to compare the allele frequencies obtained in this study to previously published reports. It seems that a study of a larger cohort will be required in the future in order to obtain a more accurate frequency distribution.

Population: Kurds from Tbilisi, Republic of Georgia

Authors: N. Bendukidze and E. Ivaskova

Authors' address: Institute for Clinical and Experimental Medicine (IKEM), Prague, Czech Republic

The Kurds (kûrds, koords) are a non-Arab Middle Eastern minority population that inhabits the region known as Kurdistan, a 191,660 sq. km region that includes parts of eastern Turkey, north-eastern Iraq, northwestern Iran and smaller sections of northeastern Syria and Armenia. Today, many Kurds live in the former Soviet republics of the south Caucasus (Georgia, Armenia and Azerbaijan), as well as in the republics of Central Asia (Kazakhstan and Turkmenia). The name 'Kurd' was only officially used in the Soviet Union, and is now officially used in Russia. The Kurdish people call themselves 'kurmandij'. In Turkey, they are known as 'Turkish Highlanders', and as 'Persian Highlanders' in Iran.

As of the late 1990s, there were estimated to be more than 30 million Kurds, about half of whom live in Turkey near the Iranian border around Lake Van, as well as in the vicinity of Diyarbakir and Erzurum. These Turkish Kurds constitute more than 25% of the population of Turkey. The Kurds in Iran,

who constitute some 10% of its people, live principally in Azerbaijan and Khorasan, with some in Fars. The Iraqi Kurds, about 23% of Iraq's population, live mostly in the vicinity of Mosul, Kirkuk, and Sulaimaniyah (1). In the Caucasus, Kurds live in enclaves among the main population. Armenian Kurds can be found in Aparan and 10 different regions that include Talin and Echmiadzin. Azerbaijani Kurds can be found in the regions of Lyaki, Kelbadjar, Kubatly and Zangelan. Georgian Kurds have largely settled in towns, although some communities are scattered in the eastern regions. The oldest Kurdish habitat is in the south of Turkmenia along the Iranian border, and many Kurds live in the capital town.

Thought to be ethnically close to the Iranians, the Kurds were traditionally nomadic herders but are now mostly semi-nomadic or sedentary. It is not known when the Kurds first appeared in the Caucasus, but they may represent the descendants of shepherds that had occupied the Kurdistan highlands since the Neolithic. From the 10th to the 12th century, the area between the rivers Kura and Araks was ruled by a dynasty of Kurdish ancestry. A more extensive resettlement of Kurds, from Kurdistan into the Caucasus, began in the beginning of the 19th century, when the Caucasus was incorporated into Russia. Historical persecution of Kurds in Persia and Turkey led some Kurdish tribes and smaller groups to seek refuge in Russia, and the number of Kurds in the Soviet Union grew from 76,000 in 1939 to 130,000 in 1983. The Kurds of the Caucasus and Central Asia have been isolated for a considerable period of time, and their development in Russia and then in the Soviet Union has been somewhat different. For example, the Kurds of the Caucasus speak the northern kurmandij dialect and have their own script, which was originally based on the Latin and later on the Russian alphabet (2). In this light, the Kurds of the former Soviet republics may be considered to be a distinct ethnic group in their own right.

In general, Kurds practice Islamic religions, but Georgian Kurds can be distinguished as either Muslim or Yezidi Kurds, and these groups constitute separate ethnic groups, with distinct religious and cultural practices. As a result of religious persecution, Yezidi Kurds have lived in isolation from other Kurdish groups. The number of Yezidis in the Soviet Union was recorded as 15,000 in the census of 1926. The Kurds in contemporary Georgia are predominantly Yezidies, and tend to congregate in or near the capital town of Tbilisi, where they number approximately 34,000. They speak an Indo-European language from the Indo-Iranian family, closely related to Farsi (Persian).

Genetic studies with classical markers, Y-chromosomal STR loci and mtDNA have shown that Kurds are close to other

Middle Eastern populations, while remaining genetically distant from Europeans (3, 4, 5).

In order to study the HLA Class I allele polymorphism in the Kurdish population, we collected blood samples from unrelated Georgian Kurds in the capital city of Tblisi for the 12th IHWS in 1995 (6, 7). In addition, thirty samples were typed using 13th IHWG Reverse Live System (RLS) method and SSP low-resolution commercial kit (8).

HLA-A*2402 and HLA-A*2601 (allele frequency (a.f.)=0.11667 for both), HLA-B*5101 (a.f.=0.12069), HLA-Cw*0401 (a.f.=0.22414) were the most frequent alleles in Kurds.

HLA-A*1101 – HLA-Cw*1202 – HLA-B*5201 was the

most frequent Class I 3 locus haplotype (haplotype frequency (h.f.)=0.05357) and HLA-A*0302 – HLA-B*4402 as well as HLA-A*1101 – HLA-B*5201 (h.f.=0.05172) were the most frequent two locus haplotypes. Strong linkage disequilibrium was shown for B:C haplotype ($p < 0.0005$).

Significant negative normalized deviates of homozygosity (Fnd) were calculated for HLA-A (Fnd=−1.5137, p -value 0.0009), but were non significant for the HLA-B (Fnd=−0.9689) or Cw (Fnd=−0.2792) loci. Overall, these observed homozygosity values are lower than those expected under neutrality, and are consistent with the operation of balancing selection.

References

1. The Columbia Encyclopedia, 6th Edition. 2001. (<http://www.bartleby.com/65/ku/Kurds.html>).
2. The Kurds of the Caucasus. <http://www.ibn-net.org/pdf/Kurdish.pdf>.
3. Comas D, Calafell F, Bendukidze N, Fananas L, Bertrapetit J. Georgian and Kurd mtDNA sequence analysis shows a lack of correlation between languages and female genetic lineages. *Am J Phys Anthropol* **112**(1), 5–16, 2000.
4. Brinkmann C, Forster P, Schurenkamp M, Horst J, Rolf B, Brinkmann B. Human Y-chromosomal STR haplotypes in a Kurdish population sample. *Int J Legal Med* **112**(3), 181–183, 1999.
5. Kirk RL, Keats B, Blake NM, McDermid EM, Ala F, Karimi M, Nickbin B, Shabazi H, Kmet J. Genes and people in the Caspian Littoral: a population genetic study in Northern Iran. *Am J Phys Anthropol* **46**(3), 377–379, 1977.
6. Bendukidze N, Meschishvili E, Ivaskova E, Cerna M, Churadze T, Guladze R, Shalamberidze K. Determination of HLA Class I Alleles in Georgian, Kurdish and Czech Populations by PCR-ARMS. *Human Immunology* **47**(1–2), 54, 1996.
7. Alexeev L, Khaitov R, Boldyreva M, Trofimov D, Guskova I, Akopian A, Fedoseeva T, Unsrisov I, Pospelov L, Bendukidze N, Meschishvili E, Ivaskova E, Cerna M, Churadze T, Guladze R, Shalamberidze K. HLA in some ethnic groups of the former Soviet Union (FSU). In: Genetic diversity of HLA Functional and Medical Implication. Ed. D. Charron 1996, Vol. 1, 364–374.
8. Arnaiz-Villena A, Karin M, Bendukidze N, Gomez Casado E, Moscoso J, Silvera C, Oguz FS, Sarper Diler A, De Pachó A, Allende L, Guillen J, Martinez Laso J. HLA alleles and haplotypes in the Turkish population: relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens* **57**(4), 308–317, 2001.

Population: Tamil from Natal Province, South Africa

Authors: M. G. Hammond, D. Anley

Authors' address: Transplantation Unit. National Blood Transfusion Services Pinetown, South Africa

The founding of an Asian Indian community in South Africa began in 1860 when a considerable number of Indians were brought to South Africa to work as contract laborers on the sugar farms in Natal. There are today more than a million Indians in the province of Natal. They can be divided into two groups; those originating from southern India (nearly all of the Hindu religion) can be further subdivided by language into Tamil or Telegu. The northern Indians can be divided by religion into Hindu or Muslim. Roughly half the Indians in Natal had their origins in southern India and half from northern India. The majority (about 80%) spoke Tamil.

Samples from 50 random, unrelated Tamil blood donors were typed at the HLA-A, B, and C loci using IHWG Reverse Line Strips. The allele frequencies observed in this dataset are consistent with those observed in previous studies of this population.

Population: Turkish from Marmara, Turkey

Authors: G. Saruhan-Direskeneli, F.A. Uyar

Authors' address: I.U. Istanbul Medical Faculty Department of Physiology, 34093 Istanbul, Turkey

The Turkish population sample has been collected from the region of Marmara including Istanbul, Turkey's largest city, where considerable inward population movements from the remainder of the country have occurred. The population is

from Anatolia (Asia Minor), which is the major part of Turkey.

Anatolia's first known inhabitants appeared in the Mediterranean region as early as 7500 BC. Beginning with the Hittites from the Middle Bronze Age, Persians, Alexander the Great, and Romans ruled Anatolia. Thereafter Byzantine Empire, with Constantinople (aka, Istanbul) as its capital, reconquered Italy, the Balkans, North Africa and also Anatolia. The Great Seljuk Turkish Empire of the 11th century was the first to rule what is now Turkey, Iran and Iraq. Later, the Ottoman Empire ruled Anatolia from 13th to 19th century. In the late 19th and early 20th century, various ethnic populations regained independence. The Turkish Republic was founded in Anatolia in 1923.

Ethnographically, Turks speak Turkic languages, are Muslims by religion, live almost entirely in the western half of Asia and fall within the Arabic, and to some extent the European, spheres of influence. Turkish-speaking populations were apparently quite nomadic and adventurous. Starting from the central regions of China, they reached India, Algeria and as far as Vienna in Europe. Since the first appearance of Seljuk Turks in Anatolia, intermarriage and conversions have taken place to such a large extent that the modern Turks are almost entirely European in appearance, and the contemporary Turkish population can be described as the end-result of centuries-long genetic admixture between the ancestral Turks, native inhabitants of Anatolia and all other populations Turks have had close contact during the expansions and contractions of Ottoman Empire.

The number of the ethnic groups that take part in today's Turkish population is estimated to be around 50. Based on

geographical relations in Anatolia, the major ethnic groups are Turks, Kurds, Circassians, Laz people of the northern coast, Caucasians, Georgians, Bosnians, Armenians and Albanians. The majority of these ethnic groups have lost their ethnic identity within the unity of Anatolia. However, there are some who still continue to preserve and nurture their identities, traditions and language (1, 2).

Turkish samples were collected from 250 unrelated healthy donors partially from the blood bank, as a control group for disease association studies (3). It is assumed that Istanbul's central role in Turkey qualifies the population of this city as representative of the entire country.

Based on the present population data, genetic associations of autoimmune diseases such as multiple sclerosis, rheumatoid arthritis and insulin dependent diabetes mellitus with HLA class II alleles have shown features in the Turkish population that are comparable to other Caucasian and Mediterranean populations (4, 5, 6). Behçet's Disease, a highly prevalent disease of Turkey, is also strongly associated with HLA class I, in particular HLA-B alleles (7).

HLA-DRB1, -DQA1 and -DQB1 loci were typed using 12th IHWG SSOP reagents. No deviations from expected Hardy-Weinberg proportions were observed. Estimations of neutral expectations under the Ewens-Watterson model have revealed that F_{nd} values at all three loci were consistent with the operation of balancing selection (DRB1: $F_{nd} = -1.50$, DQA1: $F_{nd} = -1.36$, DQB1: $F_{nd} = -1.099$). The evaluation of this data in the context of anthropological setting will bring more information and extend our knowledge of admixture between European and Asian ethnicities.

References

1. Rosser ZH, Zerjal T, Hurles ME, et al. Y-Chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67**, 1526–1543, 2000.
2. Arnaiz-Villena A, Gomez-Casado E, Martinez-Laso J. Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens* **60**, 111–121, 2002.
3. Saruhan-Direskeneli G, Uyar FA, Bakar S, Eraksoy M. Molecular analysis of HLA-DRB, -DQA1 and -DQB1 distribution in Turkey. *Tissue Antigens* **55**, 171–174, 2000.
4. Saruhan-Direskeneli G, Esin S, Baykan-Kurt B, Örnek I, Vaughan R, Eraksoy M. HLA-DR and -DQ associations with multiple sclerosis in Turkey. *Human Immunol* **55**, 59–65, 1997.
5. Istanbul Rheumatology Study Group and Saruhan-Direskeneli G. Shared epitope homozygosity is strongly associated with rheumatoid arthritis in Turkey. *Br J Rheumatol* **37**, 1126–1128, 1998.
6. Saruhan-Direskeneli G, Uyar FA, Bas F, et al. HLA-DR and -DQ associations with insulin-dependent diabetes mellitus in a population of Turkey. *Hum Immunol* **61**, 296–302, 2000.
7. Gül A, Uyar FA, Inanç M, Ocal L, et al. Lack of association of HLA-B*51 with a severe disease course in Behçet's disease. *Rheumatology* **40**, 668–672, 2002.

South-East Asia

Populations: Atayal from Wulai/Chenshih/Wufen, Taiwan

Toroko from Hsiulin, Taiwan

Saisiat from Wufen/Nanchuang, Taiwan

Bunun from Hsin-I/Taitung, Taiwan

Tsou from Tapang, Taiwan

Rukai from Wutai, Taiwan

Paiwan from Lai-I, Taiwan

Ami from Hualien/Taitung, Taiwan Puyuma from Peinan, Taiwan

Tao from Lan Yu, Taiwan

Pazeh from Fengyuan/Puli/Liyutan, Taiwan

Siraya from Tanei/Tsochen, Taiwan

Thao from Yuchih, Taiwan

Minnan, Taiwan

Hakka from Hsinchu/Pintung, Taiwan

Ivatan from Bantanes, Philippines

Authors: C. C. Chu, J. Trejaut, H. L. Lee, S. L. Chang, M. Lin

Authors' address: Medical Research Department, Mackay Memorial

Hospital, Taipei, Taiwan

Introduction

The island of Taiwan lies 160 kilometers off the southeastern coast of China, on the edge of the East Asian continental shelf. It is approximately 36,000 square kilometers, and extends 380 kilometers from north to south and 150 kilometers from east to west at its broadest points. The populations of Taiwan comprise the Minnan and Hakka people (about 90%) (also referred to as the “Taiwanese”), the Mainland Chinese (7.5 %) and the indigenous people (1.8%). The “Taiwanese” are the descendants of early settlers who came 400 years ago from the Fukien and Kwangton provinces on the south east coast of China. These early settlers are speakers of Sino-Tibetan languages and often married with Taiwan indigenous women. The other Mainland Chinese immigrated after World War II and were from many different ethnicity and provinces of China.

The indigenous people of Taiwan inhabited most of the island before the 17th century. They were speakers of Austronesian languages and were classified into “plains” and “mountain” tribes according to their habitats and degree of sinicisation. The indigenous plains tribes lived in the large western plains of the island and gradually disappeared through wars and marriages with the “Taiwanese”. Only a few such tribes have succeeded in preserving their original languages and population identities. Two of the ten plain tribes, the Pazeh and the

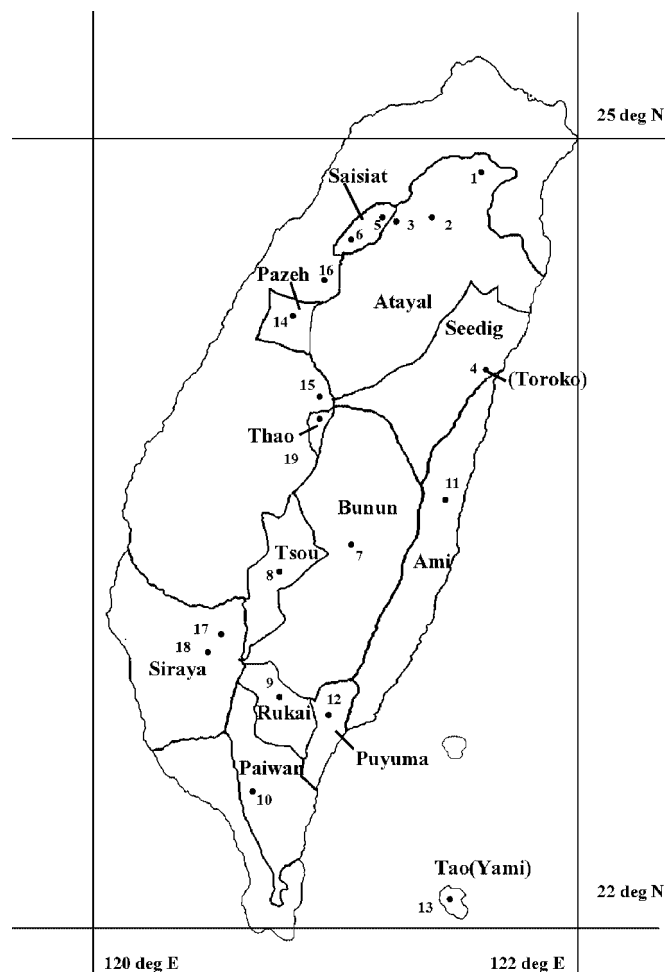


Figure 1. Geographic distribution of the indigenous tribes in Taiwan. The black dots indicate the sampling locations within each indigenous region: Atayal (1-Wulai, 2-Chenshih, 3-Wufen); Toroko (4-Hsiulin); Saisiat (5-Wufen, 6-Nanchuang); Bunun (7-Hsin-I); Tsou (8-Tapang); Rukai (9-Wutai); Paiwan (10-Lai-I); Ami (11-Kuangfu); Puyuma (12-Peinan); Tao (13-Lanyu); Pazeh (14-Fengyuan, 15-Puli, 16-Liyutan); Siraya (17-Tanei, 18-Tsochen); Thao (19-Yuchih).

Siraya, were sampled in this study (Figure 1). “The indigenous mountain tribes” comprised nine tribes. The Atayal, Saisiat, Bunun, Tsou, Rukai, and Paiwan groups reside in the rugged internal mountain ranges, the Puyuma and Ami live along the eastern coast, and the Tao (previously called Yami) live on the remote island of Lanyu off the east coast of Taiwan. The mountain tribes have continuously been isolated from the early Minnan and Hakka people and have therefore retained their distinct cultural, social and genetic heritage till the present day. A few tribes in this study were not well categorised. For example the Toroko are considered to belong to a branch of the Atayal tribe, but the Toroko believe they are members of a more distinct group called the Seediq. Likewise, the Thao tribe has been in contact with the Minnan and Hakka settlers for more than 100

years, and has been considered a subgroup of the Tsou. They have now been recognised as the 10th official indigenous tribe in Taiwan. Based on linguistics, cultural comparisons and oral history, it is believed that the Tao immigrated to the island of Lanyu (Taiwan) from the Batan Islands (Philippines) about 1000 years ago. Batan is a group of 10 small islands covering 88.8 square kilometers, and holds a population of 15000 Ivatans (1990 census).

The peopling of Taiwan and the Austronesian expansion from Taiwan throughout Oceania remains a subject of debate (1–6). Archaeological records have shown that Taiwan was inhabited more than 15000 years B.P. by Palaeolithic cultures that remained along the east coast till 5000 B.P. (7). Subsequently, distinct Neolithic cultures were introduced (<6000 B.P.). These cultures were very similar to the cultures seen in south/central China, insular southeast Asia (Vietnam, Lao and Cambodia) and the Philippines Islands. Moreover, the Neolithic cultures on the east coast of Taiwan differed markedly from the Neolithic cultures of the west plains (8). The relationship between archaeological cultures and those seen in the present day populations remains uncertain. Two models have been proposed to explain the presence of the great degree of cultural diversity in Taiwan. A “population migration model” proposes that intertribal differences are the result of diverse origins for Taiwan’s indigenous populations. An “internal differentiation hypothesis” proposes that the tribes shared a common origin and differentiated through geographic isolation (4, 7). Furthermore, the process by which the genetic structure of present day Taiwan’s indigenous people was acquired (replacement from demic diffusion, admixture, or conservation of the ancestral gene pool) remains to be determined.

HLA studies were previously performed on these populations using serological methods for the HLA-A, B and Cw loci (6,9) and using PCR-DNA based techniques for the HLA-A and DRB1 loci (10). The antigen distribution differed between the tribes, but a high degree of genetic homogeneity was generally observed within the tribes. These distributions were associated with both northern and southern Asian origins, and particularly to Oceanian origins. The distribution profiles seen within the tribes of the central mountain ranges and the tribes of the east-coast of Taiwan were very different and suggested that the ancestry of the genetic material seen among the Taiwan indigenous tribes must have different origins.

Population samples

As part of the Anthropology/HLA diversity component of the 13th international histocompatibility workshop, HLA typing was tested in the following groups of populations: (I) Minnan

(n=102) and Hakka (n=55); (II) “Indigenous mountain tribes”: Atayal (n=106), Toroko (n=55), Saisiat (n=51), Bunun (101), Tsou (n=51), Thao (n=30), Rukai (n=50), Paiwan (n=51), Ami (n=98), Puyuma (n=50) and Tao (n=50); (III) “Indigenous plain tribes”: Pazeh (n=55) and Siraya (n=51). (IV) Ivatan (n=50). All samples were obtained from unrelated individuals whose parents belonged to the same ethnic population. The Minnan people were sampled from individuals enrolled for paternity testing at the Mackay Memorial Hospital; the Hakka samples came from two regions (29 individuals from Hsinpu in the Hsinchu county, and 26 from Neipu in the Pingtung county). Most of the indigenous people were sampled on their tribal locations, but part of the Ami, the Bunun and the Atayal were sampled in city centers. The geographic locations and sampling sites of the indigenous populations studied are shown in figure 1.

Methods and Analysis

Samples were typed at the HLA-A, C and B loci using the Reverse Line Strip (RLS) typing system (Roche Molecular Systems) as described in the IHWG protocol. When genotyping was difficult to obtain, assignments were referred back to previous serological results. For some HLA-B assignments, sequence based typing (SBT) was also performed. All DRB1 typing were performed using SBT on exon 2 with an in-house protocol and primer sets (10). Nei’s Da genetic distance between the studied populations was calculated and a dendrogram was constructed using the neighbor-joining method with bootstraps test using Poptree computer program.

Results and Discussion

Hardy-Weinberg equilibrium test

The exact test of Guo and Thomson was used to evaluate deviation from expected Hardy-Weinberg equilibrium proportions (HWE). In the Ivatan population, only HLA-DRB1 deviated from HWE ($p=0.0147$). Analysis of the Ivatan population revealed an excess of DRB1*1502 homozygous genotypes. No other alleles were found after re-testing all DRB1*1502 homozygous samples with a commercial RLS-SSO kit (Dynal). Five HLA-A-B-DRB1 haplotypes containing DRB1*1502 were seen with frequency greater than 3%. They were A*3401-B*4002-DRB1*1502 (7%), A*0201-B*1801-DRB1*1502 (5%), A*1101-B*1301-DRB1*1502 (4.8%), A*2402-B*4801-DRB1*1502 (4%), and A*2407-B*3505-DRB1*1502(3%). Consequently, directional selection could be operating on DRB1*1502 in the Ivatan population. How-

ever, the deviation from neutrality expectation under the Ewens-Watterson model ($F_{nd}=0.1317$) was not significant ($p=0.6658$), and a deviation from the HW proportion caused by sampling, inbreeding or admixture could not be excluded.

Neutrality test for selective forces

HLA-A*2402 was the most common allele seen among the Taiwan indigenous tribes. Frequencies ranged from 44.5% in the Toroko to 86.3% in the Paiwan. Normalized F values (F_{nd}) greater than 0.5 were seen in the Paiwan, Tsou, Puyuma and Rukai populations ($F_{nd}=2.17, 0.87, 0.86$ and 0.68 respectively). Of these, F_{nd} in the Paiwan approached significance ($p>0.95$, two-tail test). In the Paiwan, the following six HLA A-Cw-B haplotypes containing A*2402 had a frequency greater than 6%: A*2402-Cw*0304-B*1301 (25.5%), A*2402-Cw*0304-B*4001 (12.3%), A*2402-Cw*0303-B*4002 (9.8%), A*2402-Cw*0801-B*4801 (7.8%), A*2402-Cw*0401-B*4001 (7.8%), and A*2402-Cw*0702-B*4001 (6.8%). The high F_{nd} value seen in the Paiwan ($p>0.95$) appeared to be associated with A*2402 (86.3%), and may indicate that directional selection or a founder event with strong genetic drift has operated on the Paiwan tribe in the past. Likewise, high F_{nd} values were previously reported in Papua New Guinea highlands (PNG) and Nusa Tenggara populations ($F_{nd}=1.0$ with $p<0.3$, and $F_{nd}=0.5$ with $p<0.3$ respectively) (11). Interestingly, these high values were also associated with a high frequency of A*2402. In addition, as was observed for the PNG highlanders, most of the HLA-B antigens seen in Taiwan carry the Bw6 epitope; only B*1301 and B*2704 antigens carry the serologically weak form of the Bw4 epitope. Conversely, the A*2402 molecule carries the common form of the Bw4 epitope. Since the Bw4 epitopes are shared by the HLA-B and HLA-A, and the Bw6 epitope is shared with HLA-B and HLA-Cw (12), it is possible that overdominance selection (heterozygote advantage) has operated on the Bw4 and Bw6 epitopes. This may be the case with the distribution of HLA-B alleles, where the Bw4 epitope is apparently deficient but is actually complemented by the presence of the A*2402 molecule carrying the Bw4 epitope.

HLA-A:

A total of 17 different alleles were seen within the populations studied. The diversity of HLA-A was higher in the Minnan and the Hakka than in Taiwan's indigenous mountain tribes. The number of allele types with frequency greater than 1% was 5-6 in most indigenous mountain tribes, and 11 in the Minnan and Hakka. HLA-A2, A11 and A24 were the most

common antigens in Taiwan. HLA-A*2402 was the most frequent allele found in the indigenous mountain tribes. While HLA-A*0201, *0203, *0206 and *0207 were mostly seen among the Minnan and Hakka, HLA-A*0201 and HLA-A*0206 were commonly seen in the mountain tribes. Two HLA-A11 subtypes were observed; HLA-A*1101 was mostly common in the Minnan and Hakka, but HLA-A*1102 had higher frequencies in the Saisiat, the Ami, the Puyuma and the Pazeh. Alleles with noticeable differences were HLA-A*3303, which was seen only in the Minnan, the Hakka and the two indigenous plain tribes, and HLA-A*2601, which was seen in the mountain tribes with high frequency, but which was not found in the Ami or the Tao. HLA-A*3401 was restricted to the eastern coastal tribes (Ami and Puyuma). Finally HLA-A*2407 was only seen in the Tao and the Ivatan.

HLA-Cw

Twenty alleles were detected using RLS. HLA-Cw*0102, *0303, *0304, *0401, *0702, *0801 were common in all populations. Cw*0302 was not seen in the mountain tribes. Cw*0403 was common in most indigenous tribes but was rare in the Minnan and Hakka. Among the indigenous tribes, Cw*1202 was distributed among the northern tribes and Cw*1502 in the southern tribes. A new Cw variant (Cw*08012) was found in the Puyuma tribe with a frequency of 5% (13).

HLA-B

A total 42 HLA-B alleles were detected. The number of alleles observed with a frequency greater than 1% was: 5 in the Toroko, 6 in the Saisiat, 7 in the Ami, 8 in the Bunun, the Rukai and the Tao, 9 in the Atayal and the Tsou, 10 in the Puyuma and the Thao, 15 in the Pazeh, the Siraya and the Minnan and 18 in the Hakka. More polymorphism was observed in the Minnan, the Hakka and the indigenous plain people than in the indigenous mountain tribes. All alleles found in the mountain tribes were present in the Minnan and Hakka populations. Among the indigenous mountain tribes, the most common antigens observed were B13, B15 (B62 and B75), B27, B38, B39, B40 (B60 and B61), B48, B55 and B56, of which HLA-B*4001 and B*4801 were solely distributed. HLA-B*1525 occurred with high frequency in the Tao (11.7%). This allele was commonly seen in the mountain and plains tribes, and was rare among the Minnan and Hakka people. HLA-B*1502 and B*3802 were found in the Tao, the Puyuma and the Ivatan but were absent in other Taiwan mountain tribes. Conversely, B*3901 and B*5502 were widely distributed among the indigenous people from the central mountain ranges but were absent in the Tao. The distri-

bution of some alleles was mostly confined to populations of the same neighborhood. For example, B*1301 was frequently seen in the southern region of Taiwan: the Bunun, the Tsou, the Rukai, the Paiwan, the Puyuma, the Thao, the Pazeh and the Siraya. HLA-B*1501 was also restricted to the southern regions (Rukai and Paiwan). HLA-B*4002 was detected in the Rukai, the Paiwan and the Siraya with high frequencies that decreasing northwards. In the north HLA-B*2704 was commonly seen in the Atayal and the Pazeh. HLA-B*5502 was found with high frequencies in the Atayal, the Bunun and the Tsou, with a southward decline. Finally HLA-B*5601 was frequent in the Ami, common in the Bunun, and the frequency declined towards the south and the north. HLA-B13, B15, B60, B46 and B58 were common in the Minnan and the Hakka people. Interestingly, B*4601 and B*5801 were the most frequently seen alleles in the “Taiwanese”, and were rare or absent in the indigenous mountain tribes. B*5101, B*5102 and B*5401 were commonly found in the Minnan and Hakka people but were rarely seen among the mountain indigenous people. Finally, the B15 antigen was the most polymorphic group, as it comprised 10 allele types. HLA-B*1501 and B*1502 were the main subtypes seen in the Minnan and Hakka populations.

HLA-DRB1

Sequence based typing identified 31 alleles. As for the HLA-B locus, the Minnan, Hakka, and indigenous plain tribes had more diversity than the indigenous mountain tribes. Among the mountain tribes, 3-5 alleles were sufficient to represent more than 60% of the total diversity. Overall the most common alleles were HLA-DRB1*0403, *0404, *0405, *08032, *09012, *11011, *12021, *1401, *1405, *15011, *15021 and *16021. Of these, DRB1*09012, *11011 and *1401 were the most widely distributed. Very distinct DRB1 distributions were obtained from the central mountain and east coast tribes. DRB1*0404 and *0405 were seen in the Ami tribe with high frequencies, but were rare or absent in the central mountain tribes. Similarly, HLA-DRB1*1502 was rare in the central mountain tribes and had high frequencies among the east-coast tribes. Conversely, DRB1*08032 was frequently seen among the central mountain tribes and was rare or absent in the east-coast tribes. HLA-DRB1*03011 and *1302 were common in the Minnan and Hakka, and were seldom seen in indigenous mountain tribes.

Population dendrogram

Nei's Da genetic distance was calculated from the allele frequencies of HLA-A, Cw, B and DRB1. The neighbor-joining

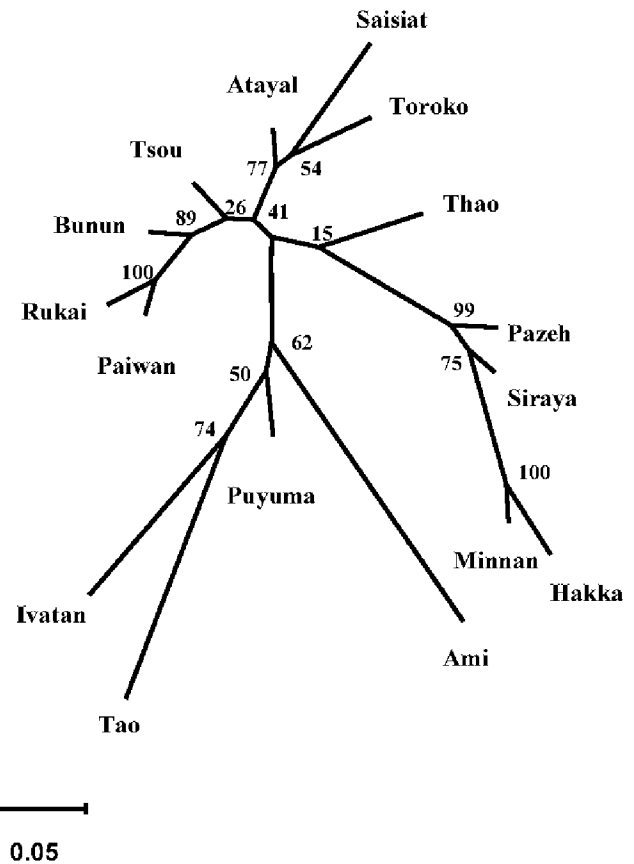


Figure 2. Neighbor-joining population dendrogram of 15 Taiwan populations and the Ivatan in the Philippines. Bootstrap values are indicated before the nodes.

(NJ) method was used to construct the population dendrogram (Figure 2). Four clusters were obtained: (I) The first cluster comprised the “Taiwanese” and the indigenous plain tribes. The genetic affinity between the Minnan, the Hakka, the Pazeh and the Siraya supported a historical interpretation suggesting population admixture over the past 400 years. (II) The second cluster comprised the indigenous tribes of the northern mountain ranges (Atayal, Toroko and Saisiat). The inclusion of the Toroko in this cluster substantiates previous anthropological hypotheses proposing that the Toroko tribe was a branch of the Atayal group. (III) The south and central mountain ranges cluster consist of the Bunun, Rukai, and Paiwan, with a more distant inclusion of the Tsou (bootstrap value 26%). (IV) Finally, the east-coast cluster (bootstrap value 62%) was dispersed, and included the Ami, Puyuma and Tao tribes along with the Ivatan. The Ivatan were closer to the Tao tribe than to Taiwan's other indigenous tribes.

References

1. Bellwood P. The Austronesian dispersal and the origin of languages. *Sci Am* **265**, 88–93, 1991.
2. Oppenheimer S, Richards M. Fast trains, slow boats, and the ancestry of the Polynesian islanders. *Science Progress* **84**, 157–181, 2001.
3. Melton T, Peterson R, Redd A, et al. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am J Hum Genet* **57**, 403–514, 1995.
4. Melton T, Clifford S, Martinson J, Batzer M, Stoneking M. Genetic evidence for the proto-Austronesian homeland in Asia: mtDNA and Nuclear DNA variation in Taiwanese aboriginal tribes. *Am J Hum Genet* **63**, 1807–1823, 1998.
5. Capelli C, Wilson J, Richards M, et al. A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular southeast Asia and Oceania. *Am J Hum Genet* **68**, 432–443, 2001.
6. Lin M, Chu CC, Lee HL, et al. Heterogeneity of Taiwan's indigenous population: possible relation to prehistoric Mongoloid dispersals. *Tissue Antigens* **55**, 1–9, 2000.
7. Tsang CH. One hundred years of Taiwan archaeology. In: The 17th congress of Indo-Pacific prehistory association. Academia Sinica, Taipei, Taiwan: 2002: 285–291.
8. Kano T. Outline review of the Taiwan archaeology and ethnology. Taipei: The Historical Research Commission of Taiwan, 1955.
9. Lin M, Chu CC, Chang SL, et al. The origin of Minnan and Hakka, the so-called "Taiwanese", inferred by HLA study. *Tissue Antigens* **57**, 192–199, 2001.
10. Chu CC, Lin M, Nakajima F, et al. Diversity of HLA among Taiwan's indigenous tribes and the Ivatans in the Philippines. *Tissue Antigens* **58**, 9–18, 2001.
11. Bugawan TL, Mack SJ, Stoneking M, Saha M, Beck H, Erlich H. HLA class I allele distribution in six Pacific/Asian populations: evidence of selection at the HLA-A locus. *Tissue Antigens* **53**, 311–319, 1999.
12. Marie A, Ennis P, Parham P, Holmes N. The primary structure of HLA-A32 suggests a region involved in formation of the Bw4/Bw6 epitopes. *J Immunol* **137**, 3671–3674, 1986.
13. Chu CC, Lee HL, Trejaut J, Lin M. A new HLA-Cw allele found in one of Taiwan indigenous tribes with a frequency of 5%. *Tissue Antigens* **59(Suppl.)**, 51, 2002.

Population: Chinese, Malay and Thai from Singapore

Authors: Steven J, Mack^{1,2}, Andrea J. Jani¹, Jennifer E. Below^{1,2}, Nilmani Saha³, Henry A. Erlich^{1,2}

Authors' addresses: ¹Children's Hospital Oakland Research Institute, Oakland, CA, ²Roche Molecular Systems, Berkeley, CA, ³Department of Molecular Medicine, University of Pittsburgh, Pittsburgh, PA

Singapore is an island-nation located at the southern tip of the Malay Peninsula (1 degree 24 minutes north latitude, 103 degrees 51 minutes east longitude) that has been a center of trade and commerce for at least the last 1700 years. For much of its history, the peninsula has been a crossroads for trade between the Middle East, India, China and Europe, and Indian and Chinese groups have been present in the peninsula (and on Singapore) since at least the 3rd and 14th centuries, respectively. In its early history, the region was in contention between Siam (modern Thailand) and Java. Along with Siam, the Portuguese, Dutch and British vied for control of the region through the 16th, 17th and 18th centuries. Britain established a trading port in Singapore in 1819. The island remained under (primarily) British control until the middle of the 20th century, and Singapore has been independent from Malaysia since 1965 (1, 2).

As of 2003, Singapore's resident population numbered 3.4 million individuals. At the time of the 2000 census, this population was 77% Chinese, 14% Malay, 8% Indian, with 1% comprised by other ethnic groups. These proportions were roughly the same in 1901, when population growth

was largely due to immigration from China and India. Historically, individuals from other Asian nations, Europe, the Middle East and Africa have been present in Singapore at low frequencies (1%–3% of the population). The predominant language spoken in Singaporean homes is Mandarin, (spoken by 35% of the population) followed by other Chinese dialects (e.g., Hokkien, Chiu Chow, Hainanese, Cantonese, and Hakka) (24%), English (23%), Malay (14%) and Tamil (3%) (3–5). The variety of Chinese dialects spoken suggests Singapore's Chinese population is ultimately derived from a variety of coastal regions of southern China.

The Malay people are native to the Malay Archipelago (the Malay Peninsula, Singapore, Borneo, Sumatra, Java and the Philippines), and those Malay people native to southern Thailand, the Malay Peninsula, Singapore, eastern Sumatra and the Riau Islands are sometimes known as Riau Malays. The Malay language is an Austronesian language. The Malay individuals sampled here are citizens of Singapore, and the class II data included here for the Malay population have been reported previously (6).

The Thai people are thought to have originated in southwestern China perhaps as long as 3000 years ago, migrating south into the region of modern-day Thailand from that time until the 13th century. The Thai language is part of the Tai-Kadai branch of the Austric language family (5). The Thai individuals sampled here live and work in Singapore.

In this study, subsets of 93 Singaporean Chinese individuals were typed at the HLA-A and B loci, subsets of 131 Singa-

porean Malay individuals were typed at the HLA-A, C, B, DRB1, DQB1, and DPB1 loci, and subsets of 99 Singaporean Thai individuals were typed at the HLA-A, C and B loci. All sampled individuals were unrelated. Class II genotypes were generated using PCR-SSOP methods (6) and class I genotypes were determined using the IHWG RLS typing system. These data were analyzed by the IHWG Biostatistics core.

A deviation from expected Hardy-Weinberg equilibrium proportions was observed at the HLA-A locus in the Singaporean Chinese population sample (p -value=0.0320). A low, but non-significant, value of homozygosity (as measured by the Ewens-Watterson homozygosity statistic, F) suggestive of the action of balancing selection, was observed in the Malay population at the HLA-C locus (F =0.0968, p -value=0.0406).

In general, genetic diversity (h) was lower at the HLA-A locus than at the C and B loci in all populations. A-locus diversity was 0.86 in the Chinese and Malay and was 0.87 in the Thai population. In contrast, C-locus diversity was 0.90 in the Malay and 0.91 in the Thai population, and B-locus diversity was 0.93 in the Chinese and Thai populations, and 0.95 in the Malay. Diversity levels at the class II loci (DRB1, DQB1, DPB1) were similar to the A-locus diversity in the Malay population, and ranged from 0.82 at DQB1 to 0.87 at DPB1.

Similar high-frequency (allele frequencies (a.f.) greater than 0.06) HLA-A alleles were observed in all three populations, and of these, the A*1101, *2402, and *3303 alleles

were at high frequencies in all three groups. A*1101 was the highest frequency allele in each population (a.f.=0.26 in the Thai, 0.22 in the Chinese, and 0.18 in the Malay).

Of the high-frequency HLA-C alleles in the Malay and Thai populations, only Cw*0801 and *0701 (a.f.=0.19 and 0.08 in the the Malay and Thais respectively, and 0.07 in both) were observed at high frequencies in both groups. Similarly, B*1502 was the only high-frequency HLA-B allele at a frequency greater than 0.06 in all three groups (a.f.=0.12 in the Chinese, 0.08 in the Malay, and 0.06 in the Thai population).

The Thai and Malay populations share several high-frequency (haplotype frequency (h.f.) greater than 0.03) HLA-A:B haplotypes. A*3303:B*5801, *1101:*1301 and *3303:*4403 were seen in both groups at frequencies ranging from 0.03 to 0.07. Of these, only A*3303:B*5801 was observed at high frequencies in the Chinese population (h.f.=0.05). The A*0207:B*4601 haplotype was seen in both the Thai and Chinese populations at high frequencies (h.f.=0.10 and 0.08, respectively). A*1101 alleles in the Chinese population were observed in different haplotypes (A*1101:B*4001 and *1101:B*1502 (a.f.=0.06 and 0.05)) from A*1101 alleles in the Malay and Thai populations. In addition, the A*3303:B*4404 haplotype common in the Malay and Thai populations was observed in only one Chinese individual. These patterns of haplotype sharing suggest that there has been little admixture between the Chinese and Malay (or Thai) populations in Singapore.

References

1. Jayapal, Maya. Old Singapore. Singapore: Oxford University Press, 1992.
2. Chew, Ernest CT, Edwin Lee (ed.). A History of Singapore. Singapore: Oxford University Press, 1991.
3. Singapore Census of Population 2000, Statistical Release 1 – Demographic Characteristics. Singapore Department Of Statistics, Singapore, 2001.
4. Singapore Census of Population 2000, Statistical Release 2 – Education, Language and Religion. Singapore Department Of Statistics, Singapore, 2001.
5. Comrie B, Matthews S, Polinsky M. The Atlas of Languages: The Origin and Development of Languages Throughout the World. London: Quarto Publishing plc, 1996.
6. Mack SJ, Bugawan TL, Moonsamy PV, Erlich JA, Trachtenberg EA, Paik YK, Begovich BA, Saha N, Beck HP, Stoneking M, and Erlich HA. Evolution of Pacific/Asian populations inferred from HLA class II allele frequency distributions. *Tissue Antigens* **55**, 383–400, 2000.

Population: Southern Han Chinese from People's Republic of China

Authors: E. Trachtenberg¹, M. Vinson¹, E. Hayes¹, M. Hsu¹, K. Houtchens¹, Y. Hsia², H. A. Erlich^{1,3}, W. Klitz^{1,4}

Authors' addresses: ¹Children's Hospital and Research Center at Oakland, CA, ²Dept of Genetics, University of Hawaii, Honolulu, HI ³Dept of Human

Genetics, Roche Molecular Systems, Alameda CA, ⁴School of Public Health, University of California, Berkeley, CA

The Chinese people are composed of one major group, the Han, comprising ~94% of the total population, and 55 minor nationality groups. Although originally found in the Yellow River area, the Han soon migrated throughout China.

The Han originated as the “Xia” population during the Xia dynasty (2205-1766 BC) and then were called the Han during the Han dynasty (206 BC-220 AD). They eventually expanded into the Huaxia population after the Shang dynasty (1766-206 BC). Dendograms from previous studies on class I and II HLA antigens from a large number of Chinese populations divide the Chinese broadly into Northern and Southern Chinese groups (1). The Northern Chinese have historically had more contact with Central Asians, Siberians, and Caucasians. The Southern Chinese are more similar to Southeast Asian populations, which have similar HLA distributions, with the frequencies of shared antigens diminishing as the geographical distance from China increases. In this study, HLA class I (HLA-A, -B and -C) allelic distributions and haplotypes were examined in individuals of the Han group from the Canton region of Southern China.

Blood samples from 270 individuals from Southern (Canton) China were collected in association with the University of Hawaii, John Burns School of Medicine, as part of a study to analyze genetic factors involved in the hereditary anemias. None are presumed to be first-degree relatives. Genomic DNA was extracted and the samples typed at the HLA class I (HLA-A, B, C) using the IHWG RLS reagents, to analyze variability in the hypervariable regions of the second exon and third exons of the class I loci. Hardy-Weinberg equilibrium (HW) was tested with the exact test of Guo and Thompson (2), and

class I haplotypes were estimated using the iterative Expectation-Maximization (EM) algorithm of Excoffier and Slatkin (3).

The most common alleles at the HLA-A locus included A*1101 and *2402, together comprising ~45% of the total allele frequencies, and both are common to other Asian groups. A slight deviation from expected HW proportions was found at HLA-A due to a small excess of A*1101:3303 genotypes. At the HLA-C locus, Cw*0702 and *0102 are the most common alleles, and a slight deviation from HW was found due to minor overall genotype deviations, with no particularly obvious allele or genotype deviation. The diversity in the HLA-B locus alleles was striking and in HW expected proportions. Of the 51 alleles found, HLA-B*4001 and B*4601 were the two most common alleles, and are common in other Asian populations. The three most common A-B-C haplotypes A*3303-B*5801-Cw*0302, A*0207-B*4601-Cw*0102, A*1101-B*1502-Cw*0801 are common to other Asian groups as well. The overall deviations from expected HW proportions, although nominally significant by this test, were small given the relatively large sample size (270). The reason for the slight deviation may be because the population as ascertained is not from a general population but from a population that was drawn to participate in a hereditary anemia program.

References

1. Lee TD, Zhao TM, Mickey R, Sun YP, Lee G, Song CX, Cheng DZ, Zhou S, Ding SQ, Cheng DX, et al. The polymorphism of HLA antigens in the Chinese. *Tissue Antigens* **32**(4), 188, 1988.
2. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**(2), 361, 1992.
3. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* **12**(5), 921, 1995.

Population: Chinese from Hong Kong, People's Republic of China

Authors: D. Middleton¹, B.R. Hawkins², F. Williams¹

Authors' addresses: ¹Northern Ireland Histocompatibility and Immunogenetics Laboratory, City Hospital, Belfast, Northern Ireland, ²Department of Pathology, The University of Hong Kong, Hong Kong

Hong Kong is located to the south of Mainland China and has a population exceeding 6 million people, more than 98% of whom are Chinese. Of these, the majority were either born in Guangdong Province in southern China or are the progeny

of people born there. Politically, Hong Kong was established as a Special Administrative Region of the Peoples' Republic of China on 1 July 1997 after more than 100 years of British colonial rule.

The HLA class I and II distribution of Chinese in Hong Kong was reported previously (1) based on a large sample of bone marrow donors on the Hong Kong Marrow Match Foundation donor registry. However, the study by Chang and Hawkins (1997) was based predominantly on serological testing for class I specificities.

The Hong Kong Marrow Match Foundation maintains a registry of potential unrelated bone marrow donors primarily

of southern Chinese origin. EDTA blood was collected from a total of 592 consecutive new donors joining the registry and the samples were tested for HLA-A and B alleles. Of the initial 592 samples, 572 were selected for detailed analysis, 20 having been excluded from the study because the surnames were non-Chinese or could not be distinguished from western surnames, or because the subject's Hong Kong identity card number indicated a non-Chinese origin.

HLA typing was performed by SSOP (2, 3).

With respect to HLA-A, the allelic distribution was not significantly different from that obtained in a previous study of Singapore Chinese (2), with 23 alleles detected in the Hong Kong group and 22 alleles in the Singapore Chinese. Nearly

twice as many HLA-B alleles were detected in the Hong Kong group ($n=60$) compared to the Singapore Chinese group ($n=35$) (3), with the greatest variation detected within the HLA-B*15 allele family.

At the HLA-A locus, 23.7% of donors were homozygous at the allele family level (e.g. genotype 02, 02) and 7.3% of donors were heterozygous at the allele level (e.g. genotype 0201, 0203), resulting in a true HLA-A homozygosity rate of 16.4%. Most of the heterozygosity was contributed by combinations of A*02 alleles. At the HLA-B locus, 10.9% of donors were homozygous at the allele family level and 3.0% were heterozygous at the allele level. Most of the HLA-B heterozygosity related to combinations of B*15 alleles.

References

1. Chang YW, Hawkins BR. HLA class I and class II frequencies of a Hong Kong Chinese population based on bone marrow donor registry data. *Human Immunology* **56**, 125–135, 1997.
2. Williams F, Meenagh A, Maxwell AP, Middleton D. Allele resolution of HLA-A using oligonucleotide probes in a two stage typing strategy. *Tissue Antigens* **54**, 59–68, 1999.
3. Middleton D, Williams F, Hamill MA, Meenagh A. Frequency of HLA-B alleles in a Caucasoid population determined by a two-stage PCR-SSOP typing strategy. *Human Immunology* **61**, 1285–1297, 2000.

Population: Kinh from Hanoi, Vietnam

Authors: An-Vu-Trieu¹, Chinh Tran-Thi¹, Thuy Ngyuyen-Thanh¹, S. Djoulah², I. Le Monnier de Gouville², J. Hors², A. Sanchez-Mazas³

Authors' addresses: ¹Dept of Immunology Medical College of Hanoi Viet Nam, ²LNH, Dept of Immunology Saint Louis Hopital Universite Paris VII, France, ³Laboratory of Anthropology, Genetics and Peopling History, University of Geneva, Switzerland

Vietnam encompasses a region of the Indochina peninsula of South-East Asia (SEA) ranging from 8.34 to 23.5 degrees north latitude and 102.2 to 109.5 degrees east longitude. In the 1999 general census, the Vietnamese people numbered 76.3 million. The majority of this population, approximately 85%, belongs to the Viet Kinh ethnic group (abbreviated here as Kinh), with the remaining 15% representing about 54 minorities living throughout the territory.

The origin of the Vietnamese population is still controversial. According to legend, the old Vietnam territory spread farther to the North into the delta of the Blue River (Yangzi). Current hypotheses regarding the origins of the Kinh people hold that they are autochthonous, having lived for a very long time in the Red River delta. Archaeological data reveals the succession of several civilizations, among which were the an-

cient Hoa Binh (12,000 BC) and the Phung Nguyen, or Red River, civilization (5,000 BC) from which modern Vietnamese are believed to originate. Historically, the region of modern Vietnam was under Chinese control for a period of almost 10 centuries (from 221 AD to 968 BC). The main religion in modern-day Vietnam is Buddhism. Culturally, Vietnam ascribes to a modernized form of Confucianism (1).

From a linguistic point of view, the Vietnamese (or Kinh) language belongs to the Mon Khmer branch of the Austroasiatic family (3, 4, 5). Mon-Khmer languages are also spoken in Laos, Cambodia, Thailand, Burma, Malaysia, the Nicobar Islands, and one such language is spoken in northeast India (4).

Blood samples from 103 unrelated Kinh donors, mostly from areas surrounding Hanoi, were collected at the Medical College of Hanoi and were typed at the DRB1 and DQB1 loci using INNOLIPA PCR-SSO kits at the Saint-Louis Hospital in Paris, France.

Previous results obtained for this population sample (6) indicate that the two most frequent DRB1 alleles observed are *1202 (29.4%) and *0901 (13.6%), and the two most frequent DQB1 alleles, *0301 (34.3%) and *03032 (16.2%). A significant deviation from HWP is observed for DQB1. Three haplotypes account for 49% of total haplotype frequencies: 1202-0301 (29.2%), 0901-03032 (13.8%), and 1502-0501 (6%). In comparison to 13 other East Asian populations, the

Kinh have some of the lowest values of heterozygosity for the DRB1 and DQB1 loci. Therefore, they represent one of the most homogeneous populations in this geographical area, which is in agreement with the observation that they exhibit three DRB1-DQB1 haplotypes at high frequencies. According to estimated genetic distances among the 13 populations, they are closest to the Thai, for DRB1, and unexpectedly, to the Manchu from Northern China, for DQB1 (very high *0301 and *0303 frequencies).

According to IHWG biostatistical analyses (7), no locus-wide deviations from expected Hardy-Weinberg equilibrium proportions (HWP) are observed in these data. However, in agreement with previous work, DQB1*0301 homozygotes are observed in significant excess (p -value=0.0482), which may reflect an undetected DQB1 allele.

Preliminary comparisons of Kinh DRB1 allele frequency distributions with those of neighboring populations studied in the present workshop indicate that the Kinh are undifferentiated, genetically, from the Malay, and are close to some Sino-Tibetan-speaking populations (southern Han, Taiwanese, Minnan and Hakka from Taiwan), Austronesian-speakers from Taiwan (Paiwan, Bunun), and Manchu. At the DQB1 locus, the *0301 allele, which reaches a frequency of 37% in the Kinh, is not observed in excess of 25% in neighboring populations (e.g., Chinese), and the Vietnamese are observed to be close, genetically, to some Han Chinese populations. However, given that an excess of DQB1 *0301 homozygotes is observed, relative to expected HW proportions, the actual frequency of this allele may be lower.

References

1. Tan Ha-van. Qua trinh hinh thanh van hoa Viet Nam. Chuong trinh KX 06.
2. Frelus M. Vietnamien et protoViet-Muong 1975 *Asie du Sud Est* 6,2.
3. Haudricourt AG. La place du Vietnamien dans les langues austroasiatiques. 1953 *Bulletin de la Societe de Linguistique de Paris* 49-122.
4. Ruhlén M. A guide to world language, Vol. 1. Classification. Edward Arnold London, 1987.
5. An Vu-Trieu, Chinh Tran-Thi, Thuy Nguyen-Thanh, Djoulah S, Le Monnier De Gouville I, Hors J, Sanchez-Mazas A. HLA-DR and DQB1 DNA polymorphisms in a Vietnamese Kinh population from Hanoi. *Europ J Immunogenetics* **24**, 345–356, 1997.
6. Results of data analysis performed on the “VNHAN-VietKinh population” by Biostatistics core of the IHWG. See Appendix C2.

Population: Muong from HoaBinh province, Vietnam

Authors: An-Vu-Trieu¹, Khoa Pham-dang¹, Huy Ho-quang¹, M Busson², AM Bouteiller², H Bleux², D Charron², J Hors³

Authors' addresses: ¹Dept of Immunology Medical College of Hanoi VietNam, ²INSERM U396 Saint Louis Hospital Paris France, ³LNH, Dept of Immunology Saint Louis Hospital Universite Paris VII, France

The Viet Muong (abbreviated here as Muong) are a minority ethnic group of Vietnam. While some Muong populations can be found in some provinces of north Vietnam, the majority of Muong live in the region of HoaBinh (70 km from Hanoi) at approximately, 20 degrees 5 minutes north latitude and 105 degrees 5 minutes east longitude. The general census of 1995 determined this group to number about 874,000 individuals.

Current hypotheses regarding the origins of the Muong hold that they are closely related to the majority population of Viet Nam, the Viet Kinh. Linguistically, the Viet and the

Muong languages are thought to be derived from Proto Viet-Muong, a branch of the Mon-Khmer language (5, 6). Biometric studies have found no significant differences between Muong and Viet Kinh in cranial dimensions (Penrose index $C2H=0.28$. similarity with $p=0.95\%$) (7). Historically, the Muong are believed to have split from the Kinh after the Chinese invasion (2nd century AD to 10th century BC) when indigenous Viet groups retreated into the mountains near the Red River delta. Isolated from Chinese culture, these groups developed into the actual Muong (2, 8). Archaeological evidence characteristic of the Neolithic Viet civilization DongSon (e.g., the renowned copper drum) has been found in this region (2). The Muong practices and customs are similar enough to those of the Kinh that Jeanne Cuisinier (4) has concluded that they were 'cousins'.

Blood samples from 83 volunteer donors were collected in the KimBoi district of Hoa Binh province, and were typed at the DRB1 and DQB1 loci at the Saint Louis Hospital in Paris, France using INNOLIPA PCR-SSO kits.

No deviation from expected Hardy-Weinberg equilibrium proportions was observed at the DRB1 or DQB1 loci; the heterozygosity of these loci in the Muong was higher than in the Kinh dataset submitted to the 13th Workshop (1).

Marked differences were observed between DRB1 and DQB1 allele distributions of the Muong with the Kinh, namely: DRB1*1601 and 1602 ($P_c=10^{-5}$), *1202 ($P_c=10^{-4}$), 1401, 1404 and 1405 ($P_c=10^{-3}$), and DQB1 0301, 0302 and 0303 ($P_c=10^{-10}$) (2).

A nonsignificant positive normalized deviate of homozygosity was observed for DQB1, while a nonsignificant negative value was observed for the DRB1 locus. While these values are suggestive of the weak influence of directional and balancing selection on these loci, they are also consistent with a pattern of neutral evolution (1).

Overall, molecular anthropology has shown many differences between the Muong and the Kinh (as cited above). An attempt to calculate the genetic distance of Arnold (a variant of Edwards) for several South-East Asian populations suggests that the Muong are closest to the DaiLue and Buyi, followed by the Northern Thai, the Viet Kinh and finally the Chinese. Moreover, some studies have shown Glucose 6 Phosphate Dehydrogenase deficiency to be much more frequent in the Muong than in the Kinh population (more than 20% versus less than 1%), although the relative risk coefficient was only 1.08 (9). So, the question of the common origin for the Muong and Kinh people must be dealt with in additional detailed investigations.

References

1. Results of data analysis performed on the "VNHNAN-VietMuong population" by Biostatistics core of the IHWG. See Appendix C2.
2. An Vu Trieu, et al. Distribution of HLA-DRB1 and DQB1 alleles in Muong ethnic of VietNam. *Tạp chí nghiên cứu khoa học 2001* Đại học Y hanoi Viet Nam.
3. Edwards AWF. Distances between populations on the basis of genes frequencies. *Biometrics* **27**, 873–881, 1971.
4. Cuisinier Jeanne. Les Muong. *Geographie humaine et Sociologie*. Paris Institut d'Ethnologie. Musée de l'Homme. Palais de Chaillot, Place du Trocadero, 1948.
5. Pham-Duc-Duong, Ha- van Tan. *Vẻ ngon ngu tieng Viet Muong*. *Dan toc hoc so 1* 1978.
6. Herlious M. Vietnamien et Proto-Viet-Muong. *Asie du Sud-Est et Monde indien (ASEMI)* 1975, 6,21.
7. Nguyen-dinh Khoa. *Quan he Viet Muong*. *Dan toc hoc So 3*, 1976.
8. Ha van Tan. *Qua trinh hinh thanh van hoa Viet Nam*. Chuong trinh KX-06. Ch. III, tr.37-53.
9. Chinh Tran-thi, et al. G-6PD deficiency and HLA class II of Muong and Kinh ethnies in Hoa Binh and Hanoi region. 1st Intern. Workshop on Histocompatibility in Viet Nam November 13, 2002 Hanoi.

Oceania

Population: Filipino from Luzon Island, Philippines

Authors: H. A. Erlich¹, M. Alejandrino², P. Pozzili³, A. Pabelo⁴, T. L. Bugawan¹

Authors' addresses: ¹Department of Human Genetics, Roche Molecular Systems, Alameda, CA, ²Children's Hospital Oakland Research Institute, Oakland, CA, ³University of Rome "La Sapienza", Libera Universita Campus Biomedico, Rome, ⁴Institute of Studies on Diabetes, Manila, Philippines

The Philippine Archipelago borders on the Pacific Ocean, and the Sulu and South China Seas, and has been inhabited for approximately 30,000 years. Extensive contacts with Chinese, Muslim (Borneo), Hindu (Java), and Arab cultures have occurred over the past 1500 years. The Philippine islands were

under Spanish control from the 16th century to the end of the 19th century, when the region came under the control of the United States. The Republic of the Philippines has been an independent nation since 1946. The current population numbers approximately 77 million and is comprised by more than distinct 75 ethnic groups.

This general population sample was collected as part of a case-control study of Type 1 Diabetes (T1D) in the Philippines (1). The Filipino sample presented here constitutes the control sample from this study. All samples were collected from the Southern region of Luzon Island. A total of 94 normal Filipino subjects lacking a family history of diabetes were collected for this study.

DNA was extracted from 200 ml of frozen blood using QIA Amp blood kits. Between 50-100ng of genomic DNA was PCR amplified, and typed for HLA class II (DRB1, DQB1, DPA1 and DPB1) loci using sequence-specific oligonucleotide (SSO) hybridization probes. HLA-DRB1 and DQB1 typings

were performed using the DYNAL RELI SSO reverse lineblot typing kit. High resolution DRB1 typing was accomplished using group specific amplification of DR1, 2, 3/5/6, and, 4 alleles. High resolution DPA1 and DPB1 typing was carried out using the forward dot blot method and non-radioactive SSO oligonucleotide probes. The HLA-A, B and C loci were typed using the IHWG RLS method. Ambiguities in the genotype assignments were resolved by additional group specific amplification followed by dot-blot hybridization with HRP-labeled sequence specific oligonucleotide (SSO) probes.

The Filipino population has a distinctive frequency distribution of HLA class II alleles as well as linkage disequilibrium patterns (1, 2). As in previous studies, (2) the frequency of the DR2 allele, DRB1*1502 is very high in this population (45%), with a common extended class II haplotype of DRB1*1502-DQB1*0502-DPB1*0101 ($f=17.4\%$). The most common A-B-DR haplotype is A*2402-B*3802-DRB1*1502 (5.7%). A DR-DQ haplotype, DRB1*0405-DQB1*0503, found at 2.6% appears to be unique to the Filipino population (1–3). This haplotype is interesting with respect to T1D susceptibility because it contains a DRB1*0405 allele positive-

ly associated with T1D in other populations, and DQA1 and DQB1 alleles (*0101-*0503) that are negatively associated in other populations. This haplotype, which is unique to Filipinos, appears to be “neutral” with respect to association with T1D, suggesting that the susceptible DRB1 and protective DQB1 may “cancel each other out” (1). In addition, the DRB1*1105 allele also appears to be unique to Filipinos (3).

A deviation from expected Hardy-Weinberg equilibrium proportions was observed at the DQB1 locus ($p\text{-value}=0.0031$) in this population sample, although all other loci were in expected H-W proportions. This DQB1 deviation is characterized by an overall excess of homozygote genotypes (24 observed/15 expected, $p\text{-value}=0.166$), as well as a deficiency of DQB1*0503 heterozygotes (5 observed/12 expected, $p\text{-value}=0.0412$) and DQB1*0501/*0502 genotypes (2 observed/8.5 expected, $p\text{-value}=0.0246$). Based on the analysis of DR-DQ haplotypes, we think that typing error is an unlikely cause for this observation. This deviation may reflect Type 1 statistical error or possibly some deviation from randomness in the sample collection.

References

1. Bugawan TL, Klitz W, Alejandrino M, Ching J, Pabello A, Solfelix CM, Petrone A, Buzzetti R, Pozzilli P, Erlich HA. The association of specific HLA class I and II alleles with type 1 diabetes among Filipinos. *Tissue Antigens* **59**, 452–469, 2002.
2. Bugawan TL, Chang JD, Klitz W, Erlich HA. PCR/oligonucleotide probe typing of HLA class II alleles in a Filipino population reveals an unusual distribution of HLA haplotypes. *Am J Hum Genet* **54**, 331–340, 1994.
3. Apple RJ, Bugawan TL, Griffith R, Chang JD, Erlich HA. A new DRB1 allele and a novel DR4 haplotype found in a Filipino family. *Tissue Antigens* **41**, 51–54, 1993.

Populations: Javanese Indonesians from Singapore, Moluccans and Nusa Tenggara from Indonesia and East Timor, Highlanders and Lowlanders from Papua New Guinea

Authors: Steven J. Mack^{1,2}, Teodorica L. Bugawan², Justin A. Erlich¹, Mark Stoneking³, Nilmani Saha⁴, Henry A. Erlich^{1,2}

Authors' addresses: ¹Children's Hospital Oakland Research Institute, Oakland CA, ²Roche Molecular Systems, Alameda CA, ³Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, and ⁴Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA

In general, geographic, linguistic and cultural factors can be used to distinguish the peoples of the Pacific islands (Oceania) into four broad groups. Western Austronesia (aka Island

Southern Asia) includes the peoples of western Indonesia, Borneo, Singapore, and the Philippines. Melanesia includes the peoples of New Guinea, Fiji and the intervening islands. Micronesia includes the populations of the islands north of Melanesia (with the exception of the Polynesian Tuvaluan islands). Polynesia includes the peoples of New Zealand and all of the islands east of Melanesia and Micronesia extending to Hawaii and Easter Island.

Humans have been present in the Pacific since the Paleolithic. Australia and the PNG Highlands are thought to have been inhabited 50,000 years ago, while Polynesia was only colonized as recently as 1000–2500 years ago. Linguistic and genetic studies have suggested that the peoples of the PNG highlands may share an ancient common ancestry with Australians (2, 3, 4, 5), although this relationship is not supported by all genetic studies (6). Studies of globin genes have indicated

a close relationship between Polynesian and Melanesian populations (6), as well as between Polynesians and southern Asian groups (7, 8). Studies of mitochondrial DNA sequences (mtDNA) have suggested that Austronesian, and Melanesian groups spread south from continental Asia and migrated through Taiwan and the Philippines, and that Polynesian groups have their origins in the eastern regions of Austronesia, subsequent to this expansion (9, 10). In addition to these mtDNA studies, a recent study of Polynesian Y-chromosome polymorphism suggests extensive mixing of Polynesians with Melanesians as they migrated eastward through Melanesia from Asia (11). The different evolutionary histories suggested by these studies (as well as by studies of HLA, see below) may perhaps be due to differences in the relative ages and modes of inheritance of these markers.

Fifty Javanese Indonesians were sampled in Singapore (1 degrees 24 minutes north latitude, 104 degrees east longitude), 51 individuals from the Molucca islands (Ternate and Hiri) in eastern Indonesia (0 degrees latitude, 125 degrees east longitude), 97 individuals from the Nusa Tenggara islands (Alor, Flores, Timor, and Roti) in East Timor and eastern Indonesia (7 degrees south latitude, 125 degrees east longitude), 92 individuals sampled in multiple locations of the Papua New Guinea Highlands (12) (PNG Highlanders) (5 degrees south latitude, 145 degrees east longitude), 96 individuals from the city of Wosera in Papua New Guinea (the PNG Lowland MV population) (5 degrees south latitude, 150 degrees east longitude), and 48 individuals from coastal regions of Papua New Guinea (the PNG Lowland population) (10 degrees south latitude, 147 degrees east longitude) were genotyped as part of this study.

Genotypes at subsets of the class I and class II loci were determined in the Javanese Indonesian (A, C, and B), Moluccan (A, DRB1, DQA1, DQB1, DPA1, DPB1), Nusa Tenggara (A, DRB1, DQA1, DQB1, DPA1, DPB1), Papua New Guinea (PNG) Highlands (A, B, DRB1, DQA1, DQB1, DPA1, DPB1), and two PNG Lowlands populations (PNG Lowland MV at A, DRB1, DQB1, and PNG Lowland at DQA1, DQB1, DPA1, DPB1). HLA class II genotyping (DRB1, DQA1, DQB1, and DPB1) was carried out as previously described (13, 14). HLA class I genotyping (A, C and B) was done using the IHWG RLS typing system. The IHWG biostatistics core analyzed these data.

The data presented here for the Molucca, Nusa Tenggara, and Papua New Guinea populations has been reported previously, along with class II data for the Indonesian population (13, 14). In these reports, close relationships were identified between PNG Highland population and Aboriginal Australian

populations, and Polynesian populations were found to be more closely related to continental Asian groups than to western Austronesian or Melanesian groups. A high value of homozygosity (F_{nd}), consistent with the action of directional selection, was observed for a different Javanese population, and the action of selective forces other than balancing selection was inferred for populations in a region of the Pacific ranging from Java to Fiji.

Deviations from expected Hardy-Weinberg equilibrium proportions (HWE) were observed in several of the populations presented here (with the exception of the Moluccas). The Javanese Indonesian population deviated at the HLA-C locus (p -value=0.0441), the Nusa Tenggara population at the HLA-A locus (p -value=0.0103), the PNG Lowland population at the DQB1, DPA1 and DPB1 loci (p -values=0.0219, <0.0001, and=0.0114, respectively), and the PNG Highland population at the HLA-B, DRB1, DQA1, DQB1, and DPB1 loci (p -values=0.0361, 0.0166, 0.0188, 0.0073, and 0.0388, respectively). Given a set of twenty-nine allele frequency distributions, 1.45 (5%) would be expected to deviate from HWE. However, 34% (10 out of 29) of the allele frequency distributions of these populations deviate from expected HWE. While it is possible that the deviations in the PNG Highland population are due to this population sample being comprised of individuals sampled from multiple locations in the PNG Highlands, this does not explain the deviations seen at the other populations, nor does it explain similar deviations from expected HWE observed in the Filipino population (15).

The genetic diversity (h) values in these populations are uniformly low in comparison to continental Asian populations (e.g., Siberian, Indian, and Han Chinese groups), where diversity values routinely range from 0.9 to 0.95. The Javanese Indonesian group shows the highest diversity with values ranging from 0.89 to 0.94, while diversity values were decreased in eastern Indonesia (mean h =0.72 in Nusa Tenggara and 0.70 in Molucca), and were even lower in Papua New Guinea (mean h =0.68 in PNG Lowland, 0.66 in PNG Highland, and 0.60 in PNG Lowland MV). Diversity was lowest at the HLA-A locus in the PNG Highland population (h =0.36), and this low value has been attributed to the action of selection on the HLA-A*2402 allele (14). Low diversity values are observed at the HLA-DPA1 locus in this population as well as in the Molucca, PNG Lowland and Nusa Tenggara populations (h =0.36, 0.49, 0.54, and 0.63, respectively). Similarly low diversity values are characteristic of the Filipino population, where the mean value of diversity is 0.77 and the lowest value of diversity is seen at the DPA1 locus (h =

0.46) (15). These values could be consistent with either the action of selection or genetic drift (founder effects and bottlenecks). However, given the evidence for directional selection inferred at the DRB1 for in region of Oceania ranging from Java to Fiji, and at the HLA-A locus in the PNG Highlands, it

may be the case that directional selection is active in these (and other) Pacific populations at multiple HLA loci, skewing genotype proportions from those expected under Hardy-Weinberg equilibrium.

References

1. Comrie B, Matthews S, Polinsky M. The Atlas of Languages: The Origin and Development of Languages Throughout the World. London: Quarto Publishing plc, 1996.
2. Nichols J. Modeling ancient population structures and movement in linguistics. *Annu Rev Anthropol* **26**, 359–384, 1997.
3. Nichols J. Sprung from two common sources: Sahul as a linguistic area. In: McConvell P, Evans N, eds. Archaeology and linguistics: Aboriginal Australia in global perspective. Melbourne: Oxford University Press, 1997: 135–168.
4. Serjeantson SW. HLA genes and antigens. In: Hill AVS, Serjeantson SW, eds. The Colonization of the Pacific: a genetic trail. New York: Oxford University Press, 1989: 120–135.
5. Van Holst Pellekaan S, Frommer M, Sved JA, Boettcher B. Mitochondrial control-region sequence variation in Aboriginal Australians. *Am J Hum Genet* **62**, 435–449, 1997.
6. Tsintof AS, Hertzberg MS, Prior JF, et al. A-Globin gene markers identify genetic differences between Australian Aborigines and Melanesians. *Am J Hum Genet* **46**, 138–143, 1990.
7. Trent RJ, Buchanan JG, Webb A, et al. Globin genes are useful marker to identify genetic similarities between Fijians Pacific Islanders from Polynesia and Melanesia. *Am J Hum Genet* **42**, 601–607, 1988.
8. O'Shaughnessy DF, Hill AVS, Bowden DK, et al. Globin genes in Micronesia: origins and affinities of Pacific Island peoples. *Am J Hum Genet* **46**, 144–155, 1990.
9. Lum KJ, Rickards O, Ching C, et al. Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum Biol* **66**, 567–590, 1994.
10. Melton T, Peterson R, Redd AJ, et al. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am J Hum Genet* **57**, 403–414, 1995.
11. Manfred Kayser, Silke Brauer, Gunter Weiss, Peter A. Underhill, Lutz Roewer, Wulf Schiefenhövel and Mark Stoneking. Melanesian origin of Polynesian Y chromosomes. *Current Biology* **10**, 1237–1246, 2000.
12. Stoneking M, Jorde LB, Bhatia K, Wilson AC. Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics* **124**, 717–733, 1990.
13. Mack SJ, Bugawan TL, Moonsamy PV, Erlich JA, Trachtenberg EA, Paik YK, Begovich BA, Saha N, Beck HP, Stoneking M, Erlich HA. Evolution of Pacific/Asian populations inferred from HLA class II allele frequency distributions. *Tissue Antigens* **55**, 383–400, 2000.
14. Bugawan TL, Mack SJ, Stoneking M, Saha N, Beck HP, Erlich HA. HLA class I allele distributions in six Pacific/Asian populations: evidence of selection at the HLA-A locus. *Tissue Antigens* **58**, 311–319, 1999.
15. Bugawan, et al. Filipino from Luzon Island, Philippines. In: HLA 2004: Immunobiology of the Human MHC. Eds. Hansen JA, Dupont B. IHWG Press, 2006.

Population: Samoans from American Samoa

Authors: Laura Severson, Douglas Crews, Raymond Lang
Authors' address: Department of Anthropology and Department of Molecular Virology, Immunology and Medical Genetics, The Ohio State University, Columbus, Ohio USA

American Samoa is a group of five volcanic islands located in the South Pacific approximately 15 degrees below the equator and just east of the International Dateline. The origin of the first inhabitants of the Samoan Islands and other Polynesian islands has long been a topic of debate. The most widely held theory asserts that Austronesian-speaking people from Southeast Asia reached Samoa and Tonga around 3,300 BP carrying with them a culture known as Lapita, characterized by red slip pottery with dentate design (1). These migrants may have passed rapidly through the Melanesian archipelago and interbred little with local populations before arriving in Polynesia. This theory, often called the “Fast Train” hypo-

thesis, has been supported by archaeological and linguistic data (1, 2, 3) as well as molecular studies (4, 5). Competing theories hold that the first Polynesians may be derived from Melanesian populations (6) or may have traveled westward from the Americas (7).

A research team led by D. E. Crews from The Ohio State University collected blood samples from 264 Samoan-Americans on the island Tutuila in 1992. Samples were drawn from 115 men and 149 women whose ages ranged from 30 to 75 years. Approximately half the population was randomly selected while the other half was part of a longitudinal study begun in 1976 (8). Some of these samples were typed for HLA Class I as part of the anthropology component of the 12th IHWC. Analysis of those results indicated support for the “Fast Train” hypothesis (9).

For this study, 50 different American Samoan samples were selected. HLA-A, B, and C loci were typed using 13th IHWG RLS reagents in accordance with the 13th IHWG protocol. Genotype frequencies were found to be within Hardy-Wein-

berg expectations at all three loci. However, the Guo and Thompson exact test for deviation from Hardy-Weinberg indicated slightly less heterozygosity than expected at the HLA-A locus (p -value=0.0255). This may be due to slightly higher than expected homozygosity of the A*2402, A*0206, and A*3401 genotypes. Overall, the allele frequencies reported here are consistent with those reported in past studies (9). The most common alleles found at the A locus include A*2402, A*1101, A*0201, A*0206, and A*3401. At the

HLA-B locus the most common alleles are B*4001, B*5502, B*4801, and B*5602. For HLA-C the most common alleles found are Cw*0102, Cw*0304, and Cw*0801. These results supplement the growing body of molecular research of the Samoans and other Polynesians. HLA is of particular importance to the question of the Polynesian origins, because its high degree of polymorphism provides a level of discrimination that is necessary for determining the relationships among the Oceanic populations.

References

1. Bellwood PS. The colonization of the Pacific: some current hypotheses. In AVS Hill and SW Serjeantson (eds.): *The Colonization of the Pacific: A Genetic Trail*. Clarendon Press, Oxford, 1989.
2. Green RC, Lapita. In *The Prehistory of Polynesia* (ed. J. Jennings) Harvard University Press 1979.
3. Gray RD, Jordon FM. Language trees support the express-train sequence of Austronesian expansion. *Nature* **405**(6790), 1052–1055, 2000.
4. Serjeantson SW. HLA genes and antigens. In *The Colonization of the Pacific: A Genetic Trail* (eds. AVS Hill and SW Serjeantson). Clarendon Press, Oxford, 1989.
5. Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M. The Origin of the Polynesians: An Interpretation from Mitochondrial Lineage Analysis. *Am J Hum Genet* **57**, 1463–1475, 1995.
6. Hagelberg E, Clegg JB. Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of Ancient bone DNA. *Proceedings of the Royal Society* **252**, 163, 1993.
7. Heyerdahl T. *Kontiki: across the Pacific by raft*. Rand McNally, Chicago, 1950.
8. Bindon JR, Knight A, Dressler WW, Crews DE. Social context and psychosocial influences on blood pressure among American Samoans. *Am J Phys Anthropol* **103**, 7–18, 1997.
9. Severson LD, Crews DE, Lang RW. Application of SSP/ARMS to HLA Class I loci in Samoans. In *Proceedings of the 12th International Histocompatibility Workshop and Conference, Vol. II* (ed. D Charron). EDK, Paris, 1997.

Australia

Populations: Aboriginal Australian from Cape York, Groote Eylandt, Kimberly, and Yuendumu, Australia

Authors: Xiajiang Gao¹, Steven J. Mack^{2,3}, Bert Currie⁴, Simon Easteal², James McCluskey³

Authors' addresses: ¹National Cancer Institute Frederick Cancer Research Facility, Maryland, ²Children's Hospital Oakland Research Institute, Oakland, CA, ³Roche Molecular Systems, Alameda, CA, ⁴Menzies School of Health Research, Darwin, Australia, ⁵Immunology and Genetics Division, John Curtin School of Medical Research, Australian National University, Canberra, Australia, and ⁶Department of Microbiology and Immunology, University of Melbourne, Melbourne, Australia

The origins of the Aboriginal Australian peoples and their relationship to the modern peoples of Oceania, and in particular to the peoples of the Papua New Guinea Highlands, have been the subject of much speculation. The human presence

in Australia dates back at least 50,000 years (1, 2). During the last ice-age (70,000–10,000 years ago), Australia was part of Sahul, a single continental land mass that included the islands of Tasmania and Papua New Guinea, and which is thought to have been colonized from south-east Asia in this period. The end of the ice-age initiated a dramatic rise in sea-level that isolated the ancestors of modern Aboriginal Australian groups from the ancient inhabitants of Tasmania and Papua New Guinea. It has been speculated that dramatic environmental differences between tropical south-east Asia and semi-arid Australia resulted in strong selective pressure on the original inhabitants of Sahul. In addition, these small populations may also have experienced extensive genetic drift. As a result, the HLA allele and haplotype frequency distributions in Aboriginal groups are likely to have undergone dramatic changes in the last 50 millennia.

Aboriginal Australia was largely isolated until Europeans made contact in 1788, when the Aboriginal population is estimated to have been between 500,000 and 1 million individuals (1). This contact resulted in a rapid population de-

cline (resulting from conquest and exposure to disease), and the current population numbers approximately 250,000 individuals, many of whom live on missions and government settlements. At the end of the 18th century, an estimated 700 Aboriginal tribes spoke distinct dialects of ~250 different languages, which are classified into two major groups; Pama-Nyungan languages are spoken throughout the southern regions of the continent, while languages in the extremely diverse non-Pama-Nyungan family are spoken in the northwestern tenth of the continent. Neither language family is clearly related to any language spoken outside of Australia. Analyses of population samples from two groups of Pama-Nyungan speakers (Cape York and Yuendumu) and two groups of non-Pama-Nyungan speakers (Kimberly (Wororan and Nyulnyulan speakers) and Groote Eylandt (Anindhilyaguan speakers)) are reported here.

Blood samples from 193 volunteers were collected at the Yuendumu mission in Northern Territory. Wherever possible, first-degree relatives were not sampled. Samples were collected from 104 individuals from the Cape York peninsula of Queensland, 75 individuals from Groote Eylandt in Northern Territory, and 41 individuals from the Kimberley region of Western Australia (3, 4). The Kimberly samples were collected at the Kalumburu mission at Napier Broom Bay and the Mowanjam mission near Derby. In this latter case, the Mowanjam samples were derived from Worora, Wunambal and Ngarinjin tribal groups. In the case of each population sample, every effort was made to sample individuals in a random fashion, but given the small sizes of these communities, some related individuals were inevitably included in the study. All samples were genotyped at class I loci (A, C and B) using high-resolution SSOP methods. The individuals from the Cape York and Kimberly regions were also typed at class II loci (DRB1, DQA1, DQB1, DPB1) using high-resolution SSOP methods.

Some of allele frequency and class II haplotype data for the Kimberly, Cape York and Yuendumu populations have been published (3–11). Although little evidence for non-Aboriginal admixture has been observed in the Kimberly and Yuendumu populations (11, 12), non-Aboriginal admixture in the Cape York population sampled has been estimated at about 20% (13). Previous studies have observed reduced allele diversity at all loci in Aboriginal Australian populations, with a pronounced restriction of diversity at the class I loci, along with novel allele diversity at HLA-A, B, DRB1 and DPB1 unique to Australia (10–12, 14, 15). Significant differences in DRB1 allele frequencies between populations (8), regional restrictions in the distribution of novel HLA-A*24 alleles

(15), and DR-DQ haplotype diversity unique to Aboriginal Australian groups (8, 9, 11, 12) were also demonstrated. In addition, several HLA studies have detected an ancient connection between Aboriginal Australians and PNG Highland populations (8–10, 16) supporting the “Sahul Hypothesis” that these modern groups are the descendants of the ancient inhabitants of Sahul.

In the current study, numerous high frequency alleles were observed at all loci, with 2–4 alleles accounting for more than 60% of the allele diversity at the HLA-A, C, B, DRB1 and DPB1 loci in each population. The HLA-A*0201 and *3401 (allele frequency (a.f.)=0.11–0.17 and 0.29–0.68 respectively), Cw*0102 and *0401 (a.f.=0.19–0.37 and 0.17–0.28), and B*5601 (a.f.=0.11–0.36) alleles were common to all populations, while DRB1*0803 (a.f.=0.26–0.35) and DPB1*0401 and *0501 (a.f.=0.10–0.11 and 0.45–0.68) alleles were observed in both the Cape York and Kimberly groups. In addition, the HLA-A*3401-B*5601-DRB1*0803 haplotype was observed in the Cape York and Kimberly groups at high frequencies (haplotype frequency (h.f.)=0.06 and 0.18) as was the DRB1*0803-DQA1*0103-DQB1*0601-DPB1*0501 haplotype (h.f.=0.09 and 0.16). The A*3401-B*5601 haplotype was observed at high frequencies in the Yuendumu and Groote Eylandt populations as well (h.f.=0.06 and 0.14), suggesting the presence of an extended HLA-A*3401-B*5601-DRB1*0803-DQA1*0103-DQB1*0601-DPB1*0501 haplotype at high frequencies (a.f. >0.5) in these Aboriginal populations.

While variation was observed in the distribution of class I alleles between these groups, it may largely be due to the differences in the sizes of the sampled populations. Only two alleles (A*0101 and Cw*1203 in Cape York) were observed at high frequencies in a single population, and only one allele was observed exclusively in both populations from a given language family (Cw*0702 in Pama-Nyungan speakers). Despite the overall pattern of sharing of common class I alleles between all groups, comparative analyses of HLA-A-B and A-C-B haplotype frequency distributions (not shown) suggest a correlation between genetics and language, with the non-Pama-Nyungan speaking Kimberly and Groote Eylandt populations more closely related to each other than to the Pama-Nyungan speaking Yuendumu and Cape York groups.

Allele frequency distributions consistent with the action of selection were indicated by the Ewens-Watterson test of neutrality in only three cases; at HLA-C and B in the Yuendumu population ($F_{nd} = -1.38$ and -1.28 , p -values=0.0203 and 0.0238 respectively), and at DQA1 in the Kimberly group ($F_{nd} = -1.57$, p -value=0.0054). In each of these cases, sig-

nificantly low homozygosity values were observed, consistent with the action of balancing selection at these loci.

The Groote Eylandt population deviated significantly from expected Hardy-Weinberg equilibrium proportions (HWE) at the HLA-A, C and B loci (p-values=0.0043, 0.0036, and 0.0201), and the Cape York population deviated significantly from expected HWE at the HLA-C and B loci (p-values=0.0019 and 0.0302). In the case of Groote Eylandt, non-significant excesses of A*2402: *3401 genotypes (expected 14, observed 20), Cw*0102: *0401 genotypes (expected 9.8, observed 16) and B*1301: *5501 genotypes (expected 6.3, observed 11), and non-significant under-representation of A*1101: *3401 genotypes (expected 11.5, observed 6) may have contributed to these deviations. While no overall deviations from Hardy-Weinberg equilibrium proportions were observed for the HLA-A, C or B loci in the Yuendumu population, significant excesses of B*1301: *1525 and *1301: *4002 genotypes were observed (p-values=0.0449 and 0.0287 respectively). The genotypic proportions of all loci

in the Kimberly population were consistent with expected HWE.

These recurrent deviations from expected HWE could result from consistent genotyping error (a consistently undetected allele) or sampling error (non-random sampling, population admixture or sub-structure). An excess of homozygous genotypes would be expected in case of consistent genotyping error, and given the extent to which the allelic diversity at each locus is shared between these populations, it might also be expected that Hardy-Weinberg deviations would be observed at the same loci in all populations. Given that this deviation is most pronounced in the Groote Eylandt population, it may potentially be the result of sampling error.

We wish to thank the Yuendumu people and the Yuendumu Land Council for their generous participation in this study, along with the people of Groote Eylandt, the Kimberly region of Western Australia and the people of Cape York. We also thank Barry Boettcher for his support of the study.

References

- Horton, D. (ed.), *The Encyclopaedia of Aborigina Australia*, Australian Institute of Aboriginal and Torres Straigh Islander Studies, 1993.
- Roberts RG, Jones R, Smith MA. Thermoluminescence dating of a 50,000 year old human occupation site in northern Australia. *Nature* **345**, 153, 1990.
- Blake NM, Spargo RM. Population genetic studies in the Kimberly of Western Australia. *Human Heredity* **36**, 286–298, 1986.
- Davies R, Pace R, Jazwinska EC, Serjeantson SW. HLA DR RFLP distributions in tow groups of Aboriginal Australians. *Aust NZ J Med* **20**, 790–793, 1990.
- Terasaki PI, Gjertson DW (eds.) *HLA 1997*. Los Angeles: UCLA Tissue Typing Laboratory, 1997.
- Lester S, Cassidy S, Humphreys I, Bennett G, Hurley CK, Boettcher B, McCluskey, J. Micro-evolution in HLA DRB1 and MHC class II haplotypes of Australian Aborigines. Definition of a new DRB1 allele and distribution of DRB1 gene frequencies. *Human Immunology* **42**, 154–160, 1995.
- Gao X, Serjeantson SW. Twelve HLA-DR6-related DRB1 alleles and associated DR, DQ haplotypes in traditional Australians and other populations of Asia-Oceania. *Eur J Immunogenet* **19**(5), 263–272, 1992.
- Gao X, Serjeantson SW. Diversity of HLA-DR4-Related DR, DQ Haplotypes in Australia, Oceania, and China. *Human Immunology* **32**, 269–276, 1991.
- Gao X, Serjeantson SW. Twelve HLA-DR6-Related DRB1 Alleles and Associated DR, DQ haplotypes in traditional Australians and other populations of Asia-Oceania. *Eur J Immunogenet* **19**, 263–272, 1992.
- Linert K, McCluskey J, Bennet G, Fowler C, Russ G. HLA class I variation in Australian Aborigines: characterization of allele B*1521. *Tissue Antigens* **45**, 12–17, 1995.
- Gao X, Veale A, Serjeantson SW. HLA class II diversity in Australian Aborigines: unusual HLA-DRB1 alleles. *Immunogenetics* **36**, 333–337, 1992.
- Lester S, Cassidy, S, Humphreys I, Bennett G, Hurley CK, Boettcher B, McCluskey J. Evolution in HLA-DRB1 and Major Histocompatibility Complex class II haplotypes of Australian Aborigines. Definition of a new DRB1 allele and distribution of DRB1 gene frequencies. *Human Immunology* **42**, 154–160, 1995.
- Cullen LM, Gao X, Eastel S, Jazwinska GC. The Hemochromatosis 845 G→A and T87 C→G Mutations: Prevalence in Non-Caucasian Populations. *Am J Hum Genet* **62**(6), 1403–1407, 1998.
- Leinert K, Krishnan R, Fowler C, Russ G. HLA DPB1 genotyping in Australian Aborigines by amplified fragment length polymorphism analysis. *Human Immunology* **36**, 137–141, 1993.
- Gao X, Lester S, Matheson B, Boettcher B, McCluskey J. Three newly identified A*24 alleles: A*2406, A*2413 and A*2414. *Tissue Antigens* **50**, 192–196, 1997.
- Mack SJ, Bugawan TL, Moonsamy PV, Erlich JA, Trachtenberg EA, Paik YK, Begovich AB, Saha N, Beck HP, Stoneking M, Erlich HA. Evolution of Pacific/Asian populations inferred from HLA class II allele frequency distributions. *Tissue Antigens* **55**, 383–400, 2000.

North-East Asia

Population: Buryat from Angarsk, Russia

Authors: Akesaka T¹, Ishikawa Y², Sideltseva EW³, Tanaka H², Sideltsev VV³, Zhuravleva E³, Mack SJ^{4,5}, Akaza T⁶, Khalturin V³, Alexeev LP⁷, Juji T⁶ and Tokunaga K¹

Authors' addresses: ¹Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, ²Japanese Red Cross Tokyo Metropolitan Blood Center, Tokyo, Japan, ³Immunogenetics Research Centre of Civil Health Service, Angarsk, Irkutsk, Russia, ⁴Department of Human Genetics, Roche Molecular Systems, Berkeley, CA, USA, ⁵Children's Hospital Oakland Research Center, Oakland, CA, USA, ⁶Japanese Red Cross Central Blood Center, Tokyo, Japan, ⁷Department of Immunogenetics, Institute of Immunology, Moscow, Russia

The Buryat are a Mongoloid group residing in the area extending from northern Mongolia to the northern shore of Lake Baikal. The size of the Buryat population is estimated to be 500,000. Based on ethnological studies (1), they have been considered to be closely related to the Mongolian people

of Mongolia. The Buryat may be a heterogeneous group due to admixture with other ethnic groups, such as the Evenk and the Yakut. The Buryat population sample presented here is derived from a Western Buryat branch, and may have been significantly affected by admixture with Russian ethnic groups.

Peripheral blood samples obtained from 141 unrelated healthy Buryat adults living in Angarsk city suburbs in the Irkutsk region of East Siberia were genotyped at the HLA-A locus using PCR-SSOP techniques. The HLA-A, B, C, DR and DQ antigen polymorphisms in this population sample had been investigated previously (2). In the previous study, gene frequencies for class I antigens A2, A24, and A1 were estimated to be more than 10%. In a phylogenetic analysis, this Buryat population clustered with other Mongoloid groups in Northeast Asia. Analysis of five-locus haplotypes (HLA-A-C-B-DR-DQ) indicated that the Buryat are genetically related to European Caucasoids, Inuits and Amerindians to some extent, while there were a few haplotypes observed in the Buryat that may be typical of East Siberian populations.

References:

1. Alexeev VP, Gokhman II. *Anthropologiya Aziatskoi Chasti SSR*, Moskva, 1984.
2. Tokunaga K, Sideltseva EW, Tanaka H, Uchikawa C, Nieda M, Sideltsev VV, Zhuravleva E, Imanishi T, Itoh K, Akaza T, Takahashi K, Khalturin V, Alexeev LP, Juji T. Distribution of HLA antigens and haplotypes in the Buryat population of Siberia. *Tissue Antigens* **45**, 98–102, 1995.

Population: Korean from Seoul, Korea

Authors: M. H. Park, E.Y. Song

Authors' address: Department of Laboratory Medicine, Seoul National University Hospital, Seoul, Korea

Anthropologically, the first Koreans are thought to have been Palaeo-Asiatics who migrated from Siberia through Manchuria into the Korean peninsula during the Neolithic age, some 3,000 to 10,000 years ago. The Tungus, a second Palaeo-Asiatic group, brought Bronze culture to the peninsula about 3,000 years ago. Their language, the Tungus-Altai family, became the common language, the origin of modern Korean and Japanese. While these ancestors of modern Koreans are thought to have been distinct from the ancestors of the Chi-

nese, a more recent, minor admixture from northern China is thought to have occurred with the introduction of Iron culture. It is believed that quite a number of Koreans had migrated to Japan islands, given that Koreans and Japanese speak Tungus-Altai languages and share similar old arts. Korean and Japanese populations have shown to be closely related to each other by mitochondrial DNA polymorphisms (1) and the distribution of HLA alleles and haplotypes (2–5).

Working with the Korea Marrow Donor Program (KMDP), we were able to obtain blood samples from 200 unrelated healthy Koreans. The HLA-A, B and C loci were typed using IHWG RLS reagents, and majority (n=174) of these samples were also genotyped at the A and B loci using the PCR-single strand conformation polymorphism (SSCP) method. HLA-DRB1 genotyping was performed using SSOP and SSCP meth-

ods. All loci were under Hardy Weinberg equilibrium, with a minor deviation observed at the C locus. The Ewens-Watterson homozygosity test of neutrality revealed that all loci were consistent with either the operation of neutral drift (A locus) or balancing selection (B, C and DRB1 loci).

The number of detected alleles at each HLA locus was: A 17, B 37, C 21, and DRB1 27. Most frequent alleles with allele frequency (a.f.) ≥ 0.05 were: A*2402 (a.f.=0.230), *0201 (0.165), *3303 (0.133), *1101 (0.105), *0206, *2601, *3001; B*5101 (0.125), *1501, *4403, *3501, *4601, *5801, *2705, *4002; Cw*0102 (0.185), *1402 (0.105), *0303 (0.100), *0801, *0304, *0702, *0401, *0602, *0302; DRB1*0901 (0.121), *0405, *0101, *1501, *1302, *0701, *0803, *0406, and *1201. Most prominent allelic polymorphisms were noted for A*02 (0201, 0203, 0206, 0207), A*26 (2601, 2602, 2603), B*15 (1501, 1502, 1507, 1511, 1518, 1527), B*40 (4001, 4002, 4003, 4006), Cw*03 (0302, 0303, 0304), Cw*07 (0701, 0702, 0704), DRB1*04 (0401, 0403, 0404, 0405, 0406, 0410) and DRB1*14 (1401, 1403, 1405, 1406, 1407, 1412). These allelic polymorphisms were quite similar to those observed

in Japanese (6). Our previous study in Koreans showed that A-B-DRB1 haplotypes containing serologic A2 specificity have substantial heterogeneity in the distribution of A*02 alleles (7). Polymorphisms of HLA class I alleles, especially A*02, A*26, B*15, and B*40 alleles, have a clinical implication in Koreans, as has been reported in Japanese for matching unrelated donor-recipient pairs at the allelic level for marrow transplantation (8). The most frequent A-B-DRB1 haplotypes with frequency (h.f.) ≥ 0.02 were: 3303-5801-1302 (h.f.=0.033), 3303-4403-0701 (0.031), 1101-1501-0406, 3001-1302-0701, 2402-0702-0101, 3303-4403-1302, 0201-2705-0101, and 2402-5201-1502. These frequencies were not much different from a previous population study (n=510) based on serologically defined A and B alleles and DRB1 genotypes (4), indicating that for these frequent haplotypes, class I serologic specificities show little diversification at the allelic level. Four of these eight frequent A-B-DRB1 haplotypes in Koreans are also frequent in Japanese, and international collaboration between Korea and Japan has shown to be very effective for donor searches in unrelated marrow transplantation (9).

References

1. Harihara S, Saitou N, Hirai M, et al. Mitochondrial DNA polymorphism among five Asian populations. *Am J Hum Genet* **43**, 134–143, 1988.
2. Park MH, Juji T, Tokunaga K. HLA ethnic study of Japanese and Koreans. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991*, Vol I. New York: Oxford University Press, 674–676, 1992.
3. Park MH, Hwang YS, Park KS, et al. HLA haplotypes in Koreans based on 107 families. *Tissue Antigens* **51**, 347–355, 1998.
4. Park MH, Kim HS, Kang SJ. HLA-A, -B, -DRB1 allele and haplotype frequencies in 510 Koreans. *Tissue Antigens* **53**, 386–390, 1999.
5. Song EY, Park MH, Kang SJ, et al. HLA class II allele and haplotype frequencies in Koreans based on 107 families. *Tissue Antigens* **59**, 475–486, 2002.
6. Saito S, Ota S, Yamada E, Inoko H, Ota M. Allele frequencies and haplotypic associations defined by allelic DNA typing at HLA class I and class II loci in the Japanese population. *Tissue Antigens* **56**, 522–529, 2000.
7. Park MH, Whang DH, Kang SJ, Han KS. HLA-A*02 allele frequencies and haplotypic associations in Koreans. *Tissue Antigens* **55**, 250–256, 2000.
8. Sasazuki T, Juji T, Morishima Y, et al. Effect of matching of class I HLA alleles on clinical outcome after transplantation of hematopoietic stem cells from an unrelated donor. Japan Marrow Donor Program. *N Engl J Med* **339**, 1177–1185, 1998.
9. Oh HB, Kim SI, Park MH, Akaza T, Juji T. Probability of finding HLA-matched unrelated marrow donors for Koreans and Japanese from the Korean and Japan Marrow Donor Programs. *Tissue Antigens* **53**, 347–349, 1999.

Population: Okinawans From Hawaii, United States

Authors: D. Kumashiro¹, N. Takashige¹, S. J. Mack^{2,3}, Y. K. Paik¹, D. Horio¹

Authors' address: ¹Clinical Laboratories of Hawaii, St. Francis Medical Center, Honolulu, HI, USA, ²Roche Molecular Systems, Alameda, CA, USA, ³Childrens Hospital Oakland Research Institute, Oakland, CA, USA.

The Okinawa Prefecture consists of 49 inhabited and 110 uninhabited islands (the Ryukyu Islands), and is Japan's south-

western and most insular prefecture. It is the only prefecture that was an independent nation prior to annexation by Japan. Located between Japan's Kyushu island and Taiwan, the Ryukyu Islands can be divided into three major groups: the Okinawa group, the Miyako group, and the Yaeyama group. Okinawa is the largest island followed in size by Irimote, Isigaki and Miyako. The indigenous people of the Ryukyu Islands are alternatively referred to as Ryukyuan or Okinawans.

The origin of the Ryukyuan people is ambiguous. Continental Asian people may have arrived at the Ryukyu Islands

via three routes: A northern route proceeding from northern Asia south through the islands of Japan; a southern route proceeding from the tropical Indies or Southeast Asia through Philippines or along the coast of China to Taiwan and onward to Irimote, Ishigaki and the Miyako islands; and a central route proceeding from Mongolia or Manchuria south along the Korean peninsula to Kyushu and beyond. An anthropological study conducted in 1998 suggested that Ryukyuan belong to the East/Southeast Asian clade (1), supporting the idea that the southern route was followed during the colonization of the Ryukyuan Islands.

Ryukyuan history begins on Okinawa Island with three rival kingdoms that were united in the 15th century to form the Ryukyu kingdom (2). Proximity to major Asian cities made the Ryukyu kingdom the southern gate to China, southeast Asia and Oceania. The kingdom maintained prosperous tributary trade and cultural links to China until 1609, when Satsuma-han of Japan invaded the Ryukyu islands. Despite four centuries of Japanese influence, and major attempts to assimilate the Okinawans during the Meiji period, Ryukyuan descendants have been socially, politically and economically marginalized. Admixture between Okinawans and non-Japanese groups remains more likely than with the Japanese population.

Prior to the Japanese invasion, the Ryukyu people had established a unique culture, blending northern and southern influences. Buddhism and Confucianism were introduced, but neither religion erased existing traditional shamanistic practices. While the Okinawan language (known as Uchina guchi or Hogen) originated from the same linguistic family as Japanese, traditional Ryukyuan speech is unintelligible to the Japanese (3). The Shuri dialect is the most common dialect in Okinawa, with characteristics that include complex honorific markers differentiating class, sex and age. Morphological, phonological, and lexical differences between Okinawans and the Japanese have been attributed to a long and separate development.

Blood was collected from 105 Okinawans on the island of

Oahu in Hawaii. Six families consisting of 20 individuals and 85 other unrelated individuals were studied. Selection was based on ancestors who emigrated from Okinawa to Hawaii to work on plantations. These samples were typed at the Class I loci using the IHWG RLS protocol.

The following high frequency (allele frequency (a.f.) >0.05) alleles were observed in the Okinawan population: HLA-A, A*2402 (a.f.=0.34), *0206 (a.f.=0.14), *2601 (a.f.=0.12), *3101 (a.f.=0.11), *0201 (a.f.=0.09), *2603 (a.f.=0.09), *1101 (a.f.=0.07). HLA-B, B*3501 (a.f.=0.16), *1501 (a.f.=0.13), *5401 (a.f.=0.11), *4002 (a.f.=0.10), *5101 (a.f.=0.09), *3901 (a.f.=0.08), *5901 (a.f.=0.06), *4001 (a.f.=0.06), *4006 (a.f.=0.06). HLA-C, Cw*0303 (a.f.=0.22), *0102 (a.f.=0.21), *0702 (a.f.=0.14), *0304 (a.f.=0.12), *0801 (a.f.=0.10), *1402 (a.f.=0.07).

Comparison of Okinawan allele frequency data with other IHWG population data reveal high frequency Okinawan alleles (HLA-A*0201, *0206, *1101, *2402, HLA-Cw*0102, *0303, *0304, *0702, and *0801) are also observed in other Northeast and Southeast Asian groups at high frequencies. In addition, HLA-A*3101 is observed at high frequency in Northeast Asia (Korea, Siberia and Mongolia) and at frequencies as high as 47% in the Americas. HLA-Cw*1402 is observed at a high frequency in Koreans alone. Unlike the case with Northeast Asian groups, Okinawans have few alleles in common with the Malay population (e.g. HLA-A*1101, 2402, 0201; HLA-B*4001 and HLA-Cw*0801). While Waijin or Hondo Japanese are considered to be part of the Northeast Asian region, data from this population was not available for comparison with the Okinawans in this study. The observation of high-frequency HLA-A and HLA-C locus alleles shared between Okinawans and Korean and Mongolian groups lends support to the idea that the central route was used during the colonization of the Ryukyu Islands.

Further attempts to type Japanese and additional Northeast Asian populations and examination of class II alleles in Okinawans may prove useful in further exploring these hypotheses.

References

1. Dodo Y, Doi N, Kondo O. Ainu and Ryukyuan Cranial Nonmetric Variation: Evidence Which Disputes the Ainu-Ryuku Common Origin Theory. *Anthropol Sci* **106**(2), 99–120, 1998
2. Kerr GH. Okinawa The History of an Island People, Charles E. Tuttle Co., Inc., 1958.
3. McCarty S. Review of Multilingual Japan. *JALT Journal* **19**(1), 138–142, 1997.

Population: Ryukyuan from Okinawa, Japan

Authors: Akesaka T¹, Ishikawa Y², Hatta Y¹, Ohashi J¹, Mack SJ^{3,4}, Akaza T⁵, Juji T⁵, Tokunaga K¹

Authors' addresses: ¹Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, ²Japanese Red Cross Tokyo Metropolitan Blood Center, Tokyo, Japan, ³Department of Human Genetics, Roche Molecular Systems, Alameda, CA, USA, ⁴Children's Hospital Oakland Research Institute, Oakland, CA, USA, ⁵Japanese Red Cross Central Blood Center, Tokyo, Japan

The population of modern Japan is divided into three distinct ethnic groups: The Ryukyuan people native to the southern Okinawa islands, the Ainu people native to the northern Hokkaido island, and the Hondo Japanese people inhabiting the islands of Honshu, Shikoku and Kyushu. Anthropologically, these modern groups are thought to be derived from two distinct ancestral groups, the indigenous pre-agricultural Jomon people, and post-Neolithic Yayoi immigrants from East Asia. Under this theory, the Ryukyuan and Ainu people are the relatively pure descendants of the indigenous Jomon people while the Hondo Japanese are the descendants of both Jomon people and Yayoi immigrants from the Korean peninsula.

Peripheral blood samples were obtained from 197 unrelated healthy adults living on the main island of Okinawa (Ryukyu) in the southwestern Islands of Japan and were typed at the HLA-A locus using PCR-SSOP methods. In a previous study (1), this population sample was serologically typed at the HLA-A, B and C loci, and molecularly typed at the DRB1 locus using PCR-MPH, PCR-SSP, and PCR-SSCP methods. HLA-A*2, A*24 and A*26 alleles accounted for more than 80% of the polymorphism at this locus. In addition, the presence of the A*24-B*54-DRB1*0405 haplotype suggested gene flow from a South Chinese population into the Okinawan Islands after the divergence of the Ryukyuan from the Ainu. Phylogenetic and principal component analyses revealed close relationships between the Ryukyuan, Ainu and Hondo Japanese populations, and these populations showed the highest affinity to two Korean populations. However, Ryukyuan were shown to be closer to Hondo Japanese and Korean populations than to the Ainu.

References

1. Ohashi J, Ohashi J, Imanishi T, Kamiyama H, Iha M, Simabukuro T, Ohawa A, Tanaka H, Akaza T, Gojobori T, Juji T, Tokunaga K. HLA genes and Haplotypes in Ryukyuan suggest recent gene flow to the Okinawa Islands. *Human Biology* 71:353-365, 1999.

Population: Tuvians from Republic of Tuva, Russian Federation

Authors: Steven J. Mack^{1,2}, Andrea J. Jani³, Marina L. Sartakova⁴, Vladimir I. Kononov⁴, Henry A. Erlich^{1,2}

Authors' addresses: ¹Children's Hospital Oakland Research Institute, Oakland, CA, ²Roche Molecular Systems, Alameda, CA, ³Arizona State University, Tempe, AZ, and ⁴Institute of Clinical Immunology, Russian Academy of Medical Sciences, Novosibirsk, Russia

Located in the Tannu Mountains of southern Siberia, the Republic of Tuva (also termed Tyva) is a republic in the Russian Federation. The population of the Republic numbers 300,000, approximately two-thirds of whom are Tuvians and the remaining third Russians. One quarter of the Tuvian population is nomadic.

The Tuvian people speak a Turkic language belonging to the Siberian group of northern Turkic languages, and are thought to be related to other central Asian Turkic groups (e.g., Uighur and Khirgiz peoples). While Neolithic inhabitants of the region may have had "caucasian" features, the Tuvian people closely resemble other "mongoloid" peoples.

With the exception of a period of independence in the mid-20th century, the region of modern Tuva (also termed Tannu Uriyangqai) has been alternately controlled by Huns, Turks, Mongols, Manchu Chinese, and Russia from the 3rd century BCE until 1991, when the Republic declared its independence from Russia (1).

Because of Tuva's location in central Asia, Tuvian genetic diversity is important to an understanding of the genetic relationships in Asia, Europe and the Americas. Different genetic markers have been used to characterize the heterogenic Tuvian

an population and compare it with other populations. Sartakova et. al. (2) determined that Tuvinians differed significantly from other central and east Asian populations, being most closely related to Khazaks. Golubenlo et al. (3) determined that Tuvinians had affinities to both Siberian and central Asian groups. Uinuk-Ool et. al. (4) concluded that Tuvinians were closely related to native American and Siberian groups, as well as to central, east and southeast Asian groups. Here, we present the first study of high-resolution HLA class I and class II allele and haplotype data for Tuvinians.

Subsets of unrelated 191 unrelated Tuvinians from the capital city of Kyzyl (50 degrees North latitude, 95 degrees east longitude) were genotyped at the HLA-A (189 individuals), C (174 individuals), and B (180 individuals) loci using IHWG RLS strips, and at the DRB1 locus (189 individuals) using DRB1 general typing and subtyping methods (5). Statistical analyses were carried out by the IHWG Biostatistics Core.

Overall allele frequency distributions at the HLA-A, C, B and DRB1 loci were consistent with expected Hardy-Weinberg equilibrium proportions, although a significant deviation in the number of common genotypes was observed at the HLA-A locus (p-value=0.02).

Significantly low homozygosity ($F_{nd} = -1.34$ p-value <0.01), consistent with the action of balancing selection, was detected at the HLA-C locus using the Ewens-Watterson homozygosity test. Homozygosity values at the HLA-A, B and DRB1 loci were consistent with neutral expectations.

Allelic diversity (h) was high ($h = >0.9$) at each of the four loci typed in this study, with 31 alleles observed at the HLA-A locus ($h = 0.90$), 25 at the HLA-C locus ($h = 0.93$), 52 at the HLA-B locus ($h = 0.96$), and 41 at the DRB1 locus ($h = 0.95$). With allele frequencies (a.f.) greater than 0.06, the most frequent alleles at each locus were A*2402, *0201, *3101, *1101, *3303, *0206, *0301, and *0101 (a.f. = 0.21, 0.14, 0.10, 0.09, 0.07, 0.07, 0.07 and 0.07, respectively), Cw*0401, *0602, *0304, *0303, *0702, *0101, and *1502 (a.f. = 0.11, 0.11, 0.11, 0.09, 0.09, and 0.07 respectively), B*4002, *0702, *5801, and *5101 (a.f. = 0.10, 0.07, 0.07 and 0.06, respectively), and DRB1*1501, *0701, *1201 and *0901 (a.f. = 0.12, 0.10, 0.07, and 0.06 respectively).

In general, Tuvinian allelic diversity is representative of that

seen in other northeast Asian populations. For example, A*2402 and *3303 are common in east Asian and Oceanian populations, while A*3101 is restricted to northeast Asia and the Americas (6). In addition, several high-frequency Tuvinian class I alleles (A*2404, *3101, *0206, Cw*0304, *0401, and B*4002) are common in native Americans (7, 8), although these alleles are observed at much higher frequencies in American populations. With the exception of DRB1*0901 (common in Eskimos and South Americans), common native American DRB1 alleles (e.g., DRB1*0802 and *1602) were detected in Tuvinians at low frequencies (a.f. = 0.01 and 0.002).

Significant levels of linkage disequilibrium (LD) were observed between all pairs of loci studied (p-values <0.001). Global disequilibrium (D') values ranged from 0.63 between the HLA-A and C loci to 0.93 between the HLA-C and B loci.

In keeping with the high levels of allelic diversity at each locus, large numbers of two and three-locus haplotypes were inferred in the Tuvinian population (e.g., 981 A:B:DRB1 haplotypes). With haplotype frequencies (h.f.) greater than 0.03, the most common A:B haplotypes were A*3303:B*5801, *2402:*4002, *0201:*4402, *3101:*4801, and *3101:*4001 (h.f. = 0.062, 0.053, 0.033, 0.031, and 0.031, respectively), and the most common three locus haplotypes were A*3303:B*5801:DRB1*1302 (h.f. = 0.033), A*3303:Cw*0302:B*5801, A*2402:Cw*0304:B*4802, and A*3101:Cw*0801:B*4801 (h.f. = 0.054, 0.042, and 0.030, respectively). Overall, these frequencies suggest the presence of an extended A*3303:Cw*0302:B*5801:DRB1*1302 haplotype (at a frequency of approximately 0.03 in Tuvinians. As with the A*3303 allele, these A*3303:C:B and A*3303:B:DRB1 haplotypes appear to be common to Asian and Oceanian populations (9).

Several of these class I haplotypes are present in Native American populations (e.g., A*2402:B*4002, A*3101:B*4801, and A*3101:Cw*0801:B*4801). While much of the allele and haplotype diversity of the Tuvinian population is not observed in native American groups, the extent of allele and haplotype sharing suggests that both groups are descended from common ancestral populations, with modern Tuvini- an allele diversity derived from population admixture that occurred subsequent to the colonization of the Americas.

References

1. Alexeev VP. Brief statements of paleoanthropology of Tuva in connection with historical aspects. In: Anthropological studies in Tuva. M. Nauka, pp. 6–74, 1984.
2. Sartakova ML, et al. Frequency of the HLA-DP genes and the antigens of the HLA-A, -B, -Cw, and -DR loci in Tuvinians. *Genetika* **34** (8), 1127–1233, 1998.
3. Golubenlo MV, et al. Analysis of distribution of “mongoloid” haplogroups of mitochondrial DNA among indigenous population of the Tuva Republic. *Genetika* **37** (6), 831–839, 2001.

4. Uinuk-Ool TS, et al. Origin and affinities of indigenous Siberian populations as revealed by HLA class II gene frequencies. *Hum Genet* **110**(3), 209–226, 2002.
5. Begovich AB, et al. Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J Immunol* **148**, 249–258, 1992.
6. Meyer, et al. Single Locus Polymorphisms. In: *HLA 2004: Immunobiology of the Human HMC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.
7. Leffell et al. Yup'ik Eskimo from the delta of the Yukon and Kuskokwim rivers, Alaska, USA. In: *HLA 2005: Immunobiology of the Human HMC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.
8. Mack, et al. Cañoncito Navajo from New Mexico, Maya from Mexico, Pima from Arizona, Sioux from South Dakota, and Zuni from New Mexico. In: *HLA 2004: Immunobiology of the Human HMC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.
9. Single, et al. Haplotype frequencies and linkage disequilibrium among classical HLA genes. In: *HLA 2004: Immunobiology of the Human HMC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.

North America

Population: Lacandon Mayan Indians from Mexico

Authors: Carmen Alaez, M. Vazquez-Garcia, Angelica Olivo, Clara Gorodezky

Authors' address: Department of Immunology and Immunogenetics, InDRE, SS. Mexico City. Mexico

The Mayan family of languages is spoken by a large group of Indian populations living in Southeast Mexico and neighboring Central American countries, and is thought to have been diverging for approximately 4600 years (1). The Lacandon Indians are one of the most studied of these Mayan groups, as they have lived in geographical isolation and have retained much of their culture. Approximately 650 Lacandon individuals live in the Chiapas Jungle on the Usumacinta River (although the Lacandon are known to have inhabited the Guatemalan Peten in the 16th century) in two communities; a northern group found between the lagoons of Met Sabok and Noh-Ha, and a southern group close to the Lagoon of Lacanja (2). The southern group has maintained their traditional culture and religion, while European missionaries have influenced the northern Lacandon since the first part of the 19th century. Each group speaks a different, but mutually intelligible, Yucatan Mayan dialect, and almost all Lacandon individuals are monolingual, speaking Lacandon only. They have preserved their genetic and cultural identity, do not admix with their neighbors and are highly endogamic (3, 4). Infectious diseases (e.g., malaria and tuberculosis), parasitic diseases, and skin allergies are prevalent. A high incidence of genetic disorders (e.g., albinism, epilepsy, retinoblastoma, and ophthalmologic diseases) and spontaneous abortions is likely an effect of inbreeding. Infanticide is practiced in cases of obvious physical malformation (5).

One hundred sixty-two Lacandon individuals were included in the study, representing four generations of 15 families. EDTA blood was drawn and mixed with supplemented RPMI for

transport. Mononuclear cells were isolated and kept at –196 degrees Celsius. DNA was obtained from the cells after treatment with proteinase K. DNA was phenol/chloroform extracted and precipitated with isopropanol. The second exons of DRB1, DQA1 and DQB1 were PCR amplified and dotted onto nylon membranes. Genotyping was carried out with 12th IHW primers and probes (6). Detection was accomplished using a chemiluminescence procedure. The results and allele frequencies have been previously published (7, 8). The IHWG Biostatistics Core carried out the statistical analyses performed here.

Deviations from overall expected Hardy-Weinberg equilibrium proportions were observed at the DRB1 (p value = 0.0323), DQA1 (p value = 0.0037), and DQB1 (p value = 0.0188) loci. These deviations likely result from the inclusion of related individuals in this study. However the Lacandon group is highly inbred, and it was impossible to obtain a completely unrelated population sample.

The most frequent DRB1 alleles are DRB1*0411 (allele frequency (a.f.) = 0.57), *0407 (a.f. = 0.1), *1402 (a.f. = 0.07), and *0404 (a.f. = 0.07). DRB1*0411 has been observed at high frequencies in some South Amerindian groups (e.g., the Venezuelan Bari). Although the frequency of the DRB1*0407 allele is much lower than *0411, *0407 has been observed at high frequencies in other Mexican Indian groups (e.g., the Seri and several groups from Oaxaca). The presence of certain DRB1 alleles not usually seen in Native American populations (DRB1*0801 (n=4), DRB1*0101 (n=1) and DRB1*0301 (n=1)) is likely due to Caucasian admixture.

The extent of inbreeding is underscored by the low haplotype diversity of this population; only 10 class II haplotypes were observed in the Lacandon. The prevalent haplotype DRB1*0411-DQA1*0301-DQB1*0302 has been observed at a very low frequency in North American Indians and natives from the northern part of Mexico. Four individuals had an alternative DRB1*0411 haplotype that includes DQB1*0402. This rare haplotype may be the result of a recombination event in the Lacandon population.

References

1. Manrique CL, Demonte V, Garza Cuaron B. Pasado y Presente de las Lenguas indgenas en México. In: Estudios Lingüísticos de Espana y México. México City: UNAM & Colegio de México, AC, México D.F. 387–420, 1990.
2. Meesmacher M. El area Maya, poblacion y territorio. Evolucion y dinamica. In: Dinamica Maya. Los Refugiados guatemaltecos. Fondo de Cultura Economica. México D.F. Pp. 23–83, 1986.
3. Baer PH, Merrifield WR. Two studies on the Lacandones of Mexico. Summer Institute of Linguistics. University of Oklahoma. Norman, 1971.
4. Serrano Sanchez C. Mestizaje e Historia de la Población en México con un Esbozo Antropológico de los Lacandones de Chiapas. In: Martín Municio A, García Barreno A. Eds. Polimorfismo Genico (HLA) en Poblaciones Hispanoamericanas. Real Academia de Ciencias Exactas, Fisicas y Naturales. Madrid, pp. 172–193, 1996.
5. Gorodezky C. Genetic Difference between Europeans and Indians: Tissue and Blood Types. *Allergy Proc* 13, 243–250, 1992.
6. Bignon JD, Fernandez Vina MA, Cheneau ML, Fauchet R, Schreuder GM, Clayton J, Marsh SGE, Charron D. HLA DNA Class II Typing by PCR-SSOP: 12th International Histocompatibility Workshop Experience. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol I, pp. 21–25.
7. Olivo A, Alaez C, Debaz H, Moreno M, Vazquez-García MN, De La Rosa G, Gorodezky C. The Influence of The MHC on Survival of Human Isolates. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol II, pp. 277–279.
8. Olivo A, Alaez C, Debaz H, Moreno M, Vazquez-García MN, De La Rosa G, Gorodezky C. Mexican Lacandon Indians of Mayan Ancestry Normal. In: Terasaki, PL., Gjertson D. Eds. HLA 1997. UCLA Tissue Typing Laboratory, Los Angeles, 1997, pp. 293–294.

Population: Seri from Sonora, Mexico

Authors: E. Infante, C. Alaez, H. Flores and C. Gorodezky

Authors' address: Department of Immunology and Immunogenetics, InDRE, SS. Mexico City. Mexico

Mexico's Indian population can be classified into 68 different linguistic groups, with more than five million Mexicans speaking Indian languages. These languages are arranged into five groups (Tarasco, Macromaya, Macromixteco, Macronahua and Macroayuma) that include twelve families (1). The Seri Indians belong to the Hokkano-Coahuitleca linguistic family in the Macroayuma group, a language family distinct from the Uto-Aztec languages (of the Macronahua group) spoken by the Pima and Papago. The first missionaries encountered several native groups when they came to Sonora in 1614, including Yaquis, Mayos, Nebomes, Opatas, Pimas and Seri. During the 16th century, the Seri took refuge in such inhospitable territories as the "Island of Sharks" ("Isla de Tiburon" in Spanish) in the Gulf of California, and returned to the mainland of Sonora in 1920, where they live as semi-nomads in coastal regions never reached by the Spanish conquerors. They have an isolated lifestyle, and number only 619 individuals. Inbreeding is common. They preserve their cultural identity, but because they are close to the State Capital, Hermosillo and have many visitors, 90% of them are bilingual, and have lost their traditional religion, practicing Protestant Christianity (2). Respiratory and viral infections, parasitic diseases, hypertension, type II diabetes, rheumatic and urinary disorders are prevalent among the Seri (3).

Although we originally studied 100 individuals in nine large families, we selected 33 unrelated individuals from the Seri populations of Desemboque and Punta Chueca for this study. Peripheral blood samples were drawn and mixed with enriched RPMI medium. Mononuclear cells were isolated and stored at -196 degrees Celsius until needed. Genomic DNA was extracted using a proteinase K-phenol/chloroform protocol (4). For class I typing, DNA was amplified for exons two and three of the HLA-A and B loci using the primers described for the 12th IHW and 13IHWG (5, 6). PCR products were dot blotted onto positively charged membranes and the hybridization signal of the SSOP was detected with a chemiluminescence protocol. The second exons of DRB1, DQA1 and DQB1 were amplified using the conditions and protocols of the 12th IHW. Typing of these genes was done in a dot blot format using 12th IHW SSOP oligonucleotides and protocols (7). DPA1 genes were typed using IHWG primers and probes. The IHWG Biostatistics Core performed statistical analyses in the context of the Anthropology Component.

The class I typing performed at low-intermediate resolution of the 100 members of the 19 families was published previously (8) as were the DRB1, DQA1, DQB1 and DQ promoter polymorphisms of the same individuals (9). In general, Seri HLA diversity is low; only 6 alleles are present for the A locus, 10 for B, 6 for DRB1, 4 for DQA1, 4 for DQB1 and 3 for DPA1. No deviation from homozygosity was detected using Ewens-Watterson homozygosity test, consistent with the assumption of neutrality. A deviation from Hardy-Weinberg equilibrium as detected by the Guo and Thompson test was observed for the HLA-A locus (p value = 0.034). These

results should be taken with caution, and may result from the small sample size (n=33).

The most frequent Seri alleles are typical of those seen in other Amerindian populations (A*0201, *6801, *3101, B*3501, *4002, *5101, *4005, DRB1*0407, *0802, *1402, DQA1*0301, *0401, DQB1*0302, *0402, *0301, DPA1*0103, and *0201). When these data are compared with those of other Mexican tribes (e.g., Nahuas, Mixtecos, Tarahumaras, Lacandon or Mazahuas), distinct differences are observed for the Seri. They do not have A33 and B39. The absence of B39 may be due to genetic drift, because B39 exists in Amerindians from the United States, (e.g., Navajo, Zuni and Havasupai). The presence of B*2705 was surprising, since B27 has not been found in any other Mexican Natives or any South American Indian group. Seri do not develop

ankylosing spondylitis, suggesting that other genes besides B27 are involved in the etiology of the disease (reviewed in 8). The Seri class II alleles are similar to those seen in other Amerindian populations. Admixture (primarily Caucasian) is estimated as 1.5% for A locus, 6.0% for B, 4.5% for DRB1, 3.0% for DQA1 and 3.0% for DQB1.

Prevalent haplotypes (A*0201-B*3501-DRB1*0407 and DRB1*0407-DQB1*0302, A*0201-B*4002-DRB1*0802 and DRB1*0802-DQB1*0402, and A*0201-B*4005-DRB1*1402 and DRB1*1402-DQB1*0301) also typify the diversity previously seen in Amerindian populations. The presence of these common haplotypes in Seri and almost all other native American groups supports the idea of a common ancestral origin for these populations, and support the notion of uniqueness of Amerindian groups (reviewed in 9).

References

1. Manrique CL, Demonte V, Garza Cuaron B, Pasado y presente de Las Lenguas Indgenas en México. In: Estudios Lingüísticos de España y México. México City. UNAM and Colegio de México, AC. 387–420, 1990.
2. Terrazas BE, Perez Ruiz ML, Pueblos Indgenas de México: Seris. México City. Instituto Nacional Indigenista, 1993.
3. Las Tribus de Sonora. <http://www.folkart.com/latitude/sonora/culture.html>
4. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd Edition. New York. Cold Spring Harbor Laboratory Press, 1989.
5. Kennedy LJ, Poulton KV, Thomson W, Williams F, Middleton D, Howell WM, Tarassi K, Papasteriades C, Albert E, Fleischauer D, Chandanayinyong JM, Tiercy, Juji T, Tokunaga WER, Ollier. HLA Class I DNA typing using sequence specific oligonucleotide probes (SSOP). 12th International Histocompatibility Workshop Experience. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol I, pp. 216–225.
6. Williams F, Mallon E, Middleton D. Development of PCR-SSOP for HLA-A typing for bone marrow registry donors. *Tissue Antigens* **49**, 61–66, 1997.
7. Bignon JD, Fernandez Vina MA, Cheneau ML, Fauchet R, Schreuder GM, Clayton J, Marsh SGE, Charron D. HLA DNA Class II Typing by PCR-SSOP : 12th International Histocompatibility Workshop Experience. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol I, pp. 21–25.
8. Infante E, Olivo A, Alaez C, Williams F, Middleton D, De La Rosa G, Pujol MJ, Duran C, Navarro JL, Gorodezky C. Molecular Analysis of of HLA class I alleles in the Mexican Seri Indians: Implications for their origin. *Tissue Antigens* **54**, 35–42.
9. Alaez C, Infante E, Pujol J, Navarro JL, Gorodezky C. Molecular Analysis of HLA DRB1, DQA1, DQB1, DQ promoter polymorphism and extended class I/class II haplotypes in the Seri Indians from Northwest Mexico. *Tissue Antigens* **59**, 388–396, 2002.

Populations: Cañoncito Navajo from New Mexico, Maya from Mexico, Pima from Arizona, Sioux from South Dakota, and Zuni from New Mexico

Authors: Steven J. Mack^{1,2}, Rosario Castro¹, Andrea J. Jani², Laura N. Geyer², Gary M. Troup³, Alan Keel⁴, Edward T. Blake⁴, Robert C. Williams⁵, Henry A. Erlich^{1,2}

Authors' addresses: ¹Roche Molecular Systems, Alameda, CA, ²Children's Hospital Oakland, Research Institute, Oakland, CA, ³University of New Mexico, Albuquerque, NM, ⁴Forensics Science Associates, Richmond, CA, ⁵University of Arizona, Tempe, AZ

The American continents were first colonized 15,000 to 25,000 years ago, most likely from northeast Asia (1). Considering linguistic differences and the geographic ranges of modern Native American groups, Greenberg proposed three distinct “waves” of colonization for the Americas (2) in which the first colonists spoke so-called Paleoindian languages and populated South America and much of North America. Greenberg theorized that Na-Dené speakers migrated to northwestern regions of North America about 10,000 years ago. In addition, some Na-Dené speaking groups (e.g.,

Apache and Navajo groups) migrated south from the Athabascan regions of Canada, arriving in the American Southwest 500-1000 years ago. Finally, Inuktitut-Aleut speakers were theorized to have arrived approximately 6,000 years ago and settled in the northern coastal regions of Alaska, Canada, and Greenland.

Subsequent to the late 15th and early 16th centuries, the size of the Native American population dropped dramatically, although this drop does not seem to have been as dramatic in South America as it has been in North America. Studies of genetic diversity in the Americas have suggested that Native American groups display a uniformly reduced genetic diversity, suggestive of a population bottleneck early in the history of these groups. HLA class II diversity has been observed to be restricted to a small number of ancestral alleles. For example, DRB1*1602, *0403, *0404, *0407, *0411, *1402, *1406, and *0802 are the primary DRB1 alleles seen in the so-called Paleoindian populations of North America (3). In general, South American populations show marginally increased class II diversity in comparison to North American populations (e.g., DRB1*0901 and DPB1*1401 are common in South America, but are mostly absent in North America), and evidence of recent novel allele diversity is seen in particular groups (e.g., DRB1*0417 in Argentina, *08042 in Ecuador, *0807 in South America, and *0811 in Na-Dené speaking groups). In addition, extensive novel class I (HLA-A and B) diversity has been observed in South American groups (4).

Here, we report allele and haplotype frequency distributions for five Native North American populations; Cañoncito Navajo, Maya, Sioux, Zuni, and two Pima population samples. Of these, the Sioux, Zuni, Maya and Pima speak Paleoindian languages, while the Cañoncito Navajo speak a Na-Dené language.

Populations

The Navajo people (aka Diné, meaning “the people”) are the second-largest Native American tribal group in the United States, and speak a Na-Dené language. The Cañoncito (aka To'hajiilee) Navajo reservation is in northwest/central New Mexico (34.9 degrees north latitude, 107.3 degrees west longitude), and is physically separate from the Navajo Reservation. While 270,000 people in the US identified themselves as Navajo in the 2000 census (5), the To'hajiilee (Cañoncito Reservation) population numbers approximately 3,000 people. HLA-A, C, B, DRB1, DQA1, DQB1, and DPB1 genotypes from 41 unrelated Cañoncito individuals are presented here.

The pre-Columbian Mayan civilization covered southeastern Mexico and Guatemala, and the contemporary Mayan lan-

guage family is a diverse group of 70–80 languages spoken in Mexico, Guatemala and Belize (6). The modern Mayan population is similarly large, with several million people speaking Mayan languages. Approximately 700,000 people speak the Yucatan Mayan language. HLA-DRB1, DQA1, DQB1 and DPB1 genotypes for 15 unrelated Mayan individuals from Mexico's Yucatan peninsula (20 degrees north latitude, 90 degrees west longitude) are reported here. Because of the small sample size of the Maya population sample, this data was not included in all Biostatistic analyses described in this Joint Report.

The Sioux are the fifth-largest Native American tribal group in the United States. Historically, this group included Lakota, Dakota and Nakota language-speaking peoples. These languages are currently spoken by approximately 40,000 individuals in North and South Dakota, northern Nebraska, northeastern Montana, southern Minnesota and Canada (6). In 2000, approximately 108,000 people in the US identified themselves as Sioux, and an additional 45,000 as being of partial Sioux ancestry (5). HLA-DRB1, DQA1, DQB1, and DPB1 genotypes for one hundred Sioux individuals from South Dakota, all of whom were registered with the US Bureau of Indian Affairs as Sioux Native Americans and some of whom have Caucasian admixture, are presented here. This population sample is composed primarily of Lakota speakers, but some Dakota and Nakota speakers are included as well.

The Zuni (A'shiwi) people live along the Zuni River in northwest New Mexico near Arizona (35.4 degrees north latitude, 108.9 degrees west longitude). The Zuni reservation includes approximately 9,000 people. The Zuni language (Zunian) is unrelated to the languages of neighboring peoples, but the Zuni culture is described as being similar to the cultures of neighboring “Puebloan” peoples (e.g., Acoma and Hopi). DRB1, DQA1, DQB1 and DPB1 genotypes for 50 unrelated Zuni individuals are reported here.

The O'Odham people reside in southern Arizona and northern Mexico and are comprised by the Pima and Papago groups. In 2000, approximately 8,500 people in the US identified themselves as Pima, and almost 3000 as being of partial Pima ancestry (5). The Akimel O'Odham (Pima) and Tohono O'Odham (Papago) dialects of O'Odham are more closely related to Uto-Aztecan languages spoken in the Mexico than to other such languages spoken in the US. The Pima reservation is in Gila River Arizona (33 degrees north latitude, 113 degrees west longitude) (6), and the Pima have been studied extensively due to the high incidence of type II diabetes in their population (7). Genotypes for two Pima population samples of 17 and 99 individuals (referred to here as Pima_17 and Pima_99) are pre-

sented here. The Pima_17 sample set was genotyped at the DRB1, DQA1, DQB1, DPA1, and DPB1 loci, while the Pima_99 sample set was genotyped at the HLA-A, B, DPA1 and DPB1 loci. Because of the small size of the Pima_17 population sample, this data was not included in all Biostatistic analyses described in this Joint Report. Studies of the frequencies of class I and class II serological specificities in larger Pima population samples (of which the Pima_99 sample is a subset) have been published previously (8, 9).

Genotyping

HLA class I (HLA-A, C, and B) genotyping was carried out using the IHWG Reverse Line Strip typing system (10). Class II (DRB1, DQA1, DQB1, DPA1, and DPB1) genotyping was accomplished using previously described PCR-SSOP methods (11, 12).

Statistical Analyses

All datasets were analyzed by the IHWG Biostatistics Core. In addition, DPA1:DPB1 haplotypes were estimated using the PyPop software package developed by the Biostatistics Core (13). Deviations from overall expected Hardy-Weinberg (HW) equilibrium proportions were observed in the Cañoncito Navajo dataset at the HLA-B locus (p -value < 0.00001), in the Sioux dataset at the DRB1 and DQB1 loci (p -values = 0.0132 and 0.0045), and in the Pima_17 and Pima_99 datasets at the DPB1 locus (p -values = 0.0058 and 0.0227). The HW deviations in the Sioux dataset are due to significant excesses of homozygotes at both loci (9.33 homozygotes expected, 24 observed at the DRB1 locus, p -value < 0.00001 , and 27.96 expected, 43 observed at the DQB1 locus, p -value = 0.0045), suggesting that this population sample is significantly admixed (see above) or that the samples represent multiple, distinct populations. HW deviations at these same loci have been detected in an independent study of another Lakota Sioux population (14). In both Pima population samples, non-significant excesses of DPB1*0401 heterozygotes may be contributing to the overall deviations from expected HW proportions (6.62 *0401 heterozygotes expected, 9 observed in Pima_17, and 40.29 expected, 48 observed in Pima_99).

Most of the allele frequency distributions in these populations did not deviate significantly from those expected for populations of the same size, displaying the same number of alleles, and evolving under neutral conditions, as determined using the Ewens-Watterson homozygosity test. However, homozygosity values at the DQA1 locus in the Pima_17 population sample and the DPA1 locus in the Pima_99 population

sample were significantly higher than expected under neutrality (F_{nd} = 1.8975 and 2.1709, respectively, and p -values = 0.001 and 0.018 respectively). In addition, F_{nd} for the DRB1 locus in the Pima_17 population (2.1355) sample was close to being significantly high (p -value = 0.045). Finally, the homozygosity values in several of these Native American populations are uncharacteristic of homozygosity values observed in populations from other parts of the world; whereas most populations outside of the Americas display negative F_{nd} values at the HLA-A, C, B, DRB1, DQA1, and DQB1 loci and either very small or positive F_{nd} values at the DPB1 locus, positive F_{nd} values were observed for the DRB1, DQA1, DQB1 loci in the Pima, Maya and Cañoncito Navajo populations, and negative F_{nd} values were observed at the DPB1 locus in the Pima, Sioux, Zuni and Cañoncito Navajo populations.

In each population, significant levels of linkage disequilibrium (LD) were detected for each locus pair, with the exception of locus pairs that included the DPB1 locus. While, significant LD was observed in the DPA1-DPB1 pairs in both Pima population samples, significant LD was not detected in 10 out of the 19 pairs involving other loci and the DPB1 locus. In addition, significant LD was not detected between the DQA1 and DQB1 in the Pima_17 population sample.

Allelic Diversity

Only four HLA-A alleles (A*0201, *0206, *2402, and *3101) were identified in the Cañoncito Navajo, and these were the four most frequent alleles (with a combined frequency of 95%) in the Pima_99 population. Cañoncito Navajo HLA-C diversity was greater than that seen for HLA-A, with seven alleles identified. Of these, the four most frequent alleles (with a combined frequency of 85%) were Cw*0401, *0801, *0702, and *0304. The HLA-B*3501, *4801, and *5102 alleles were identified in the Cañoncito Navajo and Pima_99 population samples at high frequencies ($> 5\%$). In addition, the B*2705, *3906, and *4002 alleles were seen at frequencies between 3–5% in the Cañoncito Navajo, and 7–10% in the Pima. Of the remaining high-frequency alleles, the Cw*3901 and *5301 alleles were seen only in the Cañoncito Navajo (allele frequencies (a.f.) = 16% and 9%), and the *4005 allele was seen only in the Pima (a.f. = 15%).

The DRB1*0407, *0802, and *1402 alleles were observed in all five populations, and at frequencies of at least 3% and as high as 53% (*0407), 44% (*0802) and 79% (*1402). In addition, the *0403, *0411, *1406, and *1602 alleles were seen at high frequencies in at least two populations. The *1402 allele was seen at a frequency of 79% in the Pima_17 population sample. While the sample size for this dataset is low, the DR52 serologi-

cal specificity (which includes the *1402 allele) was observed at a similar frequency (74%) in a much larger ($n=349$) O'odham population sample (9). The largest number of DRB1 alleles was seen in the Sioux, and a number of alleles not usually seen in Native American populations (e.g., DRB1*0701, *0801, *1501, *1601) are present in this population, suggesting considerable non-Native American admixture in this population sample. The Cañoncito Navajo population was distinguished from the other four populations by the presence of the *0406 and *1401 alleles at frequencies of 1% and 4% respectively. The DRB1*0411 allele was seen in the Maya and Zuni (a.f.=0.13 and 0.05); this allele is primarily seen at high frequencies in Central and South American populations, and the presence of this allele in the Zuni may suggest a Central or South American origin for this population.

The DQA1*0301, *0401, and *0501 alleles were observed in all five populations, with combined frequencies ranging from 88% in the Sioux to 100% in the Pima. DQA1*0501 was observed at a frequency of 94% in the Pima. The *0103 and *0201 alleles were present in the Sioux at a combined frequency of 5%, again suggesting non-Native American admixture in this population. The DQB1*0301 allele was observed at very high frequencies (a.f. ranged from 0.13 in the Maya to 0.94 in the Pima) in all five populations, and the *0201, *0302, and *0402 alleles were observed at high frequencies in at least two populations. DQB1*0301 and *0402 were the two most frequent Cañoncito Navajo DQB1 alleles, with combined frequencies of 91%.

DPA1*0103 was the most frequent DPA1 allele in the Pima_99 population sample (a.f.=0.95) and was the only DPA1 allele detected in the Pima_17 sample. The DPB1*0402 allele was observed at high frequencies (a.f. ranged from 0.25 in the Sioux to 0.73 in the Maya) in all five populations, and the *0201, *0301 and *0401 alleles were observed at high frequencies in at least three populations. In addition to these, the DPB1*2301 allele was observed at a frequency of 15% in the Pima_17 population sample. This allele was also detected in the Sioux at a low frequency (0.6%). While a considerable number of DPA1 and DPB1 alleles that were identified in the Pima_99 population sample went undetected in the smaller Pima_17 population sample, the frequencies of the major DP alleles in both datasets (e.g., 26% and 31% for DPB1*0401, and 53% and 50% for *0402) are roughly equivalent.

Haplotypic Diversity

Haplotype estimation can result in large numbers of haplotypes, a great number of which will have only been observed

in one or two individuals. Our aim here is to report those haplotypes that appear at high frequencies in multiple populations, or which are characteristic of a particular population by virtue of their appearance in that population at a high frequency.

Sixty-five A:B haplotypes were identified in the Pima_99 population sample, while 47 were identified in the Cañoncito Navajo. The HLA-A*0201:B*5102 and *2402:*3501 haplotypes were identified at high frequencies in both the Cañoncito Navajo and Pima_99 population samples. A*0201:B*5102 was the highest frequency A:B haplotype in the Pima (haplotype frequency (h.f.)=0.12), and *2402:*3501 was the highest frequency A:B haplotype in the Cañoncito Navajo (h.f.=0.16). A*0201:B*3501, *0201:4801, *3101:4801 were observed at low to intermediate frequencies in both populations. In addition, the A*0201:B*3901 haplotype was observed at a frequency of 11% in the Cañoncito Navajo, and the *2402:4005 haplotype was observed at a frequency of 9% in the Pima.

The DRB1*1402:DQA1*0501:DQB1*0301 haplotype was identified in all five populations at frequencies ranging from 3% in the Maya to 79% in the Pima. In addition, the *0403:*0301:*0302, *0407:*0301:*0302, *0411:*0301:*0402, *0802:*0401:*0402, and *1602:*0501:*0301 were identified at high frequencies in at least two populations. The *0404:*0301:*0302 haplotype was identified only in the Sioux and at a frequency of 12%. Several of these DRB1*04 alleles were identified in haplotypes containing the DQA1*0301 and DQB1*0302 alleles, potentially suggesting selection for particular DRB1*04:DQ haplotypes, or that these DRB1*04 alleles were generated in the Americas from an ancestral DRB1*04 alleles. That the DRB1*0411 allele is associated with a different DQB1 allele perhaps suggests a different (i.e., South American) origin for this allele. We note that in many Asian populations, the predominant DR4 haplotype is DRB1*0405-DQB1*0402. As might be expected from the number of presumed non-Native American DRB1, DQA1 and DQB1 alleles identified in the Sioux, a large number of non-Native American DRB1:DQA1:DQB1 haplotypes were observed in the Sioux at intermediate frequencies (e.g., *0701:*0201:*0201 at 4% and *1501:*0102:*0602 at 3%).

Fifty-one of the seventy-one distinct DRB1:DPB1 haplotypes seen in these five populations were detected in the Sioux. Of these, only the DRB1*1402:DPB1*0402 haplotype was identified in all five populations (h.f. ranged from 0.03 in the Maya to 0.43 in the Pima). In addition, *0403:*0402, *0802:*0402, *1402:*0401, and *1602:*0402 were observed at high frequencies in multiple populations. The

*0802:*0401 haplotype was identified in the Cañoncito Navajo alone, and at a frequency of 24%. In the Pima_99 population sample, 81% of the DPA1:DPB1 haplotypes were either DPA1*0103:DPB1*0401 (h.f.=0.31), or *0103:*0402 (h.f.=0.5). The high frequencies of these DPA1:DPB1 haplotypes are consistent with those seen in other Native American populations, as described elsewhere in this volume (15).

Conclusions

The unique history of Native American populations distinguishes them from the rest of the world in several ways; class II allelic diversity is reduced, and selective forces and demographic history have shaped allele frequency distributions in unexpected ways. The identification of new HLA alleles in the Americas may represent increased selection for diversity, given the lack of diversity in the founding populations. Signi-

ficant distinctions between Cañoncito Navajo alleles and haplotypes and those from Paleoindian-speaking populations are absent, providing little support for the “three-waves” hypothesis proposing separate, distinct colonization histories for populations speaking Paleoindian languages in contrast to those speaking Na-Dené languages; while the Cañoncito Navajo population can be distinguished from the others by a few unique (usually class I) alleles and haplotypes, the number of shared alleles and haplotypes at each locus suggest either considerable recent gene-flow between the Cañoncito Navajo and their neighbors, or a shared common ancestry between this Na-Dené-speaking population and Paleoindian language-speaking populations. However, the distribution of HLA alleles in Inuktitut-Aleut (Inuit-Eskimo) groups (16), is consistent with the notion that these populations are distinct from other Native American groups.

References

1. Fitzhugh WW, Chaussonnet V, Eds. *Anthropology of the North Pacific Rim*. Smithsonian Institution Press, Washington, 1994.
2. Greenberg, Joseph H. *Language in the Americas*, Stanford University Press, Stanford, 1987.
3. Erlich HA, Mack SJ, Bergstrom T, Gyllenstein UB. HLA class II alleles in Amerindian populations: implications for the evolution of HLA polymorphism and the colonization of the Americas. *Hereditas* **127**, 19–24, 1997.
4. Cadavid LF, Watkins DI. Heirs of the jaguar and the anaconda: HLA, conquest and disease in the indigenous populations of the Americas. *Tissue Antigens* **50**, 702–711, 1997.
5. Ogunwale SU. The American Indian and Alaska Native population: 2002. Census 2000 Brief (C2KBR/01-15) United States Census Bureau. <http://www.census.gov/population/www/cen2000/briefs.html>, 2002.
6. Ethnologue: Languages of the World, 14th Edition Copyright ©2004 SIL International. <http://www.ethnologue.com/>.
7. Knowler WC, Pettitt DJ, Bennett PH, Williams RC. Diabetes mellitus in the Pima Indians: genetic and evolutionary considerations. *Am J Phys Anthropol* **62**, 107–114, 1983.
8. Williams RC, McAuley JE. HLA class II variation in the Gila River Indian Community of Arizona: alleles, haplotypes, and a high frequency epitope at the HLA-DR locus. *Hum Immunol* **33**, 29–38, 1992.
9. Williams RC, McAuley JE. HLA class I variation controlled for genetic admixture in the Gila River Indian Community of Arizona: a model for the Paleo-Indians. *Hum Immunol* **33**, 39–46, 1992.
10. Described in the Report of the IHWG Technology Core: Immunobiology of the Human MHC. Eds: Hansen JA, Dupont B. IHWG Press, 2006.
11. Mack SJ, Erlich HA. HLA class II polymorphism in the Ticuna of Brazil: evolutionary implications of the DRB1*0807 allele. *Tissue Antigens* **51**, 41–50, 1998.
12. Begovich AB, Moonsamy PV, Mack SJ, Barcellos LF, Steiner LL, Grams S, Suraj-Baker V, Hollenbach J, Trachtenberg E, Louie L, Zimmerman P, Hill AV, Stoneking M, Sasazuki T, Kononov VI, Sartakova ML, Titanji VP, Rickards O, Klitz W. Genetic variability and linkage disequilibrium within the HLA-DP region: analysis of 15 different populations. *Tissue Antigens* **57**, 424–439, 2001.
13. Alex Lancaster, Mark P Nelson, Richard M Single, Diogo Meyer, Glenys Thomson. “Py-Pop: a software framework for population genomics: analyzing large-scale multi-locus genotype data”. In: *Pacific Symposium on Biocomputing* 8 RB Altman, et al., eds. World Scientific: Singapore. Pp. 514–525, 2003.
14. Leffell MS, Fallin MD, Hildebrand WH, Cavett JW, Iglehart BA, Zachary AA. HLA alleles and haplotypes among the Lakota Sioux: report of the ASHI minority workshops, part III. *Hum Immunol* **65**, 78–89, 2004.
15. Mack SJ, Erlich HA. Ticuna from Brazil. In: *HLA 2004: Immunobiology of the Human MHC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.
16. Leffell MS, Fallin MD, Erlich HA, Fernandez-Vina M, Hildebrand WH, Mack SJ, Zachary AA. Yup'ik Eskimo from Yukon and Kuskokwim rivers, Alaska. In: *HLA 2004: Immunobiology of the Human MHC*. Eds: Hansen JA, Dupont B. IHWG Press, 2006.

Population: Yup'ik Eskimo from the delta of the Yukon and Kuskokwim rivers, Alaska

Authors: Mary S. Leffell¹, M. Daniele Fallin², Henry A. Erlich^{3,4}, Marcelo Fernandez-Vina⁵, William H. Hildebrand⁶, Steven J. Mack^{3,4} and Andrea A. Zachary¹

Authors' addresses: ¹Department of Medicine, Johns Hopkins University, Baltimore, MD, USA, ²Department of Epidemiology, Johns Hopkins University School of Public Health, Baltimore, MD, USA, ³Department of Human Genetics, Roche Molecular Systems, Alameda, CA, USA, ⁴Children's Hospital, Oakland Research Institute, Oakland, CA, USA, ⁵C.W. Bill Young DoD Marrow Program, Georgetown University, Kensington, MD, USA, and ⁶University of Oklahoma, Oklahoma City, OK, USA

Alaska Native Eskimos belong to two major linguistic groups. The Inupiat live in the far north and the Yupik live on St. Lawrence Island and in the delta of the Yukon and Kuskokwim rivers (YK Delta) in southwest Alaska. The Yupik distinguish themselves as either the Siberian Yupik of St. Lawrence Island (numbering about 1200), and the Central Yup'ik of the YK Delta (comprised of over 22,000 individuals) (1). While the term “Yupik” refers to both groups, the Central Yupik prefer the term “Yup'ik”, derived from “yuk” (person) and “pik” (real or genuine) meaning “real people”. “Yup'ik” (plural “Yupiit”) can refer to the people, language, and culture (1, 2).

Abundant resources have permitted the Central Yup'ik a settled existence in villages and fishing camps along the Bering Sea coast and the lower stretches of the Yukon and Kuskokwim rivers. Travel between villages was limited by geography and weather. Today, the villages are accessible only by air, or by boat in summer and snowmobile or dog sled in winter. The YK region first experienced significant non-native settlement in the early 1800s with the arrival of Russian traders and Russian Orthodox and Moravian missionaries (2, 3). The Yup'ik experienced an estimated 60% population reduction in some villages after a smallpox epidemic in 1838, followed by influenza epidemics in 1852 and 1861 (1, 3). According to United States 2000 Census data, 88.2% of the YK Delta population claim Alaska Native heritage (4). The Yup'ik reside in some 70 villages ranging in size from 150 to 800, with the exception of two large regional centers, Bethel and Dillingham.

HLA alleles were defined for 460 volunteer Yup'ik Eskimos from the YK Delta as part of the ASHI coordinated study among minority populations, the first comprehensive allele-level study of Yupik HLA allele and haplotype diversity (5). Portions of this report are re-printed here with permission from Human Immunology. Samples were collected from 8 villages representing the geographic distribution of the Yup'ik across the YK Delta. Full Yupik ancestry through both maternal and paternal grandparents was claimed by 81.1% of participants in this study. A subset of 252 unrelated Yup'ik adults (having no first-degree relatives in the study) typed at the HLA-A, C, B, DRB1, DQA1, and DQB1 loci using SBT, SSOP, RLS and RSCA methods was submitted to the Anthropology/Human Genetic Diversity Component for analysis.

Two previously reported novel alleles were found among the Yup'ik (6). A*24021 was found in a single individual, while Cw*0806 was observed at a frequency of 0.10. This

HLA-C allele was found in participants from all eight villages and is strong linkage-disequilibrium with B*4801.

The Yup'ik share several common alleles with other Native American populations, including: A*2402, *0206, *6801; B*1501, *2705, *3501, *4002, *4801, *5101; Cw*0202, *0304, *0401; and DRB1*1602 (5). Previous studies have noted reduced DRB1 allele diversity in Amerindian and Na-Dene Athabaskan groups (7, 8); only DR2 (DRB1*16), 4, 6 (DRB1*14), 8 and 9 allelic lineages are commonly observed in these groups. Among these DR lineages, DR4 demonstrates the most diversity (although the predominant allele varies in different populations), and common Amerindian DR4 alleles include DRB1*0403, *0404, *0407, *0408, *0410, and *0417 (7). In the Yup'ik, DRB1*0401 was the single predominant DR4 allele observed (allele frequency (a.f.) = 0.232). In addition, DRB1*1101 was observed at a frequency of approximately 0.11 in the Yup'ik. Neither the DRB1*0401 nor the *1101 allele is observed in Amerindian or Na-Dene groups. The Yup'ik class II allele distribution is similar to those reported for the Siberian Eskimos of the Chukotha peninsula on the coast of the Bering Sea (9) and the Eskimos of East Greenland (10). DRB1*0401 and *1101 were observed at high frequencies in both of these Eskimo populations, confirming a closer relationship among the Yup'ik, Siberian and Greenland Eskimos than to other Amerindians.

Analysis of genotype frequency data for the 252 unrelated Yup'ik subjects by the IHWG Biostatistics Core revealed an overall deviation from expected Hardy-Weinberg equilibrium proportions at the DQB1 locus (p -val = 0.0195). In addition, an excess of homozygotes (p -val = 0.021) and an absence of B*2703:2705 genotypes (expected = 6.33, observed = 0, p -value = 0.0119) was observed for the HLA-B locus. These deviations at the HLA-B locus likely result from difficulty resolving allele ambiguity in genotypes homozygous for the ambiguous B*2703/05 allele detected by the RLS system. Finally, a non-significant positive homozygosity value (Fnd 1.55) was observed at the HLA-A locus, resulting from the high frequency of A*2402 (a.f. = 0.58).

The success of this project is directly attributable to the Yup'ik people who supported and participated in this study. The authors gratefully acknowledge the invaluable assistance of the staff of the Yukon Kuskokwim Health Corporation and the community health aides in organizing and facilitating the fieldwork. Portions of this report were reprinted with permission from Human Immunology and the American Society for Histocompatibility and Immunogenetics.

References

1. S. Beeman, Inupiaq and Yupik people of Alaska: introduction. *Alaska Geographic* 28 (2001), p. 4.
2. A. Fienup-Riordan Boundaries and Passages: Rule and Ritual in Yup'ik Eskimo Oral Tradition. Volume 212 in the Civilization of the American Indian series, University of Oklahoma Press, Norman, OK, 1994.
3. A. Fienup-Riordan, The Yup'ik people of southwestern Alaska: continuity and change. *Alaska Geographic* 28, 8, 2001.
4. Census 2000 Redistricting Data (Public Law 94–171). Summary File <http://www.quickfacts.census.gov>.
5. MS Leffell, MD Fallin, HA Erlich, M Fernandez-Vina, WH Hildebrand, SJ Mack, AA Zachary, HLA antigens, alleles and haplotypes among the Yup'ik Alaska natives: report of the ASHI Minority Workshops, Part II. *Hum Immunol* 63, 614, 2002.
6. Ellexson-Turner ME, Leffell MS, Zachary AA, Turner S, Bennett T, Sidebot DAtom, Cao K, Fernandez-Vina M, Hildebrand WH. Sequence-based typing of HLA class I alleles in Alaskan Yup'ik Eskimo. *Hum Immunol* 62, 639, 2001.
7. Erlich HA, Mack SJ, Bergstrom T, Gyllenstein UB. HLA class II alleles in Amerindian populations: implications for the evolution of HLA polymorphism and the colonization of the Americas. *Hereditas* 127, 19, 1997.
8. Monsalve MV, Edin G, Devine DV. Analysis of HLA class I and class II in Na-Dene and Amerindian populations from British Columbia, Canada. *Hum Immunol* 59, 48, 1998.
9. Grahovac B, Sukernik RI, O'hUigin C, Zaleska-Rutczynska Z, Blagitko N, Raldugina O, Kosutic T, Satta Y, Figueroa F, Takahata N, Klein J. Polymorphism of the HLA class II loci in Siberian populations. *Hum Genet* 102, 27, 1998.
10. Welinder L, Graugaard B, Madsen M. HLA antigen and gene frequencies in Eskimos of East Greenland. *Eur J Immunogenetics* 27, 93, 2000.

South America

Population: Bari from Perija Mountain Range, Venezuela

Authors: Z. Layrisse, S. Montagnani, O. Balbas

Authors' address: Laboratorio de Fisiopatologia, Centro de Medicina Experimental, Instituto Venezolano de Investigaciones Cientificas, Caracas, Venezuela

The Bari population is one of the three Chibcha-speaking Amerindian populations living in Venezuela. They live in the tropical lowlands and humid forests southwest of Lake Maracaibo, having survived 500 years of war, slavery, hunger and disease, and have been known as Motilones since the 17th century, but as Bari only since 1960 (1). Peaceful contact between Europeans and the Bari extended from 1772 until 1818, when the 13 Spanish Capuchins mission villages were abandoned in Venezuela's Independence War (1). After the discovery and exploitation of oil and the installation of an industrial infrastructure in the region, the Bari were forced to fight with cattle ranchers who began to populate the area (2). In the 1950s, they were forced to move west to the foothills and valleys of the Perija mountain range on the frontier between Colombia and Venezuela (8 – 10 degrees north latitude, 72–73 degrees west longitude). Since the return of Catholic Capuchin missionaries in 1960 and the creation of Indian reservations by the Colombian and Venezuelan governments in 1961, the Bari population has slowly increased, reaching 1071 in 1982 in Venezuela. The Venezuelan Reserve set aside land for both the Bari and their traditional neighbors and enemies, the Yukpa. Between 1961 and 1966, an estimat-

ed 25% of the Bari died from epidemics of measles and malaria. Approximately 70% of the deaths were among very young children (3). The indigenous populations of the Perija mountain range suffered an epidemic Hepatitis B outbreak between 1979 and 1982 (4) and were subjected to HBV vaccination starting in 1983.

The Bari language has been classified as Chibcha, and is closely related to members of the same linguistic family in southern Central America and the Colombian Andes (1). As a consequence of their contact with western civilization, many of the Bari's traditional subsistence activities have been replaced. They now live in individual houses, are monogamous, and distinguish consanguineous relatives from relatives resulting from marriages.

The Chibcha-Paeza speaking populations have shown mean high frequencies of the I^o (ABO), P1, R1, Fy^a, and Hp1 genes and high variability of allele frequency for each marker. They differ from other indigenous groups by their high frequency of the D_{chi} transferrin, several private genetic variants (Alb-Warao, Alb YAN 1, Rh factor eⁱ, etc.), the presence of HLA-Cw1, and the absence or very low frequency of the Diego red cell antigen (5). Based on the gene frequencies of the erythrocytic systems it has been estimated that the Bari population has not mixed with European or African populations (6).

Very little is known about HLA Class I allele frequencies among the Chibcha-speaking populations. Serological data have been reported on 28 Bari (7), and 3 Colombian populations (Guaymi of Panama, Kuna, and Awa-Kwaiker (8–10)). High frequencies of A2, A24, B35 and B40, and an absence of A31, seem to be characteristic of these populations.

Table 1. Allelic frequencies obtained by analysis performed locally using data from 56 unrelated Bari individuals and by the IHWG Biostatistics Core on a different number of samples

Allele	N=56	N= 44	Allele	N=56	N= 44	Allele	N=55	N= 33
A*0101	0.0000	0.01136	B*1501	0.1250	0.11364	Cw*0102	0.1909	0.21212
0201	0.2321	0.21591	3543	0.1786	0.15909	0202	0.0091	0.00000
0217	0.0804	0.07955	3501	0.0179	0.01136	0304	0.4636	0.40909
0222	0.0625	0.09091	3508	0.0089	0.01136	0305	0.0182	0.0000
2402	0.6161	0.60227	3906	0.2411	0.23864	0401	0.0182	0.01515
			4002	0.3750	0.397734	0702	0.2364	0.30303
			5101	0.0089	0.01136	1502	0.0364	0.04545
			5104	0.0089	0.01136	1507	0.0091	0.01515
			5106	0.0000	0.01136			
			5201	0.0268	0.03409			
Blank	0.0089		Blank	0.0089				

The HLA class II alleles present among the Bari (11) belong to the 4 DRB1 lineages (*04, *08, *14 and *16) reported in indigenous populations of North and South America. DRB1-DQA1-DQB1 haplotypes present in decreasing order of frequency are: 1602-0501-0301; 0407-03-0302; 0411-03-0302; 0403-03-0302; 1402-0501-0301 and 0807-04010402.

Blood samples were collected in 1992–1993 from two villages: *Saimadodyi*, founded in 1968, is the major settlement within the reserve, with a population of 300 people in 1988. (3). *Campo Rosario* is the largest village outside the reserve, where 150 people now live (3). The collection of samples was organized by the Fundacion Zumaque, as part of an extensive health program to provide medical and dental assistance to the entire population of several Bari villages (12). A total of 102 Bari were selected for the study from a genealogy prepared in advance by an anthropologist, in order to have a sample of unrelated individuals, each of them with one or both parents to provide information for haplotype assignment. Data was obtained from a total of 93 individuals: 56 from *Saimadodyi* and 37 from *Campo Rosario*. These samples included 13 nuclear families from *Saimadodyi* and seven from *Campo Rosario*. 44.08% of the tested individuals were females. They were tested for class I antigens by serology, for complotypes (Bf, C4AB) and for Class II (DRB, DQA, DQB) by PCR-SSO DNA typing using 12th IHW reagents, allowing the description of extended (9 locus) haplotypes (11).

As part of the IHWG Anthropology component, we have tested the same samples for class I alleles using the Reverse Line Strip (RLS) typing kits. Analysis was performed independently in our laboratory using a subset of the 56 less-related Bari individuals and by the IHWG Biostatistics Core on a smaller number of unrelated samples (44 for A, 44 for B and 33 for C). No deviation from expected Hardy-Weinberg equilibrium proportions was observed in the Bari. The Ewens-Watterson homozygosity test of neutrality revealed no significant deviation from homozygosity under neutral condi-

tions. Allele frequencies obtained with both sets of samples are shown in Table 1. Only two lineages were observed at the A locus: A2 with 3 variants and A24 with only one variant. A larger number of variants was present at the B and C loci (5 at B and 6 at C). Most of the alleles correspond to putative ancestral Amerindian alleles, and only 3 are “novel” variants: B*3543, B*39061 and the recently described Cw*0305 (13). This last allele was seen only in our laboratory in 3 unrelated individuals who were re-typed and assigned Cw*0102, *0305 in 2 cases and Cw*0305, *0702/10 in one case by the RLS program. We have noticed that the RLS program assigned the Cw*03041 allele in homozygous form to 12 Bari individuals, giving the combination Cw*03041+Cw* 0305 as a second option. Sequencing will be necessary to find out the exact frequency of Cw*0305 in this population.

Pairwise linkage disequilibria estimates obtained for locus pairs were: 0.68052 for A: C, 0.84037 for A: B and 0.91519 for C: B.

The A-C-B haplotype frequencies obtained by our laboratory and by the IHWG Biostatistics Core are shown in Table 2, and reflect the difference in number and samples used for the analysis in each place. A large number of 3-locus haplotypes were observed (> 30), but only seven were found in two or more

Table 2. Most frequent class 1 haplotypes and selected others in the Bari population

HLA-A: C: B Haplotype	IHWG Analysis N=23 Frequency	Local Analysis N=56 Frequency
2402:0702:3906	0.23913	0.2054
2402:0304:4002	0.32609	0.3035
0201:0304:3543	0.10870	0.0804
0217:0102:3543	0.08696	0.0357
0201:0102:1501	0.06522	0.0982
0222:1502:5101	0.02174	0.0178
0201:0304:4002	0.00000	0.0089
0217:0304:4002	0.00000	0.0089
0222:0304:4002	0.00000	0.0089
0222:1507:5201	0.02174	0.0089
0222:0102:1501	0.02174	0.0089
0222:0401:3501	0.02174	0.0089
0222:1502:5104	0.02174	0.0089
0201:0702:3906	0.04348	0.0089
0222:0702:3906	0.02174	0.0000

copies. The novel Cw*0305 allele was found in association with A*2402-B*4002 and with A*2402-B*3543. The combination of Class I haplotypes described here with Class II loci variants identified previously among the same individuals, has

shown the existence of 32 different extended (A-C-B-Bf-C4-DRB1-DQA1-DQB1-DPA1-DPB1) haplotypes among the Bari (14). Twenty of these have been found in at least two members of nuclear families and 12 in four or more of the Bari tested.

References

1. Wilbert J. Historia y etnografía de las etnias Bari y Warao. In: Polimorfismo génico (HLA) en poblaciones Hispanoamericanas. (Eds) A Martín Municio, P García Barreno. Real Academia de Ciencias Exactas, Físicas y Naturales, Madrid. P. 149–170, 1996.
2. Lizarralde R. Bari settlement patterns. *Human Ecology* **19**(4), 437–452, 1991.
3. Lizarralde M, Lizarralde R. Bari exogamy among their territorial groups: choice and/or necessity. *Human Ecology* **19**(4), 453–467, 1991.
4. Hadler S, de Monzon M, Ponzetto A, et al. Delta virus infection and severe hepatitis. An epidemic in the Yucpa Indians of Venezuela. *Ann Intern Med* **100**, 330–344, 1984.
5. Salzano FM, Callegari-Jacques SM. South American Indians. A case study in evolution. Oxford, Clarendon, 1988.
6. Layrisse Z, Layrisse M, Rodríguez-Laralde A. Estudios genéticos en poblaciones de lengua chibcha. *Boletín Museo del Oro* No. 38–39: 71–85, 1995.
7. Johnson AH, Noreen H, Spees EK, et al. The distribution of HLA antigens in the Motilones Indians of Venezuela. *Tissue Antigens* **12**, 163–169, 1978.
8. Spielman RS, Migliazza EC, Neel JV, et al. The evolutionary relationships of two populations: a study of the Guaymi and the Yanomama. *Current Anthropology* **20**(2), 377–388, 1979.
9. Restrepo M, Rojas W, Montoya F, et al. HLA and malaria in four Colombian ethnic groups. *Rev Inst Med trop. Sao Paulo* **30**(5), 323–331, 1988.
10. Bernal JE, deBrigard D, Duran C, et al. HLA (A, B, C and D) antigens in the Awa-Kwaik-er Indians of Colombia. *Tissue Antigens* **37**, 141–144, 1991.
11. Layrisse Z, Guedez Y, Domínguez E, et al. Extended HLA haplotypes among the Bari Amerindians of the Perija Range: relationship to other tribes based on 4-loci haplotype frequencies. *Human Immunology* **44**, 228–235, 1995.
12. R Holmes, J Scorza (eds) Estado de salud indígena: los Bari ... Fundación Zumaque, Gerencia Regional Zulía. Caracas, Venezuela, 1994.
13. Hollenbach JA, Thomson G, Cao K, et al. HLA diversity, differentiation and haplotype evolution in Mesoamerican Natives. *Human Immunol* **62**, 378–390, 2001.
14. Layrisse A, Montagnani S, Balbas O, Arnaiz-Villena A. HLA Class I allele-level and 8-loci haplotype data of a Chibcha speaking Amerindian tribe. In: HLA 2004: Immunobiology of the MHC. Proceedings of the 13th International Histocompatibility Workshop. In press.

Populations: Guarani-Kaiowá and Guarani-Ñandeva Amerindians, Mato Grosso do Sul State, Brazil

Authors: M. L. Petzl-Erler, L. T. Tsuneto

Authors' address: Laboratory of Human Molecular Genetics, Department of Genetics, Federal University of Paraná, Curitiba, Brazil. petler@ufpr.br

Amerindians living in Brazil are very diverse. Today they number 220 distinct groups, speaking more than 180 different languages and comprising approximately 400,000 individuals (1). It is estimated that, at the time the Europeans first arrived, there were approximately 4–5 times more groups and 6–10 times as many people. The majority of the contemporary Brazilian Indians reside in 614 indigenous lands spread throughout the national territory, having a total extension corresponding to roughly 13% of the country (1). However, many of these lands are subjected to invasions.

The Guarani language belongs to the Tupi-Guarani linguistic family, which is part of the Tupi branch. The Guarani were amongst the first Amerindians contacted by the European

conquerors at the beginning of the 16th century. Part of the population was decimated, part incorporated by intermarriage into the Brazilian admixed population, and part remained isolated. The contemporary Guarani population in Brazil has been estimated to number 35,000 people, living in the Southern, Southeastern and Central regions. Guarani villages are also found in Argentina, Bolivia and especially, Paraguay. Guarani-M'byá, Guarani-Kaiowá and Guarani-Ñandeva are three sub-tribes, differing in several aspects of their culture. In spite of intense contact with neo-Brazilians and with other indigenous groups, the Guarani have preserved much of their cultural identity. Of the three groups, the M'byá have remained the most traditional.

We studied the M'byá previously for the HLA polymorphism (2–6), as well as for polymorphisms of other genes. The gene flow from non-Amerindians has been estimated as 4.5% (5) and from other Amerindians as 1.4% (2). We have first surveyed the Kaiowá and the Ñandeva, with the purpose of analyzing the polymorphism of the HLA genes (6) and of other genetic markers (7–11).

The Guarani-Kaiowá (n=144) were contacted in the indigenous lands of Amambai (23 degrees 06 minutes south latitude; 55 degrees 12 minutes west longitude; population of approximately 4,500 people among Kaiowá and Nandeva) and Limão Verde (23 degrees 12 minutes south latitude, 55 degrees 06 minutes west longitude. 460 people). The Guarani-Nandeva (n=53) were from Porto Lindo (23 degrees 48 minutes south latitude; 54 degrees 30 minutes west longitude. 1,600 people) and Amambai. These villages are located in southern Mato Grosso do Sul State, in Central Brazil.

Typing was performed according to the PCR-SSOP protocols of the IHWG (for HLA-A, B, C) and the 12th IHW (for HLA-DRB1, DQA1, DQB1).

Despite the smaller population sample size, the Nandeva presented more alleles than the Kaiowá for all loci, with the exception of HLA-B. This results from the higher gene flow from non-Amerindians observed for the Nandeva (6), which has been estimated as 14%, as opposed to 3% for the Kaiowá.

The most common alleles (frequency ≥ 0.1 in one or both sub-tribes) were (values for the Kaiowá and the Nandeva, respectively):

A*3101 (0.469, 0.356), *0201 (0.192, 0.308), *0204 (0.171, 0.058), *6801 (0.105, 0.087).

B*3505 (0.142, 0.066), *1504 (0.125, 0.198), *3905 (0.111, 0.094), *4004 (0.101, 0.104).

Cw*0702 (0.364, 0.240), *0304 (0.217, 0.192), *0401 (0.185, 0.173), *0303 (0.094, 0.192), *1503 (0.094, 0.173).

DRB1*1602 (0.351, 0.330), *1402 (0.205, 0.255), *0802 (0.108, 0.075), *0807 (0.101, 0.066).

DQA1*0501 (0.559, 0.660), *0302 (0.226, 0.113), *0401 (0.215, 0.189).

DQB1*0301 (0.559, 0.651), *0302 (0.170, 0.075), *0303 (0.097, 0.047), *0402 (0.173, 0.179).

The two sub-tribes differed significantly in allele frequencies for all loci. The differences of both the Kaiowá and the Nandeva relative to the M'byá are even more prominent. One of the most conspicuous findings is the virtual absence of allele DRB1*1413 from the Kaiowá and the Nandeva, as opposed to its high frequency among the M'byá (0.15). This allele had first been identified in the Guarani-M'bya (3) and ever since has been found only in the Ache Amerindians from Paraguay, at a frequency of 0.08 (6). This indicates a genetic link between the M'byá and the Ache, and emphasizes the uniqueness of the M'byá relative to the other Guarani sub-tribes. A deviation from expected Hardy-Weinberg proportions was noted for the Kaiowá at the HLA-A locus. This resulted from an excess of *0201 homozygotes ($p=0.0405$) and deficiency of *0201/*0204 heterozygotes ($p=0.0156$), most likely due to failure to distinguish the *0201 and *0204 alleles. The Ewens-Watterson test provided evidence of balancing selection for locus HLA-C in the Kaiowá ($F_{nd} = -1.533$, $p < 0.025$). A non-significant trend towards lower than expected homozygosity was also observed for the other loci, especially HLA-DQA1. However, among the Nandeva, this trend was seen only for the HLA-C locus.

References

1. www.socioambiental.org/ Captured on February 16, 2004.
2. Petzl-Erler ML, Luz R, Sotomaior VS. The HLA polymorphism of two distinctive South-American Indian tribes: the Kaingang and the Guarani. *Tissue Antigens* **41**, 227–237, 1993.
3. Petzl-Erler ML, McDevitt HO. Molecular analysis of the HLA-DRB genes in two tribes of Brazilian Indians. *Hum Immunol* **41**, 180–184, 1994.
4. Parham P, Arnett KL, Adams EJ, et al. Episodic evolution and turnover of HLA-B in the indigenous human populations of the Americas. *Tissue Antigens* **50**, 219–232, 1997.
5. Petzl-Erler ML, Gorodezky C, Layrisse Z, et al. Anthropology Component Report for Region Latin-America: Amerindian and Admixed Populations. In: Charron D (Ed.) *Genetic Diversity of HLA. Functional and Medical Implications*. EDK, Sèvres, France, pp. 337–345, 1997.
6. Tsuneto LT, Probst CM, Hutz MH, et al. HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché tribe. *Tissue Antigens* **62**:512–526, 2003.
7. Silva Jr WA, Bonatto SL, Holanda AJ, et al. Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. *Am J Hum Genet* **71**, 187–192, 2002.
8. Battilana J, Bonatto SL, Freitas LB, et al. Alu insertions versus blood group plus protein genetic variability in four Amerindian populations. *Ann Hum Biol* **29**, 334–347, 2002.
9. Gaspar PA, Hutz MH, Salzano FM, et al. Gene polymorphisms of CYP1A1, CYP2E1, GSTM1, GSTT1 and TP53 genes in Amerindians. *Am J Phys Anthropol* **119**, 249–256, 2002.
10. Bortolini MC, Salzano FM, Bau CHD, et al. Y-chromosome biallelic polymorphisms and Native American population structure. *Ann Hum Genet* **66**, 255–259, 2002.
11. Bortolini MC, Salzano FM, Thomas MG, et al. Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet* **73**, 524–539, 2003.

Population: Central American from Costa Rica and Panama

Authors: S. J. Mack^{1,2}, M. Stoneking³, H. A. Erlich^{1,2}

Authors' Addresses: ¹Children's Hospital Oakland Research Institute, Oakland, CA, ²Department of Human Genetics, Roche Molecular Systems, Alameda, CA, ³Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

The Central American population sample is comprised by 50 Chibchan language-speaking individuals from Costa Rica and Panama. Chibchan languages are spoken in an area that extends from northern and central Honduras down through Nicaragua, Costa Rica and Panama to western Colombia, northwestern Venezuela, and northern Ecuador, with a large number of languages spoken in Costa Rica and Panama (1). The Chibchan languages spoken by the individuals in this population sample are Bribri and Cabécar (two languages in the Talamancan group of Chibchan languages) and Guaymí (which can be further subdivided into the Ngäbere and Buglere languages).

In particular, 34 individuals from the Ujarraz and Tolari tribes, hybrid populations descended from Bribri- and Cabécar-speaking groups, from southeastern Costa Rica (9.5 degrees North latitude, 88.3 degrees West longitude), and 17 Guaymí-speaking individuals from northwest Panama (8.5 degrees North latitude, 82 degrees West longitude) were included in this population sample. More than 130,000 Guaymí-speakers live in Panama, while approximately 10,000 Bribri and Cabécar speakers live in Costa Rica. HLA-DRB1, DQA1, DQB1, and DPB1 genotypes were generated for this population using PCR-SSOP methods (2, 3). In addition to these 50 Chibchan-speaking individuals, five individuals from the Carib-speaking Macushi tribe of Guyana were inadvertently included in this population sample for the analyses carried out by the IHWG Biostatistics Core. The analyses presented in the other chapters of this Joint Report include genotypes from these Macushi samples, but the biostatistical analyses reported

here were carried out on genotypes generated for Guaymí, Ujarraz and Tolari individuals.

Genotypic proportions did not deviate from those expected under Hardy-Weinberg equilibrium. Homozygosity values (F_{nd}) were lower than expected under neutral conditions ($F_{nd}=0$) at the DRB1, DQA1, and DPB1 loci ($F_{nd}=-0.65$ to -1.28) and higher than zero at the DPB1 locus ($F_{nd}=1.23$), but none of these values deviated significantly from those expected under neutral conditions. Significant linkage disequilibrium was detected between the DRB1, DQA1 and DQB1 with D' values ranging from 0.96 to 1.0 (p -value <0.0001), as well as between the DRB1 and DPB1 loci ($D'=0.65$, p -value $=0.02$) and the DQA1 and DPB1 loci ($D'=0.47$, p -value $=0.036$).

With allele frequencies totaling 79%, DRB1*0407, *1402, *1602, and *0403 (allele frequencies (a.f.) $=0.25$, 0.24 , 0.2 , and 0.1 , respectively) were the most frequent Central American DRB1 alleles, and as suggested by the high D' values for the DRB1 and DQ loci, the four most frequent Central American DRB1:DQA1:DQB1 haplotypes are DRB1*0407:DQA1*0301:DQB1*0302, *1402:*0501:*0301, *1602:*0501:*0301, and *0403:*0301:*0302 (haplotype frequencies (h.f.) $=0.25$, 0.24 , 0.19 , and 0.1). The DRB1*0701 allele, an allele not typically observed in Native American populations, was seen at a frequency of 0.06 (along with the DRB1*0701:DQA1*0201:DQB1*0201 haplotype). Only three DQB1 alleles (DQB1*0201, *0301, and *0302) were observed. Eighty-three percent of DPB1 allele diversity was represented by the DPB1*0402 and *1401 alleles (a.f. $=0.72$ and 0.11), and the *0402 allele was found on all DRB1*1402 haplotypes (h.f. $=0.24$).

Overall, the HLA class II allele and haplotype diversity observed in this population is consistent with the reduced diversity seen in other Native American populations, although some non-native admixture is inferred from the presence of alleles not usually detected in Native American populations, such as DRB1*0701, at high frequencies (6%).

References

1. Greenberg, Joseph H. *Language in the Americas*, Stanford University Press, Stanford, 1987.
2. Mack SJ, Bugawan TL, Moonsamy PV, Erlich JA, Trachtenberg EA, Paik YK, Begovich AB, Saha N, Beck HP, Stoneking M, Erlich HA. Evolution of Pacific/Asian populations inferred from HLA class II allele frequency distributions. *Tissue Antigens* **55**, 383–400, 2000.
3. Begovich AB, Moonsamy PV, Mack SJ, Barcellos LF, Steiner LL, Grams S, Suraj-Baker V, Hollenbach J, Trachtenberg E, Louie L, Zimmerman P, Hill AV, Stoneking M, Sasazuki T, Koenkov VI, Sartakova ML, Titanji VP, Rickards O, Klitz W. Genetic variability and linkage disequilibrium within the HLA-DP region: analysis of 15 different populations. *Tissue Antigens* **57**, 424–439, 2001.

Population: Ticuna from Brazil

Authors: Steven J. Mack, Henry A. Erlich

Authors' Address: Children's Hospital Oakland Research Institute, Oakland, CA, and Department of Human Genetics, Roche Molecular Systems, Alameda, CA

The Ticuna are native to the regions near the Solimoes section of the Amazon river, in an area unevenly trisected by north-western Brazil, northeastern Peru, and southeastern Colombia (4 degrees South latitude, 70 degrees West longitude). The Ticuna population is large, with approximately 12,000 Ticuna in Brazil, 8,000 in Peru, and 4,000 in Colombia. Contact with the Ticuna was established by Europeans in the 17th century, and modern Ticuna maintain contact with the surrounding native groups. The Ticuna language is an isolate, unrelated to neighboring languages (1). The 49 Ticuna individuals in this population sample are unrelated Brazilian representatives of this group. Allele and haplotype frequencies for the DRB1, DQA1, DQB1, and DPB1 genotype data presented and analyzed here have been previously published (2). DPA1 genotype data were generated using PCR-SSOP methods (3), and this population is one of the few Native American groups for which HLA-DPA1 allele frequencies have been reported.

The genotype data generated for this population were analyzed by the IHWG Biostatistic Core. DPA1:DPB1 haplotypes were estimated independent of these analyses using the PYPOP software package developed by the Biostatistic Core (4). Significant deviations from genotypic proportions expected under Hardy-Weinberg equilibrium were detected at the HLA-DRB1 and DPB1 loci (p values = 0.0049 and 0.0056). While homozygosity values at all typed loci were all lower than those values expected under neutral conditions, homozygosity at the DQA1 locus was significantly lower than neutral expectations (F_{nd} = -1.735, p_{val} = 0.0171). Significant levels of linkage disequilibrium were detected between all pairs of loci (all p values were <0.0001 except for DQA1:DPA1, with a p value of 0.001) with D' values ranging from 0.76 for the DQB1 and DPB1 loci to 1.0 for the DRB1 and DQA1 loci.

With combined frequencies of 80%, the three most frequent Ticuna DRB1 alleles were DRB1*0411, *1602, and *0807 alleles (allele frequencies (a.f.) = 0.32, 0.25 and 0.23). The DRB1*0807 allele is a putatively novel allele generated from *0802 (observed at a frequency of 0.03 in this population) by a point mutation at amino acid position 57 (2). To

date, this allele has only been observed in South American populations (e.g., in Brazil and Argentina), and at lower frequencies than in the Ticuna (1% to 12%), suggesting that the DRB1*0807 allele may originally have been generated in western Brazil. In addition, the high frequency of this allele in this population may suggest the operation of selection, favoring *0807 over *0802. Alleles not usually seen in Native American populations (e.g., DRB1*0801, a.f. = 0.01) were also observed in the Ticuna, suggesting a low degree of non-Native American admixture.

Allelic diversity at the DQ loci was low, with three alleles accounting for 100% of the DQA1 (DQA1*0301, *0401, and *0501) and 94% of the DQB1 (DQB1*0301, *0302, *0402) diversity in this population. While allelic diversity at the DPA1 loci was similarly low, with only the DPA1*0103, *02011, and *02012 alleles observed in this population, the *0103 allele was observed at a frequency of 72%. The three most frequent DPB1 alleles were DPB1*0402, *1401 and *0401 (a.f. = 0.43, 0.18 and 0.19). Each DRB1 allele was observed in a single DRB1:DQA1:DQB1 haplotype, with corresponding frequencies, and the high frequency DRB1:DQA1:DQB1 haplotypes were DRB1*0411:DQA1*0301:DQB1*0302, *1602:*0501:*0301, and *0807:*0401:*0402. Greater haplotype diversity was observed for DRB1:DPB1 haplotypes; while the DRB1*1602 allele was only observed in a DRB1:DPB1 haplotype with DPB1*0402, DRB1*0807 was observed in 2 DRB1:DPB1 haplotypes (with DPB1*0401 and *0402), while DRB1*0411 was observed in six DRB1:DPB1 haplotypes. With frequencies totaling 76%, the three most frequent Ticuna DPA1:DPB1 haplotypes were DPA1*0103:DPB1*0402, *0103:*0401, and *02011:*1401 (haplotype frequencies (h.f.) = 0.44, 0.17, and 0.15). Each of these three DPA1:DPB1 haplotypes has been observed at high frequencies in other native Central and South American populations; for example, *0103:*0402 was observed at high frequencies (0.46 to 0.90) in the Ecuadoran Cayapa population and the Mexican Mixe, Mixtec, and Zapotec populations, *0103:*0401 was observed at a frequency of 14% in the Zapotec, and *02011:*1401 was observed at a frequency of 42% in the Cayapa (3).

Ticuna allele and haplotype frequency distributions reported here (including those for DPA1) are consistent with patterns of reduced diversity observed for other Native American populations. Future studies of Native American populations should include DPA1 genotype data and estimated DPA1:DPB1 haplotypes to permit expanded multi-locus comparisons between the Ticuna and other groups.

References

1. Ethnologue: Languages of the World, 14th Edition Copyright ©2004 SIL International. <http://www.ethnologue.com/>.
2. Mack SJ, Erlich HA. HLA class II polymorphism in the Ticuna of Brazil: evolutionary implications of the DRB1*0807 allele. *Tissue Antigens* **51**, 41–50, 1998.
3. Begovich AB, Moonsamy PV, Mack SJ, Barcellos LF, Steiner LL, Grams S, Suraj-Baker V, Hollenbach J, Trachtenberg E, Louie L, Zimmerman P, Hill AV, Stoneking M, Sasazuki T, Konev VI, Sartakova ML, Titanji VP, Rickards O, Klitz W. Genetic variability and linkage disequilibrium within the HLA-DP region: analysis of 15 different populations. *Tissue Antigens* **57**, 424–439, 2001.
4. Alex Lancaster, Mark P. Nelson, Richard M. Single, Diogo Meyer, Glenys Thomson. "Py-Pop: a software framework for population genomics: analyzing large-scale multi-locus genotype data". In: *Pacific Symposium on Biocomputing* 8. Altman RB, et al., ed. World Scientific: Singapore. 514–525, 2003.

Admixed Populations

Populations: Irish from Northern Ireland, Brazilian from Minas Gerais State, Brazil, White Cuban from Cuba, Cuban Mulatto from Cuba, Omani from Oman, Singapore Chinese from Singapore

Authors: F. Williams and D. Middleton

Authors' address: Northern Ireland Tissue Typing Laboratory, City Hospital, Belfast, Northern Ireland, United Kingdom

As noted below, the genotype data on these populations have been published previously.

Irish

The HLA- A, B, C and DRB1 loci were typed in a population sample of 1000 unrelated volunteer bone marrow repository blood donors in Northern Ireland. The sampled subjects were both rural and urban, and given that the population of Northern Ireland numbers approximately 1.5 million and that immigration to the region is low, this population sample should be considered an extremely representative sample of the general population. (1–3)

Brazilian

Samples were collected from 97 unrelated individuals in the metropolitan area of Belo Horizonte City, Minas Gerais State in Brazil for genotyping at the HLA-A and B loci. These indi-

viduals are Caucasoid and are considered typical for this area. (4, 5).

White Cuban and Cuban Mulatto

Unrelated blood donors from Havana City were randomly selected for genotyping at the HLA-A and B loci. These donors consisted of two sub-populations, Caucasoid (n=70) and Mulatto (n=42). (4, 5, 6)

Omani

Oman, situated on the southeast coast of the Arabian peninsula, borders on the United Arab Emirates, Saudi Arabia, and Yemen. The Omani population is predominantly Arab (74%), while a proportion of the remaining fraction have connections with the Indian subcontinent and Africa. The 128 unrelated Omanis sampled in this study, healthy individuals comprised of blood donors and potential bone marrow donors, were typed at the HLA-A and B loci. They resided in various areas of the country, and were considered representative of a normal population for this region (4, 5, 7). A deviation from expected Hardy-Weinberg equilibrium proportions was observed at the HLA-A locus in this population (p-value= 0.0199).

Singapore Chinese

Samples were collected from 149 unrelated blood donors of Chinese origin, living in Singapore, and were typed at the HLA-A and B loci. (3, 4, 5).

References

1. Williams F, Meenagh A, Patterson C, Middleton D. Molecular diversity of the HLA-C gene identified in a caucasian population. *Hum Immunol* **63**, 602, 2002.
2. Middleton D, Williams F, Hamill MA, Meenagh A. Frequency of HLA-B alleles in a Caucasoid population determined by a two-stage PCR-SSOP typing strategy. *Hum Immunol* **61**, 1285, 2000.
3. Williams F, Middleton D, Savage D, Gorodezky C, Wilson DW, Fitzgerald JM, Urbaniak SJ. Development of PCR-SSOP for the identification of HLA-A*02 subtypes and determination of HLA-A*02 frequencies within different ethnic populations. *Tissue Antigens* **49**, 129, 1997.

4. Middleton D, Williams F, Meenagh A, Daar AS, Gorodezky C, Hammond M, Nascimento E, Briceno I, Perez MP. Analysis of the Distribution of HLA-A Alleles in Populations from Five Continents. *Human Immunology* **61**, 1048, 2000.
5. Williams F, Meenagh A, Darke C, Acosta A, Daar AS, Gorodezky C, Hammond M, Nascimento E, Middleton D. Analysis of the distribution of HLA-B alleles in populations from five continents. *Hum Immunol* **62**, 645, 2001.
6. Williams F, Curran MD, Padadoa Perez M, Acosta A, Middleton D. Characterisation of a new HLA-B allele, HLA-B*0720, identified in the Cuban population. *Tissue Antigens* **57**, 80, 2001.
7. Williams F, Curran MD, Leheny WA, Daar AS, Middleton D. Characterization of HLA-B*3921 and confirmation of HLA-B*4415, two variant HLA-B alleles identified in the Omani population. *Tissue Antigens* **57**, 492, 2001.

Population: Brazilian (Admixed, African and European) from the Northeast region of State of São Paulo, Brazil

Authors: P. Louzada-Junior¹, N. H. Deghaide¹, M. B. Araujo², A. G. Smith³, M. H. S. Kraemer², E. A. Donadi¹

Authors' addresses: ¹Division of Clinical Immunology of Faculty of Medicine of Ribeirão Preto, University of São Paulo, Brazil, ²Department of Clinical Pathology of School of Medical Science of University of Campinas, Brazil, and ³Fred Hutchinson Cancer Research Center, Seattle, USA

Historical immigrations from Europe have distributed various Caucasian populations, mainly Portuguese, Spanish, Italian and German, across Brazil. Brazilian Amerindian populations still exist as semi-isolated tribal groups, but are also represented within modern urban populations. African-American individuals represent populations that originated from the Niger-Congo of Equatorial Africa. The language spoken in this region is exclusively Portuguese. The distribution and relative admixture of these varied racial and ethnic backgrounds, reflected in the HLA gene pool and present in a normal healthy population, can vary significantly on a regional basis. Although several serological HLA reports have been published (1, 2), few studies have been published on the HLA-class II alleles in the Brazilian population. Most studies have described normal individuals (3) and various Brazilian Amerindian populations (4, 5).

Since Amerindians are very rare in the Northeast region of the State of São Paulo, Brazil, the population presented predominant phenotypic features that were assigned as Caucasian, Black and Mulatto. A total of 100 random, normal, healthy unrelated blood donors were studied: 64 Caucasians, self-described to be of Western and Southern European ancestry; 25 Blacks, mostly of Bantu, Benin and Senegal African ancestry (6), and 11 Mulatto (Caucasian and Black admixture) individuals, all from the Northeastern region of the State of São Paulo (23 degrees 10 minutes 42 seconds south latitude, 47 degrees 48 minutes 24 seconds west longitude),

Brazil. The same investigator asked the donors about their ancestry and made the racial assignments. All individuals were submitted to a medical examination that showed no evidence of previous underlying diseases, and presented negative serology for Chagas' disease, hepatitis-B and C, and HIV infection. The Ethics Committee of the Faculty of Medicine of Ribeirão Preto, University of São Paulo, approved the study protocol and all subjects gave informed consent to participate in the study.

HLA-A, B and C loci were typed using IHWG reverse line strip (RLS) reagents, and DRB1 alleles were typed using previously described PCR/SSOP systems (7, 8).

The most frequent alleles in the sample studied were A*0201 (allele frequency (a.f.)=19.1%), B*0702 (10.1%), Cw*0701 (15.5%) and DRB1*0301 (12.1%). With respect to the number of distinct alleles, 54 alleles were found in locus B, followed by 34 alleles in each of locus A and DRB1, and 30 alleles in the Cw locus. At the HLA-A locus, the A2 group showed an extraordinary polymorphism, exhibiting eight alleles including A*0201, *0202, *0204, *0205, *0211, *0217, *0220 and *0224. In the HLA-B locus, seven B*15 (*1501, *1503, *1516, *1517, *1523, *1537 and *1538) alleles, six B*35 (*3501, *3502, *3505, *3512, *3517 and *3520) and six B*44 (*4402, *4403, *4404, *4405, *4406 and *4407) alleles were identified. The HLA-C locus showed five alleles of Cw*07 group (*0701, *0702, *0704, *0708 and *0714). Thirty-four unique DRB1 alleles were detected, and the DRB1*04 allele family was the most diverse, with seven alleles (*0401, *0402, *0403, *0404, *0405, *0407 and *0411). Most of the DRB1 alleles observed in this study have been frequently observed in many Caucasian populations (6,9,10). Although individuals of purely Amerindian heritage are very rare in the Northeastern region of the State of São Paulo, alleles of typical Amerindian populations, such as DRB1*0411, *0807 and *1402 (8, 9, 11, 12), were detected in this study among Caucasians, Blacks and Mulattos. Similarly, the DRB1*0302, *1503 and *0804

alleles, characteristic of African and North and South American Black populations (3,8,9), were found among Caucasians. The DRB1 locus presented the highest degree of observed heterozygosity (h) ($h=0.949$), followed by the A ($h=0.929$), B ($h=0.913$) and C ($h=0.896$) loci. According to the Guo and Thompson test, the HLA-A, B and Cw loci were in Hardy-Weinberg equilibrium. In addition, linkage disequilibria were observed in all tested locus pairs (A/C, A/B, A/DRB1, C/B, C/DRB1, and B/DRB1). The variability of HLA-A-C-B haplotypes was high (452), and the majority of haplotypes occurred in frequencies lower than 1%. The most frequent haplotypes were A*0101/Cw*0702/B*0702 and A*0101/Cw*0701/B*0801, at frequencies of 2.985% and 2.239%, followed by A*0101/Cw*0701/B*5801, A*0201/Cw*0401/B*0702, A*0202/Cw*0401/B*3501, A*0201/

Cw*0401/B*4402, A*3001/Cw*1701/B*4202 each with a frequency of 1.493%, and A*0101/Cw*0701/B*0805, A*0301/Cw*0702/B*0702 both with frequencies of 0.756%.

The extraordinary genetic diversity of HLA alleles seen in highly mixed populations may result from the aggregation of distinctive alleles from separate racial and ethnic groups (13). In particular, the Brazilian population of the present study included one example each of DRB1*0804 and DRB1*1401 associated with DQB1*0602, haplotypes confirmed by family studies in the Rio de Janeiro population (4). These results illustrate the heterogeneous composition of the normal Brazilian population with many individuals exhibiting a mixture of HLA alleles and haplotypes typical of Caucasian, Black and Native American populations.

References

- Petzl-Erler ML, Gorodezky C, Layrisse Z, et al. Anthropology report for Region Latin-America: Amerindian and admixed populations. In Charron D (ed) *Proceedings of the Twelfth International Histocompatibility Workshop and Conference 1997*. EDK, Paris, France.
- Mack SJ, Erlich HA. HLA class II polymorphism in the Ticuna of Brazil: evolutionary implications of the DRB1*0807 allele. *Tissue Antigens* **51**, 41, 1998.
- Fernandez-Vina M, Moraes JR, Moraes ME, Miller S, Stastny P. HLA class II haplotypes in Amerindians and in Black North and South Americans. *Tissue Antigens* **38**, 235, 1991.
- Moraes ME, Fernandez-Vina M, Salatiel I, Tsai S, Moraes JR, Stastny P. HLA class II DNA typing in two Brazilian populations. *Tissue Antigens* **41**, 238, 1993.
- Louzada-Junior P, Smith AG, Hansen JÁ, Donadi, EA. HLA-DRB1 and -DQB1 alleles in the Brazilian population of the northeastern region of the state of São Paulo. *Tissue Antigens* **57**, 158–162, 2001.
- Gjertson DW and Lee S-H. HLA-A/B and –DRB1/DQB1 allele level haplotype frequencies. In Gjertson DW and Terasaki PI (Eds), *HLA 1998*. American Society for Histocompatibility and Immunogenetics, Lenexa, USA, 365, 1998.
- Mickelson E, Smith A, McKinney S, Anderson G, Hansen JA. A comparative study of HLA-DRB1 typing by standard serology and hybridisation of non-radioactive sequence-specific oligonucleotides probes to PCR-amplified DNA. *Tissue Antigens* **41**, 86, 1993.
- Kimura A, Sasazuki T. Eleventh International Histocompatibility Workshop reference protocol for the HLA DNA-typing technique. In: Tsuji K; Aizawa M, Sasazuki T. (Eds) *HLA-1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference*. Oxford University Press, Tokyo, Japan. 1, 397, 1992.
- Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K; Aizawa M, Sasazuki T. (Eds) *HLA-1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference*. Oxford University Press, Tokyo, Japan. 1, 1065, 1992.
- Clayton J, Lonjou C. Anthropology Tables. Appendix B: HLA allele and haplotype frequencies by DNA typing: Local assignment. In: Charron D (ed) *Proceedings of Twelfth international Histocompatibility Workshop and Conference*. EDK, Paris, France, 712, 1997.
- Layrisse Z, Matos M, Montagnani S, Herrera F. The novel HLA-DRB1*0807 allele is present in a Carib Amerindian population: the Yucpa. *Tissue Antigens* **49**, 427, 1997.
- Erlich HA, Mack SJ, Bergstrom T, Gyllenstein UB. HLA class II alleles in Amerindian populations: implications for the evolution of HLA polymorphism and the colonization of the Americas. *Hereditas* **127**, 19, 1997.
- Parham P. HLA, anthropology and transplantation. *Transplant Proc* 1993; 25(1): 159.

Populations: African American, Caucasian, Hispanic, Asian Pacific Islander and Native American from the United States

Authors: K. Cao, M. A. Fernández-Viña

Affiliations: American Red Cross National Histocompatibility Laboratory; C.W. Bill Young/DoD Marrow Donor Program, Naval Medical Research

Center, Bethesda MD, USA and the Department of Oncology, Georgetown University, Washington DC, USA

Molecular typing of HLA-A, C and B alleles was performed for 1,332 unrelated subjects from five outbred North American populations living in the U.S.A. These include subjects self identified as African American (AFAM), Caucasian (CAU),

Hispanic (HIS), Asian Pacific Islander (ORI), and Native American (NA). PCR-SSOP methods were used to unequivocally type alleles of the HLA class I (HLA-A, C and B) loci (1). Genotypic ambiguities were resolved by group-specific amplification and SSOP hybridization (1). Polymorphisms in exons 2 and 3 were analyzed and alleles with differences outside these exons were not distinguished.

HLA class I alleles found in various ethnic groups: forty-six HLA-A alleles, 100 HLA-B alleles and 33 HLA-C alleles were observed in the present study. The number of alleles identified in each ethnic group ranged from 28 to 36 for HLA-A (28 in AFAM, 28 in CAU, 36 in HIS, 29 in NA, and 34 in ORI), 49 to 66 for HLA-B (49 in AFAM, 48 in CAU, 62 in HIS, 55 in NA, and 66 in ORI) and 22 to 26 for HLA-C (23 in AFAM, 22 in CAU, 26 in HIS, 22 in NA, and 24 in ORI). More alleles were identified in Hispanics and Asians. There was no apparent correlation between the number of alleles observed and the size of the population tested. In African Americans, HLA-A*2301, A*6802, A*3402, A*7401, A*0202, A*3601, A*0205, A*6601, A*6602, A*8001, and A*0102 were found at intermediate or high frequencies; these alleles were not observed or had significantly lower frequencies in other populations. In Asians, A*1101, A*2402, and A*3303 were found at significantly higher frequencies than those observed in other populations. A*0207, A*0206, A*0203, A*2407, A*3401 and A*2602 were common in Asians, but were absent or rare in Caucasians and African Americans. While no unique HLA-A allele was observed in Caucasians, the most frequent Caucasian HLA-A allele was A*0201. However, this allele is also common in other populations. HLA-A alleles observed in Caucasians were almost the same as those observed in Hispanics and Native Americans. The A*6803 and A*6805 alleles, which have been observed in Mexican Indians (2), were also observed in Hispanics.

The most frequent HLA-B alleles in Caucasians are B*4402, B*0702, and B*0801. These alleles were found in strong linkage disequilibrium with the three most common HLA-A alleles (A*0201, A*0301, and A*0101 respectively) in this population. B*5301 is the most common African American B allele; this allele is observed at much lower frequencies in Caucasians and was absent in Asians. The B*1503, B*4201, B*5802, B*8101, B*0705, B*5702, and B*1516 alleles have intermediate frequencies (allele frequencies (a.f.) range from 0.073 down to 0.009) and were rarely observed or were completely absent in Caucasians and Asians. These alleles may be considered to be unique to populations with African ancestry. B*5703 and B*7801 were found less frequently; however, they appear to be specific to African populations; some

of these alleles were found in Hispanics. This observation suggests that there was a significant flow of African genes into the Hispanic population. The most frequent HLA-B alleles in Asians are B*5801, B*4001, and B*5101; these alleles were also found frequently in other ethnic groups. The B*3802, B*4601, B*1502, B*5401, B*1301, and B*1521 alleles were found at significant frequencies in Asians; together with B*4006, B*5901, B*6701, B*1802, B*1511, B*1513, B*2704, and B*5602, these alleles appear to be unique to Asian populations. The B*3501, B*5101, B*4002, B*5201, B*1501, B*3512, B*3905, B*3901, B*3906, and B*4801 alleles were observed with a.f. >0.01 in Hispanics. These alleles are common in Native American tribes from Meso-America and South-America (2) and may reflect the contribution of Amerindians to the genetic make-up of the Hispanic population living in the U.S.A. The Amerindian contribution to Hispanics is reflected in the low frequencies of the B*3505, B*3504, B*1508, B*3517, B*3908, B*5102, B*1515, B*3902, B*1505, B*1530, B*3509, and B* alleles observed in this study. Other alleles common in Caucasians and Africans living in the U.S.A. were also observed in Hispanics.

Some HLA alleles contributed by the African and European populations to the Hispanics are different from those found in the Caucasians and Africans living in the U.S. For example, B*1801, B*4403, and B*1402 were found at higher frequencies in Hispanics than in U.S.A. Caucasians. These three alleles are more common in the European Mediterranean countries than in the Northern European countries. The lower frequencies of B*0801 and B*4402 in Hispanics (common in Ireland and other Northern European countries) seems to fit with this hypothesis. In the Native Americans B*5101, B*3501, B*2705, and B*1501 are the most frequent alleles. Together with B*4002, B*3901, and B*4801, these alleles are rather common in the North American Indian (3) and Eskimo populations (4). All these alleles have gene frequencies ranging from 0.02 to 0.11 and were observed at relatively high frequencies in Asians as well. These seven HLA-B alleles appear to represent the B locus allelic lineages found in all Amerindian tribes, and have been postulated to be the putative B-locus founder alleles of all American Indian tribes. The contribution of Northern European genes to the Native Americans appears to be reflected by the relatively high frequencies of B*4402, B*0801, and B*0702.

Fewer HLA-C alleles were identified. With gene frequencies ranging from 0.15 to 0.05 in each population, the most frequent C alleles are: Cw*0701, Cw*0702, Cw*0401, Cw*0501, Cw*0602, Cw*0304, Cw*0303, and Cw*1203 in Caucasians; Cw*0401, Cw*0701, Cw*0602, Cw*0202,

Cw*0702, Cw*1601, and Cw*1701 in Africans; Cw*0702, Cw*0102, Cw*0801, Cw*0304, Cw*0302, Cw*0401, and Cw*1402 in Asians; and Cw*0401, Cw*0702, Cw*0701, Cw*0304, Cw*0602, Cw*0501, Cw*1601, Cw*0802, and Cw*0102 in Hispanics and North American Natives.

In general, the most common HLA-C alleles in the five populations showed strong linkage disequilibria with the most frequent HLA-B alleles. For example, Cw*0701, Cw*0702, and Cw*0501 are in strong association with B*0801, B*0702 and B*4402, respectively, in Caucasians. B*5301, B*1503, and B*4201 are strongly associated with Cw*0401, Cw*0202, and Cw*1701, respectively, in Africans. In contrast to the HLA-A and B loci, HLA-C presents only a few alleles that are specific to a particular ethnic group. Examples of race-specific alleles are: Cw*1801 and Cw*0804 in Africans, Cw*0403 and Cw*1403 in Asians, and Cw*0305 in Hispanics. This last allele was also found in the Mexican Indians tribes Mixteco and Mixe Alta from Oaxaca (2).

The frequencies of HLA-A and B alleles observed in this study differed significantly in Caucasians, Africans, and Asians. There are many HLA-A and B alleles that appear almost exclusively in Africans and Asians, and are not observed in Caucasians. The fact that there are no specific Caucasoid alleles may be the result of some racial admixture with the flow of some Caucasoid alleles into African-Americans, and to a lesser extent to Asian Americans. In contrast, the HLA-A and B alleles in Native Americans and Hispanics showed a higher overlap with the alleles observed in Caucasians; however, some of the alleles differed significantly in frequency due to their original contributors. There is not a single unique allele at the HLA-A or B or C loci to be observed in Caucasians. The HLA allele composition observed in both Native Americans and Hispanics resembles an admixture between Caucasians and the North American and Mexican Indians respectively (3–4).

Allele frequency of heterogeneity and Hardy-Weinberg equilibrium

High frequencies of heterozygosity at all loci were observed; the highest rates were observed at the HLA-B locus with values ranging from 0.9094 to 0.9721 (0.9721 in AFAM, 0.9094 CAU, 0.9417 in HIS, 0.9359 in NA, and 0.9419 in ORI). The frequency of heterozygosity in HLA-C (0.9286 in AFAM, 0.8836 in CAU, 0.9146 in HIS, 0.9234 in NA, and 0.9077 in ORI) was higher than HLA-A (0.9177 in AFAM, 0.9024 in CAU, 0.8826 in HIS, 0.8599 in NA, and 0.8826 in ORI) in all populations with exception of Caucasians. The

Caucasian heterozygosity levels for HLA-A were higher than those for HLA-C. Since the number of HLA-C alleles identified in each population was lower than those of HLA-A, there is not a direct correlation between the number of alleles and the levels of heterozygosity. This may be the result of the evenness in the distributions of HLA alleles (and probably in response to) selective forces that may have acted on each locus (balancing selection versus directional selection and neutrality). In fact, a few HLA-A alleles with relatively high frequencies account for a high proportion of the overall gene count of the HLA-A locus (A*0201 with a.f.=0.2576, A*0101 with a.f.=0.1566, and A*0301 with a.f.=0.1229 in CAU; A*0201 with a.f.=0.2308 and A*2402 with a.f.=0.1255 in HIS; A*2402 with a.f.=0.1837 and A*0201 with a.f.=0.0925 in NA; and A*1101 with a.f.=0.2409, A*2402 with a.f.=0.1837, A*3303 with a.f.=0.1180 in ORI) in all populations except the African Americans. In the latter, the most common alleles are A*0201 (a.f.=0.1235) and A*2301 (a.f.=0.1098); these gene frequencies are significantly lower than those of the most common alleles in other populations.

Haplotype frequencies and haplotypic associations

The occurrence of haplotypes (HLA-A-C-B, HLA-A-B, HLA-A-C, and HLA-C-B) in each population and the global linkage disequilibrium between class I loci were evaluated. Significant linkage disequilibrium is observed between certain individual pairs of alleles at all loci in the five populations. Similar to the observations in other population studies, HLA-C-B haplotypes exhibited stronger associations than HLA-A-B or HLA-A-C haplotypes in all five populations. Some haplotypes were observed in all populations. Other haplotypes were found in only one population because ethnic specific alleles were observed in only one population. The most common HLA-A-C-B haplotypes are: A*2301-Cw*0202-B*1503, A*0101-Cw*0701-B*0801, and A*3001-Cw*1701-B*4201 [haplotype frequencies (h.f.) range from 0.03187 to 0.0259] in African Americans; A*0101-Cw*0701-B*0801, A*0201-Cw*0501-B*4402, and A*0301-Cw*0702-B*0702 (h.f. range from 0.0744 to 0.0451) in Caucasians; A*2902-Cw*1601-B*4403, A*0301-Cw*0702-B*0702, and A*2402-Cw*0702-B*3906 (h.f. range from 0.0292 down to 0.0167) in Hispanics; A*0201-Cw*0702-B*0702, A*0206-Cw*0202-B*2705, and A*2402-Cw*0304-B*4002 (h.f. range from 0.0365 to 0.0248) in Native Americans; and A*3303-Cw*0302-B*5801, A*0207-Cw*0102-B*4601, and A*1101-Cw*0801-B*1502 (h.f. range from 0.0623 to 0.0318) in Asians.

Genetic distance at HLA Class I loci

Genetic distances were estimated according to the distribution of alleles at HLA class I loci and compared with other populations (3). The largest genetic distances between populations were observed when comparing the distribution of HLA-B alleles. African Americans showed close genetic dis-

tance to Ugandans, Malians, and Kenyans. Hispanics and Native Americans appear to be genetically closer to each other than to other populations; their distances to Caucasians were smaller than those to Africans and Asians. Asians exhibited the largest distances from all the other four populations in the U.S.A.

References

1. Cao K, Chopek M, Fernández-Viña MA. High and intermediate resolution DNA typing systems for class I HLA-A, B, C genes by hybridization with sequence-specific oligonucleotide probes (SSOP). *Reviews in Immunogenetics* **1**, 177–208, 1999.
2. Hollenbach JA, Thomson G, Fernández-Viña MA, Cao K, Erlich HA, Bugawan T, Winkler C, Klitz W. HLA Diversity, Differentiation and Haplotype Evolution in Mesoamerican Natives. *Hum Immunol* **62**, 378–390, 2001.
3. Cao K, Hollenbach JA, Shi XJ, Shi WX, Chopek M, Fernández-Viña MA. Analysis of the frequencies of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* **62**, 1009–1030, 2001.
4. Leffell Mary S, Fallin M Daniele, Erlich Henry A., Fernández-Viña Marcelo, Hildebrand William H, Mack Steven J, Zachary Andrea A. HLA Antigens, Alleles and Haplotypes among the Yup'ik Alaska Natives: report of the ASHI Minority Workshops, Part II. *Hum Immunol* **63**, 614–625, 2002.

Population: Mestizos from the Mexican Highlands, Mexico

Authors: E. Infante, M. N. Vazquez, C. Alaez, C. Gorodezky
Authors' address: Department of Immunology and Immunogenetics, InDRE, SS Mexico City, Mexico

The Mexican Mestizos are an admixed population descended from Mexican Indians, Spanish and Mediterranean Europeans, Jews, Arabs and West African slaves. Their history began with the Spanish conquest of the region known today as Mexico in the early 16th century (1–5). The native Mexican population at that time is estimated to have numbered approximately 10 million, representing 33–50% of the population of the Americas. The great physical diversity of this ancestral native population is reflected today in the diversity of Mexico's 68 Indian groups and large, varied Mestizo populations.

Samples from 204 Mestizo individuals living in the Mexican highlands were collected for study. A study of part of this sample has been published previously (9–11). Approximately 50% of these subjects were born in or around Mexico City, with the remainder having moved to the area from other Mexican states. This population sample should be considered a rough approximation of the overall Mexican Mestizo population. Subsets of the 204 sampled individuals were typed at the class I loci (155 at HLA-A, 40 at HLA-B, and 29 at HLA-C), and all 204 individuals were typed at the class II loci.

High-resolution PCR-SSOP DNA typing was done using 12th IHW (6, 7) and IHWG primers, probes and protocols (Class I and Class II protocols developed by M. Fernandez Viña). DNA was extracted using a phenol chloroform protocol (8). Detection was performed with a chemiluminescence technique. The IHWG Biostatistics Core carried out statistical analyses.

No deviations from expected Hardy-Weinberg equilibrium proportions were detected. A significantly low value of homozygosity ($F_{nd} = -1.5606$, p value = 0.0066), consistent with the operation of balancing selection, was detected at the DQA1 locus. F_{nd} values for the remaining HLA loci did not deviate from neutral expectations.

Previous studies have demonstrated a strong Indian genetic contribution to highland Mestizos. For example, analyses of blood group systems have shown a prevalence of the O, Dia, k, D+, and Ms groups characteristic of Amerindian populations. A similar pattern is seen at the HLA loci, where many of the Mestizo class I (A*0201, *2402, *6801, *2601, *0206, and *3101, B* 3501, *3512, *3905, *1539, *3517, *4002, and *4005, and Cw*0401, *0303, *0802, *0304, *1203, *1204 and *1601) and class II alleles (DRB1*0407, *0802, *1406, and *1602, DQA1*0301 *0501, and *0401, and DQB1*0302, *0301, and *0402) are characteristically Amerindian. However, many alleles (A*3001 and *0101, B* 1801, *0702, and *0801, DRB1* 0701, *0301, *0102, and

*0404, DQA1*0101 and *0201, and DQB1 *0201 and *0501) are characteristic of European populations, and are likely to be of Spanish origin. Similarly, certain alleles are characteristic of Asian (A*2601, B *5201 and*4001, Cw*0101, *0102 and *0302, and DRB1 *0405) or African (A*3001, B*5301 and *5801, DRB1*1501, and DQB1*0603) populations.

In comparison to Indian groups, Mestizo allelic diversity is high at all loci, suggesting this is an outbred population, as expected of an urban group. Twenty-three HLA-A alleles, 37 B alleles, 15 C alleles 36 DRB1 alleles, 7 DQA1 alleles and 13 DQB1 alleles were observed with frequencies greater than 1%. The distribution of these allelic variants in Mexican Mestizos

is consistent with previous findings demonstrating the presence of European, Semitic, Mediterranean, Spanish, Asian, African and primarily Amerindian genes (9, 10, 11). Haplotype analyses are similar, indicating complex exchanges between the different populations, again with the Amerindian component predominant. Analyses of previously calculated genetic distances place the Mestizos between Mediterranean populations and a group including Mongolian, East Asian and Amerindian populations, suggesting that Mestizos are an outbred and admixed group. The results clearly show the high variability resulting from interethnic admixture, suggesting that this mechanism is the main factor responsible for the large diversity found in Mexico's modern urban populations.

References

1. TUC J. History of México- Affirmative Actino and Hernan Cortes (1485–1547). www.Mexconnect.com/mex.
2. Serano-Sánchez C. Mestizaje e Historia de la Población de México (con un esbozo antropológico de los lacandones de Chiapas). In: Martín Municio A, García Barreno P, Eds. Polimorfismo Génico (HLA) en poblaciones hispanoamericanas. Madrid: Real Academia de Ciencias Exactas, Físicas y Naturales, 171–194, 1996.
3. Mirambell L. El Poblamiento del Continente Americano. La Vida de sus Primeros Pobladores. In: Historia General de la Medicina en México, Eds. México Antiguo, Vol. 1. México D.F. UNAM-Academia Nacional de Medicina, 15–23, 1984.
4. Chavero A. El Virreinato. Historia de la Dominación española en México desde 1521 a 1808. In: Riva Palacio DV, Ed. México a Través de los Siglos. Tomo II. México d.F., Editorial Cumbre, 1967.
5. Aguirre Beltrán G. La población Negra de México. In: Secretaría de la Reforma Agraria Eds. Estudio Etno- Histórico. México. Centro de Estudios del Agrarismo en México, 1981.
6. Kennedy LJ, Poulton KV, Thomson W, Williams F, Middleton D, Howell WM, Tarassi K, Papasteriades C, Albert E, Fleischauer D, Chandanayinyong JM, Tiercy, Juji T, Tokunaga WER Ollier. HLA Class I DNA typing using sequence specific oligonucleotide probes (SSOP). 12th International Histocompatibility Workshop Experience. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol. I, pp. 216–225.
7. Bignon JD, Fernandez Vina MA, Cheneau ML, Fauchet R, Schreuder GM, Clayton J, Marsh SGE, Charron D. HLA DNA Class II Typing by PCR-SSOP : 12th International Histocompatibility Workshop Experience. In: Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication. EDK, Paris, 1997. Vol. I, pp. 21–25.
8. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd Edition. New York. Cold Spring Harbor Laboratory Press, 1989.
9. Gorodezky C. Genetic Difference between Europeans and Indians: Tissue and Blood types. *Allergy Proc* **13**, 243–253, 1992.
10. Petzl-Erler, ML, Gorodezky C, Layrisse Z, Klitz W, Fainboim L, Vullo C, et al. Anthropology Report of The Latinamerican Region: Amerindian and Admixed populations. In Charron D, Ed. Genetic Diversity of HLA. Functional and Medical Implication, Vol. I. Paris: EDK, 1997.
11. Gorodezky C, Alaez C, Vazquez MN, De la Rosa G, Infante E, Balladares S, Toribio R, Perez-Luque E, Muñoz L. The Genetic Structure of Mexican Mestizos of Different locations: Tracking back their origins through MHC genes, blood group systems and Microsatellites. *Human Immunol* **62**, 979–991, 2001.