

UNIVERSIDAD COMPLUTENSE DE MADRID

**FACULTAD DE MEDICINA
Departamento de Microbiología I**



TESIS DOCTORAL

**Estudio genético HLA en poblaciones iraníes: epidemiología,
antropología y farmacogenética**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Diego Federico Rey Medrano

Director

Antonio Arnaiz Villena

Madrid, 2016

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE MEDICINA

DPTO. MICROBIOLOGÍA I



**ESTUDIO GENÉTICO HLA EN
POBLACIONES IRANÍES: EPIDEMIOLOGÍA,
ANTROPOLOGÍA Y FARMACOGENÉTICA**

DIEGO FEDERICO REY MEDRANO

DIRECTOR: ANTONIO ARNAIZ VILLENA

MADRID, 2015

TESIS DOCTORAL

TÍTULO: ESTUDIO GENÉTICO HLA EN POBLACIONES IRANÍES: EPIDEMIOLOGÍA, ANTROPOLOGÍA Y FARMACOGENÉTICA

AUTOR: DIEGO FEDERICO REY MEDRANO

DIRECTOR: PROF. DR. ANTONIO ARNAIZ VILLENA

Catedrático de Inmunología y Coordinador I+D
Centro de Transfusión de la Comunidad de Madrid
Universidad Complutense de Madrid

LUGAR DE REALIZACIÓN:

- **Departamento de I+D del Centro de Transfusión de la Comunidad de Madrid,**
- **Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid**

Don Antonio Arnaiz Villena, Doctor en Medicina y Cirugía, Catedrático de Inmunología de la Universidad Complutense de Madrid y Coordinador de I+D del Centro de Transfusión de la Comunidad de Madrid,

CERTIFICA:

Que Don Diego Federico Rey Medrano, ha realizado bajo su dirección el trabajo de Tesis Doctoral que lleva por título:

“Estudio genético HLA en poblaciones iraníes: epidemiología, antropología y farmacogenética”.

Revisado el presente trabajo, considero que tiene la debida calidad para su defensa y calificación

Fdo.: Prof. Dr. Antonio Arnaiz Villena

A mis padres y a Eva

La más larga caminata comienza con un paso

Proverbio hindú

AGRADECIMIENTOS

En primer lugar deseo hacer constar mi gratitud al Dr. Antonio Arnaiz Villena, quien como director del presente trabajo de Tesis Doctoral ha sido mi mentor durante todo este apasionante periodo de investigación, por su inestimable ayuda no sólo en este proyecto sino en todos los realizados, por transmitirme la pasión por el trabajo científico y el esfuerzo que conlleva, por introducirme en el estimulante campo de la genética de poblaciones humanas y la inmunología, así como por brindarme la oportunidad de realizar este proyecto.

Al Dr. Ali Amirzargar, del Molecular Immunology Research Center de la Universidad de Teherán de Ciencias Médicas, por la obtención de las muestras en las distintas poblaciones del norte de Irán.

Al Dr. Carlos Hernando Parga Lozano, compañero de laboratorio, cuya contribución a mi formación técnica fue fundamental para la consecución de este trabajo.

Al Ministerio de Sanidad y a la Fundación Mutua Madrileña Automovilista, cuya financiación ha permitido llevar a cabo éste y otros trabajos.

A la Universidad Complutense de Madrid, concretamente al Departamento de Inmunología de su Facultad de Medicina.

A todos mis compañeros, presentes y pasados, del Departamento de Investigación y Desarrollo del Centro de Transfusión de la Comunidad de Madrid, lugar en el que he desarrollado mi doctorado, especialmente a Cristina Areces Viña y Mercedes Enríquez de Salamanca Carrascosa, por su apoyo y ayuda desinteresada, que han sido vitales durante la realización del presente trabajo.

Y por último al Centro de Transfusión de la Comunidad de Madrid y a su personal, por su amable acogida y por las facilidades provistas para desempeñar mis tareas.

DEDICATORIA

El contenido de este trabajo es el resultado de años de dedicación y esfuerzo, que me ha reportado un sinfín de buenos momentos, pero que también ha supuesto una serie de etapas difíciles que he podido sortear gracias al apoyo, la paciencia, la comprensión y las enseñanzas de grandes personas con las que afortunadamente he podido contar. Por ello me gustaría dedicárselo a los profesores que en las distintas etapas de mi vida han contribuido a formarme y a guiarme para descubrir mi vocación.

A toda mi familia, en especial a mi madre y a mis hermanos, por su apoyo incondicional de todo tipo, por haberme dado estabilidad y por hacerme saber que siempre puedo y podré contar con ellos.

A mis amigos porque, aun sin comprender muy bien a que me he dedicado durante estos años, sí han comprendido el esfuerzo que ha supuesto y me han regalado palabras y gestos de ánimo.

A Eva, por ser mi refugio cuando se avecina tormenta, por ser la mejor compañera que nunca pude imaginar, por su infinita generosidad y comprensión, por hacerme ver lo realmente importante, porque desde que está conmigo sé que soy muy afortunado y que nada es imposible.

A mi padre, porque siempre ha estado ahí incluso cuando no ha estado.

ÍNDICE

SUMMARY	1
INTRODUCCIÓN	9
I. EL COMPLEJO PRINCIPAL DE HISTOCOMPATIBILIDAD HUMANO	11
1. DEFINICIÓN Y GENERALIDADES	11
2. LOCALIZACIÓN Y ESTRUCTURA	13
2.1 Mapa genético del sistema HLA. Región MHC.	13
2.2 Grupos de genes MHC	16
3. REGIÓN MHC DE CLASE I.....	19
3.1 Genes HLA de clase I	19
3.1.1. Genes HLA de clase I clásicos o HLA-Ia.....	19
3.1.2. Genes HLA de clase I no clásicos o HLA-Ib	20
3.1.3. Pseudogenes HLA de clase I	20
3.1.4. Genes del grupo HLA de clase I no HLA	21
3.2 Estructura génica de los genes HLA de clase I.....	21
3.3 Estructura proteica de las moléculas de HLA de clase I.....	23
4. REGIÓN MHC DE CLASE II	27
4.1 Genes HLA de clase II clásicos	29
4.2 Estructura génica de los genes HLA de clase II.....	30
4.3 Estructura proteica de las moléculas HLA de clase II	31
5. REGIÓN MHC DE CLASE III	33
6. FUNCIÓN DE LAS MOLÉCULAS HLA	34
6.1 Presentación antigénica de las moléculas HLA de clase I.....	34
6.2 Presentación antigénica de las moléculas HLA de clase II.....	35
7. CARACTERÍSTICAS DEL SISTEMA HLA.....	37
7.1 Polimorfismo del Sistema HLA	37
7.1.1 Polimorfismo serológico del sistema HLA	37
7.1.2 Polimorfismo genético del sistema HLA.....	38
7.1.3 Evolución del polimorfismo del sistema HLA	39
7.1.4 Nomenclatura del sistema HLA	42
7.2 Desequilibrio de ligamento de los genes HLA	44
8. APLICACIONES DEL SISTEMA HLA	45
8.1 HLA y trasplantes	46
8.2 HLA y enfermedad.....	46
8.3 HLA y farmacogenética	49
8.4 HLA y genética de poblaciones	52

II. POBLAMIENTO Y POBLACIONES DE LA REGIÓN DEL MAR CASPIO	55
9. LA REGIÓN DEL MAR CASPIO.....	55
10. ANTROPOLOGÍA Y MOVIMIENTOS POBLACIONALES.....	56
10.1 Movimientos prehistóricos: Culturas de los Kurganes	57
10.2 Poblamiento y culturas de la región sur del Mar Caspio	59
10.2.1 Poblamiento del Golestán: turcomanos en Irán	61
10.2.2 Poblamiento del Azerbaiyán Iraní	63
10.2.3 Poblamiento del Kurdistán Iraní.....	65
10.2.4 Otros grupos étnicos de la región sur del Mar Caspio.....	67
11. ASPECTOS ANTROPOLÓGICOS Y LINGÜÍSTICOS DE LAS POBLACIONES DE LA REGIÓN DEL MAR CASPIO	69
11.1 Lenguas indoeuropeas: las lenguas kurdas	69
11.2 Lenguas altaicas: el turcomano y el azerí	70
12. MARCADORES GENÉTICOS EN EL ESTUDIO DEL POBLAMIENTO DE LA REGIÓN DEL MAR CASPIO	72
12.1 ADN mitocondrial.....	72
12.2 Cromosoma Y	74
12.3 Estudio de genomas completos: secuenciación de alto rendimiento	75
12.4 Aplicación al estudio del poblamiento de la región del Mar Caspio	76
13. ESTUDIOS DEL SISTEMA HLA EN POBLACIONES DE LA REGIÓN SUR DEL MAR CASPIO	82
 OBJETIVOS	 87
 MATERIALES Y MÉTODOS	 91
1. POBLACIÓN ESTUDIADA	93
1.1 Selección de individuos	93
1.2 Recolección de muestras	93
2. EXTRACCIÓN DE ADN GENÓMICO.....	93
3. AMPLIFICACIÓN Y GENOTIPADO DE LOS GENES HLA DE CLASE I Y DE CLASE II.....	97
3.1 Reacción en cadena de la polimerasa (PCR).....	98
3.2 Oligonucleótido-sondas específicas de secuencia (SSOP)	99
3.3 La citometría de flujo: tecnología Luminex.....	99
3.4 Metodología PCR-SSOP-Luminex	100
3.4.1 Amplificación PCR de los genes HLA-A, -B, -DRB1 y -DQB1	100
3.4.2 Control de amplificación: electroforesis en gel de agarosa.....	102

3.4.3 Hibridación con oligosondas	103
3.4.4 Reacción de la estreptavidina-picoeritrina	104
3.4.5 Medición por fluorimetría en Luminex	105
3.4.6 Interpretación de los resultados	105
4. ANÁLISIS ESTADÍSTICOS	106
4.1 Cálculo de frecuencias alélicas	106
4.2 Estimación de los haplotipos extendidos	106
4.3 Estimación del desequilibrio de ligamiento	107
4.4 Distancias genéticas y dendrogramas	108
4.5 Análisis de Correspondencias	109
4.6 Estimación de la proporción de mestizaje.....	110
 RESULTADOS	 113
1. ALELOS DEL SISTEMA HLA ENCONTRADOS EN LAS POBLACIONES DE ESTUDIO	115
1.1 Distribución de los alelos HLA de clase I en las poblaciones de estudio	115
1.2 Distribución de los alelos HLA de clase II en las poblaciones de estudio....	118
2. ESTUDIOS ANALÍTICOS BASADOS EN ALELOS HLA DE CLASE II.....	120
2.1 Análisis basados en el locus HLA-DRB1	121
2.2 Análisis basados en los loci HLA-DRB1 y -DQB1	129
3. ESTUDIOS DE HAPLOTIPOS BI-LOCUS EN LAS POBLACIONES DE ESTUDIO	136
3.1 Estudio del desequilibrio de ligamiento	136
3.2 Análisis basados en haplotipos de clase II con desequilibrio de ligamiento.	141
4. HAPLOTIPOS HLA EXTENDIDOS PRESENTES EN LAS POBLACIONES DE ESTUDIO	143
5. ESTIMACIÓN DE LA PROPORCIÓN DE MESTIZAJE EN LAS POBLACIONES DE ESTUDIO.....	149
 DISCUSIÓN	 151
1. GENÉTICA HLA DE LAS POBLACIONES DEL SUR DEL MAR CASPIO.....	153
2. RELACIONES GENÉTICAS HLA E HISTÓRICAS ENTRE LAS POBLACIONES DEL SUR DEL MAR CASPIO Y OTRAS POBLACIONES	155
2.1 Relación con poblaciones iraníes	155
2.2 Relación con poblaciones del Mediterráneo y del Cáucaso.....	156
2.2.1 Poblaciones del Mediterráneo Oriental: turcos y judíos.....	157
2.2.2 Poblaciones del Cáucaso: georgianos, esvanos y armenios	159

2.3 Relación con poblaciones de Asia Central.....	162
2.4 Análisis de haplotipos	165
2.5 Estimación del mestizaje.....	170
3. PROGRAMA DE TRASPLANTES Y ELABORACIÓN DE LISTAS DE ESPERA VIRTUALES EN POBLACIONES DEL SUR DEL MAR CASPIO	171
4. ENFERMEDADES LIGADAS A HLA Y FARMACOGENÉTICA EN LAS POBLACIONES DEL SUR DEL MAR CASPIO	173
CONCLUSIONES	177
BIBLIOGRAFÍA	181
ANEXOS	205
ANEXO 1: ABREVIATURAS MÁS FRECUENTES EMPLEADAS	207
ANEXO 2: PERFILES HLA DE LOS INDIVIDUOS DEL ESTUDIO	209
ANEXO 3: CURRICULUM VITAE	215
ANEXO 4: TRABAJOS PUBLICADOS.....	223

SUMMARY

Introduction

The HLA system is the most polymorphic genetic system described in humans. It consists of several closely linked loci encoding cell surface glycoproteins whose best known function is activating immune system response through antigenic presentation. New loci and new alleles have been described since the discovery of this genetic system and the presently available DNA typing and sequencing of these new alleles have increased the variety of HLA allelism. Due to the fact that HLA gene frequencies have a large degree of variability and a remarkable geographical correlation, HLA genes are an important and useful tool to infer genetic background and ethnical composition of modern human populations and also for tracing migration of ancient ones. In addition, certain combinations of contiguous alleles due to the strong linkage disequilibrium between HLA neighbouring loci show a characteristic frequency or are distinctive in many present day populations. Thus, HLA genetic system is a unique tool for studying the origin of relatively isolated groups, like Turkmen, Azeri and Kurd people, the populations under study, living in North Iran, in the surrounding area of Caspian Sea. Finally, HLA polymorphism is crucial for the compatibility between donor and receptor in organ transplantation and several HLA alleles have been linked to diseases and to response to drug treatments, which accomplishes relationships of certain variants with different pathologies treatment including AIDS. This is important in personalized treatments design.

Turkmen could be descendants of Oghuz tribes from Seljuq branch coming from Transoxiana region (Central Asia) contemporarily to the foundation of the Seljuk Empire in 10th century AD. Conversely, this people could belong to another group within the Oghuz, arriving to Iran five centuries later. Migrations of this people were initially developed peacefully, being vassals of the Safavid Empire, and later by violent raids. They speak a language belonging to the Turkish-Oghuz group. In Iran, Turkmen live in Golestan province, mainly in Türkmensähra (“Turkmen plain”) area and amount 1.5 million people (2% of Iranian population). Most of this people are Sunni Muslims.

Azeri people, living nowadays in Republic of Azerbaijan and Iran, are believed to be of mixed origin, whose more ancient element come from transcaucasian indigenous populations, like Mannai, who lived in northwest of modern Iran during the 1st millennium BC and underwent an assimilation cultural process and also gene input by people of Andronovo culture coming from Central Asia (Indo-Iranians). Historic, linguistic, archaeological and cultural data support the presence of an Iranian substratum

in this ethnic group, mainly because of the existence of Old Azeri language (Adari) belonging to Iranian language family that persists in some Azeri dialects. The arrival of Oghuz Turkic tribes mentioned previously since 10th century AD that conquered Southwest Asia till Anatolia Peninsula led to an uninterrupted “turkification” process in next centuries that caused the substitution of Iranian Old Azeri language for the modern Azeri Turkish, belonging to Altaic language family. In Iran, Azeris live mainly in Iranian Azerbaijan and amount between 15 and 16 million people, being the biggest Azeri community in the World. Most of Azeri people are Shia Muslims.

Kurdistan (“land of Kurds”) is a geographical and cultural region located South of Caucasus and North of the former Mesopotamia, in which Kurd people is the main ethnic group mostly living in highlands. The origin of Kurds is believed to stem from the Neolithic Euro Asiatic aboriginal tribes settled in North of Mesopotamia, like Lullubi or Gutians, in 5th millennium BC. This population probably also underwent the influence of Andronovo culture. These migrations could influence in the multiethnic origin of the Kurds forefathers and caused the substitution of proto-non-Iranian languages spoken in the region for the Iranian languages. In Iran, Kurds live mainly in Iranian Kurdistan (Northeast part of the Country) and amount between 4 and 6.5 million people. Kurds have historically been living in a nomadic fashion. Nowadays in Iran, they form both urban and rural communities, but there are also a few groups that still practice nomadism. Most of this people are Sunni Muslims.

Aims

We have undertaken the present study in order to: 1) Determine the HLA class I (A and B) and class II (DRB1 and DQB1) polymorphism in Turkmen, Azeri and Kurd people from Iran, 2) Calculate genetic distances from HLA class II allele frequencies, between these populations and worldwide populations, with greater emphasis in closer populations (from Mediterranean and Central Asia), 3) Construct Neighbour-Joining dendrograms from genetic distances so as to determine the degree of relatedness between the studied populations, 4) Carry out correspondence analyses with all studied populations for the purpose of determining the existence of a correlation between geography and HLA genetics, 5) Estimate frequencies and linkage disequilibria of both HLA bi-locus and HLA extended haplotypes characteristic of populations of study in order to compare with haplotypes defined in other worldwide population and assign the geographic provenance if possible, 6) Discuss the HLA genetic relationship between the

populations of study through the comparison of allele and haplotype frequencies, and 7) Infer the possible migratory waves and groups movements that originated the peopling of Southwest Asia. HLA genes study in these populations enables to obtain a virtual transplant waiting list and HLA profiles to pharmacogenomics and therapeutics.

Results

Thirty different HLA-A and fifty-five different HLA-B alleles were found in total. The more frequent HLA class I alleles found are: A*02:01, A*03:01, A*11:01, A*24:02, B*35:01, B*51:01, B*13:01 and B*08:01 (only in Turkmen), B*52:01 (only in Azeris) and, B*18:01 and B*50:01 (only in Kurds). Thirty-three different HLA-DRB1 and nineteen different HLA-DQB1 alleles were found. In this case, the more frequent HLA class II alleles found are: DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*07:01, DRB1*11:01, DRB1*14:01, DRB1*15:01, DQB1*02:01, DQB1*03:01 and DQB1*05:01.

Three types of analyses were done in order to compare the studied populations HLA frequencies with other World population frequencies: 1) by using DRB1, 2) DRB1 and DQB1 allele frequencies data, and 3) DRB1-DQB1 haplotype frequencies. It was not possible to carry out a study comparing HLA class I allele frequencies or HLA class I and II conjointly due to the lack of class I studies in many worldwide. Genetic distances based on these data between Turkmen, Azeri and Kurd people and populations included in the analysis have been calculated. It is shown that the three sampled populations present the closest values followed by Caucasoid European Mediterraneans, such as French, Russians, Croatians, Macedonians, Italians and Turks, populations from Caucasus, like Georgians and Svan, and western Siberians, such as Tuvinians, Chuvash, Todja and Tofalar. The NJ relatedness dendrogram based on both DRB1, and DRB1 and DQB1 allele frequencies separates populations in two well-differentiated clusters. One of them groups North and South Mediterranean, Middle Eastern, Negroid, Caucasian and Siberian populations. The second cluster grouped Orientals and Pacific Islanders. The studied populations are integrated in the first cluster, in intermediate position between Iranian populations and western Siberians (Chuvash, Mansi, Buryat) and central Siberians (Tuvinians, Todja, Tofalar). Correspondence analyses based on previous data show similar results. Also, two groups are clearly defined according to first dimension that explains most of the variability among populations. The first one is divided according to the second dimension into two subgroups. The first one includes

Siberians and the second one Orientals and Pacific Islanders. The second main group is also divided into two subgroups. First subgroup clusters Europeans, northern Mediterranean, Caucasus, Iranian populations (near to Kurds and Chuvash), while second subgroup clusters southern and eastern Mediterraneans. Turkmen and Azeris are located in an intermediate situation between these two major groups in this analysis but closer to the first one.

HLA class II haplotype frequencies were used to develop NJ dendrogram and correspondence analysis too. The NJ relatedness analysis also groups the three Caspian populations with Mediterranean, Middle Eastern, and Caucasian peoples, and separated from Siberians (second cluster), except Chuvash and Aleuts. Correspondence analysis gives identical results, since our sample is located in the cluster that groups Europeans, Mediterraneans and Caucasians, defined according to first dimension, close to the other Iranian populations, although Azeris tends to be in an intermediate situation between this main group and Siberian cluster.

The most probable two loci haplotypes (A-B, B-DRB1 and DRB1-DQB1) were calculated. These two-loci linkage disequilibrium data show that the most frequent combinations are characteristic of Siberian or Central Asian populations and they have found in several Siberian populations (A*24:02-B*35:01, DRB1*01:01-DQB1*05:01, DRB1*11:01-DQB1*03:01, DRB1*13:01-DQB1*06:03, DRB1*15:01-DQB1*06:01, DRB1*15:01-DQB1*06:02). However, several two-loci combinations found in this population in a considerable high frequency are characteristic of Middle East and Mediterranean populations (A*02:01-B*50:01, A*30:01-B*13:01, A*33:01-B*14:01, DRB1*03:01-DQB1*02:01, DRB1*16:01-DQB1*05:01). Most of HLA extended haplotypes obtained has been described as new. But eight of these more frequent haplotypes were found in populations from Mediterranean, Caucasus and Central Asia.

Finally, admixture proportions of studied populations have been also estimated. It is observed a Mediterranean substratum both in Turkmen, Kurds and Azeris. However, Centro Asiatic component is higher in Turkmen and Azeris, and practically absent in Kurds, whereas Caucasian component is predominant in this population.

Conclusions

- 1) Comparative population studies through dendrograms and correspondence analyses strongly suggest that the three studied populations have a common genetic

substratum which related them with western Mediterranean and Caucasian population, especially Kurds and Azeris.

- 2) These studies, together with admixture estimations, suggest that ancestral populations genetically close to current Altaic peoples from Central Asia could contribute to the Turkmen and slightly to the Azeri HLA genetic background.
- 3) HLA bi-locus and HLA extended haplotype analysis also place these populations in the HLA context of western Mediterranean and Caucasian populations, due to the presence of these frequent haplotypes: A*01:01-B*35:01-DRB1*03:01-DQB1*02:01, A*02:01-B*50:01-DRB1*07:01-DQB1*02:01, A*01:02-B*08:01-DRB1*03:01-DQB1*02:01 and A*24:02-B*58:01-DRB1*03:01-DQB1*02:01.
- 4) The presence of HLA extended haplotypes such as A*24:02-B*35:01-DRB1*15:01-DQB1*06:02, A*03:01-B*44:02-DRB1*04:01-DQB1*03:02 and A*03:01-B*07:02-DRB1*15:01-DQB1*06:02 in Turkmen and Azeris could demonstrate the Central Asia populations genetic contribution through migrations happened in different times, supported by other genetic and archaeological data.
- 5) It is observed a strong correlation between geography and HLA Genetics but not between these aspects and linguistics. The three studied populations, geographically close, show a close HLA relatedness, especially Kurds and Azeris. However they speak languages belonging to different families.
- 6) The hypothesis of “elite dominance” is supported by the fact that the studied populations are genetically close but they show different linguistic and cultural traits. According to this, a cultural imposition was carried out by invader peoples of Southwest Asia. This originates an “iranization” process in Kurds and a “turkification” process in Azeris, which could carry a genetic contribution in the last case.
- 7) HLA profiles obtained have a direct application in organ transplantation since enable the enforcement of a virtual transplant waiting list for the purpose of future therapeutics.
- 8) HLA profiles obtained also present therapeutic applications in pathologies linked to HLA with high incidence in these populations, such as AIDS with B*57 and B*58 alleles (supertypic HLA-B17 serotype) or Behçet’s disease, arthritis rheumatoid and malaria caused by *Plasmodium vivax*, and in pharmacogenomics, due to the presence of certain frequent alleles, like A*24:02 and DRB1*15:01, associated to adverse drug reactions.

INTRODUCCIÓN

I. EL COMPLEJO PRINCIPAL DE HISTOCOMPATIBILIDAD HUMANO

1. DEFINICIÓN Y GENERALIDADES

El Complejo Principal de Histocompatibilidad (CPH, en inglés MHC, *Major Histocompatibility Complex*) es un grupo de genes que codifica para una serie de glicoproteínas de superficie. La función principal de estas glicoproteínas es la de unir fragmentos peptídicos de origen intracelular o extracelular y presentarlo a las células efectoras del sistema inmune.

Este sistema, se encuentra en un gran número de especies, considerándose que está presente en todos los vertebrados (Klein, 1986; Humphreys & Reinherz, 1994) y se denomina Sistema HLA (del inglés, *Human Leukocyte Antigen*) en humanos, teniendo un papel fundamental en la presentación antigénica que inicia la respuesta inmunológica ante la presencia de elementos ajenos al organismo (Dausset, 1958; Payne & Rolfs, 1958; van Rood et al., 1959).

Los genes de histocompatibilidad fueron identificados por primera vez en la década de los 40, gracias al estudio del cáncer mediante el trasplante de tejidos tumorales en ratones. Ya en los inicios del siglo XX se había constatado que los tumores transplantados en ratones pertenecientes a cepas diferentes eran rechazados por el receptor (Jensen, 1903; Loeb, 1908) ocurriendo lo mismo cuando se trasplantaba tejido sano (Little & Johnson, 1922). En la década de los 30, J. B. Haldane sugirió que la destrucción de los tejidos trasplantados por parte del receptor podía tratarse de un proceso similar al de la lisis eritrocitaria y al de la hemaglutinación, cuando se pone en contacto eritrocitos de un individuo con suero de otro individuo previamente inmunizado con esos eritrocitos (caso de la lisis) o sin inmunizar (caso de la hemaglutinación). Años más tarde se identificaron varios antígenos eritrocitarios en ratones receptores de trasplante, relacionándose uno de ellos (el antígeno II) con la supervivencia del tumor trasplantado, ya que en las cepas de ratón donde este antígeno no estaba presente se correlacionaba con el rechazo al trasplante, lo que permitió elaborar una teoría inmunológica del trasplante (Gorer, 1936). Ya en los años 40, G. Snell identificó la región génica responsable de este rechazo, siendo denominados genes de histocompatibilidad, y en el caso concreto del ratón, genes H-2, en referencia al antígeno II (Snell, 1948). En lo que respecta a los genes de histocompatibilidad en humanos, fueron descubiertos por J. Dausset cuando detectó la presencia de anticuerpos capaces de provocar la aglutinación de los leucocitos del donante pero no de los del

receptor, al reconocer los antígenos en la superficie de estas células, en personas que habían recibido múltiples transfusiones sanguíneas (Dausset, 1958). Esta observación inicial, llevó a denominar a este conjunto de moléculas con las siglas HLA. En la década de los 60 se comprobó que estos antígenos leucocitarios representaban el sistema principal de histocompatibilidad (Dausset et al., 1965; Ceppellini et al., 1967). Sin embargo, actualmente se conoce que el papel biológico fundamental de las moléculas HLA consiste en la presentación antigénica mediante su interacción con los linfocitos T (Stern et al., 1994; Madden, 1995).

Este papel fundamental de las moléculas HLA involucradas en la respuesta inmune frente a antígenos fue propuesto en 1970 cuando se demostró que los linfocitos T no eran capaces de reconocer antígenos libres en la circulación, sino que necesitaban de estas moléculas para poder reconocerlos y procesarlos. La asociación necesaria entre moléculas HLA y linfocitos T puede ser de dos tipos respondiendo a dos vías distintas de procesamiento antigénico, en función del origen del antígeno presente, habiendo dos clases de moléculas HLA:

- Los antígenos de origen intracelular (péptidos producidos por degradación citosólica, procedentes fundamentalmente de virus, bacterias intracelulares y células tumorales) son presentados por las moléculas HLA de clase I y estimulan a los linfocitos T citotóxicos que expresan el marcador CD8+.

- Los antígenos de origen extracelular (bacterianos o parasitarios, introducidos en endosomas o lisosomas), son presentados por las moléculas HLA de clase II y estimulan a los linfocitos T que expresan el marcador CD4+ (Dausset et al., 1965; Bjorkman et al., 1987b; Klein, 1990; Stern et al., 1994; Madden, 1995).

Por otra parte, el estudio de la función de las moléculas HLA ha permitido conocer que su papel biológico está definido por un alto polimorfismo, que permite el reconocimiento de un gran número de elementos extraños y, a su vez, dirigido por los patógenos presentes en el medio ambiente. Además, requiere de un proceso de aprendizaje durante la embriogénesis, que ayuda al sistema inmune a diferenciar lo “propio” de lo “no propio” e inclusive elementos “propios” alterados como es el caso de los antígenos tumorales y/o los fenómenos de autoinmunidad.

Las características principales del sistema HLA son las siguientes:

- Es poligénico: está constituido por varios genes organizados en tres regiones.

- Es dialélico: cada gen que forma parte de este sistema presenta dos variantes en el individuo y su expresión es codominante. La mayor parte de los individuos de una población son heterocigotos para cada gen de este sistema.
- Es muy polimórfico: existen múltiples variantes alélicas para cada locus. Los distintos alelos difieren entre sí en la habilidad para unir y presentar con mayor eficacia diferentes antígenos proteicos a los linfocitos T.
- Presenta desequilibrio de ligamiento: diferentes alelos de loci situados de manera adyacente en el mismo cromosoma aparecen asociados con una frecuencia mayor a la esperada por una combinación aleatoria.

2. LOCALIZACIÓN Y ESTRUCTURA

El sistema HLA se localiza físicamente en el brazo corto del cromosoma 6, en la parte distal de la banda 6p21.1-21.3. Este sistema se considera una de las regiones con mayor densidad génica dentro del genoma humano, ya que presenta un elevado número de loci distribuidos en una extensión de casi 4 centimorgan, aproximadamente 4.000.000 de pares de bases (pb) (Berger et al., 1979).

2.1 Mapa genético del sistema HLA. Región MHC.

El primer mapa genético de esta región genómica se completó en 1991 (Trowsdale et al., 1991). Años más tarde se elaboró otro mapa más completo a partir de secuencias parciales en individuos diferentes, según el cual la región HLA ocupa 3.6 Mb y está compuesta por 224 genes de los cuales 128 podrían expresarse y aproximadamente el 40% tendrían función inmunológica, que se reparten en tres regiones bien definidas: región HLA de clase I, región HLA de clase II y región HLA de clase III (The MHC sequencing consortium, 1999). Este mapa fue sustituido por una secuencia de referencia de un haplotipo homocigoto obtenida de la secuenciación de una línea celular consanguínea denominada PGF (Stewart et al., 2004). La secuenciación completa del cromosoma 6 (Mungall et al., 2003) reveló la existencia de otros genes relacionados con MHC más allá de las 3.6 Mb descritas en 1999, por lo que se elaboró un nuevo mapa genético, denominado mapa MHC extendido (xMHC) (Horton et al., 2004) para diferenciarlo del mapa constituido anteriormente, denominado mapa MHC clásico (The MHC sequencing consortium, 1999). El “MHC Haplotype Project” ha confirmado la existencia de más de 44.000 variaciones, tanto sustituciones como indels (inserciones y deleciones) y la presencia de más de 300 loci en este sistema, de los cuales al menos

160 codifican para proteínas. Dentro de este grupo, 122 genes codificarían sustituciones, de las cuales 97 son no sinónimas (Horton et al., 2008).

El mapa clásico está conformado por las tres regiones mencionadas anteriormente, dispuestas en el siguiente orden de telómero a centrómero: HLA de clase I, HLA de clase III y HLA de clase II. El mapa extendido completa el anterior al añadir dos nuevas regiones: una en el extremo telomérico (región de clase I extendida) y otra en el extremo centromérico (región de clase II extendida), ampliándose así las dimensiones de la región HLA hasta las 7.6 Mb (Figura I1). Según el mapa extendido, la región HLA se sitúa entre las posiciones cromosómicas 6p22.2-6p21.3 y contiene 421 genes, de los cuales el 60% expresan proteínas funcionales, el 33% son pseudogenes (expresan proteínas aberrantes o bien no se expresan) y el 7% restante se transcriben dando lugar a moléculas de ARN transferente (ARNt), que participan en el proceso de traducción del ARN mensajero (ARNm) en el ribosoma (Figura I1 y Tabla I1). La mayor concentración de genes se observa en la región de clase III, que contiene 58 genes funcionales (23%) en 0.7 Mb, lo que la convierte además en la región de mayor densidad génica de todo el genoma humano (Horton et al., 2004), mientras que la región de clase I extendida recoge el mayor número de genes funcionales (109, 43% del total de estos genes) del sistema.

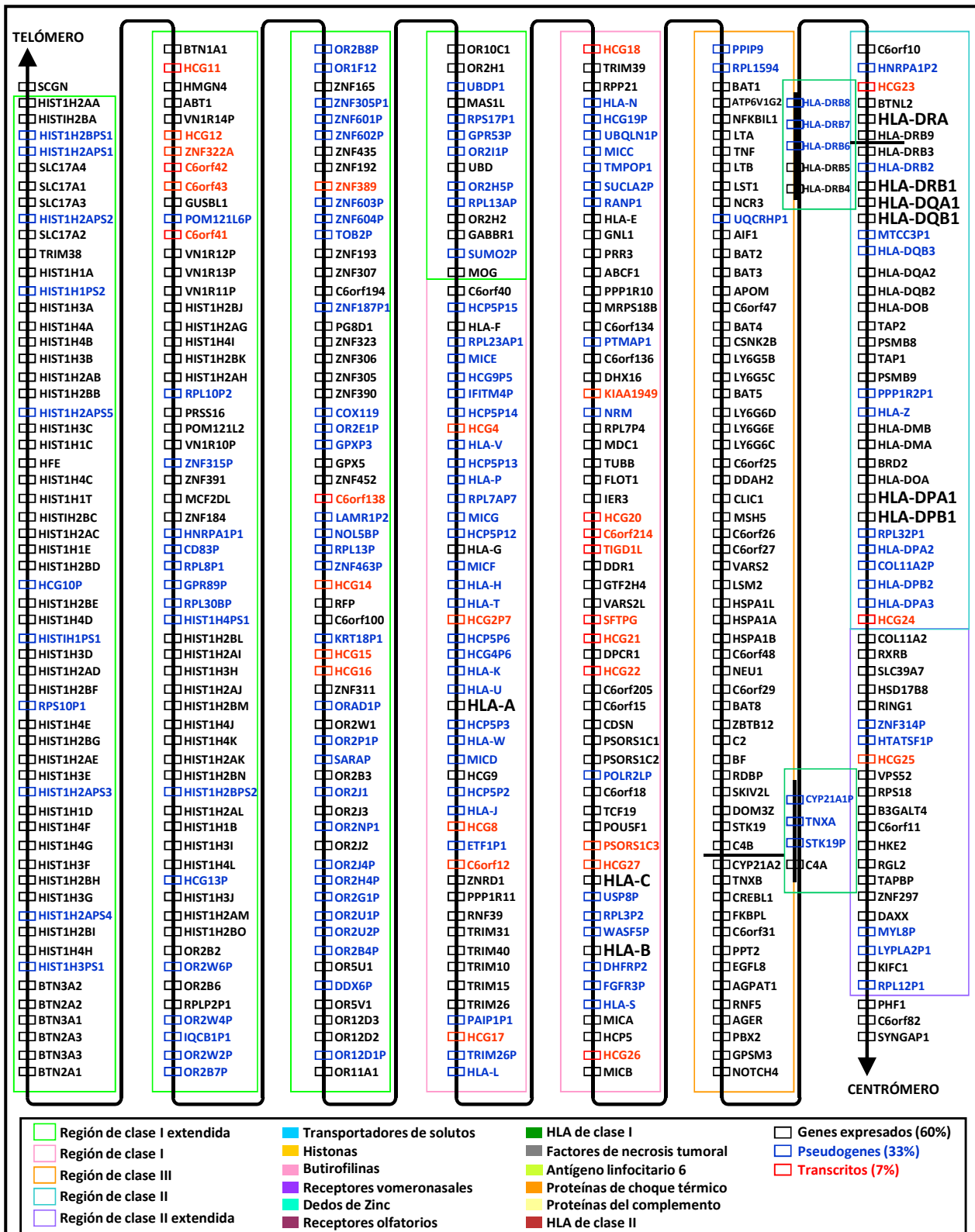


Figura II. Mapa genético de la región MHC extendida (7.6 Mb). El código de colores sirve para identificar las diferentes subregiones así como los distintos grupos de genes MHC, y si se trata de genes expresados, pseudogenes o transcritos (véase leyenda). Con letra de mayor tamaño se identifican los genes HLA de clase I y clase II clásicos (basado en el mapa genético original de Horton et al., 2004).

TABLA II. Principales grupos de genes del sistema MHC extendido en humanos. Se indica la distribución de los genes en las distintas regiones (A: clase I extendida, B: clase I, C: clase III, D: clase II, E: clase II extendida), así como el número de loci codificantes (dan lugar a proteínas funcionales) y no codificantes (pseudogenes y genes ARNt) de cada grupo (Horton et al., 2004).

Grupos de genes MHC	Región MHC	Nº loci codificados	Nº loci no codificados
Histonas	A	55	11
ARN transferente (ARNt)	A	151	6
Transportadores de solutos (SLC)	A, E	4	0
Butirofilinas (BTN)	A, D	8	0
Receptores vomeronasales (VNR)	A	0	5
Receptores olfatorios	A	14	20
Dedos de Zinc (Zinc-finger)	A, B	26	10
HLA-I	B	9	17
Factores de necrosis tumoral (TNF)	C	3	0
Antígeno linfocitario 6	C	5	0
Proteínas de choque térmico (HSP)	C	3	0
HLA-II	D	15	9

2.2 Grupos de genes MHC

Los genes de las histonas codifican proteínas involucradas en la formación del núcleo proteico del nucleosoma, alrededor del cual se pliega el ADN para conformar la cromatina. La mayor concentración de estos genes en el genoma humano se encuentra en la región HLA de clase I extendida, que contiene 55 genes que expresan histonas de las 5 clases existentes (6 H1, 12 H2A, 15 H2B, 10 H3 y 12 H4) y 11 pseudogenes distintos.

Los genes de ARNt son muy cortos (75-90 nucleótidos) y codifican moléculas de ARN implicadas en la síntesis proteica, ya que reconocen específicamente los tripletes del ARNm y aportan el aminoácido correspondiente. Son 157 genes (6 de ellos no funcionales) que incluyen todo los tipos existentes salvo Asn- y Cys-ARNt. Están distribuidos a lo largo de la región de clase I extendida y representan el mayor agrupamiento de este tipo de genes de todo el genoma humano.

Los genes transportadores de solutos (SLC, del inglés *Solute Carrier*), que aparecen tanto en la región de clase I extendida como en la región de clase II extendida, dan lugar a proteínas transportadoras de nutrientes cuyo papel fisiológico se desconoce.

Los genes de las butirofilinas (BTN), presentes mayoritariamente en la región de clase I extendida, son miembros de la superfamilia de las inmunoglobulinas. Son 8 genes cuya función es desconocida, aunque se cree que están involucrados en el metabolismo de los lípidos (BTN2A2) y en la secreción de estos en la leche (BTN1A1). Son muy similares a los genes MOG de los mamíferos (situado entre las regiones de clase I y clase III en humanos) y a los genes B-G del sistema principal de histocompatibilidad del pollo (complejo B).

Los genes de los receptores vomeronasales (VNR) se encuentran también en la región de clase I extendida y pertenecen a la familia de genes que codifican receptores de feromonas. Los únicos representantes de esta familia en humanos son pseudogenes.

Los genes de los receptores olfatorios son 34 (al menos 14 funcionales) y se encuentran en la región de clase I extendida. Presentan polimorfismo y se les atribuye una función relacionada con la supervivencia en los procesos del comportamiento.

Los genes *dedos de zinc* (del inglés *zinc-finger*) codifican proteínas muy diversas que se agrupan por la presencia de dominios capaces de unir iones de zinc, más que por la similitud de las secuencias completas. Por esto, los productos de estos genes presentan funciones muy diversas, pudiendo actuar como enzimas, factores de transcripción y proteínas de replicación. En la región de clase I existen 36 de estos genes, divididos en 3 grupos: los que tienen un dominio SCAN (13 genes), los que tienen un dominio triple denominado TRIM (8 genes) y el resto (15 genes). Diez de ellos no se expresan.

Los genes HLA-I se caracterizan por expresar proteínas estructuralmente muy similares, aunque diferentes desde el punto de vista funcional. Según esto, se clasifican en genes de clase I clásicos (HLA-A, HLA-B y HLA-C, también denominados HLA-Ia), genes de clase I no clásicos (HLA-E, HLA-F, HLA-G y 12 pseudogenes también conocidos como HLA-Ib), genes tipo HLA-I (MIC-A, MIC-B y 5 pseudogenes) y gen de la hemocromatosis (HFE). Excepto éste último, que se encuentra en el extremo telomérico de la región de clase I extendida, muy alejado del resto, todos están localizados en la región de clase I.

Los genes de los factores de necrosis tumoral (TNF) codifican 3 tipos de citoquinas (TNF, LTA y LTB) pertenecientes a la superfamilia TNF y que participan en procesos inflamatorios. Se encuentran situados en la región de clase III.

Los genes del antígeno 6 linfocitario (LY6), que codifican proteínas de membrana asociadas a glicosil-fosfatidil-inositol y pueden tener una función inmunológica. Están localizados en la región de clase III.

Los genes de las proteínas de choque térmico (HSP, del inglés *Heat Shock Proteins*) aumentan su expresión en situaciones de estrés celular, como un choque térmico, colaborando en los procesos de síntesis, plegamiento, ensamblaje, transporte y degradación de proteínas. Concretamente este grupo que se encuentra en la región de clase III está involucrado en la señalización inducida por estrés del sistema inmune para la eliminación de células dañadas, infectadas o tumorales.

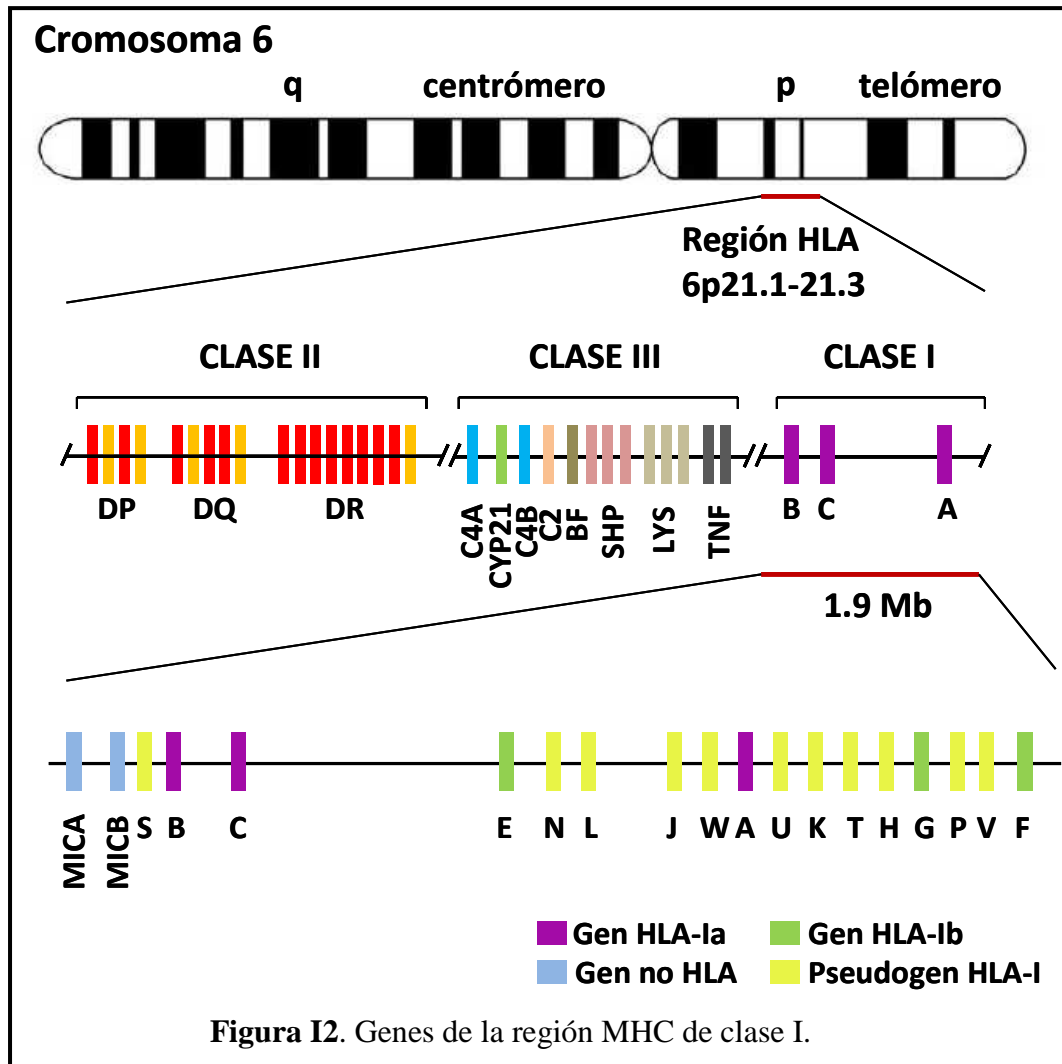
Algunos genes de complemento de la vía clásica (C2, C4A, C4B) y de la vía alternativa (Bf) también están presentes en la región de clase III.

Los genes HLA-II están ubicados en la región de clase II y dan lugar proteínas polimórficas implicadas en la presentación antigénica que se denominan HLA-DM, HLA-DO (conocidos también como genes de clase II no clásicos), HLA-DP, HLA-DQ y HLA-DR (conocidos como genes de clase II clásicos). Se han descrito numerosos genes funcionales y también pseudogenes (Figura I1). Los genes clásicos se expresan en células presentadoras de antígenos (APC, del inglés *Antigen-Presenting Cells*), fundamentalmente células dendríticas (DC, del inglés *Dendritic Cells*), macrófagos y linfocitos B. Su función es unir péptidos de origen extracelular (generalmente procedentes de bacterias) y presentarlos a linfocitos T cooperadores (Th, CD4+) para poner en marcha una respuesta inmunológica de tipo humoral, mediada por anticuerpos. Los genes no clásicos no se expresan en la superficie celular, pero forman complejos que están involucrados en el intercambio y carga de péptidos en las moléculas de clase II clásicas.

Una característica fundamental de la región HLA es que los genes que contiene están muy próximos entre sí, generando desequilibrios de ligamiento muy acentuados (la recombinación entre ellos es altamente improbable) que provocan la existencia de haplotipos (combinaciones de alelos concretos situados en loci contiguos) que se transmiten como unidades genéticas a lo largo de muchas generaciones.

3. REGIÓN MHC DE CLASE I

Esta región, junto con la región de clase I extendida, se encuentra en la parte más telomérica del MHC en el cromosoma 6, y comprende aproximadamente 1.9 Mb, donde se localizan los genes HLA de clase I (Figura I2) (Horton et al., 2004).



3.1 Genes HLA de clase I

Estos genes codifican un grupo de proteínas estructuralmente muy similares pero con diferencias significativas a varios niveles, entre ellas la función, expresión tisular o el grado de polimorfismo

3.1.1. Genes HLA de clase I clásicos o HLA-Ia

A este grupo pertenecen los genes HLA-A, -B y -C. Son los primeros descritos dentro del sistema HLA. Codifican para glicoproteínas de membrana que se expresan en todas las células nucleadas del organismo, si bien su nivel de expresión varía desde un

máximo en células pertenecientes al sistema inmune (linfocitos T, B y macrófagos) hasta un mínimo en células musculares, del sistema nervioso y fibroblastos (Wu et al., 2013). Los genes HLA-Ia son muy polimórficos y expresan proteínas que asociadas con β 2-microglobulina (β 2M) están implicadas en el reconocimiento de péptidos antigénicos de origen intracelular (por degradación de proteínas citosólicas derivadas de virus y tumores), que son presentados a los linfocitos T citotóxicos (Tc, CD8+) para poner en marcha una respuesta inmunológica de tipo celular (citotóxica) mediada por linfocitos Tc y células NK (del inglés, *Natural Killer*).

3.1.2. Genes HLA de clase I no clásicos o HLA-Ib

En este grupo se incluyen los genes HLA-E, -F y -G. Codifican para proteínas estructuralmente similares a las de los genes clásicos. Se diferencian básicamente de los anteriores por:

- Su limitada expresión tisular (HLA-F y -G): todas se expresan en las células del trofoblasto. HLA-G se expresa también en las células epiteliales del timo en adultos, en la córnea y en la matriz de las uñas; HLA-F limita su expresión a las células B y a los tejidos linfoides (Gomez-Prieto et al., 2010).
- Su bajo polimorfismo: el locus más polimórfico de este grupo de genes presenta 50 alelos distintos descritos hasta la fecha, un número mucho menor si lo comparamos con los 2553 alelos descritos para el locus menos polimórfico de los genes clásicos Ia, HLA-C (Robinson et al., 2015).
- Su función: aunque poco conocida, se postula que tienen una función tolerogénica al tratarse de moléculas inmunosupresoras a varios niveles de la respuesta inmune, y que atendiendo a su expresión tisular, tendrían un papel relevante en evitar el rechazo del feto durante el embarazo (Gomez-Prieto et al., 2010).

3.1.3. Pseudogenes HLA de clase I

Además en esta región se han descubierto toda una serie de secuencias de ADN que mantienen cierta homología estructural con los genes HLA de clase I no clásicos (grupo de genes HLA-I). Estas secuencias reciben el nombre de pseudogenes, ya que presentan deleciones que conllevan desplazamientos del marco de lectura ocasionando codones de terminación prematuros y por ello no codifican para proteínas funcionales. También se considera a algunos como genes truncados o fragmentos génicos que podrían haber sido procesados.

Se han descrito doce que se encuentran localizados distribuidos a lo largo de toda la región de clase I, éstos se denominan según la última nomenclatura y dispuestos en orden desde telómero a centrómero HLA-V, -P, -H, -T, -K, -U, -W, -J, -L, -N, -S, y -X (Horton et al., 2004; Shiina et al., 2009; Gray et al., 2013).

3.1.4. Genes del grupo HLA de clase I no HLA

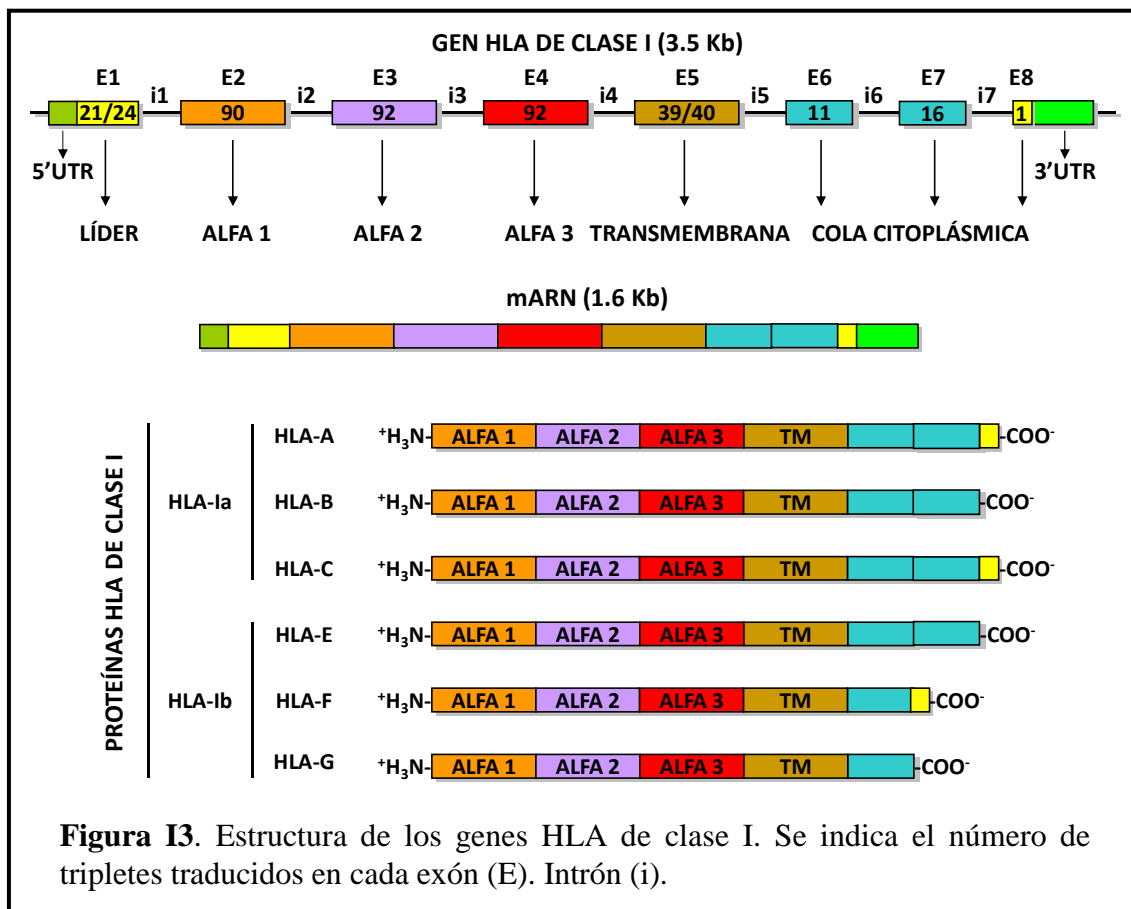
Otros genes pertenecientes a este grupo pero que no comparten ninguna función relacionada con la presentación antigénica son los genes relacionados con la cadena de MHC de clase I (MIC-A y MIC-B) y el gen HFE. Este último se encuentra fuera de la región de clase I, y aunque su producto se asemeja en muchos aspectos a las proteínas de HLA de clase I clásicas, incluido en su asociación con la cadena $\beta 2M$, está implicado en el transporte del hierro.

Por su parte, los productos de los genes MIC-A y MIC-B no se asocian a $\beta 2M$ y se expresan en varios tipos celulares incluyendo células endoteliales y monocitos (Zwirner et al., 1999). Los genes MIC expresan proteínas que son reconocidas por los receptores activadores NKG2D de las células NK y que pueden tener una función en el sistema inmune de la mucosa intestinal. La posibilidad de que MIC-A pueda ser una diana durante las respuestas inmunes anti-injerto en pacientes trasplantados de riñón fue demostrado por el trabajo de Stastny *et al.* (Zwirner et al., 2000). También ha sido demostrada la importancia de estos anticuerpos en trasplantes de corazón (Suarez-Alvarez et al., 2007). Se han descrito cinco pseudogenes que mantienen cierta homología estructural con los genes MIC: MIC-E, -G, -F, -D, y -C (Horton et al., 2004).

3.2 Estructura génica de los genes HLA de clase I

Los genes HLA de clase I están constituidos por 7 intrones y 8 exones (Malissen et al., 1982; Jordan et al., 1985; Figura I3). Estos últimos codifican para diferentes regiones (dominios) de la molécula, de tal manera que: el exón 1 codifica para la región 5'UT (del inglés *UnTranslated*) y para un péptido líder de 24 aa (aminoácidos, 21 para HLA-E y -F) que contiene un núcleo hidrófobo característico de las proteínas que migran desde el retículo endoplasmático (RE). Los exones 2 y 3 son los más polimórficos y codifican respectivamente para los dominios $\alpha 1$ y $\alpha 2$ de la molécula a los que se une el péptido antigénico, mientras que el exón 4 es menos polimórfico y codifica para el dominio $\alpha 3$. En cuanto a los genes HLA Ia, el polimorfismo de los exones 2 y 3 es el responsable de la especificidad de unión al repertorio de péptidos

antigénicos para la clase I. Así mismo, el exón 5 codifica para el dominio transmembrana hidrofóbico.



Por último, el dominio citoplasmático en los genes HLA Ia está codificado por los exones 6 y 7 más parte del 8 (para las moléculas HLA-A y HLA-C). El exón 8 codifica además para la región 3'UT. La diferencia en el tamaño del dominio citoplasmático se debe a la situación que ocupa el codon de terminación, estando en el exón 8 para HLA-A y HLA-C y en el exón 7 para HLA-B.

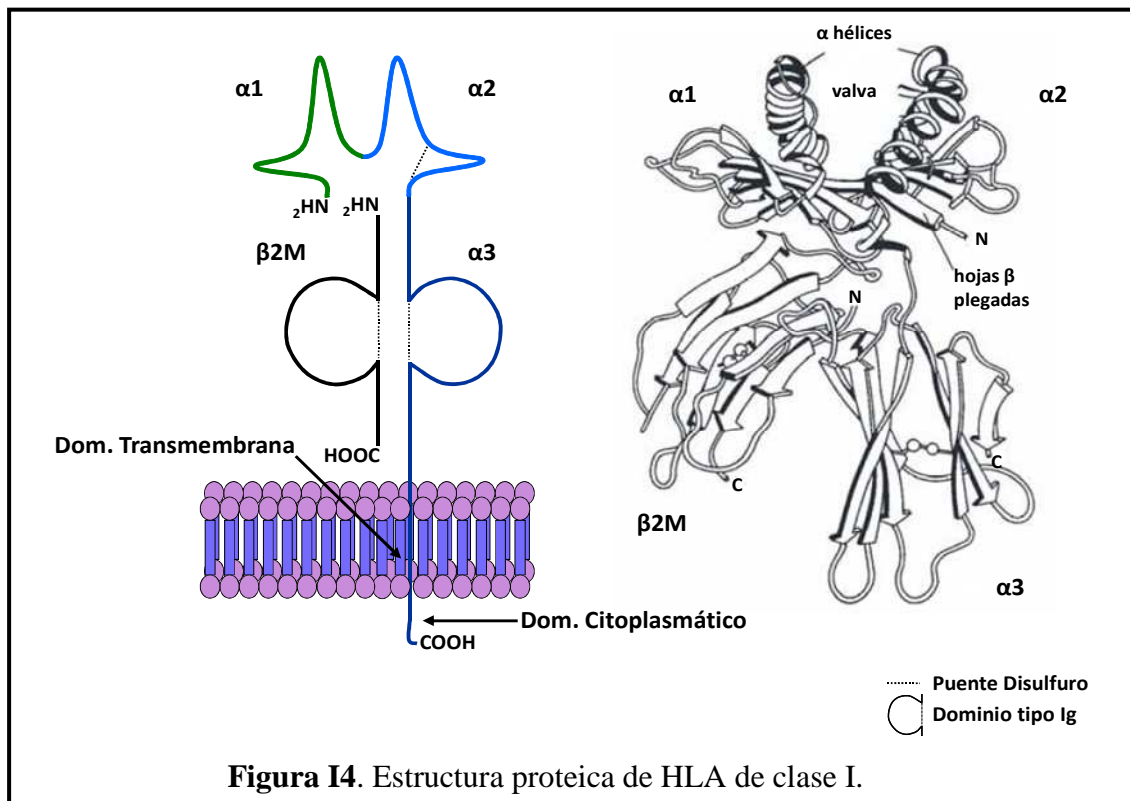
En cuanto a HLA-E, el dominio citoplasmático está codificado por lo exones 6 y 7, aunque este presenta una delección de cinco nucleótidos que origina un codon de terminación por lo que este dominio es más corto en comparación con el de HLA-B (un aminoácido menos). En HLA-F, este dominio esta codificado por el exón 6 y parte del 8, ya que elimina el exón 7 por "splicing" (maduración del ARNm) debido a una mutación puntual en la señal de splicing en sentido 3' del intrón 6. HLA-G presenta un dominio citoplasmático corto formado por 6 aa, cinco codificados por el exón 5 y uno codificado por el exón 6, ya que el segundo triplete del exón 6 es un codon stop.

3.3 Estructura proteica de las moléculas de HLA de clase I

Moléculas HLA de clase I clásicas

Las moléculas HLA de clase I clásicas (HLA-A, HLA-B y HLA-C) son glicoproteínas heterodiméricas de superficie pertenecientes a la superfamilia de las inmunoglobulinas, compuestas por dos cadenas polipeptídicas: la cadena pesada (α) de 44 kilodalton (Kd) codificada en el sistema HLA y la cadena ligera (β) de 12 Kd β 2M (Grey et al., 1973; Peterson et al., 1974), cuyo gen se encuentra fuera del sistema HLA, en el cromosoma 15 (Ploegh et al., 1981). La cadena pesada, es el único miembro del heterodímero que atraviesa la membrana celular. El extremo aminoterminal de esta cadena está orientado hacia el exterior de la célula. Ambas cadenas se unen no covalentemente en su porción extracelular (Figura I4).

La cadena α está formada por 338 aa para el locus HLA-B y 341 aa para los loci HLA-A y HLA-C. La porción extracelular de esta cadena, se divide en tres dominios globulares bien diferenciados de unos 90 aa cada uno, que pueden escindir-se de la superficie celular bajo la acción proteolítica de la enzima papaína. Estos dominios son: α 1 (aa 1-90, porción N-terminal de la cadena α), α 2 (aa 91-182) y α 3 (aa 183-247), codificados cada uno por los exones 2, 3 y 4 respectivamente (Malissen et al., 1982). La porción transmembrana, forma una estructura en α -hélice, de unos 25 aa, que continúa con un pequeño tallo citoplásmico de unos 30 aa rico en tirosinas y serinas.



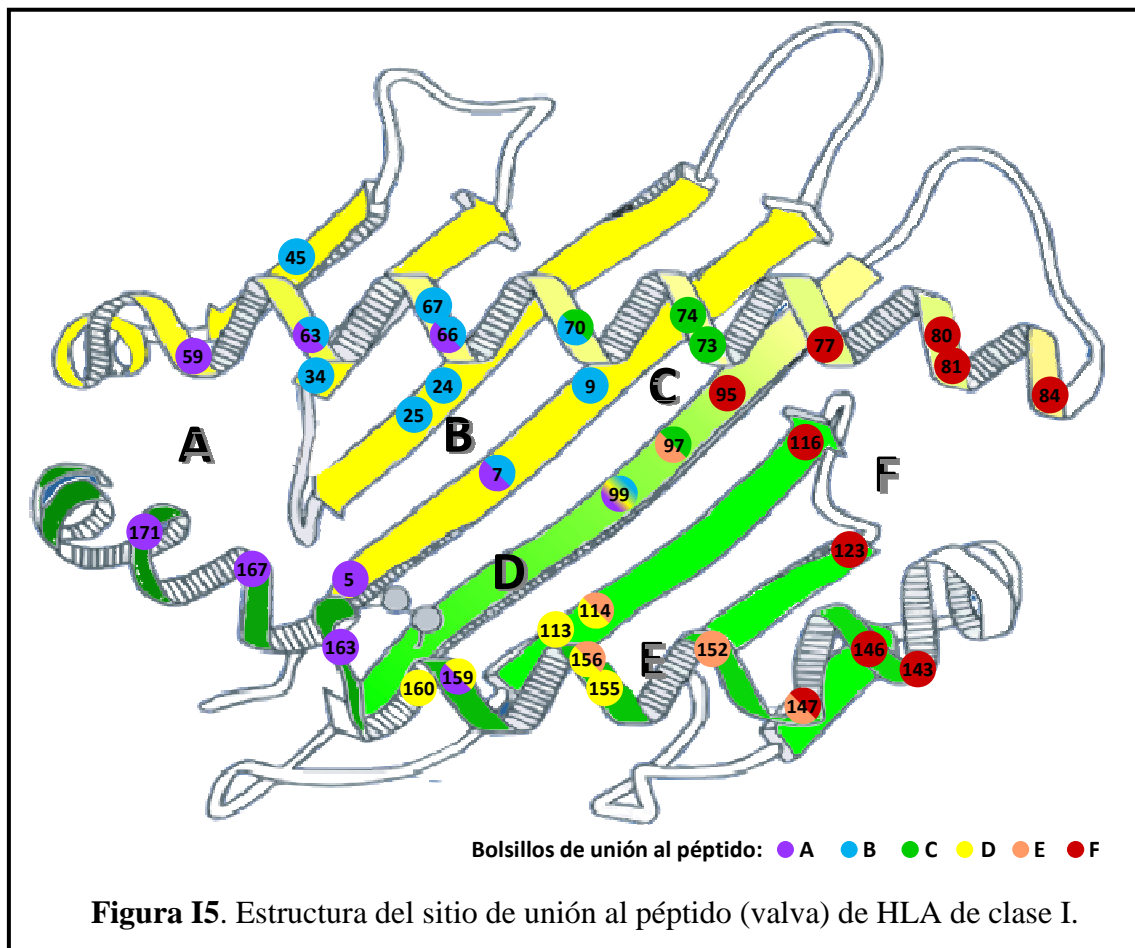
En 1987, Bjorkman y sus colaboradores determinaron mediante cristalografía por rayos X, la primera estructura tridimensional de una proteína HLA, concretamente al antígeno HLA-A2 (Bjorkman et al., 1987a; Saper et al., 1991). Posteriormente, se ha conseguido cristalizar otros antígenos como HLA-A3, HLA-A68, HLA-B27, HLA-B35, HLA-B53, HLA-B62 y HLA-Cw4 (Garrett et al., 1989; Madden et al., 1991; Fan & Wiley, 1999; Roder et al., 2006; McMahon et al., 2011). Todos los cristales obtenidos presentaban una región distal formada por los dominios $\alpha 1$ y $\alpha 2$ de la cadena pesada y una región proximal a la que pertenecen los dominios $\alpha 3$ y $\beta 2m$, que presentan alta homología de secuencia y de estructura entre ellas y con las regiones constantes de las inmunoglobulinas. Los dominios distales ($\alpha 1$ y $\alpha 2$) presentan una estructura compuesta por dos láminas β formadas cada una por cuatro hebras antiparalelas y una región de α -hélice localizada en el extremo carboxi-terminal de las láminas β (Figura I4). Las láminas β de ambos dominios, forman la base de una valva flanqueadas por las α -hélices orientadas hacia el exterior. Esta valva, formada por los dominios $\alpha 1$ y $\alpha 2$ de las moléculas HLA de clase I, constituye la estructura para la unión de los péptidos procesados y para la interacción con los receptores de las células T (TCR, del inglés *T Cell Receptors*) (Bjorkman et al., 1987a; Bjorkman et al., 1987b). Este espacio es lo suficientemente grande para que quepan péptidos de 8 a 11 aa (Rammensee et al., 1993). Además, el dominio $\alpha 2$ contiene una hélice corta adicional en posición carboxi-terminal para unirse con el dominio $\alpha 3$.

La región proximal a la membrana está formada por el dominio $\alpha 3$ y la $\beta 2M$ y cada uno de ellos, al igual que los dominios constantes de las inmunoglobulinas, están estructurados en dos láminas β antiparalelas, una formada por cuatro hebras y la otra por tres. Las hebras de cada lámina se conectan entre sí por puentes de disulfuro internos formando una estructura similar a un “sándwich”. La $\beta 2M$ se sitúa por la parte de debajo de la valva formada por los dominios $\alpha 1$ y $\alpha 2$, de tal manera que existe un mínimo contacto entre estos dos dominios y el dominio $\alpha 3$.

El gran polimorfismo que muestran estas moléculas, se encuentra concentrado en la valva (dominios $\alpha 1$ y $\alpha 2$) y la variabilidad de los residuos sirve para interactuar con los distintos péptidos y/o con el TCR. De los 20 residuos aminoacídicos de alta variabilidad identificados en las moléculas de HLA (Parham et al., 1988), 16 constituyen residuos de contacto con el péptido (posiciones 9, 24, 45, 62, 66, 67, 70, 74, 77, 80, 95, 97, 114, 116, 156 y 163) y tres de posible contacto con el TCR (residuos 65, 69 y 76) (Parham et al., 1988; Li & Bouvier, 2004; Figura I5). Los residuos

polimórficos relacionados con la unión al antígeno, se encuentran principalmente en el dominio $\alpha 1$ mientras que en el dominio $\alpha 2$ se encuentran mayoritariamente los residuos polimórficos relacionados con la unión al TCR (Parham et al., 1988).

El dominio $\alpha 3$ interacciona con el correceptor de la célula T, CD8. Esta interacción requiere la presencia de aminoácidos ácidos, que se encuentran en las posiciones 223, 227 y 229, ocupando una zona localizada en un lazo del dominio $\alpha 3$ en una cavidad justo debajo de la valva (Connolly et al., 1990; Salter et al., 1990). Por otra parte, se ha demostrado que el dominio $\alpha 3$ no es imprescindible para el mantenimiento de la conformación nativa de la valva, al menos en lo que se refiere a su unión con el péptido a presentar (Collins et al., 1995).



La variabilidad que presentan las moléculas HLA de clase I se traduce en cambios topológicos en los sitios de unión al péptido. Las características químicas y la exclusiva configuración de la estructura de cada valva (o *Peptide-Binding Groove*), explican cómo pueden unir gran variedad de péptidos (Garrett et al., 1989; Madden et al., 1992). Saper *et al.* identificaron una serie de depresiones a lo largo de la valva, que son los llamados

“*Peptide-binding Pockets*” o bolsillos de unión a péptidos (Saper et al., 1991). Los péptidos unidos a la valva, pueden acomodar una o más de sus cadenas laterales aminoacídicas en estos bolsillos. Se han identificado seis bolsillos denominados con letras (de la A a la F), los cuales se encuentran localizados en las uniones de las láminas β con las dos α -hélices (bolsillos B, C, D y E) o en los extremos de las dos α -hélices (bolsillos A y F). Los dos últimos bolsillos, contienen aminoácidos más conservados que los otros cuatro (Barber & Parham, 1993).

Moléculas HLA de clase I no clásicas

Estas moléculas presentan la misma composición polipeptídica que las moléculas HLA Ia (cadena α dividida en tres dominios y cadena β 2M), aunque con una variabilidad en la región distal (valva) muchísimo menor.

HLA-E

La estructura cristalográfica de la molécula de HLA-E ha sido determinada formando complejo con un ligando (O'Callaghan et al., 1998). Presenta dos posiciones variables, 107 y 157, ambas en el dominio α 2 (Gomez-Prieto et al., 2010). Su valva interacciona con péptidos derivados de secuencias líder de otros HLA de clase I con un repertorio restringido (Lee et al., 1998). Sorprendentemente, a pesar de que estos péptidos líder están en el RE, lugar donde ocurre el ensamblaje de las moléculas de HLA y la unión del péptido antigénico, se ha demostrado que la expresión en superficie de HLA-E es dependiente de la presencia en la célula de proteínas TAP y tapasina (Braud et al., 1998). Tanto TAP como tapasina son unas proteínas implicadas en el procesamiento antigénico, se encargan de introducir los péptidos a presentar al interior del RE (véase apartado 6.2). Esto implicaría que es necesario que los péptidos líderes, sustrato para HLA-E, deben abandonar el RE y sufrir una fragmentación en el citoplasma celular que origine péptidos de tamaño adecuado, y posteriormente deben ser introducidos de nuevo. Esta es la vía normal del procesamiento antigénico para todas las moléculas HLA de clase I.

HLA-F

La molécula HLA-F es más corta que las moléculas HLA Ia debido a la ausencia del exón 7 en el ARNm, lo que resulta en una proteína con un dominio citoplasmático corto. Presenta cuatro posiciones variables: -9, 50, 118 y 251, localizados en la región 5'UT, dominio α 1, α 2 y α 3 respectivamente (Gomez-Prieto et al., 2010). No se han

encontrado péptidos de unión específicos para esta molécula, y solamente se ha obtenido la estructura terciaria para la cadena pesada y la $\beta 2M$, sin ligando.

HLA-G

La molécula HLA-G presenta 13 posiciones variables repartidas en los tres dominios α , lo que contrasta con las moléculas HLA Ia, que concentran su mayor variabilidad en la región distal (dominios $\alpha 1$ y $\alpha 2$). Esta molécula une un repertorio limitado de péptidos procedentes de un número restringido de proteínas, principalmente derivados de citoquinas. HLA-G presenta 7 isoformas generadas mediante splicing alternativo del transcrito primario, cuatro de ellas unidas a membrana y tres solubles (Donadi et al., 2011). HLA-G1 es la isoforma completa con una estructura similar a la de una molécula de HLA I clásica, asociada a $\beta 2M$. Mediante un ARNm más corto (900 pb) se originaría la isoforma HLA-G2, con la ausencia del dominio $\alpha 2$ al carecer del exón 3 (Ishitani & Geraghty, 1992). Un ARNm aún más corto (600 pb) daría lugar a la isoforma HLA-G3 que se caracteriza por la ausencia de los dominios $\alpha 2$ y $\alpha 3$ (exones 3 y 4). Una cuarta forma de splicing alternativo daría lugar a la isoforma HLA-G4 que carece de dominio $\alpha 3$ (exón 4) (Kirszenbaum et al., 1994). Las isoformas solubles HLA-G5 y HLA-G6 presentan los mismos dominios extraglobulares que HLA-G1 y HLA-G2 respectivamente, en las que el marco de lectura se extiende hasta parte del intrón 4 donde un codon stop eliminaría la porción transmembrana, originando un extremo carboxi-terminal de 21 (Fujii et al., 1994). Por último la isoforma HLA-G7 presenta únicamente el dominio $\alpha 1$ unido a dos aminoácidos codificados por el intrón 2 (Paul et al., 2000).

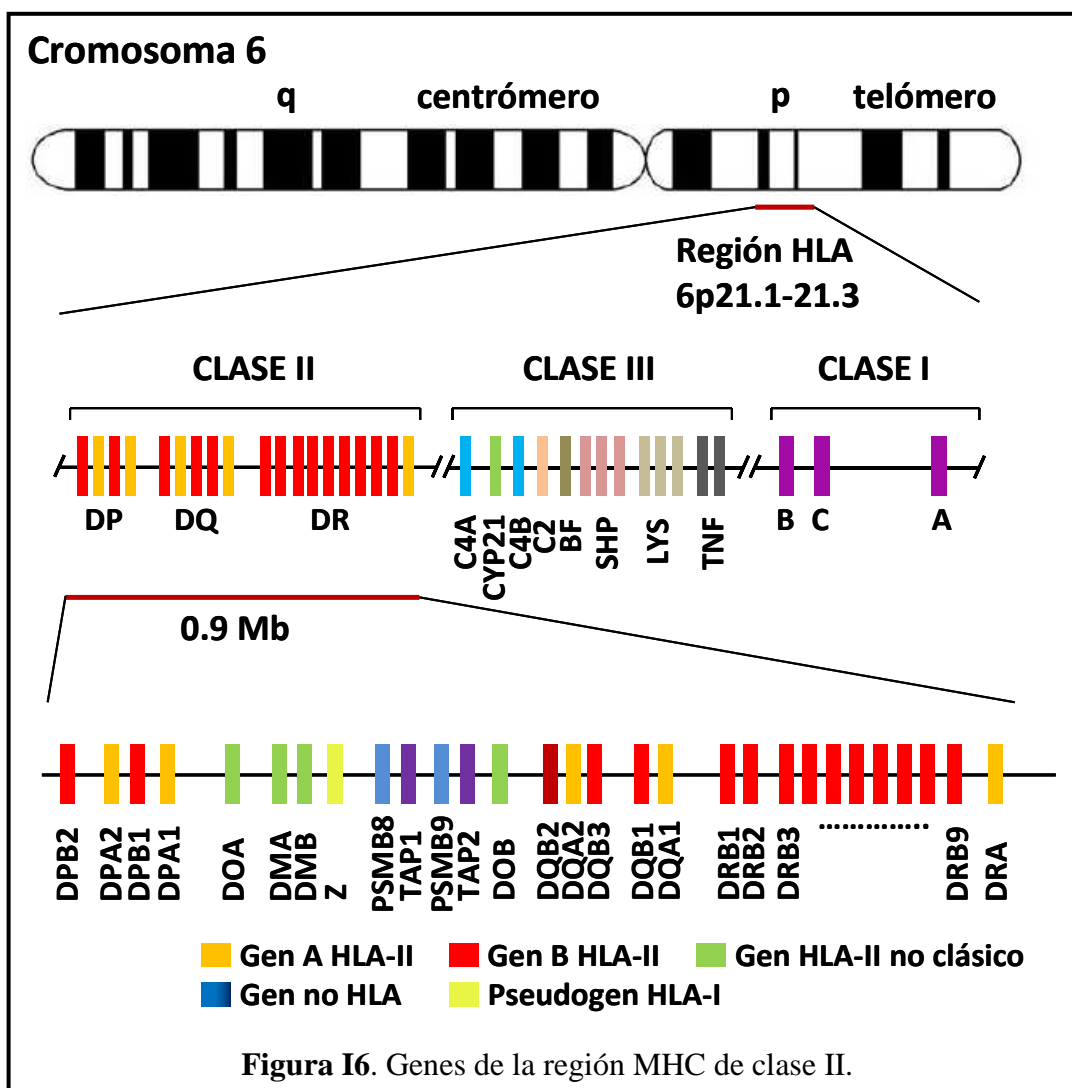
Recientemente se ha determinado la estructura cristalográfica del dímero de HLA-G (Shiroishi et al., 2006). La presencia de un residuo de cisteína (Cys) en la posición 42 del dominio $\alpha 1$ es una característica única de la molécula HLA-G que le permite formar dímeros mediante puentes de disulfuro Cys42-Cys42, de tal forma expone dos sitios de unión para los receptores con los que interactúa: CD8 y LILR (del inglés *Leukocyte Ig-like Receptor*).

4. REGIÓN MHC DE CLASE II

Esta región, junto con la de clase II extendida, se encuentra en la parte más centromérica de MHC en el cromosoma 6, y comprende aproximadamente 0.9 Mb (Horton et al., 2004), donde se localizan los genes HLA de clase II clásicos, que

nombradas por la posición que ocupan en el cromosoma desde el centrómero son: HLA-DP, -DQ y -DR (Figura I6).

Estos genes codifican para proteínas heterodiméricas (α/β) que están implicadas en fenómenos de restricción del reconocimiento antigénico mediado por linfocitos T cooperadores CD4+. La distribución tisular de estas moléculas queda restringida fundamentalmente a células del sistema inmune presentadoras de antígenos: macrófagos, células DC, Linfocitos B, y cuando están activados también linfocitos T y células NK. Igualmente se expresan en alta densidad en otras células que actúan como presentadoras, tales como células endoteliales (Wu et al., 2013).



Las cadenas α están codificadas en genes denominados 'A' (HLA-DPA, HLA-DQA, etc.) y las cadenas β están codificadas en genes denominados 'B' (HLA-DPB, HLA-DQB, etc.). Se han descrito en algunos casos varios genes 'A' y también varios genes 'B' codificantes para la misma proteína de clase II.

Dentro esta región, se encuentran también genes que codifican para proteínas estructuralmente similares a los genes de las subregiones mencionadas previamente, los genes HLA de clase II no clásicos: HLA-DM (HLA-DMA y -DMB) y HLA-DO (HLA-DOA -previamente denominado DNA- y -DOB) (Trowsdale & Kelly, 1985; Tonnellet et al., 1985). Al igual que los genes HLA II clásicos presentan genes A y B que codifican para las cadenas α y β respectivamente. Los genes HLA-DM codifican para una proteína heterodimérica implicada en la unión de los péptidos del endosoma a las moléculas HLA de clase II clásicas (véase apartado 4.3; Sloan et al., 1995), mientras que los genes HLA-DO codifican para heterodímeros que posiblemente actúen como moduladores negativos de la presentación antigénica por parte de las moléculas HLA de clase II clásicas, mediante la interacción con HLA-DM (Denzin et al., 1997; Figura I6).

Además de los anteriores, se localizan en esta región genes implicados en la presentación antigénica de las moléculas HLA de clase I: los genes TAP1 y TAP2 que codifican para proteínas transportadoras de péptidos citosólicos degradados (véase apartado 5.1; Trowsdale et al., 1990; Spies et al., 1990) y los genes anteriormente conocidos como LMP, PSMB8 y PSMB9, que codifican para subunidades del inmunoproteosoma que interviene en el procesamiento y degradación de los péptidos antigénicos (véase apartado 5.1; Glynn et al., 1991; Schmidt et al., 1999). También se localiza en esta región un pseudogen homólogo a los genes HLA de clase I, HLA-Z (Beck et al., 1996; Figura I6).

4.1 Genes HLA de clase II clásicos

Existen 5 genes HLA-DP, tres genes A (HLA-DPA1, HLA-DPA2 y HLA-DPA3) y dos genes B (HLA-DPB1 y HLA-DPB2). Los genes HLA-DPA1 y -DPB1 codifican para cadenas α y β respectivamente, conformando un heterodímero de clase II funcional. Sin embargo, los genes HLA-DPA2, -DPA3 y -DPB2, a pesar de tener una secuencia genética intacta, no se expresan, considerándose pseudogenes.

En relación a los genes HLA-DQ, se han identificado dos genes A (HLA-DQA1 y HLA-DQA2) y tres genes B (HLA-DQB1, HLA-DQB2 y HLA-DQB3). Todos se expresan menos HLA-DQB3, que es considerado un pseudogen. El gen HLA-DQA1 presenta la estructura típica de los genes A de clase II mientras que el gen DQB1 tiene un dominio citoplasmático más corto al eliminarse el exón 5 durante la maduración de su ARNm (Radley et al., 1994).

Los genes HLA-DR se caracterizan por presentar un único gen A (HLA-DRA) y 9 genes B (de DRB1 a DRB9), que pueden ser funcionales y altamente polimórficos (DRB1), funcionales y con menor polimorfismo (DRB3, DRB4 y DRB5) o pseudogenes (DRB2, DRB6, DRB7, DRB8 y DRB9; Figura I6). En relación a los distintos genes B se observa un fenómeno denominado polimorfismo de dosis génica. Significa que no todos los individuos poseen todas las copias descritas de genes codificantes de cadenas beta de las moléculas HLA-DR ni de pseudogenes. Si bien todos los individuos poseen el gen HLA-DRA (que es de copia única) y el gen HLA-DRB1, la presencia o no de los otros depende del alelo expresado por el gen HLA-DRB1, constituyendo así distintos grupos de haplotipos: unos con un único gen DRB funcional (DRB1), los grupos DR1 y DR8, y otros con dos genes DRB funcionales (DRB1 más otro), grupos DR51, DR52, y DR53, ambos asociados o no a los distintos pseudogenes (Klein et al., 1991; Tabla I2).

TABLA I2. Grupos de haplotipos y haplotipos HLA-DR. Entre paréntesis se indica el tipo de cadena β codificada por el gen funcional que forma parte del haplotipo. Todos los haplotipos presentan el gen HLA-DRB1. Los que presentan otros genes funcionales distintos a este pueden formar moléculas de clase II con la cadena α del gen HLA-DRA. De esta manera, los individuos que poseen el haplotipo HLA-DR2 tienen y expresan el gen HLA-DRB5, por lo que además de proteínas HLA-DR2 ($\alpha + \beta_1$) tienen proteínas HLA-DR51 ($\alpha + \beta_5$). También son portadores de los pseudogenes HLA-DRB6 y HLA-DRB9.

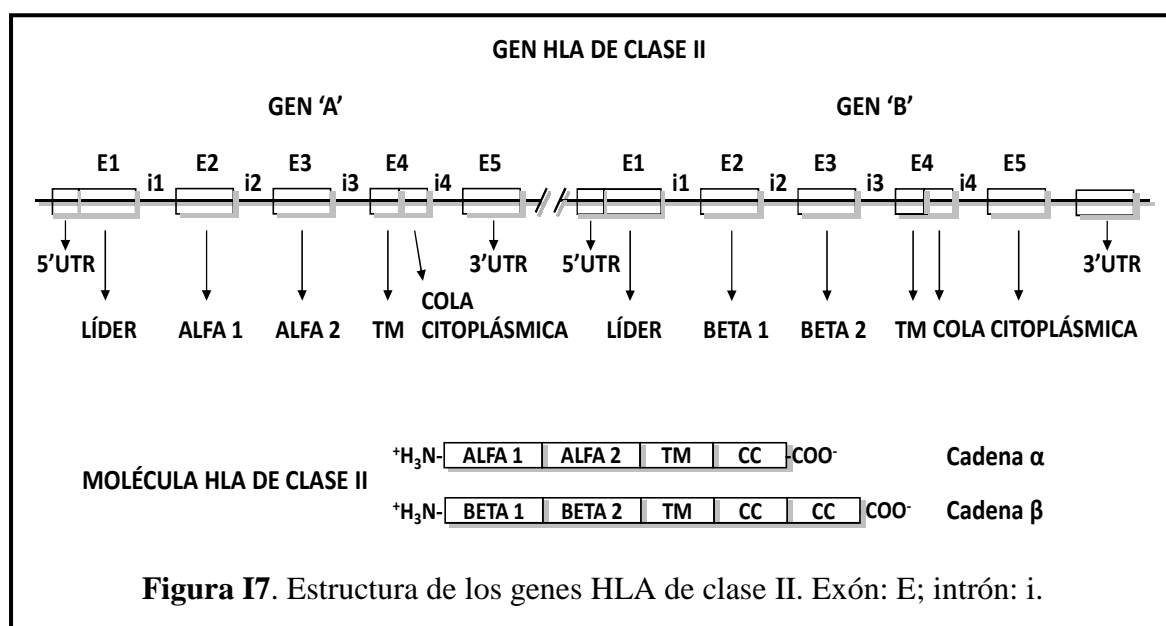
Grupo de Haplotipos	Haplotipo	Gen DRB funcional	Pseudogen DRB
DR1	DR1, DR10	DRB1 (β_1)	DRB6, DRB9
DR8	DR8	DRB1 (β_1)	
DR51	DR2	DRB1 (β_1), DRB5 (β_5)	DRB6, DRB9
DR52	DR3, DR5, DR6	DRB1 (β_1), DRB3 (β_3)	DRB2, DRB9
DR53	DR4, DR7, DR9	DRB1 (β_1), DRB4 (β_4)	DRB7, DRB8, DRB9

4.2 Estructura génica de los genes HLA de clase II

Las moléculas HLA de clase II están formadas por la unión de dos cadenas proteicas, cada una de ellas codificadas por un gen diferente. Estos genes son de dos tipos, un gen A y un gen B (véase apartado 4). Los productos de estos genes se denominan cadena pesada α y cadena ligera β de la molécula HLA de clase II.

Las cadenas α se sintetizan a partir de genes de tipo A formados de 5 exones que incluyen la región 5'UT, el exón 1 que codifica para el péptido líder, los exones 2 y 3 que lo hacen para los dominios $\alpha 1$ y $\alpha 2$ respectivamente, el exón 4 que da lugar al péptido conector, la región transmembrana y el dominio citoplasmático. La región 3'UT está codificada en parte por el exón 4 y el resto por el exón 5. Las cadenas β se sintetizan a partir de los genes B formados de 6 exones. El péptido líder está codificado por el exón 1, los exones 2 y 3 lo hacen para los dominios $\beta 1$ y $\beta 2$ respectivamente, el exón 4 codifica para el péptido conector, el dominio transmembrana y parte del dominio citoplasmático, y el resto de este dominio está originado por al menos un exón más que puede variar en los diferentes genes de esta región (Figura I7).

Los genes HLA de clase II son altamente polimórficos, pero este polimorfismo se restringe de forma casi exclusiva al exón 2 de los genes de tipo B (HLA-DRB1, HLA-DQB1 y HLA-DPB1), mientras que los genes del tipo A apenas presentan polimorfismo, en especial HLA-DRA

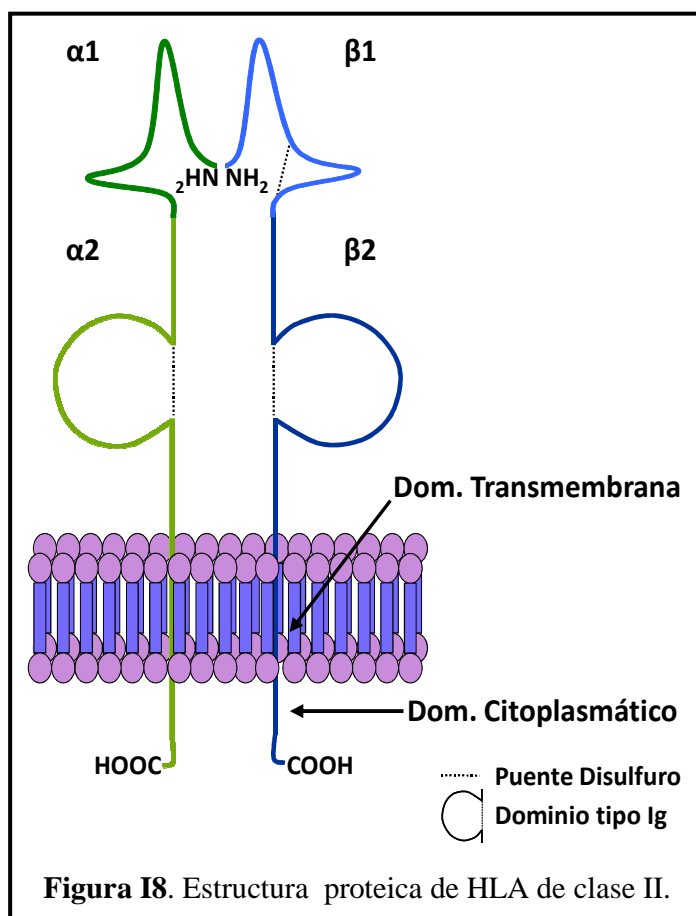


4.3 Estructura proteica de las moléculas HLA de clase II

Las moléculas HLA de clase II (HLA-DP, HLA-DQ y HLA-DR) son heterodímeros transmembrana formados por dos cadenas, una cadena α de 33-35 Kd y una cadena β de 26-28 Kd asociadas no covalentemente y ancladas a la membrana plasmática. Estas cadenas se orientan de forma que sus extremos amino-terminal aparecen en el exterior celular y sus extremos carboxi-terminales se sitúan en la porción citoplasmática de cada cadena (Figura I8).

Ambas cadenas presentan dos dominios extracelulares de 90 a 100 aa denominados $\alpha 1$, $\alpha 2$ (en el caso de la cadena pesada) y $\beta 1$, $\beta 2$ (en el caso de la cadena ligera). Los dominios $\beta 1$ son todos polimórficos y en algún caso los $\alpha 1$ también (HLA-DQA, HLA-DPA). Los dominios $\alpha 2$ y $\beta 2$ son homólogos a las regiones constantes de las inmunoglobulinas.

Un péptido conectivo hidrofóbico de 10 a 12 aa une los dominios extracelulares de ambas cadenas a las respectivas porciones transmembrana (20 a 25 aa) y a los segmentos citoplasmáticos (8 a 15 aa) (Kaufman et al., 1984).



Las proteínas HLA de clase II se asocian en el interior de la célula con un tercer elemento proteico no polimórfico de 31 Kd, denominado Cadena Invariante (Ii, proteína de membrana de tipo II invertida) y codificado fuera del sistema HLA. El complejo Ii-dímero α/β se forma en el RE y su función es permitir el transporte de las moléculas de clase II a los compartimentos endocíticos para unirse a los péptidos antigénicos. En estos compartimentos Ii es

degradada en un conjunto de aminoácidos denominado CLIP (del inglés *Class II-associated Invariant Chain Peptides*) que ocupan la valva de unión al antígeno (Ghosh et al., 1995). Sin embargo, este complejo se disocia antes de su expresión en la superficie celular (Brocke et al., 2002) mediado por moléculas HLA-DM (Sloan et al., 1995; Taxman et al., 2000).

Stern y sus colaboradores (Stern et al., 1994), usando técnicas de difracción de rayos X, consiguieron determinar la estructura tridimensional de las moléculas de clase II concretamente de HLA-DR1. Descubrieron que ésta era muy similar a la de clase I descrita anteriormente, y eran casi superponibles. Los dominios $\alpha 1$ y $\alpha 2$ de HLA-DR1

se asemejaban al dominio $\alpha 1$ y $\beta 2M$ de las moléculas de clase I y los dominios $\beta 1$ y $\beta 2$ de HLA-DR1 eran superponibles respectivamente a los dominios $\alpha 2$ y $\alpha 3$. Los cristales obtenidos de la molécula HLA-DR1 también evidenciaron la tendencia que presentan a formar dímeros $(\alpha/\beta)_2$, lo que parece tener importancia en la cascada fisiológica de transducción de señales al interior celular. El punto de interacción con CD4 se ubica en el dominio $\beta 2$ de la molécula de clase II (Konig et al., 1992; Cammarota et al., 1992).

Los dominios distales de la proteína, $\alpha 1/\beta 1$, forman la valva de unión del péptido que consiste en una lámina β formada por ocho hebras antiparalelas y flanqueadas por dos regiones en α -hélice, de la misma manera que sucede en las moléculas HLA de clase I entre los dominios $\alpha 1$ y $\alpha 2$. La región helicoidal del dominio $\alpha 1$ presenta un extremo amino-terminal en cadena extendida y su extremo carboxi-terminal se pliega hacia la base de la valva. Esta peculiaridad hace que la zona de unión al péptido muestre los extremos abiertos. En esta valva abierta por los extremos se puede acomodar un péptido grande de entre 12 y 24 aa. Este péptido puede permanecer en la valva con una conformación extendida permitiendo que sus extremos amino y carboxi-terminales sobresalgan por fuera de la molécula (Rudensky et al., 1992; Chicz et al., 1992; Stern et al., 1994; Rammensee, 1995).

5. REGIÓN MHC DE CLASE III

Esta región se sitúa entre las regiones de clase I y de clase II, y comprende 0.7 Mb (Horton et al., 2004). Se trata de la región con mayor densidad génica del genoma humano, de ahí el altísimo desequilibrio de ligamento que sufren los genes de esta región.

Los principales genes que se localizan en esta región son genes que expresan proteínas con función inmunológica como los genes TNF (TNF-A y -B), los genes de complemento de la vía clásica, C4A, C4B y C2; y de la vía alternativa, gen del factor B del complemento, y genes del LY6. También se localizan los genes HSP, que expresan proteínas que intervienen en procesos de síntesis, plegamiento, ensamblaje, transporte y degradación de proteínas, aumentando su expresión en condiciones de estrés celular (véase apartado 2.2) y el gen CYP21A2 que codifica para la esteroide 21-hidroxilasa, miembro de la superfamilia de enzimas del citocromo P450 que catalizan reacciones involucradas en la síntesis del colesterol, esteroides, incluyendo el cortisol, y otros lípidos.

6. FUNCIÓN DE LAS MOLÉCULAS HLA

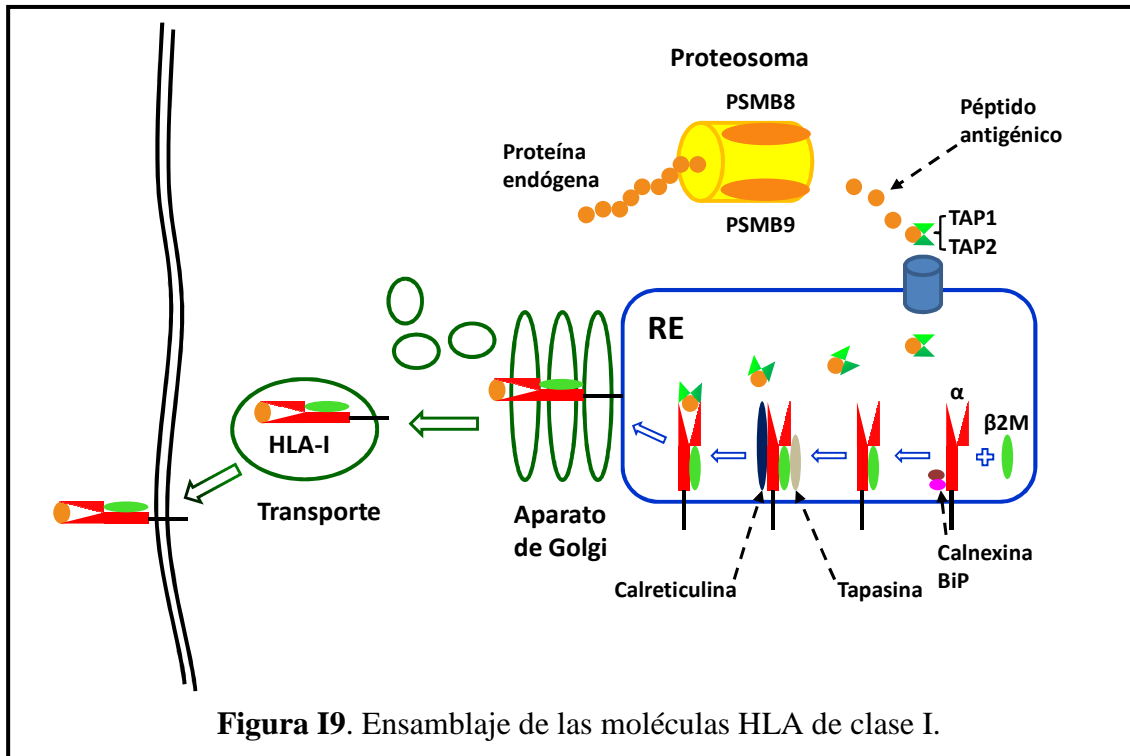
6.1 Presentación antigénica de las moléculas HLA de clase I

Las moléculas HLA de clase I actúan tanto como receptores de péptidos endógenos propios, normales y/o alterados (tumoraes), como también de péptidos virales, para hacerlos accesibles a la detección por parte de las células T que llevan en su superficie celular el marcador CD8 (mayoritariamente linfocitos citotóxicos).

Las funciones de unión y presentación peptídica, implican que cada molécula de clase I pueda formar complejos con un amplio espectro de péptidos. Sin embargo, cada una de ellas une un grupo determinado, lo cual indica que estas moléculas presentan selectividad alelo-específica (Falk et al., 1991). La estructura cristalina de la molécula HLA con el péptido unido, ha demostrado que un único producto alélico puede unir un extenso abanico de péptidos con alta afinidad (Madden, 1995).

El único requerimiento imprescindible para que el péptido se adapte bien a la valva es que el aminoácido carboxi-terminal de éste sea hidrofóbico o cargado (todos excepto la prolina y la glicina). Las cadenas laterales de los residuos aminoácidos P1, P2, P3, P6, P7 y P9 interactúan con los “bolsillos” A, B, D, C, E y F respectivamente (Figura I5). Para péptidos de 10 u 11 aminoácidos, los extremos se adaptan a los “bolsillos” A y F y los demás se curvan hacia fuera de la valva. Estos aminoácidos se unen a los “bolsillos” mediante puentes de hidrógeno y la configuración de la valva permite que sus propios residuos interactúen con el péptido y puedan adoptar variedad de conformaciones para mejorar la unión. Además, la mayor parte de estas interacciones confieren estabilidad tanto a la molécula como al péptido y los hacen más accesibles al contacto con el TCR (Falk et al., 1991; Rammensee et al., 1993; Madden, 1995).

Los péptidos proceden principalmente, como se ha explicado anteriormente, de proteínas del interior celular degradadas en el proteosoma (orgánulo consistente en un complejo proteico dependiente de ATP; Koopmann et al., 1997). En este orgánulo participan dos subunidades LMP, actualmente denominadas PSMB8 y PSMB9, cuyos genes se localizan en la región HLA de clase II. Éstas, tras ser inducidas por el interferón gamma (INF- γ), reemplazan a dos subunidades constitutivas del proteosoma, de manera que orientan el procesamiento hacia un patrón distinto y, probablemente específico de los requerimientos de las moléculas HLA de clase I (Fruh et al., 1994; Belich et al., 1994).



Los péptidos resultantes pasan al citoplasma, donde son transportados a través de la membrana del RE por los transportadores de péptidos TAP1 y TAP2, cuyos genes también se localizan en la región HLA de clase II. Esta introducción en el RE es dependiente de ATP (Abele & Tampe, 2004). En este mismo orgánulo, la cadena pesada α se asocia a la proteína BiP (del inglés, *Binding Immunoglobulin Protein*) y a la calnexina. Posteriormente se ensambla con la $\beta 2M$, lo que supone la liberación de las dos proteínas anteriores. A continuación, al complejo $\alpha/\beta 2M$ recién formado se le unen dos proteínas chaperonas (proteínas que intervienen en el plegamiento de otras proteínas recién formadas): la calreticulina y la tapasina. Posteriormente se les une el heterodímero TAP1/TAP2. Una vez formado este complejo, la molécula de clase I es cargada con el péptido proporcionado por TAP1/TAP2, así, adquiere la conformación madura, se libera, y es transportado hacia la membrana celular a través del cis- y trans-Golgi siguiendo la ruta secretora (Koopmann et al., 1997; Figura I9).

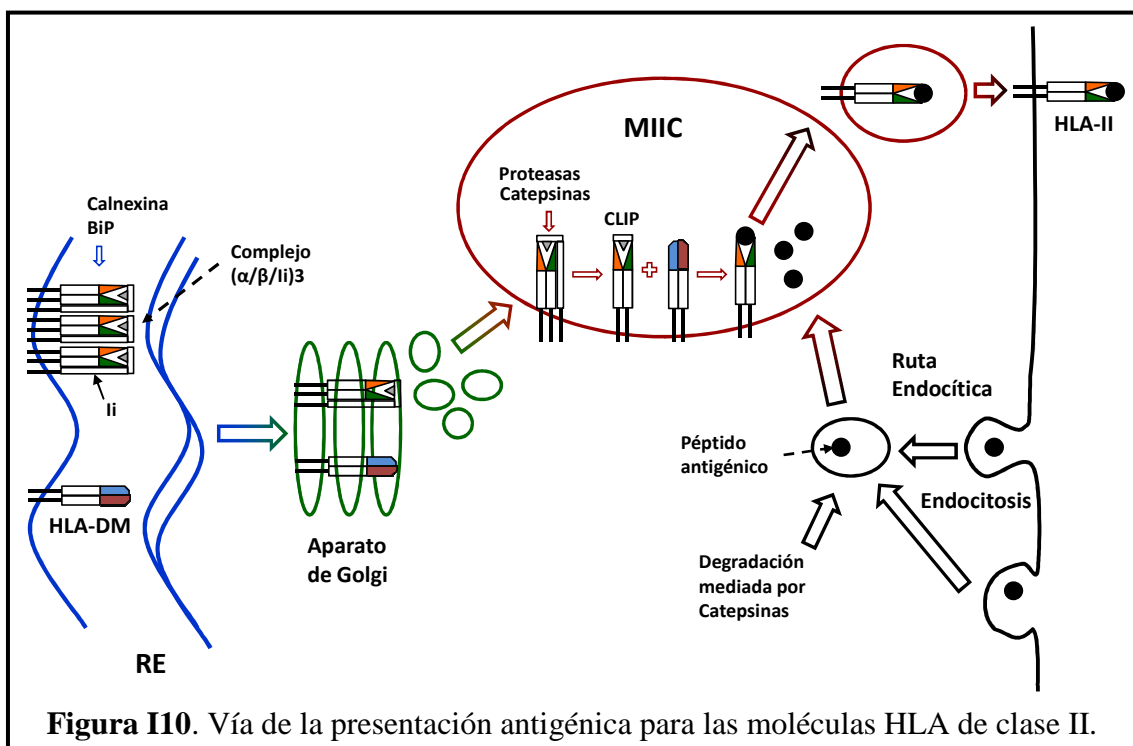
6.2 Presentación antigénica de las moléculas HLA de clase II

Las moléculas HLA de clase II unen un grupo heterogéneo de péptidos procedentes en su mayor parte del exterior celular (bacterianos y/o parasitarios) y se encargan de “exponerlos” en la superficie de las células presentadoras de antígenos para ser “analizados” y reconocidos por el TCR de los linfocitos T con marcador CD4

(CD4+, generalmente cooperadores). La mayor parte de los péptidos presentados por las moléculas de clase II provienen de la degradación endocítica de proteínas exógenas o bien, en menor medida, de proteínas endógenas que acceden a los endosomas mediante la acción de proteasas como la catepsina B (Katunuma et al., 1994).

El ensamblaje de la molécula de clase II con el péptido ocurre en unos compartimentos especiales llamados MIIC (del inglés, *MHC Class II Compartment*) (Tulp et al., 1994; Pieters, 1997). Mientras los péptidos se forman, las cadenas α y β de los antígenos HLA de clase II son sintetizados en el RE e inmediatamente el heterodímero naciente se asocia a la cadena invariante Ii.

Esta cadena, además de intervenir en el ensamblaje de las cadenas α/β , también bloquea la hendidura del heterodímero impidiendo la unión en el RE. Ii interacciona con el complejo α/β gracias a la acción de la proteína BiP y la calnexina (Marks et al., 1995). De esta forma se forman complejos nonaméricos $(\alpha/\beta/Ii)_3$ (Roche et al., 1991). Estos complejos son rápidamente exportados hacia los compartimentos endocíticos a través del aparato de Golgi guiados por las señales de la cola citoplasmática de la cadena invariante (Pieters, 1997).



Cuando el complejo nonamérico entra en el MIIC, la cadena invariante sufre una proteólisis secuencial, mediada por proteasas y catepsinas entre otras, que la reducirá a un pequeño péptido de unos 20 aminoácidos denominado CLIP y que ocupará la

hendidura. En este proceso es cuando los complejos nonaméricos se disocian en dímeros α/β . En este punto, interviene otro heterodímero de clase II, HLA-DM (DMA/DMB), que actúa como catalizador, ya que: promueve la separación del CLIP, estabiliza la molécula de clase II libre de péptido y favorece la unión del péptido antigénico, reacción que se ve favorecida por el pH ácido del compartimento MIIC (Sloan et al., 1995; Taxman et al., 2000; Figura I10).

7. CARACTERÍSTICAS DEL SISTEMA HLA

Como ya se mencionó en el primer apartado de esta introducción, el sistema HLA se caracteriza por ser poligénico y dialélico con expresión codominante. Pero las características más importantes de este sistema que lo hacen destacar con respecto a otros sistemas génicos son su elevado nivel de polimorfismo y su alto grado de desequilibrio de ligamento.

7.1 Polimorfismo del Sistema HLA

La gran cantidad de variantes alélicas que presenta el sistema HLA hacen de este sistema el más polimórfico de los sistemas genéticos conocidos en la actualidad en el ser humano. Esta característica implica que dos individuos, sin ser parientes, van a tener muy pocas probabilidades de ser HLA idénticos. No todos los loci son igualmente polimórficos, concentrándose su mayor grado en aquellos cuya función principal es la presentación antigénica, HLA de clase I y clase II clásicos. Actualmente, están descritos 3760 alelos del locus más polimórfico, HLA-B (Robinson et al., 2015).

Existen varios mecanismos genéticos que pueden contribuir a la generación del polimorfismo HLA como son las mutaciones puntuales y los eventos de conversión génica asociados a fenómenos de recombinación entre regiones de secuencias homólogas.

7.1.1 Polimorfismo serológico del sistema HLA

El polimorfismo a nivel serológico consiste en el número de tipos serológicos (especificidades serológicas) distintos asociados a cada una de las variantes alélicas de los loci que conforman el sistema HLA. En varios casos, especificidades serológicas que se pensaba que eran individuales han resultado ser grupos de antígenos estrechamente relacionados, con una especificidad más estricta, que proceden del tipo

serológico original (Hegland, 1997). A estos se les conoce como especificidades subtípicas, y a los originales, especificidades supertípicas.

Los distintos equivalentes serológicos se detecta mediante técnicas serológicas como son los test de microlinfocitotoxicidad. Esta técnica consiste en enfrentar poblaciones de linfocitos (antígeno) con una batería de anticuerpos monoclonales específicos para cada uno de los antígenos posibles y añadir complemento que produce lisis celular cuando hay reconocimiento entre el antígeno HLA y el anticuerpo respectivo (Terasaki & McClelland, 1964; Danilovs et al., 1980). El uso de anticuerpos monoclonales se debe a que permiten una mayor rapidez y exactitud para la determinación del subtipo HLA serológico.

Los comités de Nomenclatura HLA de la Organización Mundial de la Salud, reunidos en los Talleres Internacionales de Histocompatibilidad, han definido las especificidades serológicas para los diferentes loci HLA. En la actualidad, se reconocen 28 especificidades para los genes HLA-A, 62 especificidades para HLA-B, 10 especificidades para HLA-C, 24 especificidades para los genes HLA-DR, 9 especificidades para HLA-DQ y finalmente 6 especificidades para los genes HLA-DP (Marsh et al., 2010).

7.1.2 Polimorfismo genético del sistema HLA

El desarrollo de las técnicas de amplificación y secuenciación permitió demostrar que el polimorfismo a nivel serológico es sólo una parte del que encontramos a nivel genómico. El polimorfismo genético viene determinado por la secuencia de ADN presente en cada grupo de alelos o en cada variante alélica en particular. Se detecta mediante el uso de varias técnicas de biología molecular como la Reacción en Cadena de la Polimerasa (PCR), la hibridación con cebadores específicos de alelos y/o grupos de alelos (SSP) o de sondas (fluoróforos) específicas de alelo (SSO o SSOP) y la secuenciación de ADN (directo o por clonaje). Todo esto ha permitido caracterizar con mayor detalle los genes del sistema HLA, lo que ha incrementado el conocimiento del polimorfismo de estos genes. Así, tenemos que para los genes HLA de clase I, el polimorfismo mayoritariamente reside en los exones 2 y 3 que codifican para las cadenas $\alpha 1$ y $\alpha 2$ de la molécula respectivamente (véanse apartados 3.2 y 3.3). Mientras que para el caso de los genes HLA de clase II, la variabilidad alélica, reside mayoritariamente en el exón 2, ya sea de la cadena α (en el caso de los genes DQA1 y

DPA1) como de la cadena β (para los genes DRB1, DQB1 y DPB1), siendo éstos últimos mucho más polimórficos (véase apartado 4.2).

Con el conocimiento de este nuevo polimorfismo genético, el número de alelos HLA se ha incrementado en cada uno de los loci, de tal forma que desde 1998 (año de publicación de la base de datos de inmunopolimorfismo IMGT/HLA) hasta la actualidad el número de variantes descritas ha aumentado en 11578 alelos, siendo el aumento más significativo el encontrado en los loci de clase I con más de 2400 alelos cada uno (Robinson et al., 2015). Actualmente las cifras oficiales incluidas en la última publicación de IMGT/HLA (3.19, enero 2015) sobre el polimorfismo genético HLA ascienden a 12542 alelos, de los cuales 12494 son de genes funcionales (Tabla I3).

TABLA I3. Número de alelos descritos para cada locus funcional HLA (IMGT/HLA v. 3.19, enero 2015). En cursiva los loci no clásicos, mucho menos polimórficos.

HLA de clase I		HLA de clase II	
Locus	Nº de alelos	Locus	Nº de alelos
HLA-A	2995	HLA-DPA1	38
HLA-B	3760	HLA-DPB1	489
HLA-C	2553	HLA-DQA1	52
<i>HLA-E</i>	17	HLA-DQB1	734
<i>HLA-F</i>	22	HLA-DRA	7
<i>HLA-G</i>	50	HLA-DRB1	1638
		HLA-DRB3	58
		HLA-DRB4	15
		HLA-DRB5	21
		<i>HLA-DMA</i>	7
		<i>HLA-DMB</i>	13
		<i>HLA-DOA</i>	12
		<i>HLA-DOB</i>	13

7.1.3 Evolución del polimorfismo del sistema HLA

Los mecanismos moleculares involucrados en la generación del polimorfismo en este sistema, son debidos de una forma general a varias fuerzas que actúan en la naturaleza:

1. Variabilidad genética. Genera diferentes variantes en el material genético a través de mecanismos como:

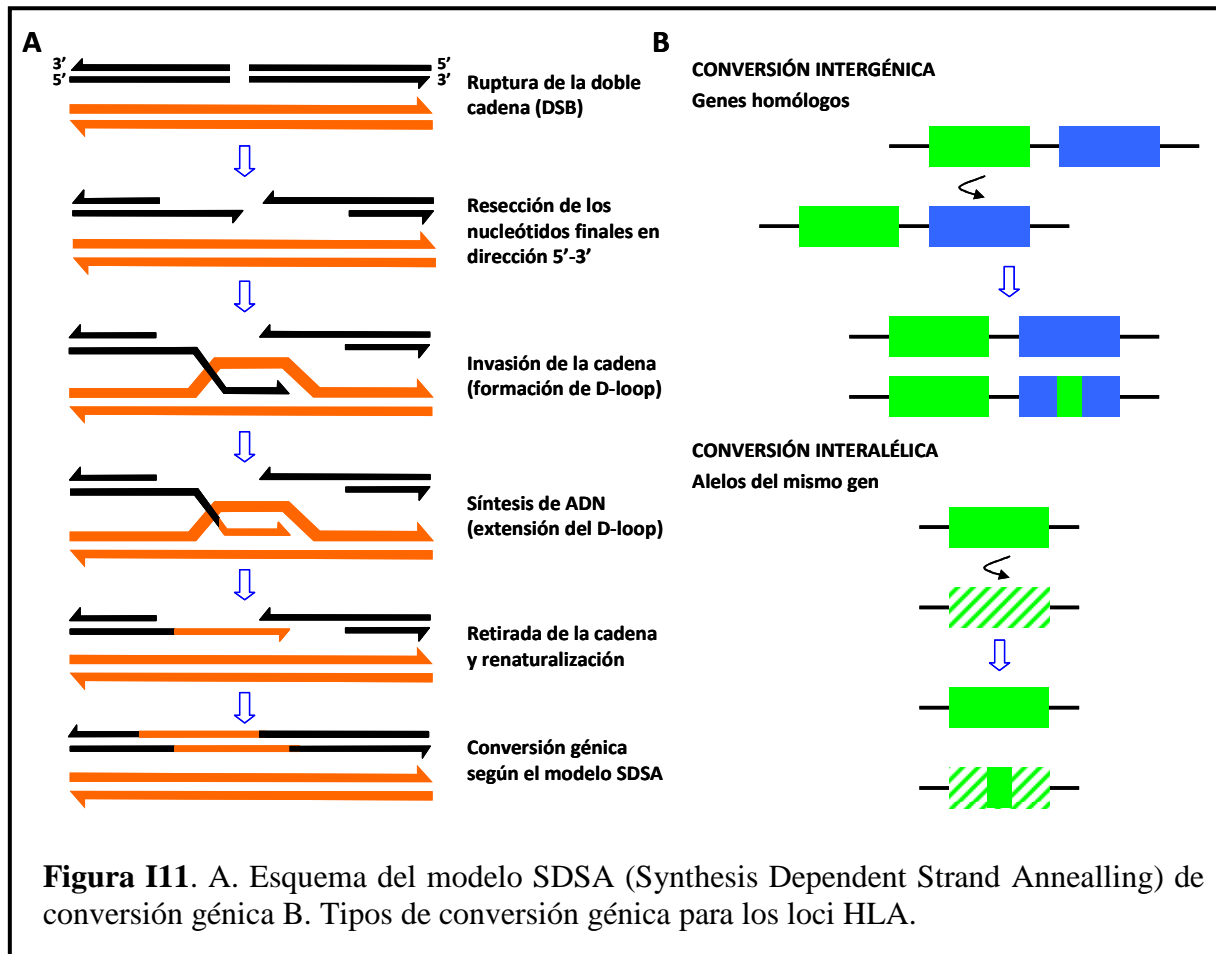
- **Mutaciones puntuales:** modificaciones en la secuencia de ADN por sustituciones (cambio de un nucleótido por otro) o por cambio en el número de nucleótidos, tales como deleciones (perdida de uno o varios nucleótidos) o inserciones (introducción de nucleótidos adicionales en la cadena).
 - **Conversión génica:** se produce por la sustitución de un segmento de la secuencia por otro segmento homólogo, sucediendo durante la meiosis o como mecanismo de reparación de ADN dañado. Estos procesos de recombinación son frecuentemente entre alelos del mismo locus (conversión interalélica) y con menos frecuencia entre alelos de diferentes genes homólogos HLA (conversión intergénica) (Parham & Ohta, 1996; Figura I11).
2. **Selección natural:** las variantes naturales más ventajosas se seleccionan frente a las que peor cumplen su función.
 3. **Deriva genética:** cambio de las frecuencias alélicas en una población como consecuencia de la aleatoriedad en la reproducción, lo que conlleva a la pérdida de alelos por azar.

El gran número de alelos HLA formados mediante procesos de conversión hace pensar a algunos autores que este fenómeno es el factor principal de la variabilidad de este sistema (Parham et al., 1995). Sin embargo, para otros investigadores, la acumulación de mutaciones puntuales durante millones de años de evolución entre especies sería la principal causa del elevado polimorfismo (hipótesis trans-especies) (Klein, 1987; Klein et al., 1993; Ayala, 1995). Consideran que los mecanismos de recombinación no generan diversidad genética por sí mismos, pero una vez que se ha acumulado un gran número de sustituciones, la conversión génica puede aumentar sustancialmente el nivel de polimorfismo (Walsh, 1987; Ohta, 1990; Ohta, 1991).

La existencia de este elevado polimorfismo en los genes HLA de clase I y de clase II clásicos, se encuentra en estrecha relación con la función inmune de las proteínas que codifican, ya que esta variabilidad se acumula principalmente en la región de unión al péptido, y por lo tanto una mayor heterocigosidad permitiría acumular un mayor repertorio para la presentación. La selección disruptiva o equilibradora (aquella que mantiene la variabilidad alélica en el conjunto de genes con frecuencias distribuidas de manera más uniforme de lo que cabría esperar en un modelo de evolución neutral) podría ser la responsable del origen y mantenimiento del polimorfismo en este sistema. Estudios recientes confirman el papel de esta fuerza selectiva de forma muy significativa para los loci HLA-C y HLA DQB1 (Solberg et al., 2008). Esto sería debido

a la ventaja de la heterocigosis o sobredominancia, ya que para llevar a cabo una respuesta inmunológica efectiva contra un amplio abanico de patógenos, los heterocigotos pueden presentar el doble de péptidos que los homocigotos. Esta diferencia debería resultar en una respuesta más eficaz por parte de los linfocitos T. Por lo tanto el mantenimiento de la heterocigosidad favorecería la formación de distintos conjuntos de especificidades de unión a péptidos a través de la acumulación de mutaciones puntuales y fenómenos de conversión génica (Klein et al., 1993; Parham & Ohta, 1996). Además, la recombinación intergénica da lugar a nuevas combinaciones de alelos de los distintos loci (haplotipos) favoreciendo una mayor variabilidad, que se podría traducir en una mejor respuesta inmunológica. De hecho las frecuencias de los haplotipos HLA muestran una mayor variación entre grupos étnicos que las frecuencias de alelos individualmente debido a la variación en cuanto al desequilibrio de ligamento entre poblaciones (Bodmer & Bodmer, 1978).

Además de la ventaja de la heterocigosis, podría existir una selección dirigida por el patógeno (PDBS, del inglés *Pathogen-Driven Balancing Selection*), también conocida como selección dependiente de la frecuencia, en poblaciones sometidas a enfermedades epidémicas bajo una gran presión selectiva por un o varios patógenos. Bajo estas condiciones cualquier alelo que conlleve una mayor inmunidad contra ese patógeno debería aumentar su frecuencia, pudiendo incluso dominar el locus. Sin embargo, los patógenos pueden mutar las secuencias de sus péptidos antigénicos, y de esta manera adaptarse al alelo o alelos más frecuentes que confieren inmunidad, lo que supondría una ventaja adaptativa para los alelos más raros (Takahata & Nei, 1990; Parham & Ohta, 1996). En este modelo de selección los alelos raros no tienen por qué desaparecer de la población por selección natural o por deriva genética ya que las frecuencias alélicas fluctuarían en función de los péptidos antigénicos de los patógenos. Por todo esto, esta selección sería transitoria generando un polimorfismo más dinámico, a diferencia de la selección disruptiva, más estable (Slade & McCallum, 1992). Un estudio reciente muestra cómo las poblaciones que están expuestas a una mayor riqueza patogénica presentan una mayor diversidad HLA de clase I, de forma muy significativa para el locus HLA-B (Prugnolle et al., 2005), para el que ya se había demostrado el efecto considerable que la selección disruptiva tiene sobre él (Satta et al., 1994). La conjunción de estas fuerzas selectivas podría ser la responsable de que se trate del locus más polimórfico (Takahata & Nei, 1990).

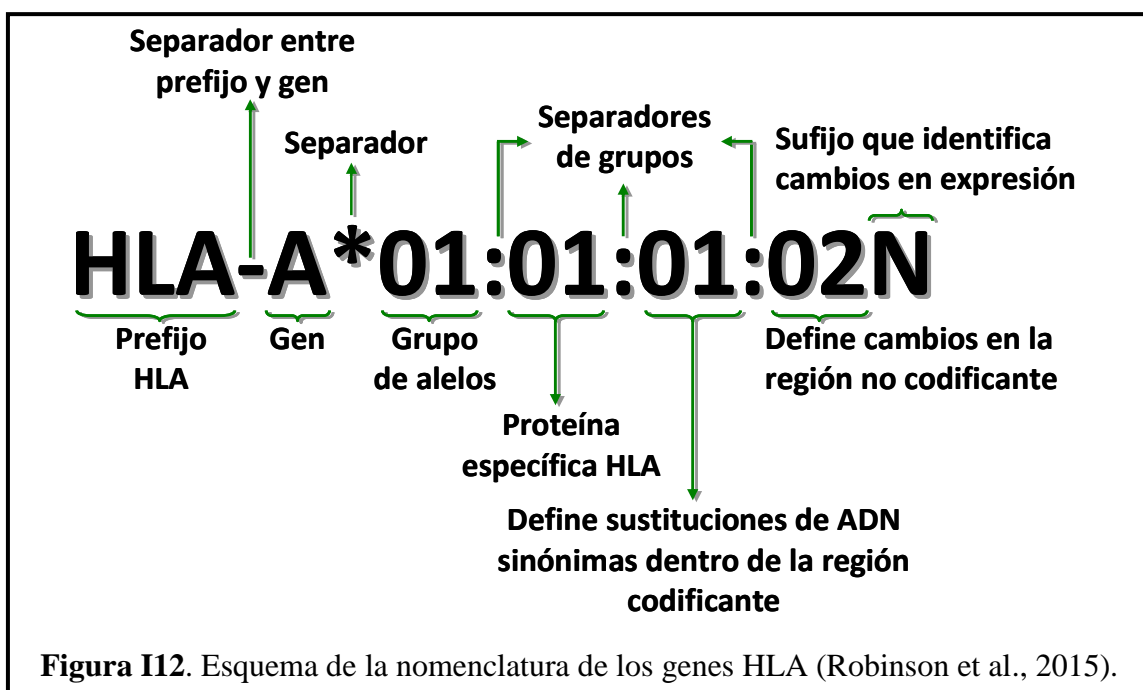


7.1.4 Nomenclatura del sistema HLA

El comité de nomenclatura para los factores del sistema HLA coordinado por la Organización Mundial de la Salud (OMS), mediante los Talleres Internacionales de Histocompatibilidad y a través de sus informes periódicos desde su primera publicación en 1968 (WHO Nomenclature Committee, 1968) ha creado un sistema de identificación y enumeración de los diferentes alelos HLA, mediante el uso de caracteres alfanuméricos asignados a aquellos patrones de reacción serológica que correlacionan con secuencias de ADN, previo análisis de los patrones que presentan reacción cruzada entre diferentes antígenos que comparten algunos epítomos (Bodmer et al., 1997).

El último informe de 2010 (Marsh et al., 2010) supuso un cambio en la designación de los alelos HLA como se describe a continuación. En primer lugar los alelos se identifican por las letras correspondientes a su locus (A, B, C, DRA, DRB1,...). A continuación, cada alelo recibe un número único que corresponde a un máximo de cuatro grupos de pares de dígitos separados por dos puntos. Todos los alelos reciben al menos un nombre de cuatro dígitos, correspondientes a los dos primeros grupos (Figura I12).

Los dos primeros números indican la especificidad serológica definida mediante anticuerpos o bien identifican una secuencia homóloga a la que codifica para ese tipo serológico. El segundo grupo de dos dígitos indica el número correspondiente dentro del grupo de alelos del locus al que se refiere y se asigna en orden cronológico y consecutivo de acuerdo a su descripción. Estas moléculas son proteínas diferentes. Para aquellos alelos que presentan un cambio en la región codificante sin producir cambios en la proteína resultante (mutaciones sinónimas), se agregan un tercer conjunto de dos números. Además, para los alelos que presenten un cambio en la secuencia del ADN en el intrón o las regiones 5' o 3'UT (regiones no codificantes), se distinguen con un cuarto conjunto de dos dígitos. Además del número de alelo se pueden añadir sufijos a continuación que dan información sobre su expresión (Figura I12). De tal forma que la letra 'N' (del inglés *Null*) identifica a los alelos con expresión nula, la letra 'L' (del inglés *Low*) a los alelos con una baja expresión celular, la letra 'S' (del inglés *Secreted*) a los alelos que codifican para proteínas solubles secretadas, la letra 'C' (del inglés *Cytoplasm*) a alelos que codifican para proteínas citoplasmáticas no presentes en la superficie celular, la letra 'A' (del inglés *Aberrant*) a los alelos aberrantes de los que se duda si expresan o no proteínas, y por último la letra 'Q' (del inglés *Questionable*) a los alelos cuya expresión se cuestiona ya que se ha comprobado que su mutación afectaría a su normal expresión.



En cuanto a la nomenclatura de los antígenos HLA (tipos serológicos), el primer comité de nomenclatura para los factores HLA formado en el segundo Taller Internacional de Histocompatibilidad estableció las primeras pautas para estandarizar dicha nomenclatura, utilizando el termino HL-A para indicar el sistema de antígenos leucocitarios en humanos (Varios Autores, 1968). El nombre HLA sustituyó a HL-A en 1975 cuando se constató la presencia de más de un gen en este sistema (WHO Nomenclature Committee, 1975). Para designar los distintos tipos serológicos se identifican con la letra correspondiente a la molécula o gen (A, B, C,...) seguido de un número asignado por orden cronológico y consecutivo de acuerdo a su descripción. Las relaciones entre las especificidades subtípicas con su tipo serológico original se identifican mediante un paréntesis en las primeras. Por ejemplo, A25(10) y A26(10) son subtipos del original A10.

7.2 Desequilibrio de ligamento de los genes HLA

Definición de haplotipo

Es la combinación de determinantes antigénicos individuales que son controlados positivamente por un alelo. Este término lo expresó por vez primera Cepelini y sus colaboradores (Ceppelini et al., 1967). Aplicado para los genes HLA se puede derivar en “una combinación de alelos de varios loci en un único cromosoma de un determinado individuo”.

Supongamos que tenemos dos loci (A y B) cercanos en el mismo cromosoma, cada uno de ellos con dos alelos (A1, A2, B1 y B2). En el doble heterocigoto resultante A1B1/A2B2, la combinación A1B1 es el haplotipo 1 y la combinación A2B2 es el haplotipo 2, este sería el ejemplo más sencillo de haplotipo. Lo que ocurre es que el sistema HLA no es tan sencillo, donde un haplotipo llamado “extendido” consta de varios loci, cada uno de ellos con multitud de posibilidades alélicas.

Desequilibrio de ligamiento

El desequilibrio de ligamiento presente en el sistema HLA está representado por la presencia de haplotipos con una frecuencia mayor a la esperada por el azar. En condiciones de equilibrio, el total de haplotipos esperados para una combinación de dos loci sería el producto de las frecuencias génicas de los alelos de esos loci, lo cual no se ha encontrado cuando se realizan estudios de población (Imanishi et al., 1992b; Clayton & Lonjou, 1997).

Este fenómeno se denomina “desequilibrio de ligamiento” y representa la diferencia entre las frecuencias haplotípicas esperadas y las observadas (Mattiuz et al., 1971). El conocimiento de los haplotipos HLA, se puede obtener a partir de estudios en familias o genealógicos (Hiller et al., 1978; Martínez-Laso et al., 2011), o también a partir de estudios de poblaciones (Cavalli-Sforza et al., 1994) y son útiles para estudios de genética de poblaciones, programas de trasplante de órganos y tejidos, asociación a enfermedades y estudios epidemiológicos.

Parámetros que miden el desequilibrio de ligamiento

- D o Δ - desequilibrio de ligamiento
- D' o Δ' - desequilibrio de ligamiento relativo, es el cociente de D entre la D máxima teórica para las frecuencias alélicas observadas (D_{\max}).
- DPA (del inglés, *Disequilibrium Pattern Analysis*), proporciona el patrón de todos los haplotipos bi-locus (de dos loci) que comparten un alelo determinado.

Origen del desequilibrio de ligamiento

Se desconoce la razón por la que se deriva la existencia de haplotipos y su desequilibrio de ligamiento, pero se podría explicar por dos causas distintas:

- 1.- La presión selectiva: ya que se ha sugerido que ciertas combinaciones de alelos HLA de diferentes loci podrían ser necesarias para el desarrollo de una respuesta inmune correcta a patógenos concretos encontrados en un área geográfica determinada, especialmente en épocas con frecuentes epidemias (véase apartado 7.1.3).
- 2.- El efecto fundador por el que una determinada población procedente de un lugar y población específicos se establece en otro, modificando la base genética de la población existente en ese lugar.

8. APLICACIONES DEL SISTEMA HLA

El conjunto de genes y moléculas que conforman el sistema HLA tiene importantes aplicaciones a varios niveles. Debido a que este sistema fue inicialmente identificado como “antígenos de histocompatibilidad”, su aplicación clínica más extendida ha sido en programas de trasplante de órganos y tejidos. Sin embargo, el posterior descubrimiento de su función biológica en la presentación de antígenos aumentó los campos de estudio y sus aplicaciones clínicas, investigando su posible papel como genes de susceptibilidad y/o resistencia a enfermedades, su participación en

el desarrollo de patología autoinmunes o su relación con reacciones de hipersensibilidad ante determinados fármacos (Nepom & Erlich, 1991; Erlich et al., 1993; Horton et al., 2004; Becquemont, 2010; Profaizer & Eckels, 2012). Su alto grado de polimorfismo y la elevada correlación a nivel geográfico de las frecuencias alélicas poblacionales también hacen de este sistema un buen marcador genético poblacional (Imanishi et al., 1992c).

8.1 HLA y trasplantes

Se ha comprobado con estudios realizados tras seguimiento de pacientes con órganos trasplantados, que la compatibilidad entre donante y receptor en las moléculas HLA es necesaria para la supervivencia del injerto en los casos de riñón, páncreas, pulmón, corazón, córnea, hígado e intestino. Esta compatibilidad, debe ser al menos con los antígenos pertenecientes a los loci HLA-B y -DR. La compatibilidad debe ser máxima cuando hablamos de trasplante de médula ósea para evitar que se produzca el llamada “enfermedad de injerto contra huésped” (GVHD, del inglés *Graft Versus Host Disease*), que surge debida a que las células linfocitarias del tejido trasplantado reconocen como extrañas las moléculas HLA (u otros antígenos menores de histocompatibilidad) de las células del donante atacándolas y causando la enfermedad; un único cambio en un aminoácido es suficiente para que se produzca este rechazo.

8.2 HLA y enfermedad

La primera vez que se estableció relación entre el sistema HLA y una enfermedad fue en pacientes con linfoma de Hodgkin, que presentaban una frecuencia de determinados antígenos HLA-B por encima de la media (Aimel, 1967). Desde entonces, numerosas variantes alélicas de genes HLA han sido asociadas con enfermedades autoinmunes (principalmente), inflamatorias e infecciosas. Sin embargo, el carácter multifactorial de muchas de estas enfermedades sumado al elevado polimorfismo de algunos de estos genes y el fuerte desequilibrio de ligamiento existente entre ellos dificulta en la mayoría de los casos la determinación de la variante alélica implicada así como su grado de participación en la enfermedad (Horton et al., 2004). Además, el estudio de estas enfermedades en grupos antropológicamente distintos muestra con frecuencia asociaciones diferentes.

HLA y autoinmunidad

Muchas enfermedades autoinmunes están asociadas a factores HLA, aunque los mecanismos responsables de esta asociación no están claros. La idea más extendida es que se rompe la tolerancia inmunológica hacia proteínas propias cuando determinadas moléculas HLA de clase II (fundamentalmente) activan linfocitos T auto-reactivos al presentarles péptidos propios o extraños. Una de las asociaciones más fuerte es la que relaciona la espondilitis anquilosante (inflamación crónica que afecta principalmente a la columna vertebral) con el grupo de alelos B*27. Sin embargo no todos los alelos de este grupo presentan asociación a esta enfermedad, como HLA-B*27:06 y B*27:09, y el estudio de esta enfermedad en distintos grupo étnicos muestra diferencias en la distribución de los alelos asociados (HLA-B*27:05 es el principal subtipo asociado en caucasoides, mientras que en chinos es el HLA-B*27:04) (Khan, 2013).

TABLA I4. Algunos ejemplos de enfermedades asociadas a antígenos HLA. Entre paréntesis figuran los alelos supuestos de predisposición.

Antígeno HLA	Enfermedad	Antígeno HLA	Enfermedad
A29	Coriorretinopatía en perdigonada	Cw6 (C*06:02)	Psoriasis
B27 (B*27:05,B*27:04)	AS	DQ6 (DQA1*01:02-DQB1*06:02)	Narcolepsia
B27	Uveitis	DR3/DR4 (DRB1*03:01/DRB1*04:01)	DM1
	Artritis reactiva	DR4 (DRB1*04:02) DQ8(3) (DQB1*03:02)	PV
B51(5) (B*51)	Síndrome de Behcet	DQ2 (DQA1*05-DQB1*02)	Celiaquía

AS: espondilitis anquilosante (del inglés *Ankylosing Spondylitis*); DM1: diabetes mellitus tipo 1; PV: pénfigo vulgar.

Otras enfermedades que tienen el sistema HLA como principal factor asociado son la esclerosis múltiple (inflamación crónica del sistema nervioso central asociada a variantes alélicas del antígeno HLA-DR2), el lupus eritematoso sistémico (proceso autoinmune generalizado asociado a variantes alélicas de los antígenos HLA-DR2 y HLA-DR3), la colitis ulcerosa y la enfermedad de Crohn (procesos inflamatorios intestinales asociados a diversos alelos y haplotipos HLA de clase II) y la artritis reumatoide (proceso inflamatorio que afecta a las articulaciones de forma generalizada asociado a las variantes alélicas del locus HLA-DRB1). La Tabla I4 recoge más

ejemplos de enfermedades asociadas a factores HLA. Sin embargo, la mayoría de estas asociaciones se han establecido casi exclusivamente en poblaciones de origen europeo y el número de genes analizados dentro del sistema HLA no supera la veintena (Hill, 2001; Fernando et al., 2008).

HLA y patógenos

La asociación entre el sistema HLA y enfermedades infecciosas se debe a la distinta capacidad que tienen los alelos de los genes implicados en la presentación antigénica para activar respuestas eficaces contra los patógenos. De ahí se deriva que unos alelos (o haplotipos) confieran resistencia y otros sean factores de susceptibilidad. Nuevamente, el elevado polimorfismo de estos genes y los desequilibrios de ligamiento entre alelos hacen muy difícil atribuir estas cualidades de resistencia o susceptibilidad a variantes concretas. Además, la eficacia de unas proteínas de histocompatibilidad puede estar condicionada por otras que coexpresan con ellas.

Los casos más estudiados de asociación están relacionados con la malaria, la hepatitis B y el SIDA. Los individuos afectados de malaria por *Plasmodium falciparum* (enfermedad producida por este protozoo parásito que invade los eritrocitos causando fiebre y anemia) son más resistentes a la enfermedad cuando poseen las variantes alélicas HLA-B*53:01, DRB1*13:02 (Hill et al., 1991) y B*35 (Gilbert et al., 1998; Young et al., 2005). La hepatitis B (causada por una infección viral que, una vez establecida de forma crónica, puede derivar en cirrosis y carcinoma hepatocelular) es prevenida por los antígenos HLA-DR13, mientras que los alelos HLA-DRB1*11, DRB1*12 y DQB1*03:01 actúan como factores de susceptibilidad (Singh et al., 2007).

En el caso del SIDA (inmunodeficiencia provocada por la invasión del virus VIH de células T CD4+) está comprobado que los individuos con mayor probabilidad de supervivencia frente a la infección por el virus VIH-1 son, en general, los que presentan mayor grado de heterocigosidad en los genes HLA, la presencia de los alelos HLA-B*57:01 y B*27 (Trachtenberg & Erlich, 2001) y ausencia de los antígenos HLA-B35 y HLA-Cw4 (Carrington et al., 1999); también se ha constatado la importancia del papel que desempeñan los factores HLA de clase I no clásicos en la resistencia al virus, ya que la presencia, tanto en homocigosis como en heterocigosis, de los alelos HLA-E*01:03 y HLA-G*01:05N se ha asociado con un menor riesgo a infección por VIH-1 (Lajoie et al., 2006; Tripathi & Agrawal, 2007); en cualquier caso, se han identificado numerosos alelos y haplotipos de clase I y de clase II que confieren resistencia o

susceptibilidad frente a la transmisión del virus y frente a la progresión de la enfermedad en individuos infectados tanto por VIH-1 (variante más virulenta) como por VIH-2 (variante menos virulenta) en distintos grupos humanos de estudio (Trachtenberg & Erlich, 2001). Un estudio reciente ha identificado más de 300 polimorfismo de una sola base (SNP, del inglés *Single Nucleotide Polymorphism*) en la región MHC asociados a la acción del virus VIH-1, sosteniendo que son determinados cambios en aminoácidos en la valva de unión al antígeno de la molécula HLA-B los que dan a los alelos implicados carácter protector o de riesgo (International HIV Controllers Study, 2010).

8.3 HLA y farmacogenética

La farmacogenética estudia el efecto que tiene la variabilidad genética de un individuo sobre la respuesta (adversa o no) a un determinado fármaco. Debido a que un porcentaje considerable de las reacciones adversas a los fármacos (ADRs, del inglés *Adverse Drug Reactions*) están relacionados con reacciones inmunoalérgicas, ya desde los años 1970 se propuso que algunas de estas ADRs podrían estar asociadas a determinadas variantes HLA. Pero fue a partir del año 2002, momento en el cual dos grupos independientes observaron que la hipersensibilidad causada por el fármaco abacavir (inhibidor de la transcriptasa inversa) se restringía a los portadores del alelo HLA-B*57:01 (Hetherington et al., 2002; Mallal et al., 2002), cuando esta área de investigación comenzó a dar resultados prometedores. A partir de entonces se describieron otras asociaciones entre alelos o antígenos HLA y el desarrollo de ADRs por fármacos, permitiendo así la predicción precoz de estas manifestaciones en función de que los individuos sean portadores o no de ciertas variantes alélicas.

La hipótesis más aceptada para explicar las reacciones inmunoalérgicas mediadas por HLA es la teoría del hapteno/prohapteno (hapteno es una molécula de pequeño tamaño que al unirse con otra más grande puede resultar inmunogénica). Según esta teoría, el fármaco (o un metabolito del mismo) sería procesado por una APC y se comportaría como un hapteno que al conjugarse con una molécula HLA de clase I o de clase II desencadenaría una respuesta inmune mediante su presentación como antígeno, siguiendo las vías ya explicadas. Según esta teoría solamente determinadas variantes HLA reconocerían estos fármacos como haptenos (Becquemont, 2010).

TABLA 15. Factores de riesgo HLA para efectos adversos a fármacos.

Fármaco	Uso	Efecto adverso	Antígeno o alelo HLA asociado
Abacavir	Inhibidor de la Transcriptasa (VIH-1)	SHS ^{c,d}	B*57:01
Alopurinol	Gota e Hiperuricemia	SJS ^{a,b,c}	B*58:01
Aminopenicilina	Antibiótico	SHS retardada ^c	A2, DR52
Aspirina	Analgésico	Urticaria ^a	DRB1*13:02- DQB1*06:09- DPB1*02:01
		Asma ^a	DPB1*03:01
Carbamazepina	Anticonvulsivo	SJS	A*24:02 ^a /A*31:01 ^{a,c} B*15:02 ^{a,b} /B*15:11 ^a B*15:18 ^a /B*59:01 ^a
		Rush cutáneo	A*31:01 ^{a,c}
		SHS	A*31:01 ^{a,c} A11 ^a , B51 ^a
Clozapina	Antipsicótico	Neutropenia ^c	DRB5*02:01
		Agranulocitosis ^c	B38, DR4, DQ3
Co-amoxiclav	Combinación de antibióticos	Daño hepático ^c	A*02:01/DRB1*15:01 DRB1*15:01- DQB1*06:02
Efavirenz	Inhibidor de la Transcriptasa (VIH-1)	Rush cutáneo ^c	DR1
Fenitoína	Antiepiléptico	SJS ^a	B*15:02/B13:01 C*08:01/DRB1*16:02
Feprazona	AINE	Erupción por fármaco ^c	B22
Flucloxacilina	Antibiótico betalactámico	Daño hepático ^c	B*57:01 DRB1*01:07- DQB1*01:03
Hidralazina	Antihipertensivo	LEIF ^c	DR4
Lamotrigina	Antiepiléptico	SJS/NET ^c	B*38:01 DRB1*13:01/DQB1*06:09

TABLA 15. Continuación.

Lapatinib	Quimioterapia	Daño hepático	DRB1*07:01- DQA2*02:01- DQB1*02:02
Levamisol	Antihelmíntico	Agranulocitosis	B27
Metazolamida	Glaucoma	SJS ^a	B*59:01/C*01:02 A*24:02-B*59:01- C*01:02
Nevirapina	Inhibidor de la Transcriptasa (VIH-1)	Rush cutáneo	B*35:05 ^b /Cw4/DR15 ^c
		SHS	B*35:05 ^{a,c} /B*58:01 DRB1*01:01/02 ^{c,d} Cw8-B14 ^{a,c}
Oxcarbazepina	Antiepiléptico	SJS ^a	B*15:02/B*15:18
		Rush cutáneo ^a	B*15:02
Oxicam	AINE	SJS/NET ^c	B*73:01
D-penicilamina	Artritis reumatoide	Proteinuria ^c	B8, DR3
		Miastenia gravis ^c	DR1
Procainamida	Antiarrítmico	LEIF ^c	DR4
Sulfametoxazol	Antibiótico	SJS/NET ^c	B*38:02
		Erupción por fármaco ^c	A30-B13-Cw6
Sulfamidas	Antibiótico	SJS ^c	A29, B12, DR7
Ticlopidina	Antiplaquetario	Daño hepático ^a	A*33:03/B*44:03 DRB1*13:02/DQB1*06:04
Ximelagatran	Anticoagulante	Daño hepático ^c	DRB1*07:01

SJS: síndrome de Stevens-Johnson (consistente en erupciones eritematosas múltiples); SHS: síndrome de hipersensibilidad (consistente en fiebre, lesiones cutáneas como rash cutáneo y daños en órganos tales como hepatitis); NET: necrolisis epidérmica tóxica (consistente en ampollas y lesiones exfoliativas de la piel); AINE: antiinflamatorio no esteroideo; LEIF: lupus eritematoso inducido por fármacos. (a) en orientales, (b) en sudasiáticos, (c) en caucasoideos, (d) en africanos (Becquemont, 2010; Gonzalez-Galarza et al., 2011; Pavlos et al., 2012).

Una de las asociaciones más estudiadas es la que relaciona las reacciones cutáneas secundarias a la toma de carbamazepina (anticonvulsivo) con la presencia del alelo HLA-B*15:02 en población asiática, pero no así en población europea. Esto puede deberse a que otro alelo en desequilibrio de ligamento con HLA-B*15:02 en asiáticos sería el verdadero responsable (Profaizer & Eckels, 2012). Otra asociación muy estudiada es la del alelo HLA-B*58:01 como posible cofactor de los efectos adversos secundarios a alopurinol. Sin embargo, estos estudios, tanto en población asiática como en europea, no han dado resultados lo suficientemente consistentes como para que la Agencia Europea del Medicamento haga una recomendación formal en relación a análisis prospectivos de este alelo (Profaizer & Eckels, 2012). En la Tabla I5 se detallan más casos de marcadores HLA que actuarían como factores de riesgo a padecer efectos adversos a determinados fármacos en función del grupo étnico al que pertenecen.

8.4 HLA y genética de poblaciones

La variación genética es un requisito fundamental para la diferenciación de las poblaciones y la evolución. Esta variación es el resultado de la mutación génica, que da lugar a nuevos alelos, y de la asociación aleatoria de estos alelos. El estudio de esta diversidad genética nos ayuda a conocer el lugar y el tiempo de origen de las poblaciones humanas, así como su patrón de dispersión por las regiones del mundo. Este estudio requiere el conocimiento de determinados marcadores genéticos de alta variabilidad, utilizando tanto datos de frecuencia de sus variantes como datos de secuencias directas. De esta forma se pueden medir las diferencias interpoblacionales e intrapoblacionales, determinar genealogías basadas en genes mediante análisis filogenéticos y extraer conclusiones sobre patrones de dispersión de esos marcadores. Entre los marcadores genéticos empleados, tales como el ADN mitocondrial (ADNmt) o el cromosoma Y, ambos marcadores de linaje, se encuentran el sistema HLA.

En el sistema HLA, la tasa y el número de sustituciones de nucleótidos que generan nuevos alelos, tanto para la clase I como para la clase II, es bastante elevada si la comparamos con la del ADNmt y del cromosoma Y. Además, los loci HLA de clase I y clase II presentan una elevada heterocigosidad (80-90%) (Panneerchelvam & Norazmi, 2012), lo que los convierte en unos marcadores óptimos para los estudios filogenéticos. Algunos de los linajes del sistema HLA, especialmente de HLA-DRB1, se retrotraen a 35 millones de años, el tiempo de divergencia entre los homínidos y el resto de primates. La diversidad haplotípica, las frecuencias alélicas y los análisis de

desequilibrio de ligamiento entre los loci HLA permiten descifrar la relación entre las poblaciones, así como conocer la cantidad de variación acumulada entre dos poblaciones relacionadas. El análisis de la variación de los loci HLA de clase I y de clase II aumenta la capacidad de analizar la estructura genética y la historia de las poblaciones ya que cada locus puede tener una trayectoria independiente según los eventos de mutación, deriva genética y migración ocurridos.

Las investigaciones relacionadas con el estudio del perfil genético HLA de las poblaciones se llevan desarrollando durante los 40 últimos años, desde los trabajos de Degos y Dausset sobre el desequilibrio de ligamento entre genes HLA en procesos migratorios con mezcla de poblaciones (Degos & Dausset, 1974). Los resultados obtenidos son de vital importancia puesto que son aplicados al desarrollo de programas nacionales de trasplante (compatibilidad HLA) y al estudio de enfermedades infecciosas en las poblaciones de estudio, ya que como hemos dicho anteriormente, el sistema HLA, debido a su elevado grado de polimorfismo y desequilibrio de ligamento, es una herramienta muy útil para poder diferenciar poblaciones o grupos étnicos. Mediante el conocimiento de los perfiles genéticos HLA se pueden establecer relaciones genéticas entre poblaciones e inferir en el origen y en los movimientos migratorios de estas mediante el apoyo de otras disciplinas científicas, como la antropología, la lingüística, la arqueología o la historia.

Desde el 5° Taller Internacional de Histocompatibilidad, en el que se analizaron los perfiles genéticos HLA de 49 poblaciones del mundo, estos encuentros de expertos han servido para avanzar en el análisis antropológico de las poblaciones humanas mediante este marcador, en especial en los talleres del 11° al 15°, gracias a la comparación de los datos de más de 100 poblaciones de todo el mundo cubriendo todas las regiones geográficas procedentes de los laboratorios de inmunogenética participantes; para la estandarización de los métodos de obtención de perfiles genéticos HLA de clase II y posteriormente de clase I; y para estandarizar los análisis estadísticos para el estudio de las relaciones. Gracias a esta estandarización y al interés creciente en esta disciplina promovido por estos talleres, se han desarrollado gran cantidad de estudios poblacionales, aunque no todos rinden resultados con el mismo nivel de caracterización (serológico y genético) y son, por tanto, difíciles de comparar.

Las frecuencias de los alelos HLA correlacionan con la existencia o ausencia de flujo genético entre poblaciones y este puede ser “medido”, puesto que sus características HLA no se han diluido con el tiempo (si tenemos en cuenta la selección

equilibradora como generadora de su polimorfismo) mediante el estudio de sus correspondientes distancias genéticas y la presencia de haplotipos HLA con fuerte desequilibrio de ligamiento (Imanishi et al., 1992a; Imanishi et al., 1992c; Clayton & Lonjou, 1997), aunque cuanto más lejana sea la relación en el tiempo entre las poblaciones se acumula mayor variación en las frecuencias de haplotipos, debido a las variaciones en el desequilibrio de ligamiento y la formación de nuevas combinaciones alélicas (Bodmer & Bodmer, 1978). Las poblaciones orientales, caucasoides y negroides muestran alelos HLA y diferencias en sus frecuencias que sirven como marcadores genéticos poblacionales.

Otro ejemplo, son las poblaciones nativas americanas (Amerindios) que poseen un restringido espectro de alelos HLA (v. g.: HLA-A*24, -A*68, -B*15, -B*35, -B*39, -B*40) respecto a las otras poblaciones del mundo, con frecuencias alélicas muy diferentes a las otras poblaciones del mundo, y con determinados subtipos de alelos específicos o cuasi-específicos de estos grupos, principalmente para HLA-B (Gomez-Casado et al., 1997; Gomez-Casado et al., 1999). Esto sugiere una evolución diferencial del sistema HLA, que podría estar asociada a su aislamiento geográfico desde la colonización del continente americano, a su estructura poblacional y/o a una fuerte selección por exposición a nuevos patógenos encontrados durante las migraciones o llevados por los nuevos colonizadores después de 1492 (Parham et al., 1995; Gomez-Casado et al., 1997; Fernandez-Vina et al., 1997; Erlich et al., 1997).

Por otra parte, el estudio genético HLA de los diferentes grupos étnicos permite analizar aquellos genes que en las diferentes poblaciones están participando en la fisiopatología de los procesos infecciosos y de enfermedades autoinmunes, lo cual favorece la inserción de programas de medicina preventiva y programas de vigilancia epidemiológica para el control de estas patologías.

El elevado polimorfismo, el estrecho ligamiento, la asociación aleatoria de alelos y la perpetuación de linajes alélicos a lo largo del tiempo hacen del sistema HLA una herramienta de gran valor para deducir el pasado de los humanos. La cantidad y la distribución de la variación de este sistema en las diferentes poblaciones del mundo nos permiten correlacionar los perfiles genéticos de estas con la geografía y el posible patrón de migraciones humanas con el objetivo de determinar su origen.

II. POBLAMIENTO Y POBLACIONES DE LA REGIÓN DEL MAR CASPIO

9. LA REGIÓN DEL MAR CASPIO

El Mar Caspio es la masa de agua interior más grande sobre la superficie terrestre, situándose entre la región sur oriental del continente europeo y la región sur occidental del continente asiático, entre las latitudes 47.07° y 36.33°N y las longitudes 45.43° y 54.20°E. Aproximadamente tiene una longitud de 1030 km, unas dimensiones este a oeste que van desde los 196 Km. en su parte más estrecha a los 435 km en su parte más ancha y una profundidad máxima de 1025 m. Al tratarse de un lago no tiene conexión con aguas oceánicas, está alimentado por 130 ríos, entre los que se encuentra el Volga (el más importante y que aporta el mayor volumen de agua), el Ural y el Kurá, y se encuentra a 26.5 m por debajo del nivel del mar.

Las aguas del Mar Caspio son ligeramente saladas (tres veces menos salinas que las de los océanos), debido a que es un remanente del océano Tetis de la Era Mesozoica, que conectaba los océanos Atlántico y Pacífico hace 60-50 millones de años. Desde hace 5.5 millones de años, cuando Tetis desapareció por la deriva continental, este mar interior no tiene salida al océano.

El nombre de este lago salado deriva del nombre del pueblo Caspio, que habitaba en la orilla suroeste junto a la desembocadura del río Kurá, mencionado por Herodoto como una de las provincias del Imperio Aqueménida del rey Darío I el Grande (Herodoto, 1989). Estrabón también los menciona en su Geografía como una tribu desaparecida que poblaba el territorio al que dio nombre, Caspiane, perteneciente a la Albania Caucásica (Estrabón, 1991c). Fueron los romanos los que extendieron el uso de este término para nombrarlo.

Sin embargo, en el periodo clásico, este mar interior recibió el nombre de océano Hircanio, en referencia a Hircania, la región histórica, parte del Imperio Aqueménida, situada en la costa sur oriental del lago (parte de las actuales Irán y Turkmenistán). No obstante, a lo largo de la historia, ha recibido multitud de nombres dependiendo de las lenguas o dialectos de las numerosos grupos poblacionales que habitaron y habitan sus orillas: los turcos se refieren a este mar como Khazar, los turcomanos como Hazar, los árabes como Bahr-e (mar) Qazvin, los iraníes farsi, tanto en la antigüedad como ahora, como Mazandarán o Fars Bahr or Hoz (Mar Persa), los habitantes del Golestán como Jorjan (Gorgan), los rusos como Khavilinsk, y en avéstico (antigua lengua persa) como

Farakhkart (el gran mar), así hasta más de 35 nombres diferentes registrados (Ajam, 2013).

Desde la disolución de la Unión Soviética, las orillas del Mar Caspio son frontera natural de los países de Azerbaiyán, Irán (provincias de Guilán, Mazandarán y Golestán), Turkmenistán (provincia de Balcan), Kazajstán y Rusia (Repúblicas de Daguestán y Kalmukia, y Óblast de Astraján, Figura I13).



Figura I13. Mapa de la región del Mar Caspio. Incluye las poblaciones iraníes donde se han tomado las muestras.

10. ANTROPOLOGÍA Y MOVIMIENTOS POBLACIONALES

Los primeros restos de homínidos encontrados en las proximidades del Mar Caspio proceden del yacimiento arqueológico de Dmanisi, en Georgia, y corresponden a

la subespecie que hoy se conoce como *Homo erectus georgicus* (anteriormente *Homo georgicus*), descubiertos entre 1991 y 2005. Estos restos están datados en 1.8 millones de años de antigüedad, siendo hasta la fecha los primeros representantes de esta especie hallados fuera de África (Gabunia et al., 2000).

Se han encontrado restos pertenecientes a un individuo pre-Neandertal en la cueva de Azykh, en el territorio de Nagorno Karabaj, en Azerbaiyán. Se trata de un hueso maxilar del que se conoce como Hombre de Azykh, datado en 300.000 años de antigüedad, y que por tanto sería uno de los restos más antiguos de esta especie en esta región (Kasimova, 2001).

Existen evidencias de la presencia de poblaciones humanas en la región del Mar Caspio ya en el paleolítico inferior (desde 700.000 años de antigüedad), basadas en el hallazgo de útiles líticos pertenecientes a la cultura pre-achelense (en la cueva de Azikh) y achelense (cueva de Darband, en la provincia de Guilán, Irán). Sin embargo, el hallazgo más significativo es el descubrimiento de restos óseos pertenecientes a tres hombres en la cueva de Huto, próxima a la ciudad de Behshahr, en la provincia de Mazandarán, Irán, que datan de hace más de 11.000 años, de los que se supone que eran cazadores de focas, ya que se hallaron junto con hojas de pedernal y huesos de estos animales (Coon, 1952). La presencia de restos óseos de perro (*Canis lupus familiaris*) datados en el mismo periodo suponen una de las primeras evidencias de la domesticación de este animal por el hombre (Reed, 1959).

10.1 Movimientos prehistóricos: Culturas de los Kurganes

En 1956, la arqueóloga Marija Gimbutas elaboró la hipótesis de los kurganes, combinando estudios arqueológicos y lingüísticos basándose en el conjunto de culturas prehistóricas Kurgan, para dilucidar el origen de los pueblos protoindoeuropeos, en la que tiene especial relevancia la región del Mar Caspio. Esta cultura recibe su nombre del término ruso, tomado del turco, que sirve para denominar túmulos funerarios, método de enterramiento común a los pueblos que se supone pertenecen a esta cultura (Gimbutas, 1977; Gimbutas, 1980).

Esta hipótesis, con sucesivas variaciones, sugiere una expansión gradual de los pueblos que desarrollaban los kurganes como principal motivo cultural, desde la región entre el Dnieper (actual Ucrania) y el medio y bajo Volga, hasta abarcar todas las estepas del Mar Negro, Balcanes y Europa Central hacia el este y Mar Caspio (estepas

pónticas), Anatolia, Irán, Asia Central e Indostán hacia el suroeste, en varias y sucesivas oleadas, desde el V milenio a. C.

La cultura Yamna o Iamna (en ruso “hoyo”), que data de la Edad de Cobre tardía y de la Edad de Bronce temprana (3200-2500 a. C.), es la primera cultura conocida en llevar a cabo estos túmulos funerarios, con sepulcros tipo hoyo (de ahí su nombre) en los que se introducían los cuerpos en posición de decúbito supino con la rodillas dobladas. Esta cultura era predominantemente nómada, aunque pudo practicar la agricultura en las regiones próximas a los cauces de los ríos donde llegó a agruparse en fortificaciones. Un factor importante que evidencia la expansión de esta cultura es la aparición de restos de ruedas e incluso de carros enteros en algunos enterramientos, como el kurgán Storozhova Mohyla (Dnepropetrovsk, Ucrania). El carro se ha visto tradicionalmente como uno de los requisitos previos para una explotación de la estepa abierta, ya que era una propiedad necesaria para una economía móvil y nómada. Esta cultura desarrolló principalmente herramientas y armas de piedra, cerámica con decoración impresa de cuerdas o a peine y objetos de cobre (Mallory, 1997c).

La cultura Yamna daría paso a la cultura de Andrónovo (por la localidad rusa del mismo nombre donde se hallaron los primeros túmulos), cultura inicialmente nómada y más tarde practicante de la agricultura y ganadería sedentaria y trashumante (principalmente equina), que data de la Edad de Bronce tardía (2300-1000 a. C.) y que reúne una serie de culturas indo-iránicas que se extendieron de este a oeste desde los Urales hasta los montes Altai y de norte a sur desde las estepas de Asia Central y Siberia hasta los actuales Kazajistán y Tayikistán (Mallory, 1997a). Está muy relacionada con el contemporáneo Complejo Arqueológico Bactria-Margiana, que es como se denomina a las culturas de la Edad de Bronce de Asia Central (actuales Turkmenistán, Afganistán y Uzbekistán), predominantemente sedentarias y con una elaborada cultura material (incluye arquitectura monumental, herramientas de bronce, cerámica y joyería), que podría suponer otra rama indo-iránica separada de los protoindoeuropeos (Mallory, 1997b). La cultura de Andrónovo daría lugar a otras culturas indo-iránicas que seguirían llevando a cabo la práctica de los túmulos kurganes, situadas en las orillas del Mar Negro y del Mar Caspio, tales como la Cultura de Srubna (del ruso *sруб*, madera, ya que de este material construían los túmulos) del 1500 al 800 a. C., y la cultura escita del 700 a. C., ya en la Edad de Hierro.

10.2 Poblamiento y culturas de la región sur del Mar Caspio

A lo largo de los milenios III y II a. C. se desarrollaron dos culturas bien diferenciadas de pueblos anteriores a los indo-iranios, en la región sur del Mar Caspio (actual orilla iraní de este lago), la cultura de Gorgán, al sureste, y la cultura Yanik, al suroeste (Bermejo, 2013).

La cultura de Gorgán se extendió por la actual provincia iraní del Golestán, llegando incluso a la meseta central (yacimiento de Tappeh Hesar), desarrollando principalmente vasijas de arcilla adornadas con líneas en relieve entrecruzadas, objetos de mármol, herramientas y armas de cobre, y orfebrería. Esta cultura estaba ubicada en medio de la ruta comercial que unía Oriente con Occidente, por lo que mantenían contacto con las culturas procedentes del valle del Indo al este, así como con las de Anatolia al oeste, como demuestran sus manufacturas. La cultura de Gorgán desapareció a lo largo del II milenio a. C. debido a las incursiones de los pueblos indo-iranios procedentes de Asia Central (posiblemente de la cultura de Andrónovo o del Complejo Arqueológico Bactria-Margiana) y desde la parte más oriental, ya que las culturas de los yacimientos en esa zona fueron las primeras en desaparecer (Bermejo, 2013).

A mediados del III milenio a. C. pueblos procedentes de Anatolia y del Cáucaso se establecieron en la actual Azerbaiyán y la región noroeste de Irán, desarrollando la cultura Yanik (por el principal yacimiento Yanik Tappeh, próximo a la ciudad de Tabriz). Los vestigios de esta cultura comprenden vasijas de arcilla más toscas que las de la cultura de Gorgán, con adornos geométricos en relieve sustituidos en época posterior por marcas de fuego, y restos de edificios de planta circular. Algunos autores sugieren que esta cultura pudo ser antecesora de los pueblos hurritas, que se extendieron por los actuales Irán, Irak y Turquía (Abdi, 2012).

Estas culturas, principalmente la de Gorgán, desaparecieron con la llegada de los pueblos iraníes, procedentes de las tribus indo-iranias de Asia Central, que colonizaron la región sur del Mar Caspio, entre otras, y pasó a conocerse como Hircania (topónimo griego, calco de la voz Verkâna del persa antiguo, que significa “Tierra de lobos”). Estos pueblos conformaron los sucesivos imperios persas iraníes: el Imperio Aqueménida, Imperio Seléucida, Imperio Parto e Imperio Sasánida (siglo VI a. C – siglo VII d. C). Hircania formó parte de estos imperios como una satrapía (provincia de los imperios persas, gobernada por un sátrapa), hasta la conquista árabe de Persia en el

siglo VIII, a partir de la cual se constituyó como un estado zoroástrico independiente, volviendo a formar parte del Imperio Persa a partir el siglo XV.

Ya en los tiempos de los sasánidas el topónimo Gorgan o Gurgan ya se utilizaba para designar tanto la provincia como la ciudad, como se deduce de las series de monedas encontradas en los yacimientos originarios de este periodo. No obstante, algunos autores suponen que la capital de la provincia sasánida se ubicó originariamente en el yacimiento arqueológico islámico de Gorgán, próximo a la actual ciudad Gonbad-e Qabus, aunque faltan evidencias directas sobre esto. De ser así, este topónimo se habría transferido a la capital actual, a 80 km al sureste, conocida anteriormente como Astarabad o Asterabad (Kiani, 2002). Sin embargo la capital de la antigua Hircania, sobre la que marchó Alejandro Magno cuando conquistó Persia a los aqueménidas, era Zadracarta. Algunos autores la han identificado con la Gorgán actual, sin embargo, es mucha más plausible que su localización se corresponda con el yacimiento de Qal'a-ye Kandān, un tepe (“montículo”, palabra turca que sirve para designar yacimientos arqueológicos en Oriente Próximo y Persia) localizado al sudoeste de la capital (Kiani, 2002).

La Gran Ruta de la Seda originada durante el siglo II a. C. debido a las relaciones comerciales establecidas entre la China de la dinastía Han con las civilizaciones de Asia Central y Oriente Próximo, entre las que se encontraba el Imperio Parto y que seguía rutas ya utilizadas desde el paleolítico, permitió no solo el comercio de este material y de muchos otros bienes (piedras y metales preciosos, porcelana, tejidos, vidrio, especias, etc.) sino también el transporte de personas y el desarrollo de una red de difusión cultural y religiosa entre los distintos pueblos por los que transitaba hasta su desintegración en el siglo XV. Esta ruta supuso el principal eje de conexión Oriente-Occidente, cuando el Imperio Romano en el siglo I a. C. comenzó a interesarse por la seda, además de otros productos asiáticos, al adquirirla de los partos, con los que ya habían establecido relaciones comerciales (Estrabón, 1991a). De la misma manera que durante el florecimiento de la cultura de Gorgán, la región sur del Mar Caspio se ubicaba en el centro de esta ruta comercial, y las ciudades que allí se encontraban (como la antigua Hecatompileos, capital del Imperio Parto próxima al Golestán, o la actual Tabriz) actuaron como grandes centros de intercambio cultural y de mercancías en diferentes épocas. Esta región fue arrasada varias veces y su población masacrada por las incursiones del pueblo mongol durante la conquista por parte de este imperio, provocando su devastación y la pérdida de su influencia geopolítica (Bosworth, 2002).

10.2.1 Poblamiento del Golestán: turcomanos en Irán

Contemporáneamente a la pérdida de influencia de toda esta región en las rutas comerciales por la desintegración de la Ruta de la Seda, la región sureste iraní del Mar Caspio (Golestán) comienza a sufrir la migración de tribus de origen túrquico. Se trata de los turcomanos o turkmenos, pueblo que habla una lengua túrquica de la rama oghuz oriental (el término oghuz deriva del nombre de su ancestro común, Oghuz Khan), el turcomano o turkmeno, y que originariamente procedían de la región Transoxiana en Asia Central (Gonem, 1996). Hay varias teorías sobre el significado de turcomano. De acuerdo con la etimología popular esta palabra deriva de “*turco*” más el vocablo iraní “*manand*”, y significaría “que parecen turcos”. Según esta hipótesis la palabra originaria para designar a este pueblo y su lengua sería *turkmanand*, que derivó en *turkmen* o *turcomano*. Sin embargo, otros autores han propuesto que el elemento “*-man/-men*” actúa como un intensificador, por lo que la palabra se traduciría por “más turco que los turcos” (Sabery, 2008). Se cree que los turcomanos originariamente practicaban el shamanismo, como otros pueblos de Asia Central, pero fueron islamizados en el siglo XVI, y actualmente la mayoría son musulmanes suníes (Askarikhaneha & Sharifkamali, 1995; Sabery, 2008).

Las primeras migraciones turcomanas tuvieron lugar en el siglo X d. C., produciéndose los primeros asentamientos turcomanos en la región con la conquista de Persia por parte de los selyúcidas, dinastía de origen túrquico oghuz procedente del Mar de Aral (Askarikhaneha & Sharifkamali, 1995; Luther, 2001). Otros autores argumentan que los turcomanos, descendientes de tribus oghuz no selyúcidas, no tomaron parte en estas migraciones, y fue a partir del siglo XV d. C. cuando se establecieron en el Jorosán iraní (Johanson, 2006). Estas migraciones se desarrollaron inicialmente de manera pacífica, siendo vasallos del Imperio Safávida, y más tarde mediante incursiones violentas. A partir del siglo XVII estos grupos constituyeron la base del poder de la dinastía Qajar de origen turco asentada en Astarabad (Gorgán), que gobernó Persia desde 1785 hasta 1921 (Bosworth, 2002). La situación en esta región durante este periodo estuvo marcada por la inestabilidad debido no solo a las incursiones de los turcomanos y de uzbekos para hacerse con el potencial económico de esta región, sino por las guerras y revueltas de las tribus ya asentadas en épocas anteriores y los intereses de otras potencias en ampliar sus territorios, como el Imperio Ruso, desde que Pedro el Grande aprovechara la decadencia de los safávidas para establecer un puesto comercial por un breve periodo de tiempo. La instauración de la

dinastía Pahlavi en el gobierno persa en 1921 conllevó la estabilización de la región, mediante, entre otras cosas, el desarme y la reorganización de las tribus turcomanas (Nejestani and Elr, 2002).

Actualmente un porcentaje considerable de la región del Golestán está compuesta por turcomanos, sin embargo no hay datos fiables de su número total, ni de su distribución, aunque se estima que la población turcomana en Irán asciende a 1.5 millones (2% de la población iraní), encontrándose la mayoría en esta región (Askarikhaneh & Sharifkamali, 1995).

Los turcomanos del Golestán se asientan en lo que ellos denominan Türkmenşähra (“llanura de los turcomanos”, zona norte del Golestán, Figura I15) y pertenecen a dos tribus principalmente: los Yomut y los Goklan, que aún conservan su estructura tribal. Los primeros constituyen el grupo mayoritario y están divididos en dos subgrupos formados cada uno por doce clanes y su territorio se extiende desde la ciudad de Gorgán hacia el norte con la frontera de Turkmenistán y hacia el este hasta la ciudad de Bandar Torkaman, a orillas del Mar Caspio. Los segundos tienen una estructura tribal similar a los Yomut y su territorio se centra en la ciudad de Gonbad-e Kavus, y se extiende hacia el norte hasta el río Atrak en la frontera con Turkmenistán y hacia el oeste hasta el desierto Yomut (Zanjani, 2002). Existe una tercera tribu turcomana, los Takeh, pero es minoritaria (solo el 3% de los turcomanos iraníes) y se encuentran fundamentalmente en la vecina provincia de Jorosán del Norte (Sabery, 2008).

Además de los turcomanos y de la población de origen persa, que es mayoritaria, otros grupos étnicos que se ubican en Golestán son:

- Azeríes: proceden originariamente de la región noreste de Irán (Azerbaiján Oriental y Azerbaiján Occidental), hablan azerí y son de religión musulmana chiíta. La mayoría de los azeríes de esta región habitan en la zona sur formando incluso territorios étnicos (véase apartado 10.2.2; Sabery, 2008).
- Sistaníes: grupo étnico procedente de Sistán en el sudeste de Irán. Hablan sistaní, dialecto del persa, y son de religión musulmana chiíta. La mayoría habita en zonas rurales del sur y del centro de la región, aunque algunas minorías lo hacen en las ciudades (Sabery, 2008).
- Baluchis: grupo étnico originario de Baluchistán al sudeste de Irán. Hablan baluchi, una lengua irania occidental del norte y al igual que los turcomanos son principalmente musulmanes suníes. La mayoría habita en zonas rurales del sur y del centro de la región, como los sistaníes (Sabery, 2008).

10.2.2 Poblamiento del Azerbaiyán Iraní

El topónimo Azerbaiyán deriva del nombre Atropates, sátrapa de la provincia de Media (región comprendida entre el Mar Caspio y Mesopotamia, donde se localizan el actual estado de Azerbaiján y la región iraní del mismo nombre) al final del Imperio Aqueménida y durante y después del Imperio de Alejandro Magno, mencionado ya por historiadores clásicos como Diodoro Sículo o Estrabón (Estrabón, 1991b; Diodoro Sículo, 2001). Atropates es la traducción en griego clásico del nombre compuesto en persa antiguo que significaría “protegido por el fuego sagrado” (en relación al ritual del fuego de la religión mazdeísta o zoroastrismo, religión oficial de los imperios persas preislámicos y de la que se supone que su origen está en esta región), y que derivó en persa medio a Aturpatakan, posteriormente a Adurbadagan, y a Adarbaiyan en persa moderno (Schippman, 1987).

El origen del pueblo azerí o azerbaijano, que habita mayoritariamente en el estado de Azerbaiyán y en las provincias azerbaijanas de Irán, ha sido profusamente discutido. Algunos académicos hablan del origen caucásico de este grupo étnico, y proponen evidencias de que, a pesar de las repetidas invasiones y migraciones, existe un sustrato de antiguos pueblos del Cáucaso, como los Mannai, que habitaban principalmente el noroeste del actual Irán en el I milenio a. C., y que fueron asimilados por los pueblos iraníes procedentes de Asia Central (medos y persas; Encyclopaedia Britannica, 2014). Estudios basados en datos históricos, lingüísticos, arqueológicos y culturales, demuestran un sustrato iraní en este grupo étnico, apoyado principalmente por la existencia de un idioma antiguo azerí de ascendencia iraní que subsistiría en algunos dialectos tati azeríes y en la toponimia de la región, por yacimientos arqueológicos y por la gran cantidad de rasgos culturales comunes (Yarshater, 1988; Paul, 1998; Planhol, 2004; Frey, 2004). La llegada de los turcos oghuz con la dinastía selyúcida a partir del siglo X d. C. conllevó un proceso de “turquificación” de toda esta región y por consiguiente del pueblo azerí, que supuso no solo una mezcla poblacional, sino también la paulatina suplantación de la antigua lengua azerí por la actual lengua azerí, de origen túrquico (altaica) y perteneciente a la rama oghuz, como el turcomano (véase apartado 10.2.1). Según la enciclopedia británica, los azeríes son un pueblo con un origen étnico mixto, cuyo elemento más antiguo proviene de las poblaciones indígenas transcaucásicas (los Mannai), y probablemente de los medas del norte de Persia, siendo “persianizados” durante la dinastía sasánida (siglos III-VII d. C.) y más tarde “turquificados” después de la conquista de la región por los turcos selyúcidas.

Con la llegada de la dinastía timúrida (siglo XIV d. C.), la región se convirtió en una provincia importante del imperio, al hacer pasar la Ruta de la Seda por Tabriz, convertida en capital de la región (actual capital de la provincia iraní Azerbaiyán Oriental), lo que hizo que surgiera de ahí la dinastía safávida, que gobernó Persia durante el siglo XVI d. C. (Frey, 2004). Tras las guerras ruso-persas, en el primer cuarto del siglo XIX, la población azerí quedó dividida en lo que hoy es el estado de Azerbaiyán y la región del Azerbaiyán iraní en el noroeste de Irán.

En la actualidad esta región se divide en cuatro provincias: Azerbaiyán Oriental, Azerbaiyán Occidental, Ardebil y Zanyán (Figura I14). En todas ellas la comunidad azerí es mayoritaria, siendo la más grande del mundo (entre 15 y 16 millones, entre el 16 y el 24 % de la población iraní, dependiendo de las fuentes consultadas) La mitad de la población es urbana residiendo en ciudades de mayoría azerí como Tabriz, Urmía, Ardebil y Zanyán (Frey, 2004; International Business Publication, 2009). La población azerí también se encuentra en otras partes de Irán como las provincias de Hamadán, Qazvín, Markazí, Kurdistán, y Guilán al noroeste y Golestán al nordeste (véase apartado 10.2.1), además existen comunidades en grandes ciudades como Teherán o Karaj (International Business Publication, 2009). Son mayoritariamente musulmanes chiítas, desde que la dinastía safávida impuso esta rama del Islam como religión oficial del imperio. Otros grupos étnicos principales que habitan en la región son:

- Kurdos: el segundo grupo étnico en número de habitantes. Se localizan fundamentalmente en la frontera de Azerbaiyán Occidental con Irak, Turquía y el Kurdistán iraní. Hablan la lengua kurda (sorani) y son musulmanes suníes. Forman comunidades tanto rurales como urbanas en esta región (véase apartado 10.2.3; Tapper, 1988; International Business Publication, 2009).
- Armenios: pueblo de origen caucásico y de lengua armenia (de origen indoeuropeo) que practican predominantemente el cristianismo gregoriano, y que a lo largo de su historia se mantuvieron como nación independiente, o formando parte de distintos imperios persas y del Imperio Otomano, entre otros. En esta región son mayoritariamente una comunidad urbana, siendo la más grande la de Tabriz (International Business Publication, 2009).
- Asirios: pueblo de origen mesopotámico y de lengua neosiriaca o neoaramea (procedente del antiguo arameo) que practican cristianismo siríaco (iglesia asiria de oriente). Históricamente descienden del antiguo Imperio Asirio que fue absorbido por los distintos imperios persas y por el Imperio Otomano. Se

localizan principalmente al oeste del lago Urmía (Azerbaiján Occidental), y la mayor comunidad se encuentra en la ciudad del mismo nombre (International Business Publication, 2009).

10.2.3 Poblamiento del Kurdistán Iraní

El Kurdistán (“tierra de los kurdos”) es una región geográfica y cultural localizada al sur del Cáucaso y al norte de la antigua Mesopotamia, en la que el pueblo kurdo es mayoritario (Figura I14). Para algunos autores el término *Kurd* que aparece en documentos del Imperio Persa islámico servía para nombrar las tribus nómadas dedicadas al pastoreo más que para designar el grupo étnico concreto (MacKenzie, 1964; Asatrian, 2009). Para otros autores estos documentos sí que reconocían el origen étnico de este término al tratase de una población reconocida anteriormente (Izady, 1992). Sin embargo, su etimología es aún discutida, aunque ha sido relacionada con diferentes topónimos, como la antigua región de *Karda* o *Qarda* al sur del lago Van (este de Turquía) recogida en textos sumerios del III milenio a. C. (Wixman, 1984), por lo que podría proceder del término troncal sumerio *kur*, que significa “montaña”, de tal forma que los kurdos serían los “habitantes de la montaña” o simplemente “montañeros” (Hennerbichler, 2012).

El origen del pueblo kurdo se remonta a las poblaciones neolíticas euroasiáticas que se asentaron en el norte de Mesopotamia, como los gutíes o los lulubíes, en el V milenio a. C (Mallory, 1989). Durante los últimos milenios a. C. estas poblaciones sufrieron un proceso de “iranianización” por los pueblos indo-iranios procedentes de Asia Central en varias oleadas distintas: una primera entre el 2200-1600 a. C. desde el norte por las estepas rusas hasta el norte de Mesopotamia y los montes Zagros, otras posteriores a través de la meseta iraní hacia el noroeste iraní dando lugar al reino de Mitanni (1500-1300 a. C.) y a los Imperios Medo y Persa (800-600 a. C.), y por último desde las estepas pónticas (escitas, 700 a. C.). Estas migraciones dieron lugar al origen multiétnico y multicultural de las poblaciones antepasadas de los kurdos, y supusieron la suplantación de las lenguas que se hablaban en la región (lenguas proto-no-iranias) por las lenguas iranianas (lenguas kurdas) (Hennerbichler, 2011; Hay, 2014).

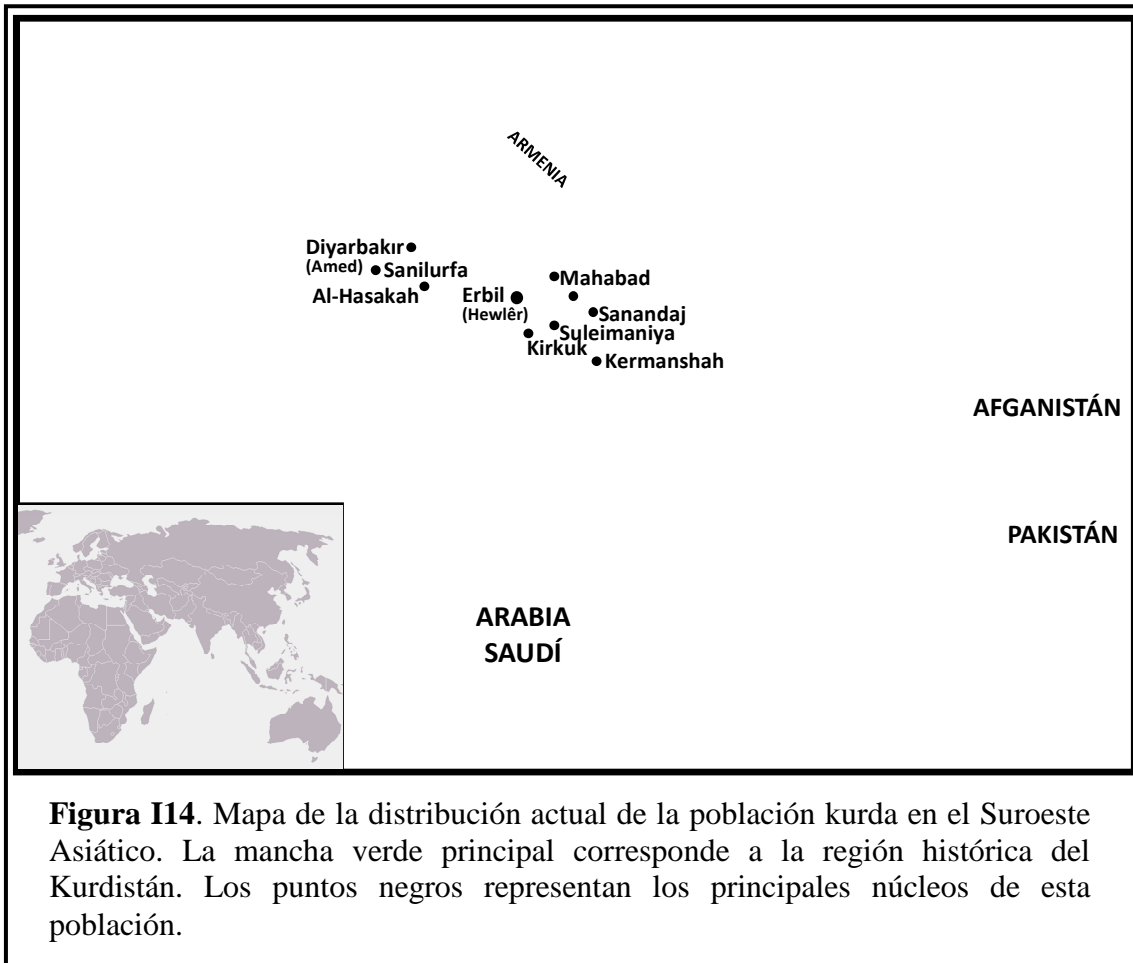


Figura I14. Mapa de la distribución actual de la población kurda en el Suroeste Asiático. La mancha verde principal corresponde a la región histórica del Kurdistán. Los puntos negros representan los principales núcleos de esta población.

El pueblo kurdo ha sufrido un destino similar al de otros pueblos del Cáucaso y Mesopotamia, como el asirio o el azerí, siendo absorbidos por los imperios persas preislámicos e islámicos, con algunos periodos de semi-independencia (como vasallos de califatos) de algunos principados kurdos, durante la baja Edad Media. Llega a tener su máximo esplendor durante la dinastía ayubí iniciada por Salah al-Din Yusuf (Saladino) en el siglo XII d. C., de origen kurdo, que entre otros muchos territorios gobernaba el sur del Kurdistán. En la Edad Moderna, el Kurdistán queda dividido entre el Imperio Safávida persa y el Imperio Otomano turco, lo que supuso una serie de rebeliones y levantamientos, primero contra los persas, y más tarde contra los otomanos, ya que estos les permitieron conservar su relativa autonomía hasta el siglo XIX d. C. Actualmente el Kurdistán ocupa la parte este de Turquía, el noreste de Siria, el norte de Irak y el noroeste de Irán (Figura I14).

La parte iraní del Kurdistán esta formada por las provincias occidentales del norte de Irán donde la población kurda es mayoritaria: Kurdistán, Kermanshah y partes de las provincias de Ilam y Azerbaiján Occidental (Figura I15). Hay entre 4 y 6.5 millones de

kurdos en Irán, siendo el tercer grupo étnico en el país por número de individuos, después de persas y azeríes (9-10 % del total de la población iraní). Forman comunidades tanto urbanas (en ciudades kurdas como Sanandaj, Mahabad, Saqqez o Baneh) como rurales, quedando algunos grupos que practican el nomadismo. Muchas de las comunidades rurales mantiene una estructura tribal de organización social. La población kurda iraní también reside en otras partes del país como las provincias de Fars, Kermán, y Sistán y Baluchistán, aunque el grupo más numeroso fuera del Kurdistán es el que se encuentra en Jorosán del Norte, en la frontera con Turkmenistán. Se trata de los descendientes de los que fueron trasladados por la fuerza desde el Kurdistán iraní durante el Imperio Safávida para defender la frontera de las incursiones de uzbekos y kazajos (siglo XVII d. C.). Los kurdos son mayoritariamente musulmanes suníes, aunque en algunas comunidades hay más diversidad religiosa que en otras, como en área del sur del Kurdistán, donde practican la rama chiíta del Islam (International Business Publication, 2009). Otros grupos étnicos que habitan en la región son:

- Luros: pueblo procedente de los montes Zagros (Lorestán) que habla una lengua irania occidental del sur, el lorí, relacionada con el kurdo y el persa. Son musulmanes chiítas. En esta región son generalmente una comunidad urbana, aunque tradicionalmente eran pastores nómadas (International Business Publication, 2009). En la región también habitan poblaciones persas y azeríes.

10.2.4 Otros grupos étnicos de la región sur del Mar Caspio

Talyshi

Pueblo iranio que habita una región compartida por los actuales estados de Azerbaiján e Irán, en la costa suroeste del Mar Caspio. Hablan el idioma talyshi, que pertenece a la rama occidental de las lenguas iránicas (Lewis et al., 2014), y que al igual que los dialectos tati, se supone que procede del antiguo azerí que se hablaba en la región antes de su “turquificación”. Se cree que su origen es mixto, al proceder de la “persianización” de los Cadusii, antigua tribu originaria del Mar Caspio, aunque es difícil establecerlo (Asatrian & Borjian, 2005). Los estudios de marcadores genéticos poblacionales parecen confirmar el origen caucásico de esta población (Nasidze et al., 2009). En Irán, los Talyshi se encuentran en las provincias de Guilán (al norte, en la región de Talesh) y Ardabil (Figura I15), y se les considera una minoría étnica. Son musulmanes, tanto suníes como chiítas.



Guilakíes y Mazandaraníes

Pueblos iraníes que habitan en la costa sur del Mar Caspio, los primeros en la provincia iraní de Guilán y los segundos en la provincia de Mazandarán (Figura I15). Los montes Elburz marcan el límite meridional de estos pueblos. Hablan las lenguas guilakí y mazandaraní, respectivamente, ambas estrechamente relacionadas entre sí y pertenecientes a la rama occidental de las lenguas iraníes, dentro de lo que algunos autores denominan lenguas del Caspio (Lewis et al., 2014). Se piensa, como en el caso

de los Talyshi, que tienen un origen mixto al estar relacionados genéticamente tanto con poblaciones del sur del Cáucaso (armenios y azeríes) como con sus grupos vecinos de origen persa (Nasidze et al., 2006). Ambos grupos tienen una población en torno a los tres millones de individuos. Son musulmanes, mayoritariamente chiítas.

11. ASPECTOS ANTROPOLÓGICOS Y LINGÜÍSTICOS DE LAS POBLACIONES DE LA REGIÓN DEL MAR CASPIO

Los aspectos lingüísticos han sido utilizados para el estudio del origen de las poblaciones que hoy ocupan el “corredor iraní” que comunica Europa con Asia, ya que la propagación y difusión de las lenguas va acompañada de fenómenos demográficos, como movimientos poblaciones, invasiones y migraciones. Esto supondría fenómenos de sustitución e implantación lingüística y/o cultural por los nuevos pobladores, que podría o no conllevar una aportación genética (véase apartado 12.3). Las lenguas que se hablan en esta región pertenecen fundamentalmente a dos grandes familias lingüísticas, la indoeuropea (farsi o persa, kurdo, armenio, baluchi, lorí, talyshi, mazandaraní, guilakí) y la altaica (azerí, turcomano, turco de Jorasán, khalaj).

11.1 Lenguas indoeuropeas: las lenguas kurdas

Los pueblos que habitaban la región desde Anatolia hasta la meseta iraní, pasando por Mesopotamia, ya en el III milenio a.C., tales como los hurritas, hablaron probablemente lenguas relacionadas con las lenguas caucásicas modernas, como las lenguas caucásicas septentrionales (lenguas nakh y daguestaníes) (Arnaiz-Villena & Alonso-García, 2001). De hecho, algunos autores sostienen que las ya extintas lenguas hurrita-urartianas (las lenguas que hablaban los hurritas de los reinos Mitanni y Urarto respectivamente) pertenecen junto con las lenguas caucásicas a la macrofamilia alarodiana al tener un alineamiento morfosintáctico similar de tipo ergativo (Diakonoff & Starostin, 1986). Otros autores incluyen estas dos familias de lenguas en la macrofamilia dené-caucásica, que incluiría lenguas como el vasco, las lenguas sino-tibetanas o las lenguas na-dené (Ruhlen, 1994). Los hurritas, así como otros grupos mesopotámicos (los gutíes o los lulubíes, véase apartado 10.2.3), son considerados por varios investigadores como antepasados del pueblo kurdo, cuyo territorio actual se corresponde con la distribución de estas poblaciones. Sin embargo, los kurdos no hablan una lengua de origen caucásico, ya que la lengua kurda pertenece a la familia de las lenguas indo-iránias (dentro de las lenguas indoeuropeas), como muchas de las lenguas

habladas por poblaciones vecinas (persa, talyshi, guilakí o mazandaraní). Esto se debe a las migraciones indo-iránicas desde Asia Central (Gimbutas, 1977) en varias oleadas, entre ellas la más importante a través la meseta iraní (800-600 a. C.), que supusieron un proceso de “indoiranización” (Hennerbichler, 2012), mediante la incursión de las tribus medas que conformaron posteriormente el imperio del mismo nombre. Esto supuso la sustitución de la antigua lengua no indoeuropea (probablemente una lengua de la familia hurrita-urartiana) por las lenguas proto-iránicas que dieron lugar a las lenguas kurdas modernas.

Dentro de las lenguas kurdas se pueden distinguir dos grupos pertenecientes ambos a la rama norte occidental de las lenguas iránicas: las lenguas kurdas y las lenguas zaza-gorani. El primero presenta cuatro variantes (Lewis et al., 2014):

- Kurmanji: kurdo del norte, hablado principalmente en Turquía, Irak y en el noroeste y noreste de Irán (este de Azerbaijón Occidental y Jorosán), con un mayor número de hablantes.
- Sorani: kurdo central, hablado en Irak e Irán norte occidental (sur de Azerbaijón Occidental y Kurdistán). La variante dialectal babani se habla en las ciudades de Saqqez y Baneh (véase Materiales y Métodos apartado 1.1).
- Kermanshahi: kurdo del sur, hablado principalmente en Irak e Irán occidental (provincias de Kermanshah e Ilam).
- Laki: hablado en Irán occidental (norte de Ilam y Lorestán), el más minoritario de este grupo en cuanto al número de hablantes.

Con respecto al segundo grupo, las lenguas zaza-gorani, éste presenta seis variantes: bajelani, dimli o zazaki, gorani, kirmanjki, sarli y shabak, todas habladas por grupos kurdos fuera de Irán (Turquía e Irak principalmente), salvo el gorani que se habla en partes del Kurdistán iraní. Este, a pesar de ser una lengua iránica, presenta en algunas construcciones un alineamiento morfosintáctico de tipo ergativo. Esto podría sugerir la presencia de la huella de la lengua hurrita que se hablaba antes de las migraciones indo-iránicas (Arnaiz-Villena et al., 2002; Hennerbichler, 2011).

11.2 Lenguas altaicas: el turcomano y el azerí

En la región sur del Mar Caspio (Azerbaijón, Irán y Turkmenistán) actualmente no solo se hablan lenguas pertenecientes a las macrofamilia indoeuropea. En esta región también están presentes lenguas de la macrofamilia altaica, que recibe este nombre del macizo de Altái, en Asia Central (posible origen de este grupo). Entre las lenguas

altaicas que se hablan en esta región se encuentran el turcomano y el azerí. Ambas se incluyen dentro de las lenguas túrquicas sur occidentales, también llamadas lenguas oghuz.

Los turcomanos son considerados un pueblo túrquico oghuz tradicionalmente nómada procedente de Asia Central (véase apartado 12.3), de una región histórica conocida como Transoxiana por los griegos (“tierra más allá del río Oxus”, actual Amu Daria) y como Turan por los persas, situada entre el mar de Aral y la cordillera de Pamir (actual Uzbekistán). Se piensa que estos pueblos llegaron a la región sur oriental del Mar Caspio en diferentes momentos a partir del siglo X d. C. llevando consigo su cultura, sus tradiciones y por su puesto su lengua de origen túrquico (Askarikhanegha & Sharifkamali, 1995; Luther, 2001). La primera migración de este pueblo de la que se tiene constancia es la producida por la tribu selyúcida procedente del Mar de Aral, que ocupó Persia a finales de este siglo consolidándose en el imperio del mismo nombre. Sin embargo, esto ha sido discutido por otros autores que argumentan que las tribus que se encuentran en el noreste de Irán no son de origen selyúcida, y que llegaron más tarde, con la conquista de Persia por parte del Imperio Mongol (procedente de Asia Oriental), que empujó a las tribus turcomanas al este del Mar Caspio, y del Imperio Timúrida (procedente de Transoxiana), periodo en el que se establecieron definitivamente en la región (Johanson, 2006).

Los azeríes, al igual que los turcos de Anatolia (Turquía), también hablan en la actualidad una lengua turca oghuz de la rama occidental. Sin embargo, los pueblos que habitaron la región al sur del Cáucaso oriental, como los Mannai, a los que se considera como el sustrato más antiguo de los azeríes, hablaron probablemente una lengua hurrita-urartiana (Zadok, 2006). Estos, al ser asimilados por los pueblos indo-iranios, los medos principalmente, sufrieron un proceso de sustitución lingüístico-cultural, que supuso el desarrollo de la antigua lengua azerí, también conocida como adari, de origen iranio, extinta en la actualidad, y hablada en Azerbaiján y las provincias azeríes de Irán. Más tarde, con la irrupción de la dinastía turca oghuz selyúcida, que llegó hasta Anatolia, comenzó un proceso paulatino de “turquificación” cultural que se consolidó con la instauración del Ilkanato mongol en Persia. De este modo comenzó a imponerse el uso de la lengua turca en la región azerí, centro político de este imperio, ya que la mayoría de sus ejércitos eran de origen túrquico (Yarshater, 1988). La subida al poder de la dinastía safávida, apoyada por tribus turcas o turco-hablantes (quezelbas) procedentes de Anatolia e instaladas en Azerbaiján, terminó por hacer desaparecer el adari, y este

término (o más comúnmente, el azerí) pasó a ser utilizado para definir el turco hablado en dicha región. Sin embargo hay constancia en fuentes escritas de que esta lengua continuó en uso hasta el siglo XVII d. C. (Yarshater, 1988). Ciertos autores consideran que algunos dialectos iraníes que se hablan actualmente en Azerbaiján y el noroeste de Irán, como ciertos dialectos tati o talyshi descienden directamente del antiguo azerí (Minorsky, 1957; Yarshater, 1988), mientras que otros argumentan que son dialectos de las lenguas iraníes occidentales importados posteriormente (Henning, 1954). Además, varios autores han recogido un gran número de palabras no turcas utilizadas en el turco azerí, que podrían ser préstamos del antiguo adari, o bien del persa, aunque su origen es discutido (Yarshater, 1988).

12. MARCADORES GENÉTICOS EN EL ESTUDIO DEL POBLAMIENTO DE LA REGIÓN DEL MAR CASPIO

Los estudios antropológicos y lingüísticos nos han ayudado a conocer cuándo y cómo se pudo poblar la región del Mar Caspio. Además, con las tecnologías desarrolladas en el campo de la biología molecular, podemos continuar descubriendo datos que arrojen más información sobre este poblamiento.

Tecnologías como estudios de polimorfismos de ADN en el cromosoma Y, estudio de polimorfismos de ADNmt y análisis de las variantes alélicas de los genes HLA entre otros, han incrementado favorablemente el conocimiento de las características genéticas de los pobladores de esta región y permiten su comparación con otras poblaciones alrededor del mundo con el fin de desarrollar análisis de emparentamiento para relacionar poblaciones y finalmente hacer un estudio en conjunto con los datos históricos, lingüísticos y antropológicos.

12.1 ADN mitocondrial

El ADNmt humano tiene características únicas que lo hacen apropiado para estudios microevolutivos. Posee una alta tasa de mutación (de 5 a 10 veces superior que la del ADN nuclear), no sufre recombinación génica y tiene una transmisión unidireccional por vía materna a lo largo de las generaciones (Brown & Wright, 1979). Fue descubierto en 1963, pero su secuencia se conoce totalmente desde el año 1981, y comprende 16569 pb (Anderson et al., 1981). Posee un total de 37 genes que codifican para los ARN ribosómicos (ARNr) 12s y 16s, para 22 ARNts, y para 13 proteínas implicadas en la maquinaria energética de la célula en la mitocondria (Wallace et al.,

1999). El ADNmt tiene muy pocas regiones no codificantes, entre las que destaca la denominada región D-loop (1122 pb), que constituye el fragmento más largo de secuencia no codificante, encargada de regular e iniciar la replicación del ADNmt. Además, la región D-loop posee un alto grado de polimorfismo, y es ahí donde se han centrado la mayoría de los estudios evolutivos recientes (Cann et al., 1987; Macaulay et al., 1999; Quintana-Murci et al., 2004).

TABLA I6. Principales enfermedades asociadas a genes mitocondriales.

Enfermedades asociadas	Genes mitocondriales afectados
<i>Encefalomiopatías</i>	
Síndrome MERRF	Genes ARNt (MTTK, MTTF, MTTH, ...)
Síndrome MELAS	Genes ARNt (MTTL1, MTTQ, MTTH, ...)
Síndrome de Kearns-Sayre	Gen ARNt MTTL1
Oftalmoplegia externa progresiva (PEO)	Deleción de ADNmt
Miopatía mitocondrial infantil transitoria	Gen ARNt MTTE
Miopatía mitocondrial infantil letal	Gen ARNt MTTT
Déficit del complejo I mitocondrial	Genes NADH deshidrogenasa, genes ARNt
Déficit del complejo IV mitocondrial	Genes citocromo c oxidasa (MTCO), genes ARNt
Déficit del complejo V mitocondrial	Genes ATP sintasa (MTATP), genes ARNt
<i>Sistema Nervioso</i>	
Genes ARNt, genes ATP sintasa, genes NADH deshidrogenasa	
Síndrome NARP	Gen MTATP6
Atrofia óptica de Leber (LHON)	Genes NADH deshidrogenasa, MTCO1, MTCO3, MTATP6
<i>Otras enfermedades mitocondriales</i>	
Diabetes mellitus y sordera (DAD)	Genes ARNt (MTTE, MTTK, MTTL1)
Síndrome de Pearson	Deleción de ADNmt (4977 pb)

MERRF: epilepsia mioclónica asociada a fibras rojas rasgadas (del inglés *Myoclonic Epilepsy with Ragged Red Fibers*); MELAS: encefalomiopatía mitocondrial, acidosis láctica y episodios parecidos a un accidente cerebrovascular (del inglés *Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episodes*); NARP: neuropatía, ataxia y retinitis pigmentaria. (McKusick, 2014).

Estos estudios están basados en dos estrategias metodológicas diferentes. La primera de ellas consiste en el análisis de polimorfismos en la secuencia del ADNmt mediante el uso de enzimas de restricción (RFLP, del inglés *Restriction Fragment Length Polymorphism*). Los resultados obtenidos a partir de la aplicación de esta metodología permitieron caracterizar sitios de restricción polimórficos propios de determinados grupos étnicos (Vigilant et al., 1991). La segunda corresponde al análisis de polimorfismos de secuencia (SNPs) en el ADNmt, mediante la secuenciación de las regiones hipervariables I, II y III del D-loop (Merriwether et al., 1996).

Los análisis evolutivos mediante este marcador genético, sólo informan de manera macrogeográfica, es decir, no tienen tanta afinidad geográfica como lo puedan tener los estudios realizados con otros marcadores como los genes HLA. Por otro lado, parece ser que sobre este genoma recaen muchas asociaciones a distintas enfermedades, con lo que, seguramente la selección natural actúe de manera efectiva para eliminar algunos haplogrupos (conjunto de haplotipos que comporten determinadas variantes alélicas) y mantener algunos otros (Tabla I6).

12.2 Cromosoma Y

El cromosoma Y, uno de los dos cromosomas sexuales, se divide atendiendo a su funcionalidad en tres regiones: las dos primeras, son las regiones pseudoautosómicas, PAR1 y PAR2 (del inglés, *Pseudoautosomal Region I, II*), llamada así porque es homóloga con ciertas secuencias del cromosoma X para poder aparearse correctamente los cromosomas sexuales durante la meiosis. La tercera región es la específica del hombre, MSY (del inglés, *Male Specific Region Y*), antes conocida como NRY (del inglés, *Non-recombining Region Y*) (Skaletsky et al., 2003). Esta región comprende el 95% del cromosoma Y. Con excepción de una región PAR1, el cromosoma Y se hereda en bloque, del padre al hijo, sin que haya recombinación, lo que representa un registro enorme de mutaciones que se han ido “grabando” en el ADN a lo largo de la evolución humana.

El cromosoma Y es un buen marcador genético para este tipo de estudios ya que no se han descrito muchas enfermedades asociadas a este marcador, las cuales están relacionadas principalmente con el aparato reproductor masculino y la espermatogénesis, como la disgenesia gonadal o la azoospermia (Tabla I7), con lo que la selección natural no ha influido en mantener diversos haplogrupos específicos.

TABLA I7. Principales enfermedades asociadas a genes del cromosoma Y.

Enfermedades asociadas	Genes afectados (localización)
Disgenesia gonadal XY (Síndrome de Swyer)	SRY (Yp11.31)
Azoospermia	USP9Y (Yq11.21)
Síndrome de células de Sertoli	Deleciones en la región AZF (Yq11)
Retinosis pigmentaria	RPY (Y)
Discondrosteosis de Leri-Weill	SHOXY (Yp11.32)
Displasia mesomélica de Langer	
Sordera ligada a cromosoma Y	DFNY1 (Y)
Síndrome de las orejas peludas	HEY (Yq)

SRY: región Y determinante del sexo (del inglés *Sex-determining Region Y*); AZF: factor azoospermico (del inglés *Azoospermia Factor*). (McKusick, 2014).

En la región MSY del cromosoma Y, sólo se han identificado 27 genes, que en su mayoría están involucrados en funciones específicas del varón, entre ellas su determinación sexual, como el gen SRY, implicado en el desarrollo testicular (Skaletsky et al., 2003; Krausz et al., 2004). Es en la región no recombinante donde se genera el polimorfismo que define los distintos haplogrupos del cromosoma Y.

12.3 Estudio de genomas completos: secuenciación de alto rendimiento

En los últimos años se han desarrollado técnicas que permiten realizar la secuenciación de forma paralela de un gran número de fragmentos génicos produciendo de miles a millones de estas operaciones simultáneamente. Basado en el método clásico de secuenciación de Sanger (las bases terminadoras de pequeños fragmentos de ADN son identificadas de forma secuencial a través de señales que emiten cuando cada fragmento es resintetizado a partir de una hebra de ADN molde; Sanger et al., 1977), esta técnica de secuenciación de alto rendimiento o NGS (del inglés *Next-Generation Sequencing*) permite secuenciar largas extensiones de ADN abarcando genomas completos siguiendo a grandes rasgos el siguiente procedimiento: el ADN genómico de una muestra es fragmentado en una genoteca de pequeños fragmentos susceptibles de ser secuenciados uniformemente en millones de reacciones paralelas; las nuevas cadenas de ADN son reensambladas utilizando un genoma de referencia, de tal forma

que la secuencia del genoma completo deriva del consenso de las secuencias obtenidas alineadas (Srivatsan et al., 2008).

Esta técnica tiene múltiples aplicaciones, entre ellas los estudios antropológicos y evolutivos. Sin embargo, en estos casos se lleva a cabo una secuenciación dirigida que permite secuenciar un conjunto determinado de regiones a lo largo de todo el genoma, donde se encuentran los SNPs que se deseen analizar, mediante la hibridación de estas regiones con sondas magnéticas específicas para cada una de las alternativas del polimorfismo, que permite fijar las regiones de interés que serán posteriormente amplificadas y secuenciadas. En este tipo de técnica se enmarca el reciente estudio genómico en muestras de ADN antiguo de europeos que habitaron 6000 y 1000 años a. C. desde la Península Ibérica hasta Europa Oriental, comparándolas con muestras de poblaciones europeas actuales. Este trabajo concluye, mediante análisis de componentes principales y de mestizaje, que las poblaciones europeas actuales presentan un componente genético ancestral procedente de los pueblos de la cultura Yamna (véase apartado 10.1) en mayor o menor proporción (menor en Europa meridional), lo que significa que hubo una migración hacia el corazón de Europa procedente de los pueblos kurganes de las estepas pónicas durante el Neolítico final (2500-2000 a. C.). Esto podría interpretarse como que fueron estos pueblos de la estepa los que introdujeron las lenguas indoeuropeas en Europa (Haak et al., 2015).

Sin embargo, esta metodología está aún en fase de desarrollo y la tecnología empleada es todavía incipiente, por lo que es bastante costosa y requiere grandes concentraciones de ADN para ser llevada a cabo. De momento, los estudios llevados a cabo no han arrojado datos o conclusiones nuevas a cerca de la historia evolutiva del hombre y sus procesos migratorios, confirmando lo que sostienen estudios previos basados en marcadores clásicos (véase apartados siguientes).

12.4 Aplicación al estudio del poblamiento de la región del Mar Caspio

Con respecto al poblamiento de la región del Mar Caspio y a los movimientos poblacionales que han tenido lugar en el “corredor iraní” que comunica Oriente con Occidente se han realizado diversos estudios a partir de estos marcadores genéticos (ADNmt y cromosoma Y).

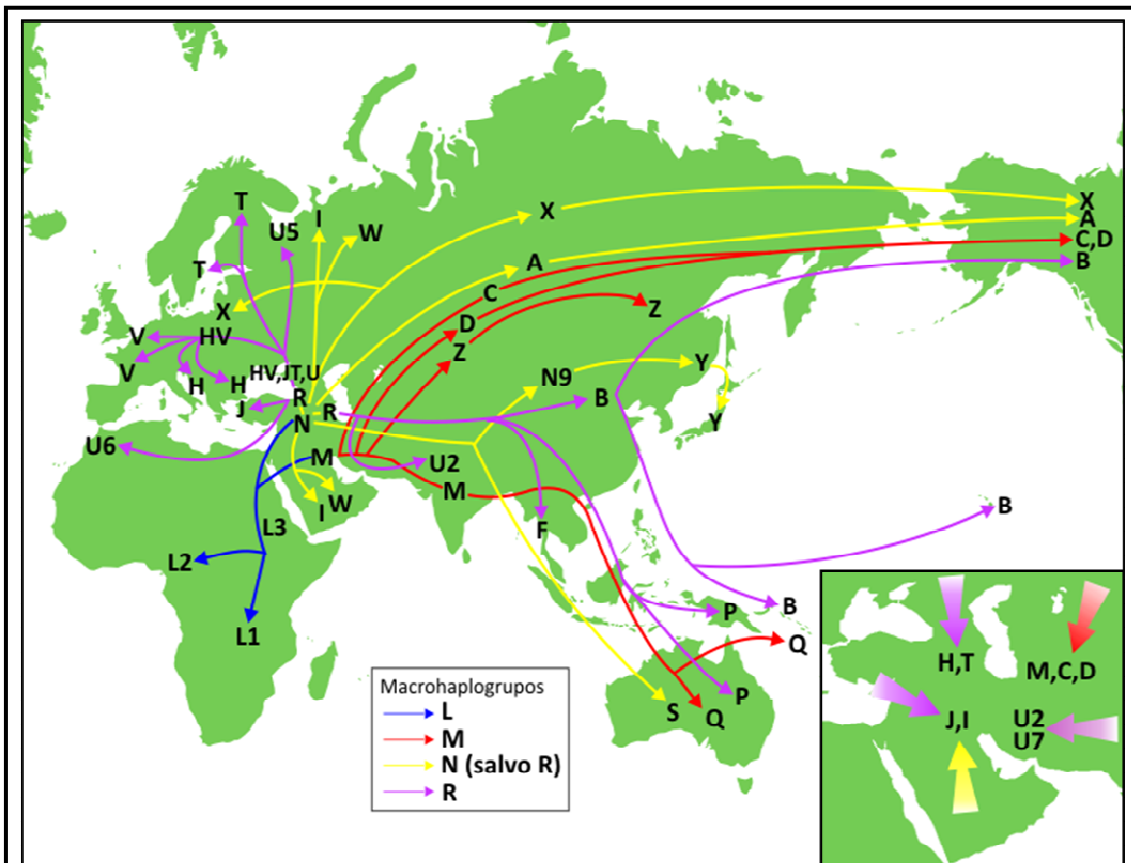


Figura I16. Mapa de las migraciones poblacionales humanas basadas en la dispersión de los haplogrupos de ADNmt. En el recuadro inferior izquierdo figuran los haplogrupos mayoritarios en las poblaciones de Irán y su posible origen (Quintana-Murci et al., 2004; Nasidze et al., 2004; Terreros et al., 2011).

Según los estudios basados en ADNmt, muchos de los pueblos euroasiáticos se caracterizan por tener haplogrupos derivados del haplogrupo N, a su vez derivado del haplogrupo L3 procedente de África (Macaulay et al., 1999). Análisis llevados a cabo mediante el estudio de haplotipos RFLP de ADNmt indican que las poblaciones situadas desde la meseta iraní hasta el Cáucaso y Anatolia presentan una alta diversidad de linajes mitocondriales pero con una composición común con poca contribución del sur de Asia o de Eurasia oriental (la mayoría de los linajes comunes o autóctonos de estas regiones están ausentes o en muy baja frecuencia), y que de hecho, las poblaciones iraníes (guilakíes, mazandaraníes, kurdos, luros y persas) presentan un llamativo grado de homogeneidad sugiriendo altos niveles de flujo génico entre poblaciones (Quintana-Murci et al., 2004; Figura I16).

El hecho de la escasa contribución genética de las poblaciones de Asia Central a las poblaciones del “corredor iraní”, Cáucaso y Anatolia, es sorprendente si atendemos a las sucesivas oleadas migratorias históricamente documentadas de las tribus altaicas

(Mongoles, Turcos oghuz) desde la baja Edad Media, que ocuparon estas regiones y llevaron a cabo del proceso de “turquificación” a nivel lingüístico-cultural. Esto concuerda con análisis de haplogrupos del cromosoma Y en poblaciones de Asia Central y Próximo Oriente, en los que se observa que el haplogrupo R-M17 (también denominado R-M198 o R1a1a), característicos de poblaciones indo-iránicas y altaicas, es bastante escaso tanto en las poblaciones iraníes actuales como en las poblaciones caucásicas, sugiriendo un origen común de estas (véase más adelante; Wells et al., 2001; Zerjal et al., 2002).

En esta misma línea los estudios llevados a cabo por Nasidze *et al.* analizan el alto grado de diversidad en la región en cuanto a los linajes de ADNmt, así como la estrecha relación de las poblaciones actuales del Cáucaso con poblaciones iraní y kurda, mostrándose mucho más próximas genéticamente que con otras poblaciones de Eurasia (Nasidze et al., 2004).

Sin embargo, un estudio reciente en población iraní contradice lo reportado por los estudios previos (Quintana-Murci et al., 2004; Nasidze et al., 2004) observándose una distribución diferente en los linajes de ADNmt en población del norte (con una distribución distinta a las poblaciones del sur del Mar Caspio) y del sur de Irán (especialmente con respecto a los haplogrupos I, J, U2 y U7), definiendo el desierto de Kavir (centro de Irán, Figura I15) como barrera natural para frenar el flujo génico desde el oeste (haplogrupos I y J) y desde el este (haplogrupos U2 y U7). Además se apunta hacia una proximidad genética con las poblaciones de origen árabe (presencia de haplogrupos J y T) (Terreros et al., 2011; Figura I16).

En cuanto a los estudios llevados a cabo a partir de marcadores del cromosoma Y, se ha constatado el alto nivel de diversidad de haplogrupos en Asia Central, lo que hace considerar a las poblaciones centroasiáticas entre las más antiguas del continente, y origen de muchos de los linajes mayoritarios en Europa, América y Suroeste Asiático (“corredor iraní”), dispersados mediante sucesivas migraciones. Entre ellos está el haplogrupo R-M17, asociado con el desarrollo y la expansión de las culturas Yamna y Andrónovo desde las estepas pónicas y rusas, y por consiguiente, de la cultura y lenguas indo-iránicas (Wells et al., 2001; Quintana-Murci et al., 2001; Figura I17). Para confirmar esta hipótesis se llevó a cabo un estudio de ADN antiguo a partir de restos óseos de individuos pertenecientes a la cultura de Andrónovo (kurganes) hallados al sur de Siberia, constatando que la mayoría de los individuos eran portadores de este haplogrupo (Keyser et al., 2009). Sin embargo, estos estudios determinan que este

haplogrupo tiene poca presencia en la población actual iraní, así como en las poblaciones turca y azerí, que parecen estar mucho más próximas a las poblaciones de Oriente Próximo y Cáucaso. Los autores hipotetizan sobre la posibilidad de que las poblaciones caucásicas y mesopotámicas asentadas antes de las migraciones indo-iránicas y túrquicas eran lo suficientemente densas como para enmascarar cualquier contribución genética y/o que la sustitución lingüística pudo deberse a un modelo de “elite dominante”, modelo según el cual una minoría recientemente inmigrada domina política, económica y culturalmente (Wells et al., 2001).

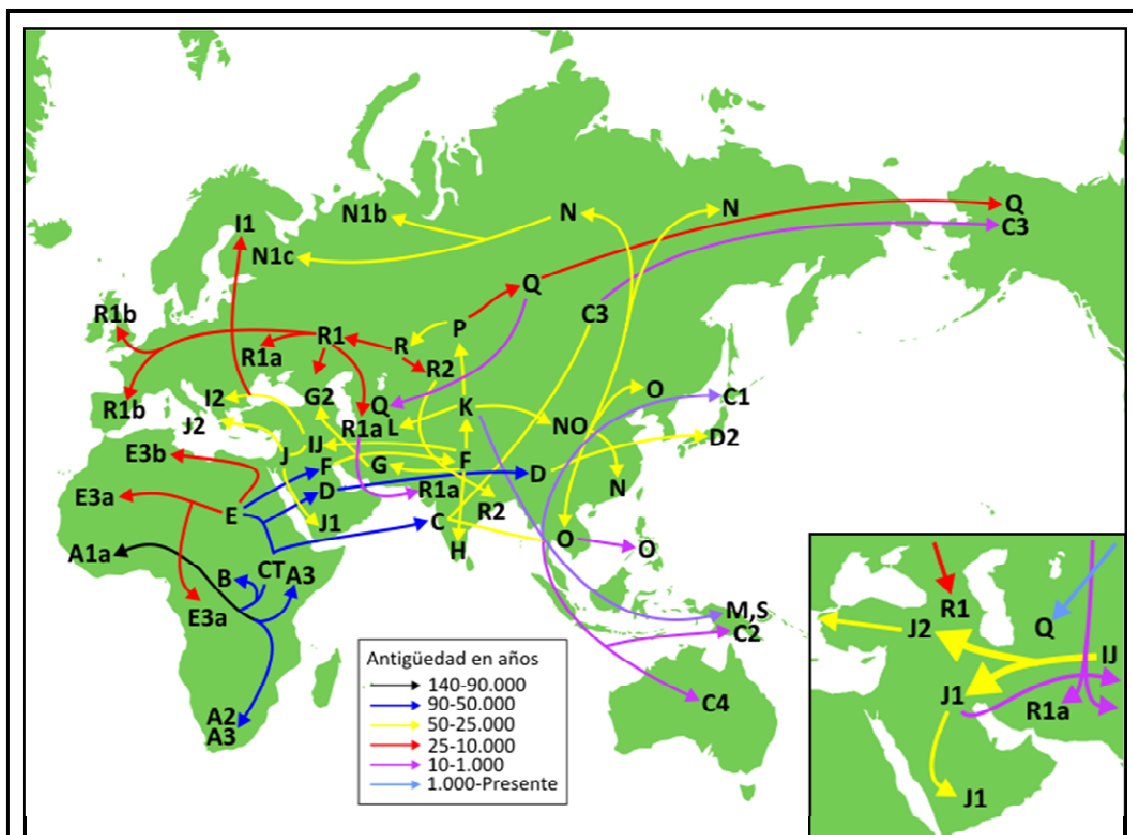


Figura I17. Mapa de las migraciones poblacionales humanas basadas en la dispersión de los haplogrupos de cromosoma Y. En el recuadro inferior izquierdo figuran los haplogrupos mayoritarios en las poblaciones de Irán y su posible origen (Semino et al., 2000; Wells et al., 2001; Grugni et al., 2012).

Sin embargo, estudios posteriores en población iraní observan una mayor diversidad en cuanto al haplogrupo R-M17 se refiere, encontrándose éste con frecuencias superiores al 18% en poblaciones urbanas de Teherán e Isfahán (Nasidze et al., 2004) o con grandes diferencias entre población del sur y del norte de Irán (15% vs. 3% respectivamente; Regueiro et al., 2006). Con respecto a la diferencia de distribución

de este haplogrupo entre las dos muestras poblacionales, los autores especulan sobre una posible migración indo-irania hacia la India que podría haber dejado su huella genética en la población del sur y en la que el desierto de Lut (sudeste de Irán, Figura I15) habría frenado su expansión a través de la meseta iraní, de forma similar a lo reportado por algunos estudios basados en el ADNmt (Terreros et al., 2011). También hacen hincapié en la importancia del “corredor iraní” en la diseminación de ciertos haplogrupos como el J1 (J-M267) y el J2 (J-M172) desde el oeste (poblaciones árabes) hacia Asia (India) (Regueiro et al., 2006). Un estudio reciente basado en 15 grupos étnicos iraníes diferentes sugiere una mayor diversidad y heterogeneidad en la distribución de los haplogrupos del cromosoma Y en estas poblaciones posiblemente debido a la aportación genética de las distintas migraciones procedentes del Cáucaso, Asia Central y Eurasia Occidental. También incide en las barreras geográficas (montes Zagros, desierto de Lut) como factores limitantes del flujo génico de determinados haplogrupos como el R-M17. Cabe destacar de este estudio la heterogeneidad en la poblaciones de la región sur del Caspio, entre ellas los azeríes (Grugni et al., 2012).

Con respecto a las poblaciones de la región sur del Mar Caspio en estudio, los turcomanos procedentes de Turkmenistán presentan una frecuencia considerable de haplogrupos de ADNmt descendientes del macrohaplogrupo M, originario de Eurasia central-oriental (del 19 al 31% según diferentes estudios), concretamente un inusual haplogrupo mixto E/G encontrado solamente en una población siberiana (Maliarchuk et al., 2002; Quintana-Murci et al., 2004). Esto indicaría la presencia de un fuerte componente túrquico-mongoloide en el sustrato genético de esta población. En cuanto a los estudios basados en cromosoma Y, los resultados son más o menos concordantes. Wells *et al.* observaron en población turcomana de Turkmenistán una presencia mayoritaria del haplogrupo R-M173 (R1) de gran presencia en Eurasia, pero cuyo origen se cree que se sitúa en Asia Central (Wells et al., 2001). Un estudio más reciente en turcomanos del Golestán identifica como haplogrupo mayoritario en esta población el Q-M25 (42.6%), grupo hermano del R-M207 (precursor del R-M173) y con un lugar de origen común (región Altái/Baikal, The International Society of Genetic Genealogy, 2014), lo que podría identificarse como una contribución de tribus altaicas (mongoles, turcos oghuz), hecho que concuerda con lo sugerido por los estudios de ADNmt. No obstante, la presencia considerable del haplogrupo R-M17 sugiere una contribución anterior por parte de las migraciones indo-irania también procedentes de esta región (Grugni et al., 2012).

Los azeríes procedentes de Azerbaiyán presentan exclusivamente haplogrupos de origen euroasiático occidental, como el H (el más frecuente), el J o el T, situándose genéticamente muy próximos a otras poblaciones del Cáucaso (Georgianos, Armenios) y de Anatolia (Turcos), lo que sugiere un origen común. También se observa relación con las poblaciones vecinas de Irán (véase apartado 10.2.2, Quintana-Murci et al., 2004). Sin embargo, en un estudio reciente se observa tanto en población turca como azerí (Azerbaiyán) haplogrupos descendientes del macrohaplogrupo M de origen en Eurasia central-oriental en un frecuencia significativa, sugiriendo cierta influencia de los grupos procedentes de esta zona posiblemente a través de la migración oghuz hacia esta región (Schonberg et al., 2011).

En lo que respecta a los estudios basados en cromosoma Y los azeríes iraníes presentan con una elevada frecuencia el haplogrupo J (principalmente el J2, del 39 al 48%, según diferentes estudios) lo que sugiere un fuerte componente genético con origen en Oriente Próximo (Wells et al., 2001; Andonian et al., 2011). Estos estudios contradicen la influencia de los grupos de origen túrquico sugerida por el ADNmt (Schonberg et al., 2011), al menos en lo que se refiere a la línea paterna. Sin embargo, otros autores observan la presencia del haplogrupo R-M17 con una frecuencia similar a los haplogrupos J2, situando a esta población más relacionada con los grupos vecinos más próximos (persas, asirios iraníes, kurdos iraníes) y con poblaciones de Asia Central próximas a las poblaciones indo-iránias que pudieron colonizar la región, apoyando así el modelo de “elite dominante” para la sustitución lingüística del antiguo azerí indoeuropeo (adari) por el actual azerí de origen altaico (Yepiskopasian et al., 2011; Grugni et al., 2012).

En cuanto a los kurdos, estudios basados en ADNmt en poblaciones procedentes de Irán, Turkmenistán, Georgia y Turquía han determinado la presencia mayoritaria de haplogrupos de origen euroasiático occidental, tales como H, J o T (como en el caso de los azeríes de Azerbaiyán; Comas et al., 2000; Richards et al., 2000; Quintana-Murci et al., 2004). No obstante los kurdos de Turkmenistán presentan una mayor frecuencia de haplogrupos originario de Asia Central (9%) (Quintana-Murci et al., 2004). También se ha observado una clara presencia del haplogrupo U5 en los kurdos procedentes de Turquía (también presente en armenios y azeríes pero ausente en poblaciones de Oriente Próximo), de gran implantación en Europa, lo que para algunos autores sugiere una migración reversa procedente de Europa (Richards et al., 2000). Los estudios de emparentamiento genético a partir de este marcador demuestran que las poblaciones

kurdas procedentes de las distintas regiones se agrupan juntas y con otras poblaciones del sur del Cáucaso, independientemente de los distintas lenguas kurdas que hablen (kurmanji o zazaki, véase apartado 11.1; Nasidze et al., 2005).

Estudios a partir de cromosoma Y en poblaciones kurdas procedentes de Turquía, Georgia y Turkmenistán, sugieren cierta influencia geográfica en las relaciones de estos grupos: los kurdos asentados en Turquía están más próximos en cuanto linajes de cromosoma Y a pesar de hablar dos lenguas kurdas diferentes (y también próximos a grupos iraníes y caucásicos), mientras que los kurdos de Turkmenistán (trasladados del Kurdistán en el siglo XVII) y los de Georgia (procedentes de Turquía en el siglo XIX) están más alejados de estos y con una menor diversidad de haplogrupos, sugiriendo un cuello de botella (Wells et al., 2001; Nasidze et al., 2005). Estos y otros estudios basados en este mismo marcador observan en estas poblaciones (incluida una población del norte de Irak) una presencia dominante de los haplogrupos J (J1 y J2) indicando que el origen de este pueblo se situaría en torno al Creciente Fértil (Anatolia, Mesopotamia o montes Zagros) y asociándolos estrechamente a judíos y armenios por la proximidad de sus haplotipos modales (Wells et al., 2001; Nasidze et al., 2005; Nebel et al., 2007; Hennerbichler, 2012). No obstante, la presencia de haplogrupos R (R-M17, R-M173/R1 y R-M124/R2) en frecuencias considerables en algunas de estas muestras (kurdos de Turquía y de Turkmenistán) y en otra procedente del Kurdistán iraní indica una posible huella genética por parte de las migraciones indo-iránias desde Asia Central, bien a través la meseta iraní, bien a través del Cáucaso, en diferentes momentos (Wells et al., 2001; Nasidze et al., 2005; Hennerbichler, 2012; Grugni et al., 2012), lo que contradice lo reportado por el ADNmt en el que no se observa haplogrupos originarios de esa región (Comas et al., 2000; Richards et al., 2000).

Las frecuentes contradicciones entre los dos marcadores empleados en estos estudios pueden deberse a un contribución asimétrica de las líneas paterna y materna, ya que están ligados al sexo, y también a que el ADNmt no tiene tanta resolución a nivel microgeográfico, como sí la presentan el cromosoma Y o el sistema HLA.

13. ESTUDIOS DEL SISTEMA HLA EN POBLACIONES DE LA REGIÓN SUR DEL MAR CASPIO

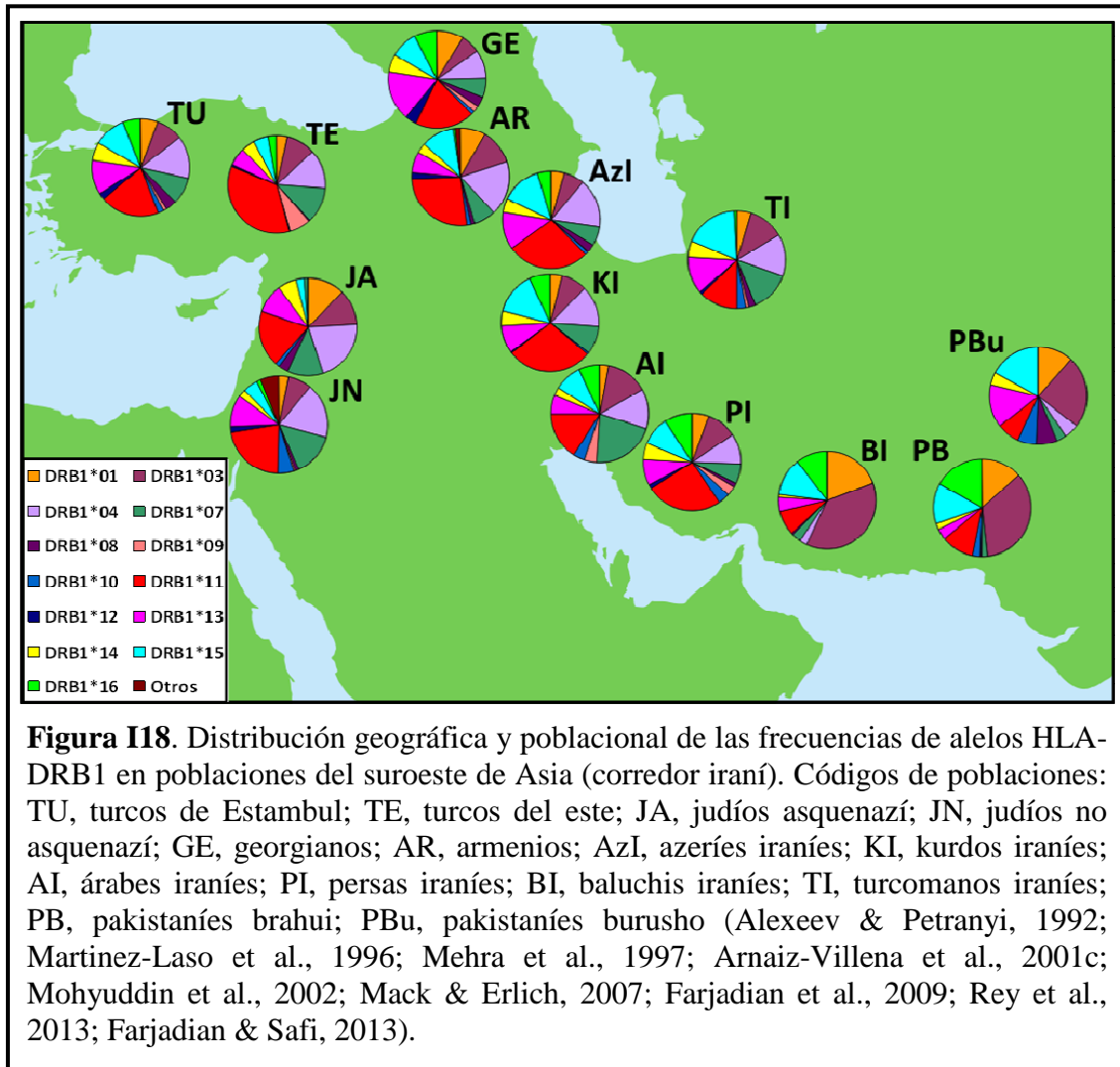
En los últimos años se han llevado a cabo diferentes estudios del polimorfismo del sistema HLA en las poblaciones asentadas en el “corredor iraní”, enfocados principalmente en conocer las relaciones genéticas de los diferentes grupos

poblacionales para arrojar algo de luz sobre el origen y las posibles contribuciones de cada una de las migraciones acaecidas en la región, así como la relación de determinados marcadores HLA con enfermedades de considerable incidencia en estas poblaciones.

Los primeros estudios poblacionales basados en HLA en esta región se realizaron gracias al impulso dado por los Talleres Internacionales de Histocompatibilidad celebrados periódicamente. Desde el 11° taller se han obtenido los perfiles HLA de poblaciones procedentes de la meseta iraní, Cáucaso y Anatolia, tales como iraníes farsi, armenios, georgianos, kurdos y turcos, aunque no todas para los mismos loci HLA. En lo que se refiere a la clase I, a nivel del locus A, las poblaciones caucásicas –georgianos y kurdos de Georgia– presentan frecuencias altas para los alelos HLA-A*03 (34.2% y 31.0% respectivamente), A*02 (31.0 % para georgianos y 12.0% para kurdos) y A*24:02 (17.1% para georgianos y 11.4% para kurdos). Con respecto al locus B, estas mismas poblaciones presentan una elevada frecuencia para el alelo HLA-B*51:01 (12.1% para kurdos y 16.2% para georgianos), siendo éste el alelo más frecuente, así como para el B*35:03 (8.6% para kurdos y 11.1% para georgianos). Por último, kurdos y georgianos presentan una frecuencia muy similar de su alelo más frecuente del locus C, HLA-C*04:01 (en torno al 22%), mientras que los primeros presentan una frecuencia notable del alelo C*15:02 (15.5%) y los segundos del alelo C*06:02 (11.7%) (Alexeev et al., 1997; Mack & Erlich, 2007).

Con respecto a la clase II, se han analizado poblaciones armenia, turca e iraní. A nivel de locus DRB1, el alelo más frecuente en estas tres poblaciones es HLA-DRB1*11 (26.4% para armenios, 25.0% para iraníes y 19.5% para turcos). En armenios e iraníes son muy frecuentes los alelos DRB1*15 (11.2% y 18.5% respectivamente) y DRB1*04 (17.8% y 10% respectivamente) mientras que los alelos DRB1*03:01 y DRB1*07:01 lo son en turcos (9.0% y 8.8% respectivamente). Referente al locus DQB1, en las tres poblaciones el alelo HLA-DQB1*03:01 es mayoritario con frecuencias similares (31.0% para iraníes, 30.1% para armenios y 25.7% para turcos), seguido de DQB1*02:01, DQB1*05:01 y DQB1*03:02 (Alexeev & Petranyi, 1992; Mehra et al., 1997; Mack & Erlich, 2007). En la figura I18 se muestra la distribución de las frecuencias de alelos HLA-DRB1 en estas poblaciones y en otras procedentes de diversos estudios. Se observa que el alelo DRB1*11 también es el más frecuente en las poblaciones de Oriente Próximo y Cáucaso (judíos y georgianos) así como en los grupos iraníes más occidentales: azeríes, kurdos y persas (del 16 al 28.5%). Sin

embargo las poblaciones de Irán situadas al este presentan alelos mayoritarios distintos: para los turcomanos es el DRB1*15 y para baluchis y poblaciones paquistaníes el DRB1*03.



Lo que se desprende de estos análisis es que la distribución del polimorfismo en estas poblaciones es similar y bastante homogénea, por lo que se puede pensar en el origen común de su sustrato genético, o en la existencia de flujo génico entre poblaciones vecinas. Algo similar se desprende de otros estudios en población turca y kurda, donde las frecuencias haplotípicas HLA y las distancias genéticas basadas en las frecuencias de alelos HLA-DR y -DQ les sitúan muy próximas entre sí y a poblaciones del Levante mediterráneo y del Cáucaso, como judíos, libaneses y armenios, sugiriendo un sustrato anatólico-mediterráneo antiguo y un proceso de “turquificación” posterior, fundamentalmente cultural, basado en un modelo de “elite dominante” (Arnaiz-Villena

et al., 2001c; Arnaiz-Villena et al., 2002). Estas conclusiones están en sintonía con las obtenidas por análisis basados en cromosoma Y (véase apartado 12.3). De la misma manera, un estudio en población azerí y kurda de Irán muestra que la distancia genética basada en las frecuencias de alelos HLA de clase II (DRB1, DQA1 y DQB1) existente entre estas dos poblaciones es bastante reducida, sugiriendo igualmente la poca contribución genética de los turcos altaicos en esta región (Farjadian & Ghaderi, 2007). Otro ejemplo de que las poblaciones iraníes pueden estar emparentadas genéticamente con poblaciones del Levante mediterráneo lo da el análisis de marcadores HLA como factores de susceptibilidad a enfermedades, tales como el pénfigo vulgar (PV). Se ha comprobado que la mayoría de los judíos asquenazí pacientes de PV (de gran incidencia en esta población) son portadores del haplotipo HLA de clase II DRB1*04:02-DQB1*03:02 (Ahmed et al., 1990; Tabla I4) mientras que otras poblaciones no judías suelen ser portadoras de otros alelos de susceptibilidad (Ahmed et al., 1991). Un estudio posterior en pacientes iraníes de PV determinó que la mayoría de estos portaban el mismo haplotipo de susceptibilidad, lo que sugiere que estas variantes alélicas tuvieron origen en un ancestro común (Mobini et al., 1997; Amirzargar et al., 2001).

En concordancia con análisis basados en otros marcadores genéticos (ADNmt y cromosoma Y), un trabajo que comparan los perfiles HLA de clase II de distintos grupos étnicos iraníes, fundamentalmente del noroeste y sur incluyendo farsis, árabes, judíos, azeríes, kurdos y baluchis entre otros, concluye que a pesar de las diferencias en la distribución de las frecuencias entre estas poblaciones, al compararlas con poblaciones europeas y asiáticas (fundamentalmente siberianas), estas poblaciones tienden a agruparse debido a su proximidad genética. Los autores deducen de esto que a pesar de los sucesivos encuentros con pueblos foráneos presentan un sustrato genético común, aunque la mezcla genética con otros grupos, como los indoeuropeos, puede ser más acusada en unas poblaciones que en otras (Farjadian et al., 2009).

Un estudio reciente en poblaciones iraníes turco-hablantes -azeríes, kashgais y turcomanos- junto con otras poblaciones de Irán está en sintonía con el modelo de “élite dominante” o “dominancia de élite” del proceso de “turquificación”. Determina que a pesar de similitudes alélicas y haplotípicas de estas tres poblaciones, no están más relacionadas genéticamente entre sí que con las otras poblaciones iraníes. Concluye que a pesar de lo que cabría esperar de poblaciones que hablan lenguas de origen común, los iraníes que hablan una lengua túrquica están relacionados con el conjunto genético del

resto de grupos étnicos iraníes y que descienden de poblaciones que ocuparon originalmente la región antes de la migraciones turcas oghuz (Farjadian & Safi, 2013).

OBJETIVOS

Las antiguas culturas turcomana, azerí y kurda que habitan en la región sur del Mar Caspio y norte de Irán, junto con las existentes en el Suroeste Asiático, pueden confirmarse como parte de la base para comprender las relaciones entre las distintas poblaciones de la región y para conocer el origen del poblamiento del “corredor iraní”, punto de unión entre Europa y Asia. Asimismo, el estudio de los genes HLA en estas poblaciones nos permite obtener una lista virtual de trasplante de las mismas y perfiles HLA para farmacogenética y terapéutica. Para ello nos hemos propuesto:

1. Caracterización del polimorfismo de los genes HLA de clase I (A y B) y de clase II (DRB1 y DQB1) en muestras representativas de las poblaciones turcomana, azerí y kurda de Irán.
2. Calcular las distancias genéticas existentes entre estas poblaciones y el resto de poblaciones del mundo, con mayor hincapié en las poblaciones geográficamente próximas (del Mediterráneo y de Asia Central) a partir de las frecuencias alélicas de los loci HLA de clase II (tipados en alta resolución).
3. Construir árboles de emparentamiento (dendrogramas) por el método de neighbor-joining (NJ, vecino más próximo), a partir de las distancias genéticas con el fin de determinar el grado de emparentamiento entre las distintas poblaciones de estudio.
4. Realizar análisis de correspondencia con todas las poblaciones estudiadas para observar las relaciones existentes entre estas poblaciones según los genes HLA y ver si existe o no correlación entre la geografía y la genética HLA.
5. Estimar las frecuencias y desequilibrios de ligamiento de los haplotipos de dos loci HLA-A-B, HLA-B-DRB1 y HLA-DRB1-DQB1, así como los haplotipos extendidos HLA-A-B-DRB1-DQB1 característicos de turcomanos, azeríes y kurdos de Irán y compararlos con otras poblaciones, principalmente del Mediterráneo y Asia Central. Asignar, si es posible, la procedencia geográfica de estos haplotipos. *Para las comparaciones se utilizan 13000 cromosomas de las poblaciones tipadas para HLA de nuestra base de datos.*
6. Discutir las relaciones genéticas HLA entre las poblaciones de estudio mediante la comparación de los alelos y haplotipos.
7. Inferir a través de todos los datos obtenidos, hipótesis sobre los movimientos migratorios que dieron lugar al poblamiento del “corredor iraní” y el Suroeste Asiático.

MATERIALES Y MÉTODOS

1. POBLACIÓN ESTUDIADA

1.1 Selección de individuos

Gracias a la colaboración del Dr. Ali Amirzargar del Molecular Immunology Research Center, de la Universidad de Teherán de Ciencias Médicas (Irán), se seleccionaron 226 voluntarios sanos como donantes de sangre para este trabajo. Estos donantes pertenecen a tres grupos étnicos distintos analizados de forma independiente: turcomanos (69 individuos), azeríes (97 individuos), y kurdos (60 individuos). Para realizar un muestreo óptimo y poder hacer un buen estudio, se cumplieron las siguientes premisas: 1) en el caso de los turcomanos, que hubieran nacido en la ciudad de Gorgan, en la región del Golestán; en el caso de los azeríes, que hubieran nacido en la ciudad de Tabriz, en la región del Azerbaiyán Oriental; y en el caso de los kurdos, que hubieran nacido en las poblaciones de Saqqez y Baneh, en la región del Kurdistán iraní (véase Introducción apartados 10.2.1, 10.2.2, 10.2.3 y Figura I13) ; 2) que hablaran la lengua característica de su pueblo (turcomano, azerí o kurdo sorani, en su caso); 3) y que sus cuatro generaciones anteriores fueran todos de las mismas zonas. Estos individuos seleccionados los hemos comparado con poblaciones pertenecientes a distintos grupos étnicos. Las poblaciones utilizadas en la comparación están registradas en la Tabla M1 y en la Figura M1, empleando 13000 cromosomas en dicha comparación.

1.2 Recolección de muestras

La extracción de las muestras de sangre periférica se realizó con un sistema de vacío en tubos de con conservante EDTA-Na⁺ como anticoagulante. Los tubos se almacenaron y se conservaron refrigerados a -20°C hasta llevar a cabo la extracción del ADN como se explica a continuación.

2. EXTRACCIÓN DE ADN GENÓMICO

Una vez que llegan las muestras al laboratorio, se mantienen a -20°C hasta que se proceda a extraer el ADN. La extracción se realizó siguiendo el método de extracción estándar de fenol/cloroformo/alcohol isoamílico (FCA) (Sambrook et al., 1989). Las muestras que se han procesado para el desarrollo de este trabajo contaban con un volumen máximo de 2 ml de sangre. Por esto se optó por utilizar un método de extracción automático, en lugar del manual, utilizando un extractor automático (Nucleic Acid Isolation System QuickGene-810, Fujifilm, Japón). De esta manera se garantiza el mayor rendimiento posible con el mínimo gasto de muestra, además de reducir

considerablemente el tiempo de procesamiento, disminuir la manipulación y el consiguiente riesgo de contaminación, y aumentar la pureza de los extractos de ADN. Para ello las muestras han de ser tratadas con anterioridad con un kit comercial (QuickGene Whole Blood Extraction Kit S, Fujifilm, Japón) indicado para la extracción de ADN a partir de sangre total. El protocolo usado se describe a continuación:

Paso previo: dejamos descongelar las muestras en dos pasos sucesivos, el primero a 4°C y el segundo a temperatura ambiente. La descongelación se realiza en dos pasos para evitar la degradación del ADN debida al cambio brusco de temperatura.

- Diluir 100 µl de sangre en 100 µl de tampón fosfato salino (PBS pH 7.4, Gibco, Invitrogen, CA, EE.UU.) en un tubo de 1.5 ml.
- Añadir 30 µl de proteasa (DNA whole blood Protease EDB-01, Fujifilm, Japón) para conseguir la degradación de las histonas.
- Añadir 250 µl de solución de lisis (DNA Whole Blood Lysis Buffer LDB-04, Fujifilm, Japón) para conseguir la degradación de las membranas celular y nuclear. Mezclar bien mediante un agitador vórtex unos segundos.
- Incubar en un baño a 56°C hasta que desaparezcan los posibles coágulos (2-10 minutos).
- Añadir 250 µl de etanol absoluto (Ethanol Absolute PRS - CH₃CH₂OH, Panreac, España) conservado a -20°C. Mezclar bien mediante un agitador vórtex unos segundos.
- Pasar la muestra a un cartucho del kit e introducirlo en el extractor automático. Este cartucho presenta una membrana microporosa, con una gran capacidad de adsorción. La muestra será sometida a tres etapas de presurización sin necesidad de centrifugación: una primera etapa de unión y adsorción del ADN a la membrana, una segunda etapa de lavado (tres ciclos) mediante un tampón de lavado (DNA whole blood Wash Buffer WDB-03, Fujifilm, Japón), y posteriormente una tercera y última etapa de elución en 200 µl de una solución de elución (DNA whole blood Elution Buffer CDB-02, Fujifilm, Japón).
- El extracto se recoge en un tubo colector suministrado también por la casa comercial.

TABLA M1. BASE DE DATOS DE POBLACIONES DEL LABORATORIO DEL PROF. ANTONIO ARNAIZ VILLENA UTILIZADA EN LA COMPARACIÓN. *Total cromosomas analizados = 13000.* Relación de las poblaciones incluidas en el estudio comparativo de los turcomanos, azeríes y kurdos. En la figura M1 se detalla la localización geográfica de cada población. N, número de individuos. (*) Poblaciones aborígenes australianas.

Población	N	Referencia	Población	N	Referencia
Ainu	50	(Bannai et al., 1996)	Kimberley*	82	(Gao et al., 1992b)
Aleutianos	85	(Moscoso et al., 2008)	Koryaks	92	(Grahovac et al., 1998)
Argelinos	51	(Arnaiz-Villena et al., 1995)	Libaneses	59	(Clayton & Lonjou, 1997)
Armenios	100	(Matevosyan et al., 2011)	Macedonios	17	(Arnaiz-Villena et al., 2001a)
Atabascos	124	(Monsalve et al., 1998)	Manchúes	50	(Imanishi et al., 1992a)
Baluchis	100	(Farjadian et al., 2004)	Madang (Papua)	65	(Gao et al., 1992a)
Bereberes	98	(Izaabel et al., 1998)	Mansi	68	(Uinuk-Ool et al., 2002)
Bubis	80	(de Pablo et al., 1997)	Marroquíes	98	(Gomez-Casado et al., 2000)
Buriatos	25	(Uinuk-Ool et al., 2002)	Mongoles Khalka	20	(Munkhbat et al., 1997)
Buyei	70	(Imanishi et al., 1992a)	Mongoles Khoton	85	(Munkhbat et al., 1997)
Cape York*	80	(Gao et al., 1992b)	Negidal	35	(Uinuk-Ool et al., 2002)
Chinos de Singapur	71	(Imanishi et al., 1992a)	Negros de Norteamérica	13	(Imanishi et al., 1992a)
Chukchis	109	(Grahovac et al., 1998)	Negros de Sudáfrica	86	(Imanishi et al., 1992a)
Chuvasio	82	(Arnaiz-Villena et al., 2003)	Nivjis	32	(Grahovac et al., 1998)
Coreanos	100	(Imanishi et al., 1992a)	Nueva Caledonia	65	(Gao et al., 1992a)
Cretenses	144	(Arnaiz-Villena et al., 1999)	Palestinos	16	(Arnaiz-Villena et al., 2001b)
Croatas	105	(Clayton & Lonjou, 1997)	Persas de Shiraz	72	(Farjadian et al., 2009)
Desierto Central*	152	(Lester et al., 1995)	Persas de Yazd	65	(Farjadian et al., 2009)
Españoles	88	(Martinez-Laso et al., 1995)	Rabaul (Papua)	60	(Gao et al., 1992a)
Esquimos	80	(Grahovac et al., 1998)	Rusos	20	(Kapustin et al., 1999)
Esvanos	80	(Sanchez-Velasco & Leyva-Cobian, 2001)	Samoanos	51	(Gao et al., 1992c)

TABLA M1. Continuación

Evenki	35	(Grahovac et al., 1998)	Sardos	91	(Imanishi et al., 1992a)
Fiyianos	57	(Gao et al., 1992a)	San (Bosquimanos)	77	(Imanishi et al., 1992a)
Franceses	179	(Imanishi et al., 1992a)	Taiwaneses	48	(Zimdahl et al., 1999)
Georgianos	119	(Rey et al., 2013)	Tlingit	53	(Imanishi et al., 1992a)
Goroka (Papua)	57	(Gao et al., 1992a)	Todja	22	(Uinuk-Ool et al., 2002)
Hotentotes	91	(Imanishi et al., 1992a)	Tofalar	43	(Uinuk-Ool et al., 2002)
Italianos	284	(Imanishi et al., 1992a)	Turcos	245	(Mack & Erlich, 2007)
Japoneses	493	(Imanishi et al., 1992a)	Tuvanos	197	(Martinez-Laso et al., 2001)
Judíos Askenazí	80	(Martinez-Laso et al., 1996)	Udegeys	23	(Grahovac et al., 1998)
Judíos Marroquíes	113	(Roitberg-Tambur et al., 1995)	Ulchi	73	(Uinuk-Ool et al., 2002)
Judíos no Askenazí	80	(Martinez-Laso et al., 1996)	Vasco españoles	83	(Martinez-Laso et al., 1995)
Keti	22	(Grahovac et al., 1998)	Yuendumu*	119	(Gao et al., 1992b)

En algunos de los casos, la adición de etanol absoluto provoca la aglutinación del ADN (formándose la denominada “medusa”, que se observa a simple vista). En estos casos hay que retirar la medusa antes de pasar la muestra al cartucho, ya que quedaría retenida en la membrana y se perdería. La medusa debe, por tanto, ser procesada manualmente como se describe a continuación:

- Añadir 600 µl etanol absoluto y dejar que evapore
- Añadir 600 µl etanol 70% (Ethanol 70% v/v BP CODEX - CH₃CH₂OH, PANREAC, España) y dejar que evapore. Se debe repetir este paso hasta que la medusa quede de un color blanquecino.
- Añaden 200 µl de agua destilada o bien de la solución de elución (DNA whole blood Elution Buffer CDB-02, Fujifilm, Japón) suministrada con el kit de extracción.

La concentración y la pureza de cada extracto de ADN fueron medidas por espectrofotometría mediante un espectrofotómetro NanoDrop (ND-1000 spectrophotometer, NanoDrop, DE, USA). Este espectrofotómetro mide la absorbancia de la muestra a 260, 280 y 230 nm. Los ácidos nucleicos absorben luz a una longitud de

onda de 260 nm. El software del NanoDrop calcula la concentración del ADN mediante la medida de la Densidad Óptica (DO) siguiendo la relación:

$$1 \text{ A}_{260 \text{ nm}} = 1 \text{ DO ADN} = 50 \text{ ng}/\mu\text{l}$$

Una vez obtenida la concentración de cada muestra, se extrae una alícuota que se ajusta a una concentración de 100-150 ng/ μl mediante evaporación en estufa a 60°C cuando precisa ser concentrada o mediante la adición de agua bidestilada y autoclavada cuando precisa ser diluida.

La medición de la absorbancia a 280 nm (máximo de absorbancia de proteínas) y 230 nm (máximo de absorbancia de solventes orgánicos) tiene como fin la obtención de la medida de la pureza, mediante la relación A260/A280 y A260/A230. Para trabajar con un extracto “puro” el primer ratio debe tener un valor por encima de 1.8 y el segundo un valor por encima de 2.



Figura M1. Localización geográfica de las poblaciones utilizadas en el estudio.

3. AMPLIFICACIÓN Y GENOTIPADO DE LOS GENES HLA DE CLASE I Y DE CLASE II

Para la obtención del perfil genético HLA, comúnmente denominado “tipaje HLA”, de cada uno de los individuos del estudio, se ha llevado a cabo una técnica basada en la combinación de la reacción en cadena de la polimerasa (PCR, del inglés *Polymerase Chain Reaction*), las oligosondas específicas de secuencia (SSOP, del

inglés *Sequence-Specific Oligonucleotide Probes*) y la citometría de flujo, el método PCR-SSOP-Luminex (Itoh et al., 2005).

3.1 Reacción en cadena de la polimerasa (PCR)

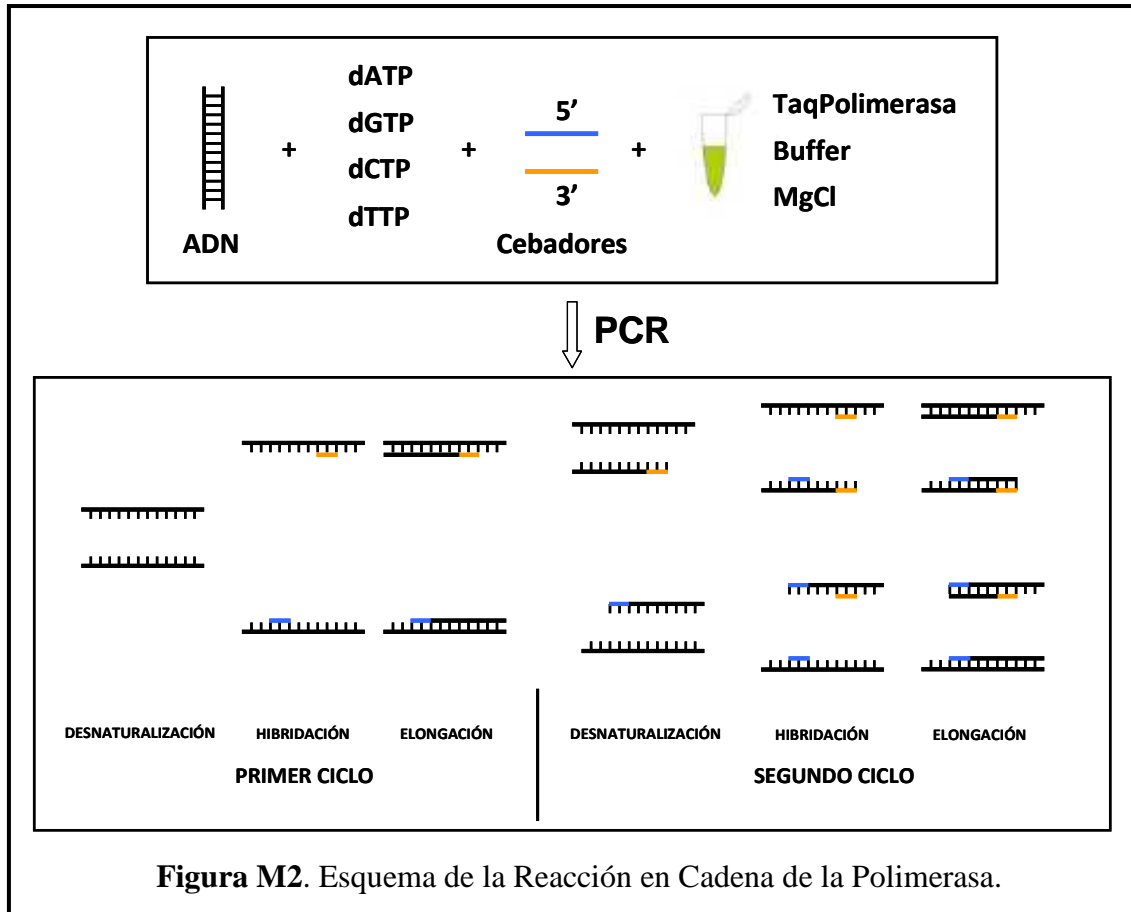
La PCR es un método de síntesis enzimática *in vitro* basado en la selección específica y concreta de una región del genoma para conseguir un aumento exponencial del número de copias de esa región genómica (Saiki et al., 1985). Partiendo de una muestra con un sustrato ADN de cualquier origen, la PCR consiste en la repetición cíclica de tres etapas en el siguiente orden:

- **Desnaturalización.** Es la separación de cada una de las dos hebras del ADN sustrato. Lo conseguimos incubando la muestra a altas temperaturas (93-97°C). Las dos cadenas permanecen separadas libres en la solución hasta que la temperatura baje lo suficiente (37-65°C) para permitir la renaturalización.

- **Anillamiento o hibridación.** Es la unión (hibridación) de los cebadores a secuencias homólogas en el ADN sustrato que se quiere estudiar. Los cebadores, son oligonucleótidos sintéticos que complementan específicamente a ciertas zonas del ADN. Para que se lleve a cabo esta unión, se requiere que haya una alta concentración de los cebadores y un descenso de la temperatura, hasta llegar a la temperatura de renaturalización (temperatura a la cual la mitad de las moléculas de ADN están en forma de una sola cadena y la mitad en forma de doble cadena). Esta temperatura depende de la longitud de la cadena de ADN sustrato y de la composición nucleotídica de la secuencia –el número de puentes de hidrógeno que unen las distintas bases, adenina con timina, y guanina con citosina, es distinto, dos y tres respectivamente, por lo para separar una cadena que tenga mayor cantidad de citosinas y guaninas se necesitara una temperatura más elevada–. Para amplificar una región de ADN “diana”, se necesita un par de cebadores, uno de ellos complementario a una de las cadenas y el otro a la cadena antiparalela, en zonas que flanqueen el segmento de ADN a amplificar. Los cebadores están definidos por la secuencia del ADN sustrato, compuestos en la mayoría de los casos por 18-28 nucleótidos y es necesario que sus temperaturas de renaturalización no excedan los 5°C de diferencia (para garantizar la unión simultánea de ambos) y que sus secuencias no sean ni autocomplementarias ni complementarias entre sí.

- **Elongación o síntesis.** Es la extensión en sentido 5' → 3' del complejo cebador-sustrato. Como enzima usamos la ADN polimerasa que incorpora los deoxinucleótidos

monofosfato, a partir de los trifosfato (dNTPs) presentes en la solución, utilizamos la enzima Taq-polimerasa (extraída de la bacteria *Thermus aquaticus*) cuya actividad catalítica es máxima entorno a los 72-75°C (Saiki et al., 1985; Figura M2).



3.2 Oligonucleótido-sondas específicas de secuencia (SSOP)

La técnica SSOP para el “tipaje HLA” está basada en la detección por fluorimetría de microesferas con colorantes fluorescentes (fluorocromos) unidas a oligonucleótidos específicas para cada polimorfismo. Los distintos alelos HLA (tanto de clase I como de clase II) se determinan por una sucesión de SNPs que los diferencian unos de otros y cada oligonucleótido es específico de una de estas variaciones. Las distintas microesferas de poliestireno presentan en su interior dos colorantes fluorescentes distintos (rojo e infrarrojo) con diferente relación de concentraciones de ambos colorantes (es decir habrá una microesfera con un 100% del primer fluorocromo y 0% del segundo, una segunda microesfera con un 99% del primer fluorocromo y 1% del segundo, etc) y que están unidas de forma covalente a oligonucleótidos que hibridarán con el ADN de la muestra (el conjugado recibe el nombre de oligosonda; Fulton et al.,

1997; Dunbar et al., 2003). Ajustando las concentraciones de ambos colorantes se pueden conseguir hasta 100 microesferas con intensidades de fluorescencia diferentes y únicas, que permitan su codificación e identificación por el citómetro Luminex.

3.3 La citometría de flujo: tecnología Luminex

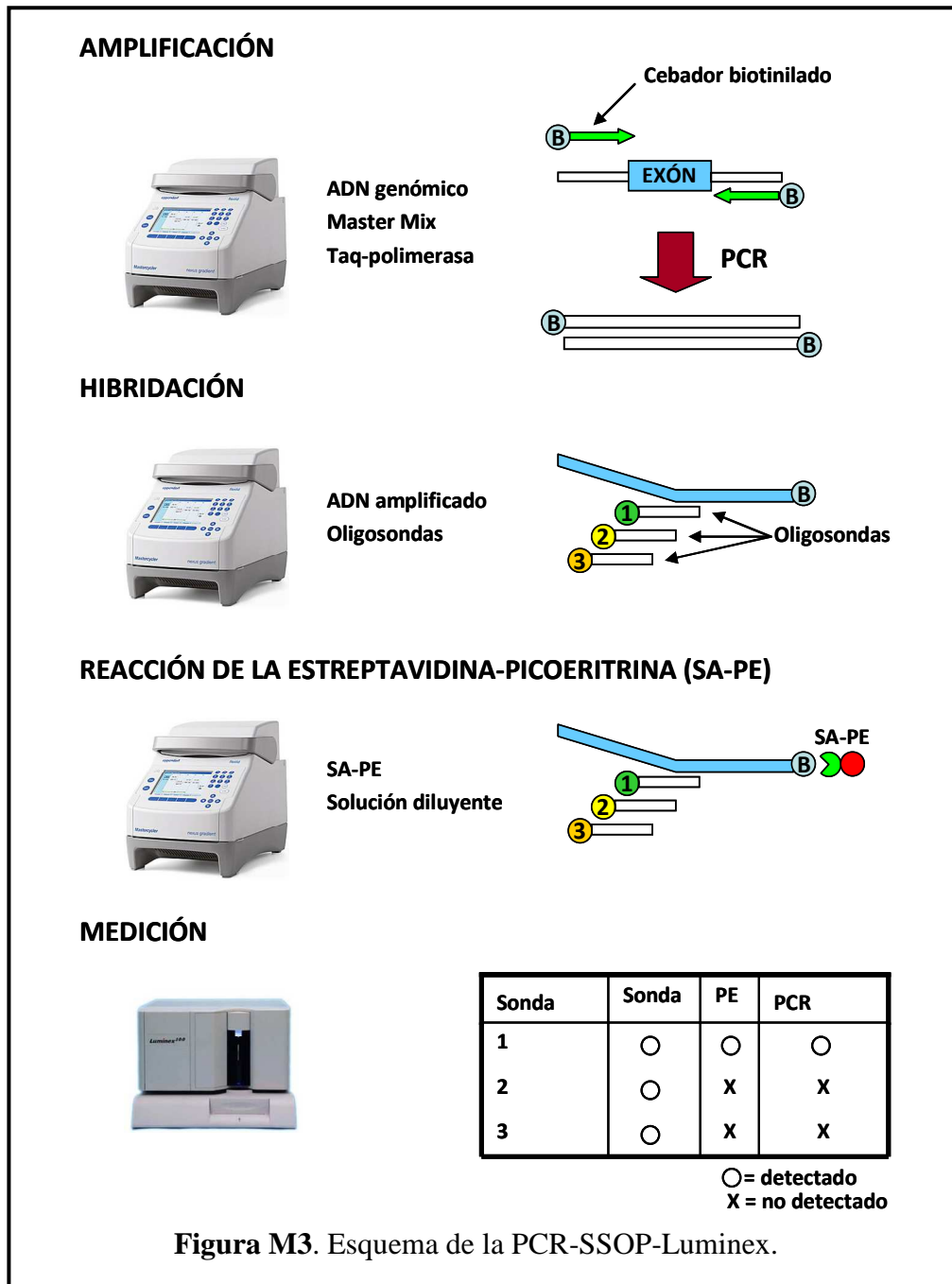
La citometría de flujo permite medir de manera sensible e inmediata múltiples parámetros celulares o moleculares (como es nuestro caso, el ADN), marcados con fluorocromos, que al excitarse emiten fluorescencia, a las que se les hace pasar dentro de un flujo laminar por una cámara de flujo sobre la que incide uno o varios láseres a longitudes de onda determinadas. Estos láseres excitan a los fluorocromos, emitiendo fluorescencia, la cual es captada por detectores que la identifican. La tecnología Luminex es una tecnología de citometría de flujo que permite el análisis de reacciones múltiples (Itoh et al., 2005), constituida por el citómetro de flujo o luminómetro Luminex 100 (Luminex Corporation, TX, USA) que incorpora dos láseres: un láser rojo que emite a 633 nm y un láser verde que emite 532 nm. El láser rojo excita los fluorocromos que se encuentran en el interior de las microesferas (véase apartado anterior), mientras que el láser verde excita la picroeritrina (véase apartado 3.4.4).

3.4 Metodología PCR-SSOP-Luminex

El genotipado HLA por el método PCR-SSOP-Luminex implica cuatro etapas que comprenden: amplificación del ADN por PCR, hibridación, reacción de la estreptavidina-picroeritrina y la medición por fluorimetría, como describe la Figura M3. A continuación se detalla cada una de estas etapas.

3.4.1 Amplificación PCR de los genes HLA-A, -B, -DRB1 y -DQB1

Esta etapa de amplificación consiste en la aplicación de la técnica PCR con la particularidad de que los cebadores empleados están biotinilados (se les ha unido una molécula de biotina en su extremo 5'). Para amplificar los genes HLA de clase I y HLA-DQB1 se utilizan cebadores diseñados por la casa comercial (Gen-Probe, CA, EE.UU.) que se unirán a las regiones intrónicas que flanquean los exones 2 y 3, basados en los diseñados por (Cereb et al., 1996). De la misma manera, para amplificar el gen HLA-DRB1 se utilizan cebadores diseñados por la casa comercial (Gen-Probe, CA, EE.UU.) que se unirán a las regiones intrónicas que flanquean el exón 2, basados en los diseñados por (Cereb et al., 1996).



Mezcla para la reacción de PCR

- Solución que contiene el buffer de amplificación, los cebadores y los dNTPs (Master Mix, Lifecodes HLA-SSO Typing Kit, Gen-Probe, CA, EE.UU.), 7.5 µl por muestra.
- Taq-polimerasa con una concentración inicial de 5U/µl (Invitrogen, CA, EE.UU.), 0.15 µl por muestra.
- ADN genómico, 1 µl por muestra.
- Agua bidestilada y autoclavada hasta obtener un volumen final de 25 µl.

Condiciones de amplificación

La casa comercial tiene estándares acerca de las condiciones de amplificación aplicadas al termociclador Mastercycler egradient S (Eppendorf, Alemania):

1 ciclo de desnaturalización inicial de 95°C durante 5 minutos.

8 ciclos de las siguientes etapas:

- Desnaturalización a 95°C, 30 segundos.
- Hibridación a 60°C, 45 segundos.
- Extensión a 72°C, 45 segundos.

32 ciclos de las siguientes etapas:

- Desnaturalización a 95°C, 30 segundos.
- Hibridación a 63°C, 45 segundos.
- Extensión a 72°C, 45 segundos.

1 ciclo de extensión final de 72°C durante 15 minutos.

Protocolo utilizado

- Se deja atemperar el Master Mix a temperatura ambiente. Se agita en agitador vortex para homogeneizar la mezcla y se centrifuga durante 10 segundos.
- Se prepara la mezcla de amplificación en un tubo de 1,5 ml añadiendo todos los componentes enumerados anteriormente salvo el ADN, para n+1 reacciones y se lleva hasta a un volumen final de 25 µl por reacción. Se agita, se centrifuga durante 10 segundos y se conserva en hielo.
- Se añade el ADN genómico en tubos de PCR.
- Se alícuota la mezcla de amplificación en los tubos que contienen el ADN (el volumen total de mezcla de amplificación y ADN debe ser igual a 25 µl por reacción).
- Se introducen los tubos de PCR tapados en el termociclador y se inicia el programa de amplificación.

3.4.2 Control de amplificación: electroforesis en gel de agarosa

La alta sensibilidad de la técnica de PCR (obtención de hasta más de un millón de copias de la región flanqueada por los cebadores utilizados), hace necesario un cuidado extremo en la manipulación de las muestras para evitar contaminaciones. Para esto, en cada reacción de amplificación, se introduce un control negativo, que consta de los mismos elementos que una reacción estándar, sustituyendo el ADN sustrato por agua bidestilada. La PCR llevada a cabo por el método PCR-SSOP-Luminex mediante los

Lifecodes HLA-SSO Typing Kits (Gen-Probe, CA, EE.UU.) está diseñada para producir tanto ADN de doble cadena como de cadena sencilla. El control de la amplificación se realizó observando las muestras (incluida el control negativo) en un gel de agarosa a una concentración del 2% con bromuro de etidio. La electroforesis se lleva a cabo en presencia de tampón TBE 1x en una cubeta de electroforesis. En cada pocillo se cargan 5 µl de la mezcla resultante de:

- 2,5 µl del producto de la PCR
- 2,5 µl de tampón de carga

Además, se incluye un marcador de peso molecular (Molecular Weight Marker XIV, Roche Molecular Biochemicals, Mannheim, Alemania), para identificar el tamaño de la amplificación de nuestras muestras. Éste se prepara tomando 2,5 µl del marcador, y 2,5 µl de tampón de carga. Una vez cargadas las muestras y el marcador en los pocillos, se someten a electroforesis durante 35 minutos a 100 voltios, visualizándose finalmente el resultado en un transiluminador ultravioleta (254 nm). El resultado se recoge mediante un sistema de fotografía. Las bandas resultantes de la amplificación de los genes HLA de clase I y clase II se recogen en la Tabla M2.

TABLA M2. Tamaño en pb de las bandas resultantes de la amplificación de los genes HLA de clase I y clase II.

	HLA-A	HLA-B	HLA-DRB1	HLA-DQB1
Doble Cadena (pb)	~420	~370, ~340	~280	~260, ~160
Cadena sencilla (pb)	~240	~200	~180	~170, ~100

3.4.3 Hibridación con oligosondas

La segunda fase consiste en la hibridación de los productos amplificados y biotinilados con las oligosondas específicas para cada SNP y que contienen las microesferas con los dos colorantes fluorescentes a distintas concentraciones, cada una específica del oligonucleótido al que va unida.

Mezcla para la reacción de hibridación

- ADN amplificado biotinilado 2,5 µl.
- Mezcla de oligosondas suministradas por la casa comercial (Probe Mix, Lifecodes HLA-SSO Typing Kit, Gen-Probe, CA, EE.UU.), 7,5 µl por muestra.

Condiciones de hibridación

La casa comercial tiene estándares acerca de las condiciones de amplificación aplicadas al termociclador Mastercycler egradient S (Eppendorf, Hamburgo, Alemania):

- 1 ciclo de desnaturalización inicial de 97°C durante 5 minutos.
- 1 ciclo de hibridación de 47°C durante 30 minutos.
- 1 ciclo de 56°C durante 10 minutos.

Finalizado este último ciclo se mantienen las muestras en el termociclador a 56°C.

Protocolo utilizado

- Se calienta la mezcla de oligosondas a una temperatura entre 55° y 60°C en baño durante 8 minutos para disolver los componentes de la mezcla.
- Se somete brevemente (15 segundos) a la mezcla de oligosondas a una limpieza ultrasónica en el baño Ultrasonic Cleaner (Nahita). Se agita en agitador vortex durante 15 segundos
- Se añade el ADN amplificado en una placa de 96 pocillos suministrada por la casa comercial
- Se añade la mezcla de oligosondas en cada uno de los pocillos con el ADN amplificado

Se introduce la placa tapada en el termociclador y se inicia el programa de hibridación.

3.4.4 Reacción de la estreptavidina-picoeritrina

En esta etapa los productos hibridados se incuban en el termociclador junto con la estreptavidina-picoeritrina. La estreptavidina presenta una gran afinidad por la biotina, por lo que se une a esta, de ahí que los productos amplificados estén biotinilados.

Mezcla para la reacción

- Estreptavidina-Picoeritrina (SA-PE, Lifecodes HLA-SSO Typing Kit, Gen-Probe, CA, EE.UU.), 0,43 µl por muestra.
- Solución diluyente (Dilution Solution, Lifecodes HLA-SSO Typing Kit, Gen-Probe, CA, EE.UU.), 100 µl por muestra.

Protocolo utilizado

- Mientras las muestras se encuentra en la etapa de hibridación, se prepara la mezcla solución diluyente/SA-PE para n+1 muestras. Esta mezcla ha de

mantenerse a temperatura ambiente y en oscuridad, ya que la estreptavidina es fotosensible.

- En la etapa final de la hibridación, se diluyen las muestras que están en el termociclador con la mezcla solución diluyente/SA-PE.
- Se saca la placa con las muestras del termociclador y se introduce en el Luminex.

3.4.5 Medición por fluorimetría en Luminex

La última fase consiste en la medición de los productos amplificados, hibridados y unidos a SA-PE en el citómetro Luminex 100. Como se ha explicado con anterioridad el Luminex presenta dos láseres que emiten luz a diferentes longitudes de onda, una 635 nm y otro a 532 nm. El primero excita las microesferas, emitiendo cada una de estas una señal de fluorescencia determinada entre el rojo y el infrarrojo que el detector del aparato identifica con un código (al determinar el ratio entre los dos colorantes) que a su vez identifica el oligonucleótido que va unido a la microesfera y por lo tanto identifica el polimorfismo. El segundo láser excita la picroeritrina sólo si está unida al ADN que a su vez está unido a las oligosondas que identifican el polimorfismo, a través de la unión estreptavidina-biotina.

3.4.6 Interpretación de los resultados

El sistema Luminex contiene un programa informático, Luminex100 IS software (Luminex Corporation, TX, USA), el cual reconoce las emisiones de fluorescencia de las oligosondas si están unidas a ADN, es decir, si la emisión de fluorescencia de éstas es simultánea a la emisión de fluorescencia de la picroeritrina, a su vez unida al ADN. De tal manera que se obtiene un patrón de sondas específico para cada alelo. Estos datos son recogidos e identificados por el software Quicktype for Lifematch v2.6.1 (Gen-Probe, CA, EE.UU.) que asigna el alelo que corresponde con dicho patrón al enfrentarlo con la base de datos Lifematch suministrada por la casa comercial empleada para este fin.

En el caso de obtener un patrón de sondas anormal, el programa informático nos notificará que no ha encontrado ningún “tipaje” de acuerdo al patrón de sondas detectado. En ese caso, llevaremos a cabo el proceso de secuenciación del ADN de la muestra problema.

4. ANÁLISIS ESTADÍSTICOS

Los “tipajes HLA” de clase I y II caracterizados en la muestra poblacional étnica estudiada, se analizaron mediante distintos programas informáticos para el cálculo de los diferentes valores estadísticos. Finalmente, estos resultados, se compararon con los estudios de población de los 11° y 12° Talleres Internacionales de Histocompatibilidad (Imanishi et al., 1992a; Clayton & Lonjou, 1997) y los reportados en la literatura.

4.1 Cálculo de frecuencias alélicas

La frecuencia de cada alelo se obtuvo mediante el programa Arlequin v2.000, que realiza un recuento de cada variante alélica (alelo) y lo divide entre el número total de variantes alélicas encontradas en nuestras poblaciones objetos de estudio (Schneider et al., 2000).

4.2 Estimación de los haplotipos extendidos

Las muestras de las poblaciones analizadas corresponden a individuos no emparentados. Por tanto, no se conocen cuales son las asociaciones alélicas o haplotipos reales formados por los alelos HLA de acuerdo a la segregación familiar. Para suplir este inconveniente, se establece un estudio estadístico que estima las asociaciones o haplotipos formados por alelos de diferentes loci.

El cálculo de haplotipos, su desequilibrio de ligamiento, la frecuencia de los mismos y el valor estadístico de significación p para cada asociación se realizó mediante el programa informático Arlequin v2.000 (Schneider et al., 2000) y v3.0b (Excoffier et al., 2005). Este programa, combina el algoritmo de expectación-maximización (EM) con el método de máxima probabilidad (maximum-likelihood). Este es un proceso iterativo que persigue estimados de máxima probabilidad de frecuencias haplotípicas, obtenidos a partir de datos genotípicos de varios loci cuando su distribución gamética es desconocida. El principio del algoritmo EM es el siguiente:

- 1) Cálculo inicial aleatorio de frecuencias haplotípicas.
- 2) Uso de estos estimados para computar las frecuencias genotípicas esperadas para cada fenotipo, asumiendo el equilibrio de Hardy-Weinberg (paso de expectación).
- 3) Las frecuencias genotípicas relativas sirven como pesos para los dos haplotipos que se constituyen en un procedimiento de conteo génico, dando lugar a nuevos estimados de frecuencias haplotípicas (paso de maximización).

- 4) Se repiten los pasos 2 y 3 hasta que las frecuencias de los haplotipos alcancen el equilibrio.

Los haplotipos extendidos HLA-A-B-DRB1-DQB1 fueron deducidos a partir de los criterios de selección de los 11° y 12° Talleres Internacionales de Histocompatibilidad (Imanishi et al., 1992a; Clayton & Lonjou, 1997):

- a) Haplotipos extendidos que se encuentren en al menos dos individuos y que hayan sido previamente definidos en otras poblaciones.
- b) Haplotipos extendidos que no habiendo sido definidos previamente en otras poblaciones, aparezcan en al menos dos individuos y que el haplotipo alternativo resultante pueda ser biológicamente posible.

4.3 Estimación del desequilibrio de ligamiento

Dos o más alelos de loci diferentes se consideran asociados o en desequilibrio de ligamiento cuando aparecen juntos en un mismo cromosoma con mayor frecuencia de la que cabría esperar por azar (Reich et al., 2001). El desequilibrio de ligamiento (D o Δ) se define matemáticamente como:

$$D(A_i B_j) = FH(A_i B_j) - a_i \cdot b_j$$

donde FH es la frecuencia del haplotipo $A_i B_j$ en la población, a_i es la frecuencia del alelo i del locus A , y b_j es la frecuencia del alelo j del locus B . Las parejas de alelos con $D \leq 0$ tienden a no estar juntos formando haplotipos, mientras que las asociaciones con $D > 0$ indican que los alelos que componen los haplotipos tienden a formar estos y no otros. En cuanto al desequilibrio de ligamiento relativo (D') -cociente de D entre la D máxima teórica para las frecuencias alélicas observadas (D_{\max}), ver apartado de introducción 7.2- los alelos con $D' \leq 0$ tienden a no estar asociados, mientras que las asociaciones con $D' > 0$ indican que los alelos que componen los haplotipos tienden a asociarse siempre entre ellos. Los haplotipos con valores de $D > 0$ pueden a su vez tener una probabilidad estadísticamente significativa para un grado de libertad de $p < 0,05$, es decir, existe un 95% de seguridad de que la asociación de alelos que forman el haplotipo es real y no se debe al azar. En este trabajo se describen aquellos haplotipos de dos loci con desequilibrio de ligamiento positivo ($D > 0$, $D' > 0$) y $p < 0.05$.

Para calcular el desequilibrio de ligamiento entre dos alelos de dos loci diferentes (haplotipos bi-locus) entre parejas de loci HLA-A-B, HLA-B-DRB1 y HLA-DRB1-DQB1, se empleó el programa Arlequin v3.0b (Excoffier et al., 2005).

4.4 Distancias genéticas y dendrogramas

Estos parámetros se calcularon mediante el uso del paquete informático DISPAN (Ota T, Universidad del Estado de Pennsylvania). Para utilizar este programa, se creó previamente una matriz con las frecuencias alélicas de los loci seleccionados (HLA-DRB1 y HLA-DQB1) de todas las poblaciones (incluidas las del estudio) que deseábamos comparar (Tabla M1). Así, mediante el subprograma GNKDST contenido en el DISPAN se calcularon las distancias genéticas D_A (Nei, 1972) mediante la comparación de las frecuencias alélicas de todos los loci de todas las poblaciones incluídas en el análisis entre sí. Para ellos se emplea el siguiente algoritmo:

$$D_{Aij} = 1 - \sum \sqrt{p_{ni} \cdot p_{nj}}$$

donde D_{Aij} es la distancia genética entre las poblaciones i y j , p_{ni} es la frecuencia del alelo n en la población i , y p_{nj} es la frecuencia del alelo n en la población j . Además de calcular las distancias genéticas, el GNKDST nos da más información, como los valores de la heterocigosidad esperada en la población total (H_T), es decir, la medida de la diversidad genética de todas las poblaciones analizadas como una sola para cada locus y para todos los loci, lo que da una idea del grado del polimorfismo del gen; el valor de la heterocigosidad media esperada en las subpoblaciones (H_S), entendiendo subpoblación como cada una de las distintas poblaciones, para cada locus y para todos los loci, lo que da una idea de la variabilidad genética de la población; y el valor de la diferenciación genética relativa entre poblaciones (G_{ST}), que se define como la proporción de diversidad genética que existe entre poblaciones para cada locus y para todo los loci estudiados (Nei, 1973). Teniendo como referencia las frecuencias alélicas de los loci estudiados y las distancias D_A que se deducen de estas frecuencias alélicas y a través del subprograma TREEVIEW se obtuvo una representación gráfica en forma de dendrograma mediante el método NJ o de “unión de vecinos” (Saitou & Nei, 1987). Este método parte de una matriz de distancias genéticas (en este caso distancias D_A) entre taxones (en este caso poblaciones) y aplica un algoritmo que comienza con un árbol sin resolver y lleva a cabo una serie de pasos con un número de iteraciones necesarias para resolverlo y conocer las longitudes de sus ramas. Los pasos son los siguientes:

- 1) A partir de la matriz de distancias genéticas se calcula la matriz Q .
- 2) Se encuentra el par de poblaciones con el valor más bajo en la matriz Q . Se añade un nuevo nodo al árbol para unir estas dos poblaciones.

- 3) Se calcula la distancia de cada una de las poblaciones del par al nuevo nodo.
- 4) Se calcula la distancia de cada una de las poblaciones analizadas (fuera del par) a este nuevo nodo.
- 5) Se comienza de nuevo el algoritmo (nueva iteración) sustituyendo el par de vecinos unidos por el nuevo nodo y usando las distancias calculadas en el paso 4 para construir la nueva matriz Q . Se repiten los pasos anteriores sucesivas veces hasta completar el árbol.

Para calcular la matriz Q se aplica la siguiente fórmula:

$$Q(i, j) = (r - 2)d(i, j) - \sum d(i, k) - \sum d(j, k)$$

donde r es el número de poblaciones, $d(i, j)$ es la distancia entre las poblaciones i y j , y k es cualquier otra población no i ni j respectivamente. Además, para cada dendrograma se realizó el test de Bootstrap (Felsenstein, 1985) que nos da idea de la bondad estadística del árbol mediante un proceso iterativo. Para las poblaciones incluidas en el análisis, se construyeron tres dendrogramas a partir de las distancias D_A calculadas teniendo en cuenta los siguientes conjuntos de datos:

- a) Frecuencias alélicas del locus HLA-DRB1 (alta resolución).
- b) Frecuencias alélicas de los loci HLA-DRB1 y HLA-DQB1 (alta resolución).
- c) Frecuencias haplotípicas de haplotipos HLA-DRB1-DQB1 (alta resolución).

En el análisis comparativo de las diferentes poblaciones, no fue posible realizar estudios con los alelos HLA de clase I individualmente y/o en conjunto con clase II, ya que alguna de ellas, no cuentan con el “tipaje HLA” de clase I (por ejemplo: Bereberes procedentes de Souss y de El Jadida, Marruecos; Judíos asquenazí, no asquenazí, y procedentes de Marruecos; Libaneses, y muchas de las poblaciones de Siberia, ver Tabla M1). Además, el “tipaje genérico HLA” (baja resolución) de clase I (el único disponible para la mayoría de poblaciones tipadas para estos loci) tiende a homogeneizar el análisis obtenido a partir de las frecuencias alélicas de HLA-DRB1 de alta resolución.

4.5 Análisis de Correspondencias

El Análisis de Correspondencias es una técnica estadística empleada para analizar desde un punto de vista gráfico las relaciones de dependencia o semejanza de un conjunto de variables categóricas a partir de los datos de una tabla de contingencia (tabla de correspondencias). Para ello asocia a cada una de las modalidades de las variables un punto en el espacio R^n (n =dimensiones) de forma que la distancia entre los

mismos refleja las relaciones de dependencia/ semejanza existentes entre ellas (Salvador Figueras, 2003).

Este análisis se realizó con el programa informático VISTA (del inglés, *Visual Statistics System*) y en este caso consiste en la representación geométrica n-dimensional de las diferencias poblacionales en base a la varianza (dispersión de los valores respecto al valor promedio) de las distancias genéticas entre todas las poblaciones analizadas (Young & Bann, 1996). Se basa en la reducción de este sistema multivariante (matriz de distancias genéticas) en unas pocas “componentes principales” o dimensiones, que cada una de ellas podrá explicar mayor o menor cantidad de esta varianza. La representación gráfica se lleva a cabo con las dos dimensiones que mayor información de la varianza puedan dar, basándonos en el valor de la inercia principal, que mide el grado total de similitud existente entre las modalidades de las variables, y sus proporciones explicada y acumulada, que ayudan a calibrar la importancia de cada dimensión a la hora de explicar las dependencias o similitudes observadas (Salvador Figueras, 2003). Se realizaron los análisis de correspondencias con las frecuencias alélicas de los loci de clase II, HLA-DRB1 y de HLA-DRB1 y HLA-DQB1 conjuntamente, así como con las frecuencias haplotípicas de HLA-DRB1-DQB1, siempre en alta resolución.

4.6 Estimación de la proporción de mestizaje

Estas estimaciones se llevaron a cabo mediante el uso del programa informático LEADMIX v1.0 basado en el método de máxima verosimilitud desarrollado por Wang (Wang, 2003) que permite calcular las proporciones de mestizaje de poblaciones a través de marcadores genéticos. Este modelo asume que una población ancestral, P_0 , se divide en d poblaciones parentales (P_k , $k=1 \sim d$) que evolucionan de forma independiente durante un número de generaciones. En este punto una población híbrida, P_h , se origina por la combinación de genes en una proporción p_k aleatoria procedente de las poblaciones parentales P_k . Durante el número de generaciones sucedidas entre la hibridación (mestizaje) y la toma de muestras de la P_h y de cada una de las P_k se supone que las poblaciones permanecen aisladas y no intercambian genes entre sí. De este modo, este modelo define la proporción de mestizaje $d-1$ (p_k , $k = 1 \sim d-1$).

El éxito de los análisis de mestizaje reside en la elección de las poblaciones parentales a través de datos de diversos campos de estudios (historia, arqueología, genética, etc.) con el fin de seleccionar poblaciones similares a esas posibles poblaciones fundadoras. En nuestro caso se han seleccionados tres poblaciones

parentales atendiendo a los tres posibles orígenes de las poblaciones de estudio: mediterráneo, caucásico y centroasiático. La población cretense se ha seleccionada como población parental mediterránea debido a su demostrado sustrato genético mediterráneo y su relación tanto con los pueblos norteafricanos como con los pueblos del mediterráneo oriental (macedonios, judíos, libaneses; Arnaiz-Villena et al., 1999; Gómez-Casado, 2002). En términos del posible ancestro caucásico de las poblaciones de estudio, los georgianos conservan el “pool” genético de los primeros habitantes de la región, relacionados con los pueblos de Mesopotamia y Anatolia desde tiempo remotos (Rosen, 1992). En lo que respecta al tercer componente, el centroasiático, la región Altái/Baikal es la que se ha postulado como lugar de origen de gran parte de los linajes genéticos euroasiáticos, incluidos los de los del suroeste de Asia (Wells et al., 2001). Por esto se ha seleccionado la población todja, grupo altaico relacionado con los grupos ancestrales indo-iranios que pudieron colonizar la región en la que se encuentran las poblaciones de estudio, debido a la presencia del haplogrupo R-M17 en esta población (haplogrupo asociado con la expansión de la cultura Andrónovo, véase Introducción apartado 12.3; Derenko et al., 2006; Figura D3). Estos análisis se han llevado a cabo utilizando los alelos de los loci HLA de clase II (DRB1 y DQB1), tanto de las poblaciones parentales como de las “poblaciones híbridas” (poblaciones de estudio). Los alelos de los loci HLA de clase I no han podido incluirse en los análisis debido a que la población parental centroasiática no cuenta con esta información.

RESULTADOS

1. ALELOS DEL SISTEMA HLA ENCONTRADOS EN LAS POBLACIONES DE ESTUDIO

Se han obtenido los perfiles genéticos HLA mediante técnicas de genética molecular en una muestra poblacional de 226 individuos sanos no relacionados. Éstos fueron seleccionados según su grupo étnico en las ciudades de Gorgan (turcomanos, 69 individuos), Tabriz (azeríes, 97 individuos), Saqqez y Baneh (kurdos, 60 individuos). Sus familias han vivido en estas ciudades desde hace varias generaciones; son por tanto, parte representativa de cada una de las poblaciones sometidas a estudio: turcomanos, azeríes y kurdos de Irán.

1.1 Distribución de los alelos HLA de clase I en las poblaciones de estudio

En la Tabla R1 se muestran las frecuencias alélicas para los loci HLA de clase I (HLA-A y -B) obtenidos en las muestras de las poblaciones analizadas. Los perfiles genéticos en alta resolución se han obtenido mediante las técnicas especificadas en el apartado de Materiales y Métodos. La mayor diversidad alélica para ambos loci la presentan los turcomanos, y la menor los kurdos.

Locus HLA-A

Se han obtenido 30 alelos para el locus A en población turcomana, 22 en población azerí y 21 en población kurda. Del total de alelos obtenidos, 35, el 49% está presente en las tres poblaciones, incluyendo los alelos más frecuentes, siendo los alelos más raros los no compartidos. Los alelos que se encuentran con mayor frecuencia en las tres poblaciones analizadas son el A*02:01 (13.4% en azeríes –el más frecuente en esta población–, 12.5% en kurdos y 11.6% en turcomanos) y el A*24:02 (17.6% en kurdos – el más frecuente en esta población–, 12.4% en azeríes y 11.6% en turcomanos), alelos ampliamente distribuidos por toda la geografía mundial. Otros alelos que compartes frecuencias significativas en las tres poblaciones son el A*03:01 (10.8% en kurdos, 9.8 en azeríes y 7.3% en turcomanos), frecuente en poblaciones europeas, y el A*11:01 (8.3% en kurdos, 6.2 en azeríes y 5.8% en turcomanos), predominante en poblaciones del Pacífico y sur de Asia. El alelo A*01:01 es bastante frecuente en azeríes y kurdos (11.9 y 8.3% respectivamente), pero poco significativo en turcomanos (2.9%), donde el alelo A*02:50, descrito raramente en otras poblaciones aparece con una frecuencia relevante (7.3%).

TABLA R1. Frecuencias alélicas de los loci HLA de clase I (HLA-A y -B) en las poblaciones estudiadas. Los alelos más frecuentes figuran en negrita.

Alelo		Frecuencia alélica			Alelo		Frecuencia alélica		
HLA-A	Turcomanos	Azerís	Kurdos	HLA-B	Turcomanos	Azerís	Kurdos		
01:01	2.9	11.9	8.3	07:02	3.6	4.6	1.7		
01:02	4.4	6.2	2.5	07:03	3.6	—	—		
02:01	11.6	13.4	12.5	07:24	—	—	0.8		
02:34	3.6	—	—	08:01	6.5	3.0	5.0		
02:48	0.7	—	1.7	13:01	7.9	3.6	3.3		
02:50	7.3	4.1	5.8	13:04	0.7	—	—		
02:52	2.9	0.4	6.7	14:01	0.7	3.1	0.8		
02:61	0.7	—	1.7	14:02	1.5	—	—		
03:01	7.3	9.8	10.8	15:01	—	2.1	—		
03:02	1.5	1.6	0.8	15:02	1.5	—	—		
11:01	5.8	6.2	8.3	15:10	0.7	0.5	—		
13:01	—	—	0.8	15:17	0.7	0.5	0.8		
23:01	2.1	0.4	2.5	15:24	1.5	0.5	0.8		
24:02	11.6	12.4	17.6	15:46	0.7	—	—		
24:03	—	1.6	0.8	18:01	1.5	5.2	5.8		
24:08	0.7	—	—	18:09	0.7	0.5	3.3		
25:01	1.5	—	—	27:01	2.1	1.6	0.8		
26:01	5.1	2.6	7.5	35:01	19.6	15.0	22.0		
26:13	1.5	—	—	35:05	—	0.5	—		
26:19	—	0.5	—	35:10	1.5	0.5	—		
29:01	5.8	3.6	—	35:19	—	0.5	—		
29:02	0.7	—	—	35:20	1.5	—	0.8		
30:01	5.1	5.2	1.7	35:27	2.9	—	1.7		
31:01	3.6	1.6	1.7	35:38	0.7	—	—		
31:02	0.7	—	—	37:01	1.5	1.6	—		
31:07	0.7	—	—	38:01	2.9	4.6	4.2		
32:01	2.1	5.7	5.0	39:01	2.9	1.0	—		
32:03	0.7	—	—	40:01	2.1	1.6	1.7		
32:04	1.5	—	—	40:02	1.5	—	—		
33:01	3.6	5.7	1.7	40:15	—	—	0.8		
34:02	—	0.4	—	40:39	1.5	—	—		
66:01	—	0.4	—	40:46	1.5	—	—		

TABLA R1. Continuación.

66:03	0.7	—	—	41:01	—	5.2	0.8
68:01	2.1	5.7	0.8	44:02	1.5	5.2	5.0
69:01	1.5	0.6	0.8	44:10	—	—	0.8
				44:18	—	0.5	—
				46:01	—	0.5	—
				48:01	1.5	—	—
				49:01	2.1	4.1	5.0
				50:01	1.5	4.1	5.8
				51:01	6.5	8.8	9.2
				51:06	3.6	1.0	0.8
				51:07	—	0.5	0.8
				51:08	0.7	1.0	—
				52:01	0.7	8.3	4.2
				53:01	0.7	0.5	2.5
				53:07	0.7	—	—
				54:01	0.7	0.5	0.8
				55:01	1.5	3.6	5.0
				56:01	—	0.5	—
				57:01	2.1	1.6	—
				58:01	1.5	2.6	5.0
				58:06	0.7	—	—
				58:08	—	0.5	—
				78:01	—	0.5	—

Locus HLA-B

Para este locus, el más polimórfico de todo el sistema HLA, se han obtenido 55 alelos en total, de los cuales el 36.4% son compartidos por las tres poblaciones. En el caso de los turcomanos, las frecuencias de sólo cuatro alelos suman más del 40% de los 42 alelos existentes en la muestra analizada (asciende al 52.1% si contamos los grupos de alelos). Algo similar sucede en la muestra de azeríes, en la que la frecuencia de seis alelos suman casi la mitad (47.7%) del total de los alelos obtenidos (38 alelos). Para la muestra de kurdos se han obtenido 29 alelos. El alelo que aparece con mayor frecuencia en las tres poblaciones es B*35:01 (22.0% en kurdos, 19.6% en turcomanos y 15.0% en azeríes), ampliamente distribuido por la geografía mundial con frecuencias discretas, salvo en Asia Central y Centroamérica, donde son más altas.

En el caso de los turcomanos, los otros tres alelos con frecuencias significativas mencionados previamente son: B*13:01 (7.9%), predominante en poblaciones del Pacífico y sur de Asia; B*51:01 (6.5%), ampliamente extendido por el sur de Asia y Oriente Próximo; y B*08:01 (6.5%), frecuente en poblaciones de Europa Atlántica y Asia Central. En cuanto a las otras poblaciones, el alelo B*51:01 también muestra una frecuencias considerables en ambas (9.2% para kurdos y 8.8% para azeríes), así como el B*52:01 en azeríes (8.3%), frecuente en Asia Oriental, y los alelos B*18:01 y B*50:01 en kurdos (ambos 5.8%), predominantes en Europa Central y Cáucaso, y en Norte de África y Oriente Próximo, respectivamente.

1.2 Distribución de los alelos HLA de clase II en las poblaciones de estudio

En la Tabla R2 se muestran las frecuencias alélicas para los loci HLA de clase II (HLA-DRB1 y -DQB1) obtenidos en las muestras de las poblaciones analizadas. Los perfiles genéticos en alta resolución se han obtenido mediante las técnicas especificadas en el apartado de Materiales y Métodos.

Locus HLA-DRB1

Se han obtenido 33 alelos para este locus, de los cuales el 36.4% son compartidos por las tres poblaciones. Los alelos que presentan las mayores frecuencias (por encima del 8%) son los mismos en todas las poblaciones, aunque los alelos más frecuentes para cada una de ellas no coinciden. En el caso de los turcomanos se han determinado 24 alelos, siendo el más frecuente el DRB1*15:01 (21.0%, en azeríes 15.0% y en kurdos 10.0%), alelo de amplia distribución en poblaciones asiáticas y europeas (especialmente en Europa Atlántica). Con respecto a la muestra de azeríes se han obtenido 20 alelos, de los cuales el más frecuente es el DRB1*04:01 (15.5%, en kurdos 12.5% y en turcomanos 10.9%), muy frecuente en poblaciones de Siberia. Se han determinado 17 alelos distintos para la muestra de kurdos, cuyo alelo mas frecuente es el DRB1*11:01 (19.1%, en azeríes 15.0% y en turcomanos 12.3%). Este alelo es muy frecuente en poblaciones de Irán y del Cáucaso (armenios, georgianos, persas, turcos, Figura I18).

Otros alelos frecuentes en turcomanos, azeríes y kurdos son: DRB1*03:01 (10.2%, 8.8% y 15.8% respectivamente) y DRB1*07:01 (13.8%, 10.8% y 6.7% respectivamente), predominantes en la cuenca mediterránea, Europa Atlántica y Asia Central; y DRB1*01:01 (5.8%, 4.6% y 7.5% respectivamente), cuyas máximas frecuencias se encuentra en Europa del Norte y Oriental con un gradiente descendiente hacia el Cáucaso y el “corredor iraní”.

TABLA R2. Frecuencias alélicas de los loci HLA de clase II (HLA-DRB1 y -DQB1) en las poblaciones estudiadas. Los alelos más frecuentes figuran en negrita.

Alelo HLA-DRB1	Frecuencia alélica			Alelo HLA-DQB1	Frecuencia alélica		
	Turcomanos	Azerís	Kurdos		Turcomanos	Azerís	Kurdos
01:01	5.8	4.6	7.5	02:01	21.0	16.5	20.8
01:02	—	—	0.8	02:02	—	—	0.8
01:03	—	—	1.7	02:03	0.7	—	—
03:01	10.2	8.8	15.8	03:01	21.0	20.1	24.2
03:02	—	0.5	—	03:02	7.3	8.8	6.7
03:03	0.7	—	—	03:03	0.7	4.6	—
03:08	0.7	0.5	—	03:04	1.5	—	—
03:13	2.2	—	—	03:05	0.7	2.1	4.2
04:01	10.9	15.5	12.5	03:07	2.8	2.6	1.7
04:15	—	—	0.8	03:12	—	0.4	—
07:01	13.8	10.8	6.7	04:01	5.0	5.7	2.5
08:01	2.9	3.0	1.7	04:02	—	—	0.8
08:20	2.2	—	—	05:01	13.0	16.5	10.8
09:01	—	1.6	—	05:02	1.5	3.6	7.5
10:01	3.6	3.6	1.7	05:03	1.5	—	—
11:01	12.3	15.0	19.1	06:01	2.2	8.3	5.0
11:17	0.7	—	—	06:02	11.6	5.2	6.7
11:20	0.7	—	—	06:03	7.3	4.1	5.0
11:28	0.7	—	—	06:04	2.2	1.5	3.3
12:01	1.5	1.6	0.8				
12:09	1.5	—	—				
13:01	3.6	4.1	5.0				
13:02	1.5	1.6	4.2				
13:03	—	1.0	—				
13:05	—	0.5	0.8				
13:06	0.7	—	—				
13:08	0.7	0.5	—				
14:01	0.7	6.7	6.7				
14:02	—	1.0	—				
14:05	0.7	—	—				
14:24	0.7	—	—				
15:01	21.0	15.0	10.0				
16:01	—	4.1	4.2				

Locus HLA-DQB1

Para el locus menos polimórfico de los analizados se han obtenido un total de 19 alelos distintos, de los que 12 son compartidos por las tres poblaciones (63.2%). Con respecto a los turcomanos se han determinado 16 alelos, de los cuales los cuatro más frecuentes suman el 66.6% del total de alelos. En el caso de los azeríes y los kurdos, se han obtenido 14 alelos distintos para este locus, y los cuatro alelos más frecuentes suman el 61.9% y el 63.3% del total, respectivamente. El alelo más frecuente en las tres poblaciones es DQB1*03:01 (24.2% en kurdos, 21.0% en turcomanos y 20.1% en azeríes), con amplia distribución en la geografía mundial, mayoritario en poblaciones siberianas y amerindias, y con frecuencias elevadas en poblaciones del Cáucaso y del “corredor iraní”. Otros alelos con frecuencias altas son: DQB1*02:01 (21.0% en turcomanos, 20.8% en kurdos y 16.5% en azeríes), predominante en Europa Atlántica, norte de África, “corredor iraní” y sur de Asia; DQB1*05:01 (16.5% en azeríes, 13.0% en turcomanos y 10.8% en kurdos), también con una amplia distribución mundial, pero predominante en población caucasoide.

2. ESTUDIOS ANALÍTICOS BASADOS EN ALELOS HLA DE CLASE II

Las distancias genéticas D_A se calcularon por el método de Nei (Nei, 1972) mediante la comparación de las frecuencias alélicas de los genes HLA en las poblaciones de estudio (turcomanos, azeríes y kurdos) respecto a otras poblaciones del mundo (Tabla M1 y Figura M1). La construcción de los dendrogramas se realizó mediante el método NJ a partir de las distancias genéticas (Saitou & Nei, 1987). Los análisis de correspondencias se obtuvieron a partir de las distancias genéticas, y proporcionan una idea global de las relaciones filogenéticas entre las poblaciones comparadas (véanse apartados 5.5 y 5.6 de Materiales y Métodos).

Estos análisis se llevaron a cabo con dos conjuntos de datos: las frecuencias alélicas del locus HLA-DRB1 y de los loci HLA-DRB1 y HLA-DQB1 de manera conjunta, ambas en alta resolución. Para estos análisis no fue posible la inclusión de los datos de las frecuencias alélicas HLA de clase I, por la falta de este tipo de información en muchas de las poblaciones mediterráneas (bereberes, judíos -asquenazí, no asquenazí y procedentes de Marruecos-, libaneses), siberianas (buriatos, tuvanos, tofalar, todja, evenki, keti entre otros) y del Pacífico (poblaciones de Papua y Samoa entre otras) empleadas en los análisis de emparentamiento genético.

2.1 Análisis basados en el locus HLA-DRB1

A partir de las frecuencias alélicas del gen HLA-DRB1 (alta resolución) se han calculado las distancias genéticas entre las distintas poblaciones de estudio y el resto de poblaciones (Tabla R3). Este análisis permite comparar un gran número de poblaciones analizadas en alta resolución para este locus y ver en conjunto la relación de las poblaciones de la región sur del Mar Caspio objeto de estudio con el resto de poblaciones con proximidad geográfica, como otras poblaciones iraníes, mediterráneas y de Asia Central, y con otras más alejadas en el espacio, como poblaciones siberianas y del Pacífico.

En primer lugar, atendiendo a las distancias genéticas D_A , se observa que las tres poblaciones están muy próximas entre sí: turcomanos con respecto a los azeríes están a una distancia de 12.02×10^{-2} , y con respecto a los kurdos a 13.20×10^{-2} , mientras que los dos últimos están a solamente 6.12×10^{-2} . En el caso de los turcomanos, las poblaciones más próximas (después de azeríes y kurdos) son los chuvasios (19.94×10^{-2}) y los rusos (20.20×10^{-2}), seguidos de poblaciones mediterráneas-europeas (franceses, vascos españoles, croatas, italianos), poblaciones del Cáucaso (esvanos y georgianos) y de Asia Central o con origen en esta zona (aleutianos, tuvanos, mansi y todja). Más alejadas están sin embargo las poblaciones persas iraníes, las poblaciones del Mediterráneo sur oriental, las poblaciones negroides, orientales, siberianas y na-dene, y finalmente las poblaciones del Pacífico y aborígenes australianos (estos últimos con distancias por encima de 90×10^{-2}). En lo que respecta a la población azerí, los grupos más próximos (después de turcomanos y kurdos) pertenecen a poblaciones europeas caucásicas (franceses -16.10×10^{-2} -, rusos -16.23×10^{-2} -, croatas, georgianos), poblaciones de Asia Central (tuvan, todja, tofalar) y poblaciones túrquicas (chuvasios y turcos). Como en el caso de los turcomanos, se encuentran más alejadas las poblaciones mediterráneas (sobre todo las poblaciones judías), las poblaciones persas iraníes, las poblaciones negroides, siberianas, na-dene y orientales, y en último lugar las poblaciones del Pacífico y aborígenes australianos. Las poblaciones más próximas a la población kurda según las distancias genéticas basadas en el locus HLA-DRB1 son los azeríes, franceses (12.41×10^{-2}), turcomanos y rusos (14.75×10^{-2}), seguidas de otras poblaciones mediterráneas-europeas (italianos, croatas, vascos españoles), poblaciones del Cáucaso (esvanos y georgianos), persas de Shiraz y poblaciones túrquicas (chuvasios y turcos). El patrón que siguen el resto de poblaciones con respecto a los kurdos es muy similar al que siguen con las otras dos poblaciones de estudio, con la salvedad de que las

poblaciones de Asia Central en este caso se encuentran más alejadas. Todo esto se puede observar con mayor detalle en la Tabla R3. La distancia genética media más alta de las tres poblaciones es la que corresponde a la población turcomana con 45.39×10^{-2} , bastante superior a la de los kurdos (40.65×10^{-2}) y a la de los azeríes (39.06×10^{-2}).

Los dendrograma (Figuras R1.A y R2), basados en las distancias genéticas del locus HLA-DRB1, muestra agrupadas a las tres poblaciones de estudio en un mismo conjunto. En el primer dendrograma (Figura R1.A), construido mediante las distancias genéticas de las poblaciones más próximas geográficamente, estas poblaciones se encuentran situadas en el primer gran conjunto de los dos en los que se divide el árbol, que agrupa a las poblaciones mediterráneas-europeas, mediterráneas sur orientales, caucásicas, persas y siberianas de Asia Central, en contraposición al otro gran grupo conformado por poblaciones siberianas y orientales. Estas tres poblaciones del sur del Mar Caspio se encuentran situadas entre poblaciones con un origen común procedentes de Siberia central, que habitan al norte del Mar Caspio y en los Urales, los chuvasios y los mansi respectivamente, y las poblaciones iraníes (persas –procedentes de Shiraz y Yazd– y baluchis), aunque la longitud de las ramas horizontales de éstas últimas indica una distancia genética relativamente grande en comparación a otras poblaciones más próximas como rusos, franceses o georgianos, aunque más alejadas en el árbol. En el segundo dendrograma (Figura R2), que incluye el resto de poblaciones que figuran en la Tabla M1, estas poblaciones se encuentran situadas, igual que en el anterior árbol, en el primer gran conjunto que agrupa a las poblaciones mediterráneas-europeas, mediterráneas sur orientales, caucásicas, persas, negroides, siberianas y na-dene, en yuxtaposición al otro gran conjunto que agrupa poblaciones orientales y del Pacífico y Australia. En este análisis, las tres poblaciones del sur del Mar Caspio se encuentran inmediatamente antes del conjunto formado por poblaciones siberianas emparentadas entre sí, como son chuvasios, aleutianos, mansi y buriatos, y el conjunto formado por persas y baluchis.

De la misma manera se realizaron análisis de correspondencias a partir de las distancias genéticas basadas en locus HLA-DRB1 únicamente con las poblaciones próximas geográficamente, y con todas las poblaciones incluidas en el estudio (Figuras R1.B y R3).

TABLA R3. Distancias genéticas (D_A) a partir de las frecuencias alélicas de HLA-DRB1 ($\times 10^{-2}$) entre las poblaciones analizadas (turcomanos, azeríes y kurdos) y el resto de poblaciones del mundo.

Turcomanos		Azeríes		Kurdos	
Población	D_A	Población	D_A	Población	D_A
Azeríes	12.02	Kurdos	6.12	Azeríes	6.12
Kurdos	13.20	Turcomanos	12.02	Franceses	12.41
Chuvasio	19.94	Franceses	16.10	Turcomanos	13.20
Rusos	20.20	Rusos	16.23	Rusos	14.75
Franceses	20.31	Aleutianos	17.67	Italianos	15.69
Vascos españoles	22.35	Croatas	17.74	Croatas	15.93
Aleutianos	23.72	Georgianos	18.93	Georgianos	15.98
Esvanos	23.88	Tuvanos	19.28	Esvanos	17.33
Tuvanos	24.74	Italianos	19.45	Vascos españoles	19.89
Mansi	25.34	Chuvasio	19.75	Persas S	20.25
Todja	25.61	Todja	21.12	Turcos	20.33
Georgianos	25.96	Turcos	21.75	Chuvasio	21.12
Croatas	26.39	Tofalar	22.51	Macedonios	21.65
Italianos	26.88	Esvanos	22.74	Españoles	21.72
Españoles	27.12	Españoles	22.98	Tuvanos	23.43
Bereberes	27.44	Mansi	23.21	Aleutianos	23.59
Hotentotes	27.48	Mongoles Khalka	24.02	Hotentotes	24.09
Mongoles Khalka	27.55	Macedonios	24.13	Cretenses	24.45
Buriatos	28.72	Vasco Españoles	24.38	Bereberes	25.62
Tofalar	29.55	Negros SA	24.93	Todja	25.66
Turcos	30.00	Hotentotes	25.83	Armenios	26.58
Negros SA	30.57	Persas S	26.25	Mansi	26.74
Persas S	32.56	Buriatos	27.04	Negros SA	27.94
Macedonios	32.62	Armenios	27.35	Mongoles Khalka	28.00
Negros NA	32.87	Cretenses	28.25	Tofalar	29.55
Cretenses	35.31	Negros NA	28.44	Baluchis	30.30
Armenios	35.42	Bereberes	29.62	Sardos	30.31
Persas Y	35.84	Manchúes	31.84	Negros NA	30.51
Mongoles Khoton	35.87	Argelinos	31.92	Argelinos	30.83
Argelinos	36.45	Palestinos	32.22	Persas Y	31.02
Marroquíes	36.58	Persas Y	34.59	Marroquíes	33.20
Coreanos	37.90	Marroquíes	35.59	Palestinos	33.32
Judíos NAs	39.34	Sardos	36.91	Judíos As	34.42
Palestinos	40.09	Coreanos	37.03	Judíos marroquíes	34.47
Baluchis	40.19	Judíos As	37.38	Buriatos	35.94

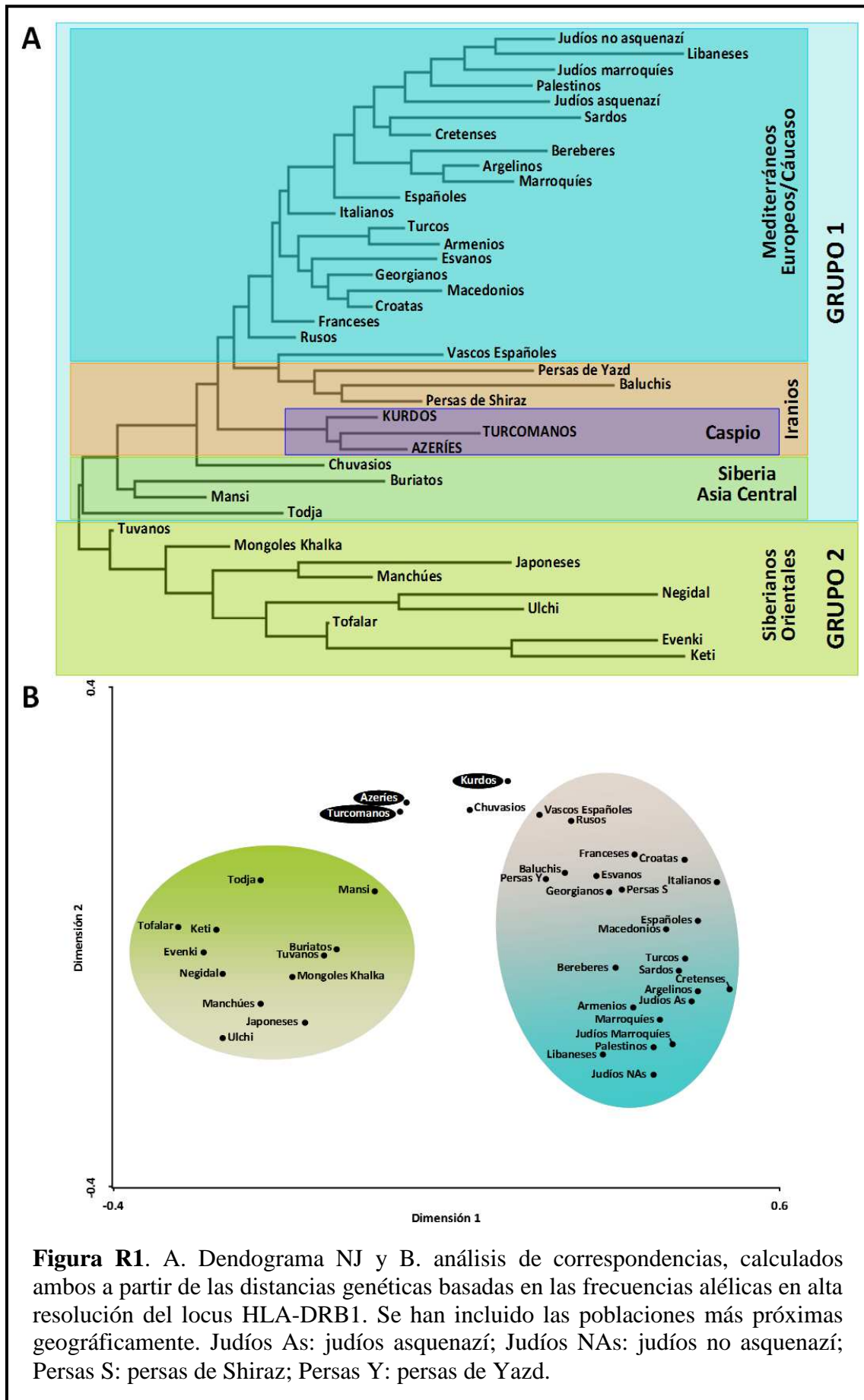
TABLA R3. Continuación.

Manchúes	41.32	Judíos marroquíes	38.26	Judíos NAs	36.57
Judíos marroquíes	41.33	Judíos NAs	38.39	Mongoles Khoton	36.74
Libaneses	42.64	Mongoles Khoton	38.46	Bosquimanos	38.03
Bosquimanos	44.50	Bosquimanos	38.53	Libaneses	38.76
Judíos As	44.61	Baluchis	38.59	Coreanos	39.47
Keti	46.27	Evenki	39.54	Manchúes	41.06
Evenki	46.53	Libaneses	39.72	Bubis	43.86
Ulchi	46.79	Ulchi	39.83	Nueva Caledonia	45.51
Sardos	47.05	Chukchis	42.75	Ulchi	47.53
Nueva Caledonia	47.40	Nueva Caledonia	43.08	Evenki	48.33
Chinos de Singapur	50.19	Negidal	44.06	Buyei	52.24
Bubis	52.16	Buyei	45.65	Negidal	52.40
Negidal	53.43	Bubis	45.86	Chinos de Singapur	52.51
Japoneses	54.40	Keti	46.18	Keti	52.81
Samoanos	55.61	Chinos de Singapur	46.64	Japoneses	53.01
Buyei	56.13	Japoneses	48.76	Samoanos	53.06
Chukchis	57.97	Samoanos	49.12	Chukchis	55.83
Fiyianos	58.61	Tlingit	50.47	Fiyianos	56.46
Tlingit	59.61	Esquimos	52.06	Madang	59.64
Madang	64.15	Atabascos	53.88	Tlingit	61.60
Goroka	65.92	Udegeys	54.35	Esquimos	61.83
Esquimos	68.89	Fiyianos	55.48	Rabaul	62.29
Atabascos	70.18	Koryaks	57.77	Atabascos	64.20
Taiwaneses	70.90	Madang	59.51	Udegeys	65.88
Udegeys	74.50	Nivjis	60.01	Taiwaneses	68.38
Rabaul	74.85	Rabaul	64.77	Ainu	69.96
Koryaks	76.55	Taiwaneses	64.82	Goroka	70.66
Nivjis	77.26	Ainu	65.72	Koryaks	71.98
Ainu	77.61	Goroka	66.86	Nivjis	74.09
Desierto Central	90.06	Yuendumu	82.75	Yuendumu	85.64
Yuendumu	92.16	Desierto Central	84.29	Desierto Central	87.29
Cape York	94.02	Cape York	88.36	Cape York	91.91
Kimberley	96.75	Kimberley	94.77	Kimberley	95.42

Judíos As: judíos asquenazí; Judíos NAs: judíos no asquenazí; Negros NA: negros de Norteamérica; Negros SA: negros de Sudáfrica; Persas S: persas de Shiraz; Persas Y: persas de Yazd.

En el primer análisis (Figura R1.B) se observan dos grandes conjuntos de poblaciones al igual que en el dendrograma, definidos por la primera dimensión que explica la mayor parte de la variabilidad entre poblaciones (proporción de inercia =

65.7%). El primer grupo incluye a las poblaciones siberianas y orientales, y el segundo está a su vez dividido en dos subgrupos definidos por la segunda dimensión (proporción de inercia acumulada = 78,5%), el primero formado por las poblaciones mediterráneas-europeas, del Cáucaso e iránicas, y el segundo por las poblaciones del mediterráneo sur y oriental. Las tres poblaciones de estudio se sitúan en una posición intermedia entre los dos grupos principales, junto con los chuvasios, donde los kurdos tienden hacia las poblaciones mediterráneas, y los turcomanos y azeríes hacia las poblaciones de Asia Central (mansí, todja). En el segundo análisis (Figura R3) la primera dimensión define tres grandes conjuntos (proporción de inercia = 63.5%), un primer grupo formado por las poblaciones siberianas y orientales (menos compacto que en el primer análisis), un segundo conjunto en posición central formado por las poblaciones negroides, y un tercer grupo formado por las poblaciones mediterráneas-europeas-caucásicas. La segunda dimensión (proporción de inercia acumulada = 78,9%) divide a su vez los dos conjuntos más grandes: en el primero las poblaciones siguen un gradiente geográfico aproximado del noroeste (Asia Central) al sureste (Australia y Pacífico), separando a las poblaciones siberianas de las orientales y del Pacífico, y en el segundo gran conjunto, las separa de la misma forma que en el primer análisis. De igual manera, las poblaciones del sur del Mar Caspio ocupan una situación intermedia (junto con chuvasios y mansí) entre siberianos y mediterráneos, pero con una mayor tendencia hacia estos últimos, en especial los kurdos.



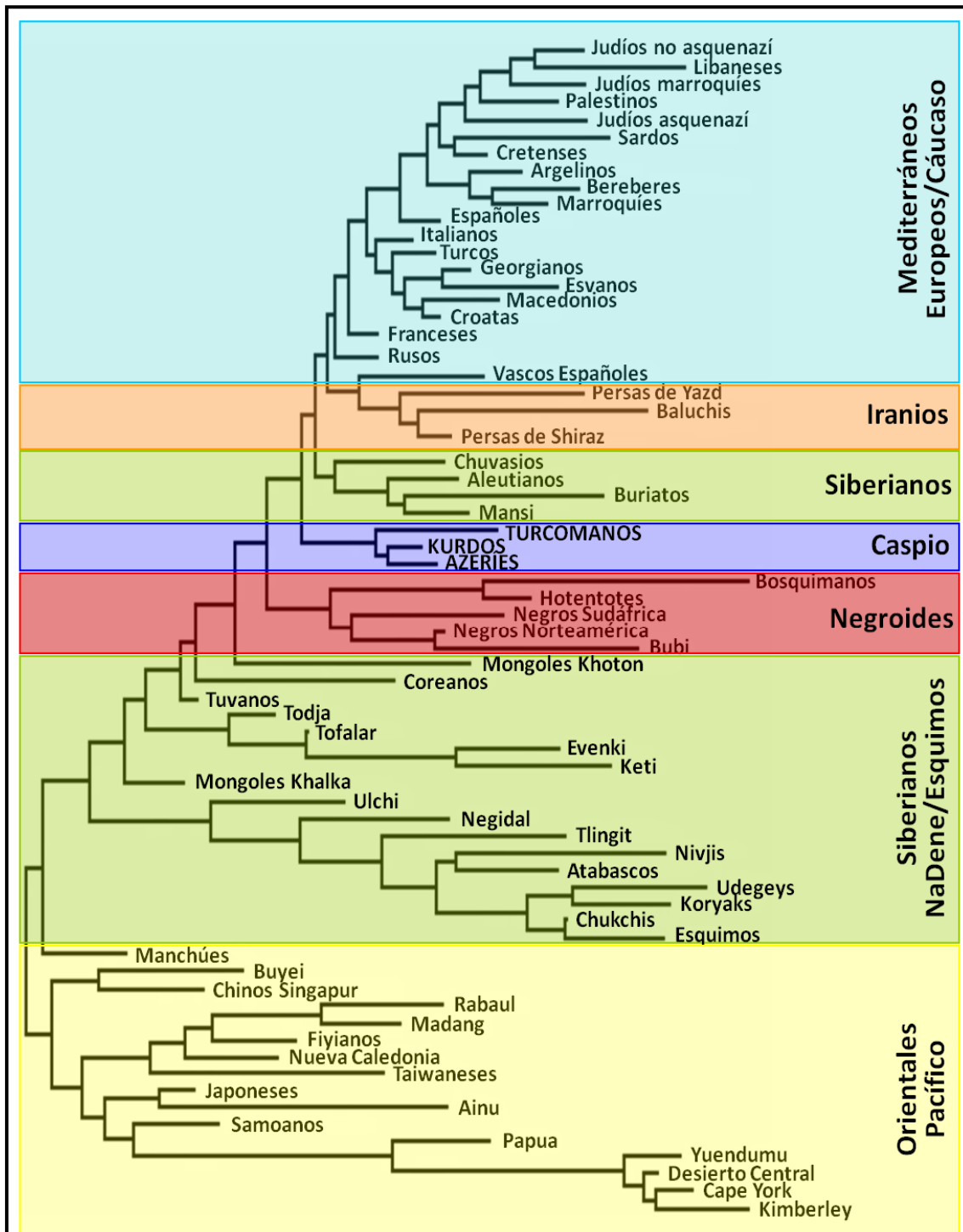


Figura R2. Dendrograma NJ calculado a partir de las distancias genéticas D_A basadas en las frecuencias en alta resolución del locus HLA-DRB1. Nos muestra las relaciones genéticas HLA entre las poblaciones de estudio (turcomanos, azeríes y kurdos) con el resto de poblaciones del mundo: mediterráneos, europeos y del Cáucaso (azul), iraníes (naranja), negroides (rojo), siberianos, na-dene, esquimos (verde), orientales y del Pacífico (amarillo) (Tabla M1).

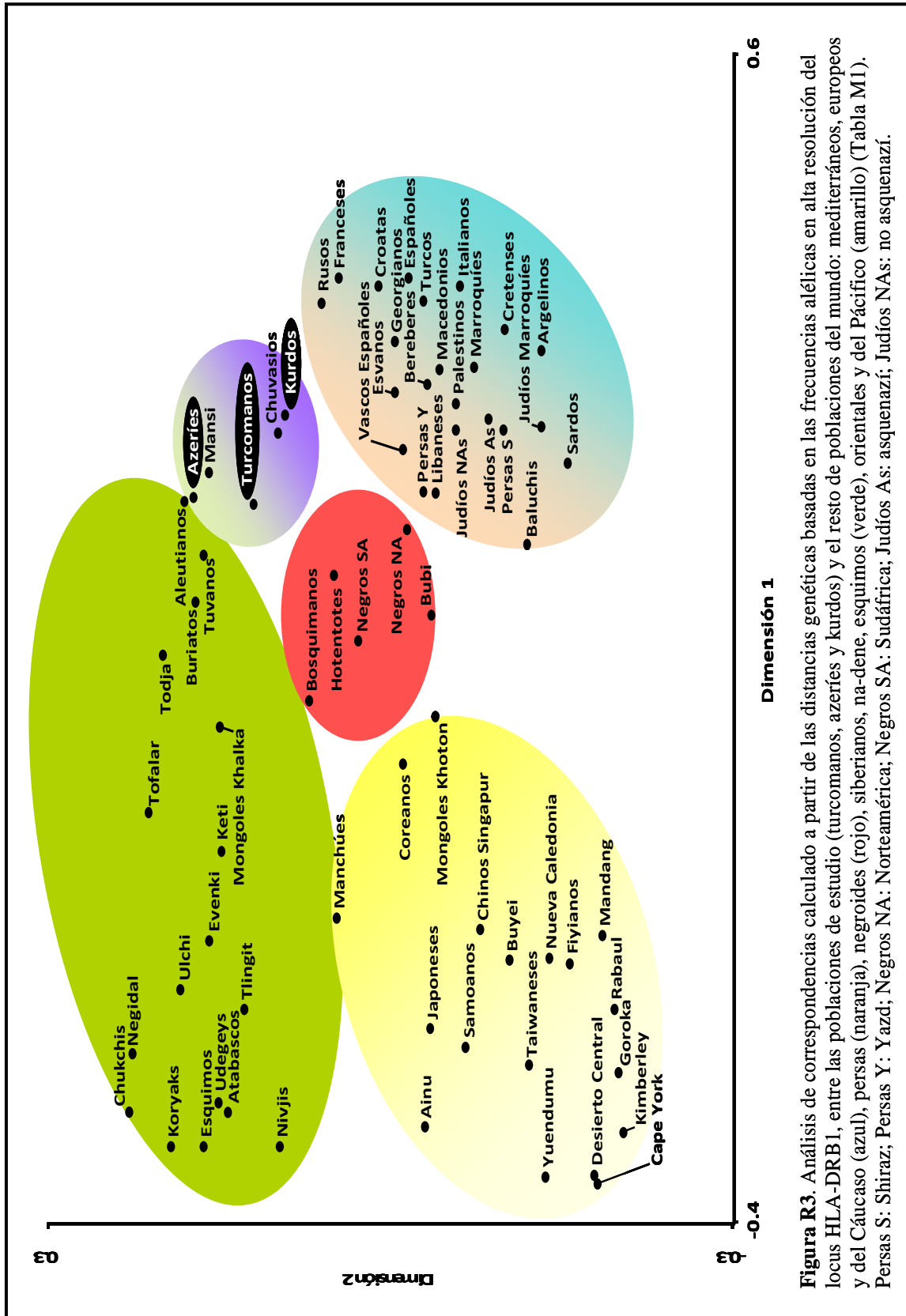


Figura R3. Análisis de correspondencias calculado a partir de las distancias genéticas basadas en las frecuencias alélicas en alta resolución del locus HLA-DRB1, entre las poblaciones de estudio (turcomanos, azerfes y kurdos) y el resto de poblaciones del mundo: mediterráneos, europeos y del Cáucaso (azul), persas (naranja), negroides (rojo), siberianos, na-dene, esquimos (verde), orientales y del Pacífico (amarillo) (Tabla M1). Persas S: Shiraz; Persas Y: Yazd; Negros NA: Norteamérica; Negros SA: Sudáfrica; Judfos As: Sudáfrica; Judfos S: asquenazí; Judfos NAs: no asquenazí.

2.2 Análisis basados en los loci HLA-DRB1 y -DQB1

La Tabla R4 muestra las distancias genéticas entre turcomanos, azeríes y kurdos y el resto de poblaciones incluidas en el estudio (Tabla M1), salvo aquellas para las que no se contaba con los datos de las frecuencias alélicas del locus HLA-DQB1.

Estos resultados muestran nuevamente que las tres poblaciones de estudio están muy próximas entre sí: los azeríes y los kurdos son los que presentan una menor distancia genética, 5.09×10^{-2} , mientras que cada uno de estas presenta una distancia mayor y similar con los turcomanos, 8.12×10^{-2} y 8.70×10^{-2} respectivamente. En lo que respecta a los turcomanos, las poblaciones más próximas (después de azeríes y kurdos) son poblaciones procedentes de Siberia central (aleutianos -15.25×10^{-2} - y mansi -15.69×10^{-2}) y poblaciones del Cáucaso (esvanos -15.66×10^{-2} - y georgianos -16.35×10^{-2}), seguido de poblaciones mediterráneas y europeas (españoles, turcos, bereberes, italianos y rusos), de forma más o menos similar a lo determinado por los análisis del locus HLA-DRB1. En el caso de los azeríes, la población más próxima (después de turcomanos y kurdos nuevamente) es la turca (14.29×10^{-2}) seguido también de poblaciones del Cáucaso (georgianos -14.50×10^{-2} - y esvanos -16.95×10^{-2}), grupos procedentes de Siberia central (aleutianos -14.84×10^{-2} -, mansi -15.92×10^{-2} - y tofalar -18.79×10^{-2}), poblaciones mediterráneas y europeas (italianos, españoles, macedonios y rusos) y persas (Shiraz, 18.98×10^{-2}). Los kurdos presentan las menores distancias genéticas con las poblaciones del Cáucaso, georgianos (12.26×10^{-2}) y esvanos (13.00×10^{-2}), y con los turcos (13.41×10^{-2}). Les siguen poblaciones mediterráneas (italianos, macedonios, cretenses y españoles), persas (Shiraz, 16.98×10^{-2}), y más alejados aleutianos y mansi. Las poblaciones con las que muestran mayores distancias genéticas en los tres casos son las poblaciones del Mediterráneo oriental, algunas poblaciones negroides, las poblaciones orientales, grupos procedentes de Siberia, del Pacífico y por último de Australia, siguiendo un patrón similar al determinado por los análisis basados en el locus HLA-DRB1 en solitario, con la diferencia de que aquí las poblaciones persas aparecen más próximas a azeríes y kurdos, y que los chuvasios, relativamente próximos a las tres poblaciones (en especial a turcomanos) en el análisis anterior, aquí aparecen con distancias superiores a 20×10^{-2} . Todo esto se puede observar con mayor detalle en la Tabla R4. Las distancias genéticas medias son bastante menores comparadas con las del anterior análisis. La más alta de las tres poblaciones sigue siendo la que corresponde a la población turcomana con 31.41×10^{-2} , pero la diferencia se reduce respecto a la de los kurdos (30.10×10^{-2}) y los azeríes (28.89×10^{-2}).

TABLA R4. Distancias genéticas (D_A) a partir de las frecuencias alélicas de HLA-DRB1 y -DQB1 ($\times 10^{-2}$) entre las poblaciones analizadas (turcomanos, azeríes y kurdos) y el resto de poblaciones del mundo.

Turcomanos		Azeríes		Kurdos	
Población	D_A	Población	D_A	Población	D_A
Azeríes	8.12	Kurdos	5.09	Azeríes	5.09
Kurdos	8.70	Turcomanos	8.12	Turcomanos	8.70
Aleutianos	15.25	Turcos	14.29	Georgianos	12.26
Esvanos	15.66	Georgianos	14.50	Esvanos	13.00
Mansi	15.69	Aleutianos	14.84	Turcos	13.41
Georgianos	16.35	Italianos	15.67	Italianos	13.57
Españoles	16.35	Mansi	15.92	Macedonios	14.57
Turcos	17.59	Españoles	16.14	Cretenses	14.90
Bereberes	17.64	Macedonios	16.55	Españoles	15.55
Italianos	18.47	Esvanos	16.95	Rusos	16.18
Rusos	19.03	Rusos	17.06	Persas S	16.98
Negros NA	19.57	Cretenses	17.83	Argelinos	18.73
Marroquíes	20.10	Tofalar	18.79	Aleutianos	18.76
Macedonios	20.20	Persas S	18.98	Mansi	18.97
Todja	20.42	Negros NA	19.57	Bereberes	18.99
Cretenses	20.76	Tuvanos	19.78	Marroquíes	19.95
Argelinos	21.22	Palestinos	20.06	Negros NA	20.50
Buriatos	21.46	Chuvasios	20.22	Tuvanos	21.79
Hotentotes	21.69	Argelinos	20.55	Sardos	21.86
Tuvanos	21.74	Buriatos	21.21	Palestinos	21.87
Negros SA	21.85	Marroquíes	21.28	Hotentotes	22.63
Mongoles Khalka	22.69	Todja	21.30	Judíos marroquíes	22.81
Persas S	22.99	Bereberes	21.81	Negros SA	23.18
Palestinos	23.16	Mongoles Khalka	22.39	Mongoles Khalka	23.28
Chuvasios	23.44	Negros SA	22.46	Persas Y	23.77
Tofalar	23.47	Persas Y	22.87	Chuvasios	23.85
Persas Y	24.60	Hotentotes	23.93	Baluchis	24.00
Judíos NAs	25.19	Judíos As	24.84	Tofalar	24.30
Judíos marroquíes	25.75	Judíos marroquíes	25.74	Judíos As	24.85
Judíos As	28.03	Judíos NAs	25.97	Todja	24.95
Libaneses	28.10	Sardos	26.83	Judíos NAs	25.26
Baluchis	29.74	Libaneses	27.59	Libaneses	25.34
Coreanos	31.63	Baluchis	27.89	Buriatos	27.17
Sardos	33.00	Ulchi	30.35	Coreanos	32.35
Keti	33.54	Manchúes	30.95	Bubis	34.11

TABLA R4. Continuación.

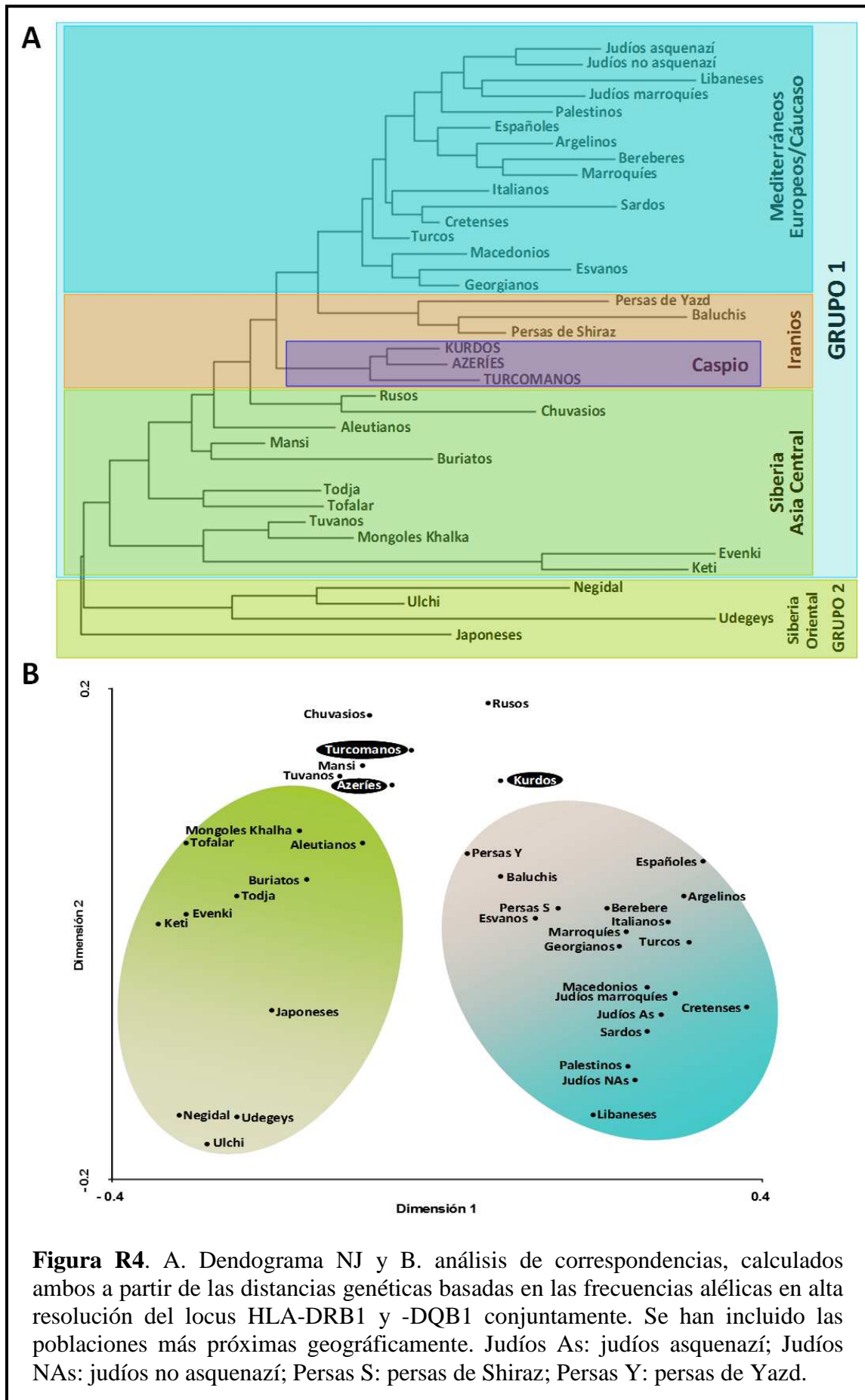
Bubis	34.21	Bubis	31.88	Manchúes	34.49
Ulchi	34.66	Japoneses	32.00	Bosquimanos	34.54
Nueva Caledonia	35.62	Coreanos	32.12	Nueva Caledonia	36.12
Bosquimanos	35.73	Keti	32.88	Mongoles Khoton	36.60
Manchúes	36.47	Negidal	33.45	Keti	36.71
Japoneses	37.26	Nueva Caledonia	33.65	Ulchi	36.80
Evenki	37.63	Evenki	33.99	Evenki	37.14
Mongoles Khoton	38.01	Bosquimanos	35.53	Japoneses	37.98
Samoanos	39.20	Samoanos	36.84	Samoanos	39.65
Negidal	39.83	Nivjis	37.27	Tlingit	41.64
Tlingit	42.15	Chukchis	37.85	Negidal	41.68
Chukchis	43.48	Mongoles Khoton	37.97	Chinos de Singapur	42.31
Chinos de Singapur	45.11	Udegeys	38.30	Udegeys	43.12
Nivjis	48.13	Tlingit	38.72	Buyei	43.84
Udegeys	48.43	Chinos de Singapur	41.45	Chukchis	44.16
Atabascos	49.19	Koryaks	42.10	Madang	44.58
Buyei	50.65	Buyei	43.47	Nivjis	44.95
Madang	50.82	Madang	44.51	Rabaul	44.97
Koryaks	51.27	Atabascos	44.73	Atabascos	49.25
Esquimos	52.70	Esquimos	46.90	Koryaks	49.25
Goroka	54.32	Rabaul	52.28	Esquimos	50.81
Rabaul	58.35	Cape York	56.77	Goroka	59.76
Cape York	60.32	Goroka	57.71	Cape York	61.58
Kimberley	76.87	Kimberley	74.02	Kimberley	76.70

Judíos As: judíos asquenazí; Judíos NAs: judíos no asquenazí; Negros NA: negros de Norteamérica; Negros SA: negros de Sudáfrica; Persas S: persas de Shiraz; Persas Y: persas de Yazd.

Los dendrogramas construidos a partir de las distancias genéticas HLA-DRB1 y -DQB1 (Figura R4.A y R5) también muestra agrupados a turcomanos, azeríes y kurdos. En el primer dendrograma (Figura R4.A), construido mediante las distancias genéticas de las poblaciones más próximas geográficamente, estas poblaciones se encuentran situadas en el primer gran conjunto de los dos en los que se divide el árbol, que agrupa a las poblaciones mediterráneas-europeas, mediterráneas sur orientales, caucásicas, persas y siberianas de Asia Central, en contraposición al otro gran grupo conformado por poblaciones siberianas orientales. Las poblaciones de estudio se sitúan entre las poblaciones iraníes, y rusos y poblaciones con un origen común procedentes de Siberia central, los chuvasios (al norte del Mar Caspio) y los aleutianos (Islas Aleutianas). En el

segundo dendrograma (Figura R5), que incluye el resto de poblaciones que figuran en la Tabla M1, estas poblaciones se encuentran situadas, en este caso junto con los esvanos, en el primer gran grupo que incluye poblaciones negroides, mediterráneas, europeas, persas, siberianas y na-dene, aunque esta vez también incluye las poblaciones orientales, mientras que las poblaciones del Pacífico y Australia se agrupan en el segunda gran conjunto. La distribución en el dendrograma de las poblaciones de estudio es similar a la que encontramos en las Figuras R1.A y R4.A, ya que se sitúan entre poblaciones procedentes de Siberia con un origen común (chuvasios, aleutianos, mansi) y el grupo conformado por las poblaciones iránias (persas y baluchis). Las poblaciones negroides también aparecen situadas, dentro de un mismo nodo, entre estos dos grupos (siberianos centroasiáticos e iránios), aunque la longitud de las ramas horizontales de estas poblaciones indica una distancia genética relativamente grande en comparación a las poblaciones mencionadas, que resultan más próximas, e incluso a otras más alejadas en el árbol como turcos y georgianos.

Los análisis de correspondencias a partir de las distancias genéticas basadas en los loci HLA-DRB1 y -DQB1 únicamente con las poblaciones próximas geográficamente, y con todas las poblaciones incluidas en el estudio (Figuras R4.B y R6) definen dos grandes conjuntos de poblaciones atendiendo a la primera dimensión. El primer grupo incluye a las poblaciones mediterráneas, europeas e iránias y el segundo está a su vez dividido en dos subgrupos definidos por la segunda dimensión: en la Figura R4.B separa a las poblaciones siberianas de Asia Central de las siberianas orientales, y en la Figura R6 separa a los siberianos y na-dene de las poblaciones orientales, del Pacífico y Australia (proporciones de inercia acumulada por las dos dimensiones = 66,7% y 69.9% respectivamente). En la Figura R6 las tres poblaciones de estudio se sitúan en el primer gran grupo definido por la primera dimensión, muy próximas a las poblaciones mediterráneas y del Cáucaso, especialmente los kurdos, pero próximos también a poblaciones siberianas procedentes de Asia Central (mansi, aleutianos y chuvasios), en especial turcomanos y azerís. Esta tendencia se hace más patente en la Figura R4.B, donde los kurdos aparecen claramente en el grupo de los mediterráneos, relativamente próximos a persas y rusos, mientras que turcomanos y azerís se agrupan con siberianos del Asia Central, muy próximos a mansi, tuvanos y chuvasios. En estos análisis (Figura R6) no se aprecia una agrupación de las poblaciones negroides como en el dendrograma (Figura R5) y en la Figura R3, y aparecen, aunque en una posición ligeramente intermedia, en su mayoría próximas a poblaciones mediterráneas.



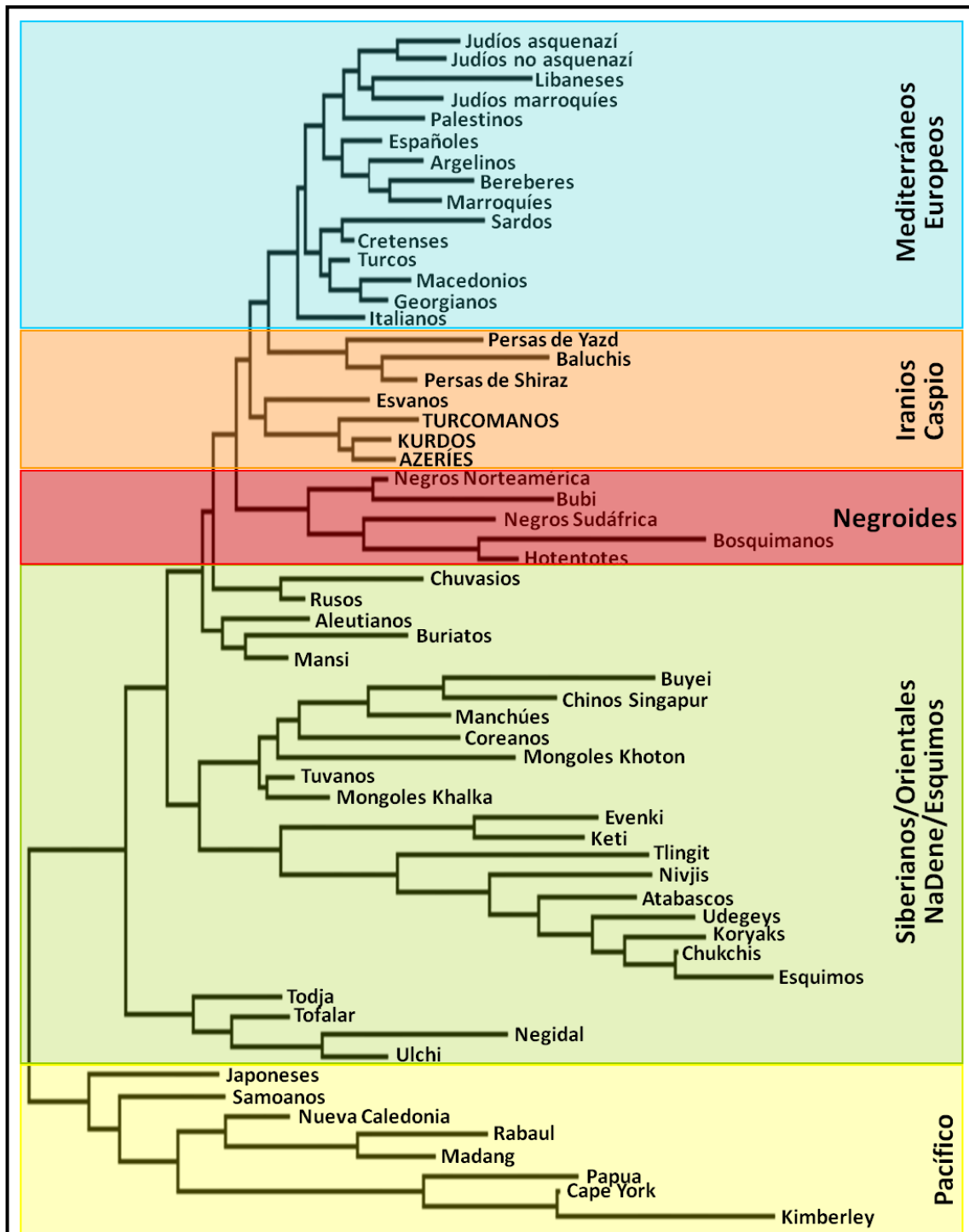


Figura R5. Dendrograma NJ calculado a partir de las distancias genéticas D_A basadas en las frecuencias en alta resolución de los loci HLA-DRB1 y -DQB1. Nos muestra las relaciones genéticas HLA entre las poblaciones de estudio (turcomanos, azeríes y kurdos) con el resto de poblaciones del mundo: mediterráneos, europeos y del Cáucaso (azul), iraníes (naranja), negroides (rojo), siberianos, na-dene, esquimos, orientales (verde), y del Pacífico (amarillo) (Tabla M1).

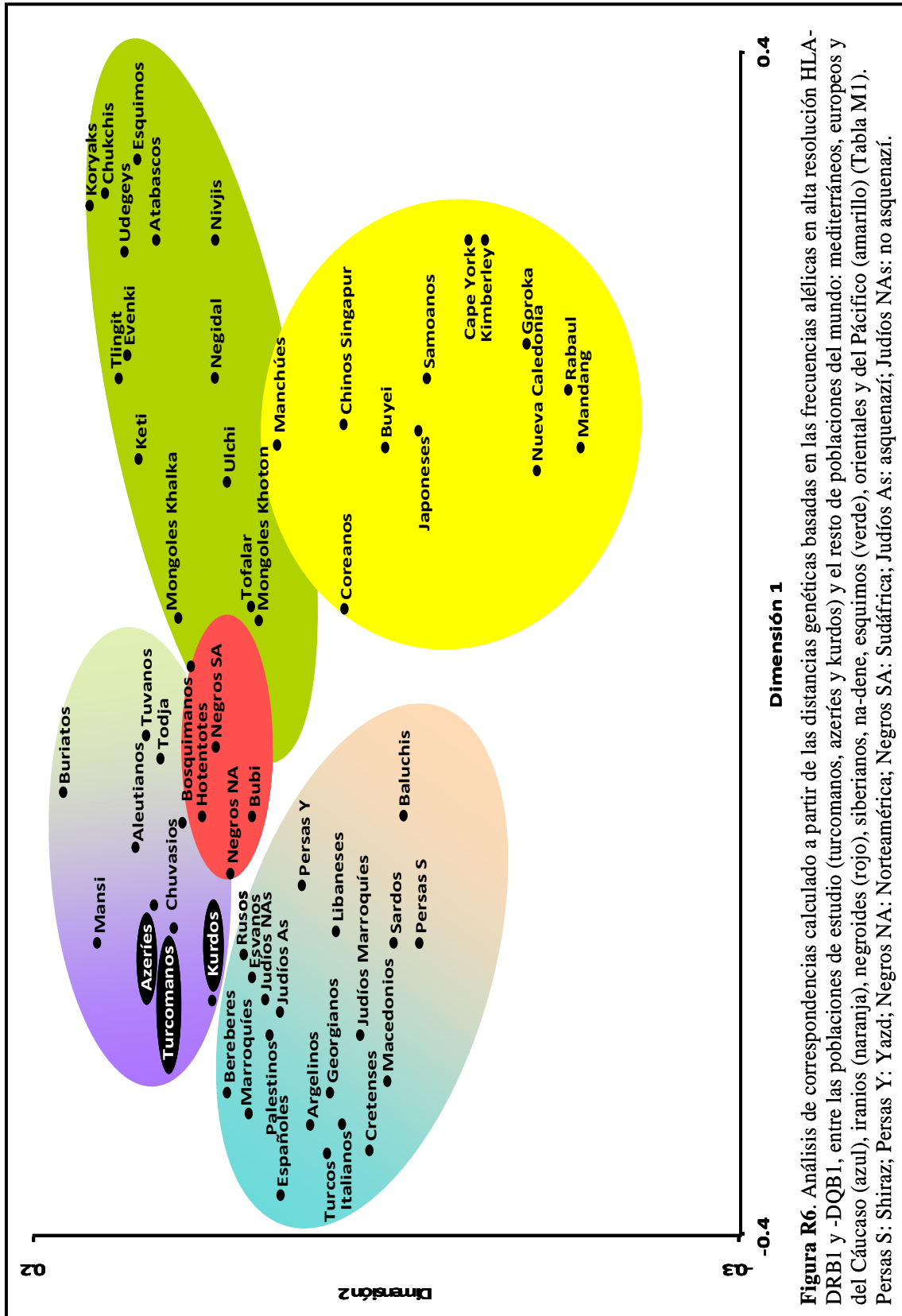


Figura R6. Análisis de correspondencias calculado a partir de las frecuencias alélicas en alta resolución HLA-DRB1 y -DQB1, entre las poblaciones de estudio (turcomanos, azeríes y kurdos) y el resto de poblaciones del mundo: mediterráneos, europeos y del Cáucaso (azul), iraníes (naranja), negroides (rojo), siberianos, na-dene, esquimos (verde), orientales y del Pacífico (amarillo) (Tabla M1). Persas S: Shiraz; Persas Y: Yazd; Negros NA: Norteamérica; Negros SA: Sudáfrica; Judíos AS: asquenazí; Judíos NAs: no asquenazí.

3. ESTUDIOS DE HAPLOTIPOS BI-LOCUS EN LAS POBLACIONES DE ESTUDIO

La Tabla R5, muestra los haplotipos bi-locus con desequilibrio de ligamiento significativo y mayor representación en las poblaciones de estudio. Estas poblaciones corresponden a grupos de individuos no relacionados, por lo que no se conoce la segregación familiar de sus haplotipos.

Mediante el programa estadístico Arlequin v 3.000 (Excoffier et al., 2005), se calcula la frecuencia de cada haplotipo (combinaciones posibles de 2-loci), desequilibrio de ligamiento y la probabilidad (p) de que cada asociación sea estadísticamente significativa. Los resultados de los haplotipos encontrados en las poblaciones de estudio, se comparan con las poblaciones descritas en la literatura y en los Talleres Internacionales 11°, 12° y 13° de Histocompatibilidad (Imanishi et al., 1992a; Clayton & Lonjou, 1997; Mack & Erlich, 2007). A continuación se describen todos aquellos haplotipos con desequilibrio de ligamiento positivo ($D > 0$, $D' > 0$), $p < 0.05$ y con frecuencias por encima del 1% (Tabla R5).

3.1 Estudio del desequilibrio de ligamiento

Desequilibrio de ligamiento entre alelos HLA-A y -B (clase I)

Se han encontrado haplotipos HLA-A-B con desequilibrio de ligamiento positivo, una asociación estadísticamente significativa y una frecuencia en la población por encima del 1% en un total de 14 para turcomanos, 20 para azeríes, y 15 para kurdos (Tabla R5). Tanto en turcomanos como azeríes el haplotipo más frecuente y con un desequilibrio de ligamiento mayor ($D=0.04$) es el A*24:02-B*35:01, haplotipo también frecuente en poblaciones del Asia Occidental y Meridional (judíos e indios) y de Norteamérica (yupik de Alaska). En cuanto a los kurdos, la mayores frecuencias la presentan los haplotipos A*11:01-B*35:01 ($D=0.01$) y A*02:01-B*35:01 ($D=0.03$), el primero frecuente en poblaciones del Sureste Asiático (chinos wa y taiwaneses), y el segundo en poblaciones del Asia Occidental y Meridional (omaneses e indios), y de Norteamérica (indios pima de Arizona). El único haplotipo común a las tres poblaciones que encontramos con un desequilibrio de ligamiento positivo y significativo es el A*33:01-B*58:01, haplotipo de gran frecuencia en poblaciones orientales y del Sureste Asiático (chinos han, buyei, taiwaneses, chinos de Singapur). Por lo general los haplotipos HLA-A-B en alta resolución no muestran frecuencias muy elevadas, debido en parte a la gran diversidad alélica que presentan estos loci.

TABLA R5. Haplotipos HLA-A-B, HLA-B-DRB1 y HLA-DRB1-DQB1 con desequilibrio de ligamiento (D, D') significativo en turcomanos, azerís y kurdos. En negrita los haplotipos más frecuentes.

Haplotipo HLA-A-B	Turcomanos			Azerís			Kurdos		
	D	D'	FH	D	D'	FH	D	D'	FH
A*01:01-B*27:01	—	—	—	0.01	0.66	1.03	—	—	—
A*01:01-B*37:01	0.01	0.49	0.73	0.01	0.66	1.03	—	—	—
A*01:01-B*53:01	—	—	—	—	—	—	0.02	0.64	1.67
A*01:02-B*08:01	—	—	—	0.02	0.65	1.95	—	—	—
A*02:01-B*07:02	—	—	—	0.01	0.23	1.55	—	—	—
A*02:01-B*03:01	0.01	0.46	2.17	—	—	—	—	—	—
A*02:01-B*15:01	—	—	—	0.01	0.71	1.02	—	—	—
A*02:01-B*15:24	0.01	1.00	1.45	—	—	—	—	—	—
A*02:01-B*35:01	—	—	—	—	—	—	0.03	0.32	5.00
A*02:01-B*49:01	—	—	—	—	—	—	0.02	0.43	2.50
A*02:01-B*50:01	—	—	—	0.02	0.57	2.58	—	—	—
A*02:01-B*51:01	0.01	0.25	2.90	—	—	—	—	—	—
A*02:50-B*07:03	0.02	0.57	2.17	—	—	—	—	—	—
A*02:52-B*13:01	0.01	0.46	1.45	—	—	—	—	—	—
A*02:52-B*51:01	—	—	—	—	—	—	0.02	0.31	2.50
A*03:01-B*07:02	—	—	—	0.01	0.26	2.06	—	—	—
A*03:01-B*08:01	0.02	0.28	1.45	—	—	—	—	—	—
A*03:01-B*18:01	—	—	—	0.03	0.56	3.10	—	—	—
A*03:01-B*35:01	—	—	—	—	—	—	0.01	1.00	1.25
A*03:01-B*35:10	0.01	1.00	1.45	—	—	—	—	—	—
A*03:01-B*44:02	0.01	0.46	2.90	—	—	—	0.03	0.63	0.73
A*11:01-B*35:01	—	—	—	—	—	—	0.01	1.00	5.83
A*11:01-B*55:01	0.01	0.47	0.73	0.01	0.24	1.03	—	—	—
A*23:01-B*49:01	—	—	—	0.01	1.00	0.52	0.02	0.65	1.67
A*24:02-B*18:09	—	—	—	—	—	—	0.03	1.00	1.25
A*24:02-B*35:01	0.04	0.38	5.67	0.04	0.36	5.53	—	—	—
A*24:02-B*35:20	0.01	1.00	1.45	—	—	—	—	—	—
A*24:02-B*41:01	—	—	—	0.01	0.32	1.02	—	—	—
A*24:02-B*52:01	—	—	—	—	—	—	0.03	0.76	3.33
A*24:02-B*58:01	—	—	—	—	—	—	0.02	0.39	1.67
A*26:01-B*07:02	0.01	0.37	1.45	—	—	—	—	—	—
A*26:01-B*08:01	—	—	—	—	—	—	0.02	0.46	2.50
A*26:01-B*38:01	—	—	—	—	—	—	0.03	0.78	3.33
A*29:01-B*58:01	—	—	—	0.01	0.38	1.03	—	—	—
A*30:01-B*13:01	0.02	0.38	2.17	0.01	0.40	1.55	—	—	—

TABLA R5. Continuación.

A*30:01-B*38:01	—	—	—	0.01	0.30	1.55	—	—	—
A*30:01-B*40:46	0.01	1.00	1.45	—	—	—	—	—	—
A*31:01-B*35:01	0.01	0.50	2.90	0.01	0.31	0.52	—	—	—
A*32:01-B*41:01	—	—	—	0.01	0.15	1.03	—	—	—
A*32:01-B*55:01	—	—	—	—	—	—	0.01	0.30	1.67
A*33:01-B*14:01	—	—	—	0.01	0.47	1.55	—	—	—
A*33:01-B*58:01	0.01	1.00	1.45	0.01	0.58	1.55	0.02	1.00	1.67
A*68:01-B*35:01	0.01	0.59	1.45	—	—	—	—	—	—
A*68:01-B*44:02	—	—	—	0.01	0.15	1.03	—	—	—
A*68:01-B*51:01	—	—	—	0.02	0.40	1.95	—	—	—
Haplotipo HLA-B-DRB1	D	D'	FH	D	D'	FH	D	D'	FH
B*07:02-DRB1*10:01	—	—	—	0.01	0.25	2.06	—	—	—
B*07:02-DRB1*15:01	—	—	—	0.02	0.48	2.06	—	—	—
B*07:03-DRB1*04:01	0.02	0.55	2.90	—	—	—	—	—	—
B*08:01-DRB1*03:01	0.03	0.51	3.62	0.03	1.00	3.10	0.04	1.00	5.00
B*13:01-DRB1*04:01	—	—	—	—	—	—	0.01	0.43	1.67
B*13:01-DRB1*07:01	0.02	0.26	1.63	0.02	0.68	2.58	—	—	—
B*14:01-DRB1*01:01	—	—	—	0.02	0.65	2.06	—	—	—
B*14:02-DRB1*01:01	0.01	1.00	1.45	—	—	—	—	—	—
B*18:01-DRB1*11:01	—	—	—	0.02	0.41	2.06	0.01	0.29	1.67
B*18:09-DRB1*11:01	—	—	—	—	—	—	0.03	1.00	2.50
B*27:01-DRB1*01:01	0.01	0.65	1.45	—	—	—	—	—	—
B*27:01-DRB1*04:01	—	—	—	0.01	0.61	1.03	—	—	—
B*35:01-DRB1*01:01	—	—	—	—	—	—	0.03	0.43	4.17
B*35:01-DRB1*11:01	0.02	0.20	2.30	0.04	0.31	6.13	0.03	0.17	7.75
B*35:01-DRB1*15:01	0.01	1.00	7.74	—	—	—	—	—	—
B*35:27-DRB1*15:01	0.02	0.68	2.17	—	—	—	—	—	—
B*37:01-DRB1*10:01	—	—	—	0.02	1.00	1.03	—	—	—
B*38:01-DRB1*13:01	0.01	0.22	0.73	—	—	—	0.02	0.37	1.43
B*38:01-DRB1*14:01	—	—	—	0.02	0.41	2.06	—	—	—
B*39:01-DRB1*08:01	0.01	0.49	1.45	—	—	—	—	—	—
B*39:01-DRB1*12:01	0.01	1.00	1.45	—	—	—	—	—	—
B*39:01-DRB1*16:01	—	—	—	0.01	1.00	1.03	—	—	—
B*40:01-DRB1*15:01	—	—	—	0.01	0.61	1.03	—	—	—
B*41:01-DRB1*07:01	—	—	—	0.01	0.22	1.55	—	—	—
B*44:02-DRB1*04:01	—	—	—	0.03	0.65	3.50	—	—	—
B*44:02-DRB1*11:01	—	—	—	—	—	—	0.02	0.38	2.50
B*49:01-DRB1*11:01	0.01	0.62	1.45	—	—	—	—	—	—
B*50:01-DRB1*07:01	—	—	—	0.02	0.44	2.06	0.03	0.54	3.33

TABLA R5. Continuación

B*51:01-DRB1*04:01	—	—	—	—	—	—	0.01	1.00	3.33
B*51:01-DRB1*10:01	0.01	0.36	1.45	—	—	—	—	—	—
B*51:01-DRB1*13:01	—	—	—	0.01	0.32	1.44	—	—	—
B*51:01-DRB1*16:01	—	—	—	0.01	0.32	1.55	—	—	—
B*52:01-DRB1*04:01	—	—	—	0.01	0.19	3.61	—	—	—
B*52:01-DRB1*15:01	—	—	—	0.02	0.34	2.58	0.03	0.78	3.33
B*53:01-DRB1*03:01	—	—	—	—	—	—	0.01	0.60	1.67
B*55:01-DRB1*15:01	—	—	—	0.01	0.33	1.55	—	—	—
B*57:01-DRB1*07:01	—	—	—	0.01	0.63	1.03	—	—	—
B*58:01-DRB1*03:01	—	—	—	0.01	0.34	1.55	0.03	0.80	3.33
Haplotipo DRB1-DQB1	D	D'	FH	D	D'	FH	D	D'	FH
DRB1*01:01-DQB1*05:01	0.05	1.00	5.80	0.04	1.00	4.64	0.05	0.75	5.83
DRB1*03:01-DQB1*02:01	0.05	0.73	7.97	0.07	1.00	8.76	0.11	0.87	14.17
DRB1*04:01-DQB1*03:02	0.04	0.66	5.07	0.07	1.00	8.76	0.04	0.71	4.88
DRB1*04:01-DQB1*03:05	0.01	1.00	0.73	—	—	—	0.03	0.77	3.33
DRB1*04:01-DQB1*04:01	—	—	—	0.01	0.25	2.06	—	—	—
DRB1*04:01-DQB1*06:03	0.01	0.22	2.17	—	—	—	—	—	—
DRB1*07:01-DQB1*02:01	0.07	0.60	10.05	0.05	0.60	7.22	0.04	0.84	5.83
DRB1*07:01-DQB1*03:03	—	—	—	0.02	0.50	2.58	—	—	—
DRB1*08:01-DQB1*04:01	0.02	0.74	2.17	0.03	1.00	3.09	0.02	1.00	1.67
DRB1*09:01-DQB1*03:03	—	—	—	0.02	1.00	1.55	—	—	—
DRB1*10:01-DQB1*04:01	0.01	0.37	1.45	—	—	—	—	—	—
DRB1*10:01-DQB1*05:01	0.02	0.54	2.17	0.03	1.00	3.61	—	—	—
DRB1*11:01-DQB1*03:01	0.09	0.93	11.59	0.11	0.91	13.92	0.15	1.00	19.17
DRB1*12:01-DQB1*03:01	0.01	1.00	1.45	0.01	1.00	1.03	—	—	—
DRB1*13:01-DQB1*06:02	—	—	—	0.01	0.21	1.03	0.02	0.46	2.50
DRB1*13:01-DQB1*06:03	0.03	1.00	3.62	0.03	0.74	3.09	0.02	0.47	2.50
DRB1*13:02-DQB1*06:04	0.01	0.49	0.73	0.01	0.66	1.03	0.02	0.74	2.38
DRB1*14:01-DQB1*05:02	—	—	—	0.03	1.00	3.61	0.05	0.73	5.00
DRB1*14:02-DQB1*03:01	—	—	—	0.01	1.00	1.03	—	—	—
DRB1*15:01-DQB1*06:01	0.02	1.00	2.17	0.07	1.00	8.25	0.03	0.63	2.50
DRB1*15:01-DQB1*06:02	0.08	0.84	10.15	0.03	0.65	3.61	0.02	0.31	2.50
DRB1*15:01-DQB1*06:03	—	—	—	—	—	—	0.01	0.26	1.67
DRB1*16:01-DQB1*05:01	—	—	—	0.03	1.00	4.12	0.01	0.33	2.50

FH: frecuencia haplotípica

Desequilibrio de ligamiento entre alelos HLA-B y -DRB1

Este estudio, nos da una idea de cómo son las poblaciones analizadas, ya que el locus HLA-DRB1 define muy bien geográficamente las poblaciones. Se han encontrado haplotipos HLA-B-DRB1 con desequilibrio de ligamiento positivo, una asociación estadísticamente significativa y una frecuencia representativa en la población (por encima del 1%) en un total de 11 para población turcomana, 23 para población azerí, y 13 para población kurda (Tabla R5), muchos de ellos asociados a haplotipos extendidos en su población respectiva (véase apartado 4). Se puede observar que tanto en azeríes como en kurdos el haplotipo más frecuente y con un desequilibrio de ligamiento mayor ($D=0.04/0.03$ respectivamente) es el B*35:01-DRB1*11:01, también encontrado en turcomanos con una frecuencia menor. Este haplotipo no presenta frecuencias significativas en ninguna población analizada hasta la fecha, como ocurre con el haplotipo con mayor frecuencia en turcomanos, el B*35:01-DRB1*15:01. Otro haplotipo común a las tres poblaciones que encontramos con un desequilibrio de ligamiento positivo y significativo es el B*08:01-DRB1*03:01, haplotipo de gran frecuencia en poblaciones del norte de Europa (irlandeses, samis de Laponia) y poblaciones de Asia Meridional como los indios golla, encontrándose en otras poblaciones de Irán en baja frecuencia, como los baluchis.

Desequilibrio de ligamiento entre alelos HLA-DRB1 y -DQB1 (clase II)

A diferencia de los análisis anteriores con las otras parejas de loci, el estudio del desequilibrio de ligamiento entre los loci de clase II arroja resultados mucho más homogéneos en las tres poblaciones analizadas, ya que muchos de los haplotipos con desequilibrio de ligamiento positivo son compartidos por estas, en especial los que aparecen con una mayor frecuencia (Tabla R5). Se han encontrado un total de 13 haplotipos en población turcomana, 19 en población azerí, y 15 en población kurda con desequilibrio de ligamiento positivo, una asociación estadísticamente significativa y una frecuencia representativa en la población (por encima del 1%), la mayor parte asociados a haplotipos extendidos en su población respectiva (véase apartado 4), y presentes fundamentalmente en poblaciones asiáticas, desde Siberia hasta el Suroeste Asiático, incluido el Levante mediterráneo. El haplotipo más frecuente y que muestra un mayor desequilibrio de ligamiento en las tres poblaciones ($D \geq 0.09$), en especial en kurdos, es el DRB1*11:01-DQB1*03:01, haplotipo que muestra sus mayores frecuencias en poblaciones del Pacífico (poblaciones de Papua), pero con frecuencias muy

representativas en poblaciones de Siberia o precedentes de esta región (nganasan, sami, yupik, evenki, keti, todja, tofalar, chuvasios) y en poblaciones del Suroeste Asiático (turcos, persas de Shiraz). Otros haplotipos con frecuencias relevantes en las tres poblaciones son: DRB1*03:01-DQB1*02:01 (muy frecuente en poblaciones del sur del Mediterráneo y del Suroeste Asiático), DRBQ*07:01-DQB1*02:01 (muy frecuente en poblaciones de Siberia, Asia Central y sur del Mediterráneo), DRB1*15:01-DQB1*06:02 (frecuente en poblaciones de Siberia, Asia Central y Suroeste Asiático), DRB1*15:01-DQB1*06:01 (frecuente en poblaciones del Asia Meridional -India- y Siberia), DRB1*01:01-DQB1*05:01, DRB1*04:01-DQB1*03:02 (ambos frecuentes en poblaciones de Siberia o procedentes de esta región -Sami- y en población persa) y DRB1*13:01-DQB1*06:03 (frecuente en poblaciones de Asia Central y del este y sur del Mediterráneo).

3.2 Análisis basados en haplotipos de clase II con desequilibrio de ligamiento

Debido al gran desequilibrio de ligamiento existente entre los loci de clase II HLA-DRB1 y -DQB1, como se observa en el apartado anterior, se ha procedido a utilizar las frecuencias de estos haplotipos bi-locus para calcular las distancias genéticas D_A entre las poblaciones de estudio y otras poblaciones del mundo (Tabla R6), así como las representaciones gráficas que se derivan de estas (Figura R7), como otra herramienta más para indagar en las relaciones genéticas de estos grupos poblacionales.

En primer lugar se observa que las tres poblaciones están de nuevo muy próximas entre sí: los kurdos y los azeríes presentan la menor distancia genética, 11.04×10^{-2} , mientras que los turcomanos con respecto a los kurdos están a una distancia de 13.09×10^{-2} , y con respecto a los azeríes a 13.88×10^{-2} . A diferencia de los análisis basados en las frecuencias alélicas, los turcomanos son los que presentan una menor distancia genética media, 29.57×10^{-2} , bastante inferior a la de los azeríes (35.08×10^{-2}) y kurdos (35.60×10^{-2}). Sin embargo, este análisis muestra un patrón similar de las tres poblaciones respecto a las distancias con el resto de poblaciones, con pequeñas diferencias. Las tres poblaciones se muestran próximas a las poblaciones europeas, mediterráneas y del Cáucaso (franceses, italianos, rusos, bereberes, turcos, georgianos, libaneses, judíos de Marruecos) y alejadas de las poblaciones siberianas más septentrionales y orientales y de los buriatos. No obstante los kurdos están más cercanos a las poblaciones iraníes, en especial a los persas de Yazd, que los turcomanos y azeríes,

que se muestran a su vez más relacionados con los chuvasios, mansi y poblaciones de Asia Central (todja y tofalar) que los kurdos (Tabla R6).

TABLA R6. Distancias genéticas (D_A) a partir de las frecuencias de haplotipos bi-locus HLA-DRB1-DQB1 ($\times 10^{-2}$) entre las poblaciones analizadas (turcomanos, azeríes y kurdos) y otras poblaciones del mundo.

Turcomanos		Azeríes		Kurdos	
Población	D_A	Población	D_A	Población	D_A
Kurdos	13.09	Kurdos	11.04	Azeríes	11.04
Azeríes	13.88	Turcomanos	13.88	Turcomanos	13.09
Franceses	18.23	Franceses	24.44	Franceses	20.36
Italianos	18.85	Rusos	26.41	Italianos	23.94
Bereberes	19.54	Chuvasios	28.49	Rusos	27.03
Rusos	19.76	Georgianos	28.59	Bereberes	27.29
Turcos	22.26	Italianos	28.85	Turcos	29.14
Georgianos	22.92	Libaneses	29.46	Georgianos	29.21
Argelinos	23.30	Turcos	30.46	Persas de Yazd	29.27
Chuvasios	24.94	Bereberes	31.04	Libaneses	30.55
Cretenses	25.88	Judíos Marroquíes	33.26	Baluchis	30.92
Judíos Marroquíes	25.97	Cretenses	34.00	Argelinos	31.90
Persas de Yazd	26.84	Mansi	35.04	Cretenses	33.13
Libaneses	27.51	Argelinos	35.41	Chuvasios	33.61
Aleutianos	29.14	Aleutianos	35.53	Judíos Marroquíes	34.93
Mansi	29.72	Tofalar	37.49	Sardos	35.76
Baluchis	29.81	Persas de Yazd	37.56	Persas de Shiraz	36.37
Todja	30.78	Todja	37.66	Mansi	39.38
Tofalar	35.28	Baluchis	38.77	Aleutianos	41.68
Sardos	35.41	Persas de Shiraz	41.99	Todja	41.68
Persas de Shiraz	36.05	Sardos	42.21	Tofalar	42.88
Buriatos	44.55	Ulchi	44.55	Evenki	53.55
Ulchi	46.25	Buriatos	47.06	Ulchi	54.26
Evenki	47.49	Negidal	49.68	Keti	58.07
Keti	48.25	Evenki	54.56	Buriatos	58.15
Negidal	53.03	Keti	54.71	Negidal	58.62

El dendrograma originado en base a las distancias genéticas de los haplotipos HLA-DRB1-DQB1 (Figura R7.A) muestra, al igual que en los análisis basados en las frecuencias alélicas, a turcomanos, azeríes y kurdos en un mismo grupo. En este árbol, estas tres poblaciones se encuentran situadas en el primer gran conjunto, que agrupa

únicamente a las poblaciones europeas, mediterráneas e iránias (con excepción de aleutianos y chuvasios), próximas a libaneses, judíos marroquíes, argelinos y bereberes. Las poblaciones siberianas conforman el otro gran conjunto, alejadas de las poblaciones de estudio, siendo los mansi lo más próximos (como se observa en las distancias genéticas, Tabla R6). El análisis de correspondencias a partir de las distancias genéticas entre las mismas poblaciones (Figura R7.B), confirma lo observado en el dendrograma, definiendo dos grandes conjuntos de poblaciones basándonos en la primera dimensión. Uno de los grupos incluye a las poblaciones mediterráneas, europeas e iránias y el otro agrupa a las poblaciones siberianas, más dispersas a lo largo de la segunda dimensión, situándose en un extremo las poblaciones más meridionales y en el otro las poblaciones septentrionales y orientales. Las tres poblaciones de estudio se sitúan en el primer gran grupo definido por la primera dimensión, muy próximas a las poblaciones mediterráneas e iránias, especialmente los kurdos.

4. HAPLOTIPOS HLA EXTENDIDOS PRESENTES EN LAS POBLACIONES DE ESTUDIO

Mediante el programa Arlequin v2.000 (Schneider et al., 2000) se han obtenido y descrito los haplotipos HLA extendidos (HLA-A-B-DRB1-DQB1) en las poblaciones turcomana, azerí y kurda. En la Tabla R7 figuran los haplotipos más frecuentes, que se han comparado con los haplotipos descritos previamente en la literatura para buscar poblaciones que los compartan y así poder inferir en el origen de los mismos.

Los haplotipos extendidos, se han deducido a partir de los haplotipos de 2, 3 y 4 loci. También se han considerado aquellos haplotipos que no fueron estimados estadísticamente por el programa informático pero que estaban presentes en al menos dos individuos y mostraban un haplotipo alternativo descrito previamente (véase Materiales y Métodos apartado 4.3).

En ninguna de las tres poblaciones se observan frecuencias muy elevadas en los haplotipos extendidos debido a que han recibido influencia genética externa y no se trata de aislados genéticos.

Haplotipos HLA extendidos frecuentes en población turcomana

La Tabla R7 incluye los ocho haplotipos extendidos más frecuentes en turcomanos. Representan el 13.4% del total de haplotipos encontrados en esta población.

Observamos que el haplotipo de clase I A*24:02-B*35:01 está presente en asociación con dos haplotipos de clase II: DRB1*11:01-DQB1*03:01 y DRB1*15:01-DQB1*06:02. Solamente estos tres haplotipos extendidos han sido descritos previamente en otras poblaciones con los alelos de clase I en baja resolución:

- A*24:02-B*35:01-DRB1*11:01-DQB1*03:01: encontrado en chuvasios (frecuencia haplotípica, FH: 1.2%) y en palestinos (FH: 1.2%).
- A*02:01-B*13:01-DRB1*07:01-DQB1*02:01: encontrado en georgianos (FH: 1.7%) y esvanos (1.3%).
- A*24:02-B*35:01-DRB1*15:01-DQB1*06:02: encontrado únicamente en tuvanos (FH: 1.4%).

El resto de haplotipos no han sido descritos en ninguna población, siendo en turcomanos la primera vez que se reportan. Sin embargo, la mayoría de los haplotipos de dos loci que componen los haplotipos extendidos coinciden con los obtenidos en los análisis de desequilibrio de ligamiento, y son característicos de poblaciones procedentes de Siberia y Asia Central (A*24:02-B*35:01, A*31:01-B*35:01, DRB1*15:01-DQB1*06:02, DRB1*07:01-DQB1*02:01) (Uinuk-Ool et al., 2002; Arnaiz-Villena et al., 2003), aunque algunos también se encuentra en una frecuencia considerablemente alta en poblaciones del Mediterráneo, particularmente del sudeste (A*02:01-B*51:01, DRB1*03:01-DQB1*02:01, DRB1*01:01-DQB1*05:01) (Izaabel et al., 1998; Williams et al., 2001; Tabla R7). En resumen, la mayor parte de los haplotipos bi-locus y haplotipos extendidos hallados en esta población tienen un origen centroasiático-siberiano principalmente, aunque también mediterráneo.

Haplotipos HLA extendidos frecuentes en población azerí

Los diez haplotipos HLA extendidos más frecuentes encontrados en azeríes son recogidos en la Tabla R7. Suman el 16% del total de haplotipos encontrados en esta población. Observamos que algunos haplotipos bi-locus que conforman estos haplotipos extendidos aparecen asociados de distintas formas: el haplotipo de clase I A*24:02-B*35:01 está presente en asociación con DRB1*16:01-DQB1*05:01 y DRB1*11:01-DQB1*03:01 (al igual que en turcomanos) que a su vez está asociado también con los haplotipos A*03:01-B*18:01 y A*02:01-B*35:01, y por último, el haplotipo de clase II DRB1*07:01-DQB1*02:01 presenta asociación con A*02:01-B*50:01 y A*30:01-B*13:01. Únicamente cuatro de estos haplotipos extendidos se han encontrado previamente en otras poblaciones del mundo, tanto en alta como en baja resolución:

TABLA R7. Haplotipos HLA extendidos más frecuentes encontrados en las poblaciones de estudio. Resaltados en verde los haplotipos compartidos por turcomanos y azeríes, y en azul, los compartidos por azeríes y kurdos.

Haplotipos descritos en Turcomanos	FH (%)	Posible Origen
A*24:02-B*35:01-DRB1*11:01-DQB1*03:01^a	2.2	CA/M
A*31:01-B*35:01-DRB1*15:01-DQB1*03:01 ^b	2.2	CA
A*01:01-B*35:01-DRB1*03:01-DQB1*02:01 ^c	1.5	M
A*02:01-B*13:01-DRB1*07:01-DQB1*02:01 ^d	1.5	C
A*02:50-B*07:03-DRB1*13:01-DQB1*06:03 ^e	1.5	CA
A*24:02-B*35:01-DRB1*15:01-DQB1*06:02 ^f	1.5	CA
A*26:01-B*07:02-DRB1*11:01-DQB1*02:01	1.5	–
A*30:01-B*13:01-DRB1*15:01-DQB1*02:01	1.5	–
Haplotipos descritos en Azeríes	FH (%)	Posible Origen
A*24:02-B*35:01-DRB1*11:01-DQB1*03:01^a	3.1	CA/M
A*01:02-B*08:01-DRB1*03:01-DQB1*02:01 ^g	2.1	M
A*03:01-B*07:02-DRB1*15:01-DQB1*06:02 ^h	1.6	CA
A*03:01-B*18:01-DRB1*11:01-DQB1*03:01 ⁱ	1.6	CA/M
A*03:01-B*44:02-DRB1*04:01-DQB1*03:02 ^j	1.6	CA
A*33:01-B*14:01-DRB1*01:01-DQB1*05:01 ^k	1.6	M-CA
A*02:01-B*35:01-DRB1*11:01-DQB1*03:01^l	1.1	CA
A*02:01-B*50:01-DRB1*07:01-DQB1*02:01^m	1.1	C/M
A*24:02-B*35:01-DRB1*16:01-DQB1*05:01 ⁿ	1.1	CA-M
A*30:01-B*13:01-DRB1*07:01-DQB1*02:01 ^o	1.1	CA/M
Haplotipos descritos en Kurdos	FH (%)	Posible Origen
A*02:01-B*35:01-DRB1*01:01-DQB1*05:01 ^p	3.3	M/CA
A*24:02-B*18:09-DRB1*11:01-DQB1*03:01 ^q	3.3	CA
A*01:01-B*35:01-DRB1*11:01-DQB1*03:01 ^r	2.5	M
A*03:01-B*44:02-DRB1*11:01-DQB1*03:01 ^s	2.5	CA
A*24:02-B*52:01-DRB1*15:01-DQB1*06:01 ^t	2.5	CA
A*24:02-B*58:01-DRB1*03:01-DQB1*02:01 ^u	2.5	M

TABLA R7. Continuación

A*01:01-B*08:01-DRB1*03:01-DQB1*02:01 ^v	1.7	M
A*02:01-B*35:01-DRB1*11:01-DQB1*03:01 ¹	1.7	CA
A*02:01-B*50:01-DRB1*07:01-DQB1*02:01 ^m	1.7	C/M
A*11:01-B*50:01-DRB1*07:01-DQB1*02:01 ^w	1.7	M/CA
A*24:02-B*51:01-DRB1*08:01-DQB1*04:01	1.7	–
A*26:01-B*08:01-DRB1*03:01-DQB1*02:01 ^x	1.7	OP-M

FH: frecuencia haplotípica; CA: centroasiático; C: caucásico; M: mediterráneo; OP: de oriente próximo; (–): sin definir; X-X: origen haplotipo clase I – origen haplotipo clase II

^aEncontrado con clase I en baja resolución en chuvasios (FH: 1.2) y palestinos (FH: 1.2).

^bHaplotipo de clase I encontrado en chinos hani (FH: 5.0). Haplotipo de clase II encontrado en evenki (FH: 1.4).

^c ^uHaplotipo de clase II encontrado en sardos (FH: 25.3), bereberes (FH: 17.3), marroquíes (FH: 14.1), mogoles khoton (FH: 13.6), persas de Shiraz (FH: 11.0), macedonios (FH: 9.2), judíos asquenazí (FH: 8.6), cretenses (FH: 7.4) y georgianos (FH: 5.8).

^dEncontrado con clase I en baja resolución en georgianos (FH: 1.7) y esvanos (FH: 1.3).

^eHaplotipo de clase II encontrado en georgianos (FH: 9.7), sami (FH: 9.2), mansi (FH: 8.1), todja (FH: 6.8), judíos asquenazí (FH: 4.7) y persas de Shiraz (FH: 3.4).

^fEncontrado con clase I en baja resolución en tuvanos (FH: 1.4). Haplotipo de clase I encontrado en yupik de Alaska (FH: 11.7), lakota sioux (FH: 4.3) y sami (FH: 3.0). Haplotipo de clase II encontrado en todja (FH: 20.5), buriatos (FH: 18.0), tofalar (FH: 14.0), mansi (FH: 10.2) y chuvasios (FH: 9.2).

^gEncontrado con HLA de clase I en baja resolución en chuvasios (FH: 1.2).

^hEncontrado en Irlanda del Sur (FH: 4.3), noroeste de Inglaterra (FH: 4.2), sami (FH: 4.2), y con HLA de clase I en baja resolución en chuvasios (FH: 4.9).

ⁱ, ^q, ^sHaplotipo de clase II encontrado en georgianos (FH: 12.6), sami (FH: 11.7), turcos (FH: 9.8), tofalar (FH: 9.3), todja (FH: 9.1), cretenses (FH: 8.9), keti (FH: 8.8), evenki (FH: 8.6) y chuvasios (FH: 4.8).

^jEncontrado en sami (FH: 1.9).

^kHaplotipo de clase I encontrado en tunecinos (FH: 3.0). Haplotipo de clase II encontrado en chuvasios (FH: 12.1), sami (FH: 9.2), tuvanos (FH: 9.0), georgianos (FH: 8.8), mongoles khoton (FH: 8.1), mansi (FH: 6.6), cretenses (FH: 6.3), aleutianos (FH: 5.6) y persas de Yazd (FH: 5.4).

^lHaplotipo de clase I encontrado en omaneses (FH: 4.9). Haplotipo de clase II como en “i”.

^mEncontrado en armenios (salvo DQB1, FH: 1.0). Haplotipo de clase I encontrado en marroquíes (FH: 3.6) y tunecinos (FH: 3.0). Haplotipo de clase II como en “o”.

ⁿHaplotipo de clase I como en “f”. Haplotipo de clase II encontrado en tunecinos (FH: 2.0).

^o, ^wHaplotipo de clase II encontrado en buriatos (FH: 22.0), bereberes (FH: 20.1), mansi (FH: 16.9), persas de Yazd (FH: 15.4), keti (FH: 11.8) marroquíes (FH: 12.1) y cretenses (FH: 10.4).

^pHaplotipo de clase I como en “l”. Haplotipo de clase II como en “k”.

^rEncontrado con clase I en baja resolución en palestinos (FH: 1.2).

^tHaplotipo de clase I encontrado en japoneses (FH: 10.7) y chinos blang (FH: 10.3). Haplotipo de clase II encontrado en marathas de la India (FH: 22.9), keti (FH: 2.9) y mongoles khoton (FH: 2.4).

^vEncontrado en Irlanda del Sur (FH: 11.5) y noroeste de Inglaterra (FH: 9.5) y con HLA de clase I en baja resolución en chuvasios (FH: 1.2).

^xHaplotipo de clase I encontrado en omaneses (FH: 4.3). Haplotipo de clase II como en “c”.

Referencias: (Martinez-Laso et al., 1996; Munkhbat et al., 1997; Grahovac et al., 1998; Izaabel et al., 1998; Kapustin et al., 1999; Gomez-Casado et al., 2000; Sanchez-Velasco & Leyva-Cobian, 2001; Martinez-Laso et al., 2001; Arnaiz-Villena et al., 2001a; Arnaiz-Villena et al., 2001b; Uinuk-Ool et al., 2002; Arnaiz-Villena et al., 2003; Moscoso et al., 2008; Johansson et al., 2008; Gonzalez-Galarza et al., 2011; Matevosyan et al., 2011; Rey et al., 2013)

- A*24:02-B*35:01-DRB1*11:01-DQB1*03:01: compartido con la población turcomana, y como sucede en esta, el más frecuente (véase apartado anterior).
- A*01:02-B*08:01-DRB1*03:01-DQB1*02:01: encontrado con HLA de clase I en baja resolución en chuvasios (FH: 1.2).
- A*03:01-B*07:02-DRB1*15:01-DQB1*06:02: encontrado en irlandeses del Sur (FH: 4.3), ingleses del noroeste (FH: 4.2), sami (lapones de Suecia, FH: 4.2), y con HLA de clase I en baja resolución en chuvasios (FH: 4.9).
- A*03:01-B*44:02-DRB1*04:01-DQB1*03:02: encontrado en sami (FH: 1.9). También se ha encontrado en baja frecuencia en turcomanos.

Los otros haplotipos más frecuentes se describen por primera vez en esta población, ya que no se han determinado en ninguna otra previamente. Sin embargo, la mayoría de los haplotipos de dos loci que componen los haplotipos extendidos coinciden con los obtenidos en los análisis de desequilibrio de ligamiento, y, al igual que en el caso de la muestra de turcomanos, son característicos de poblaciones procedentes de Siberia y Asia Central (A*24:02-B*35:01, A*03:01-B*07:02, DRB1*15:01-DQB1*06:02, DRB1*04:01-DQB1*03:02, DRB1*01:01-DQB1*05:01) (Martinez-Laso et al., 2001; Uinuk-Ool et al., 2002; Arnaiz-Villena et al., 2003). No obstante, algunos también se encuentra en una frecuencia considerablemente alta en poblaciones del Mediterráneo (A*02:01-B*50:01, A*33:01-B*14:01, DRB1*03:01-DQB1*02:01) (Izaabel et al., 1998; Gonzalez-Galarza et al., 2011; Tabla R7).

Haplotipos HLA extendidos frecuentes en población kurda

De la misma forma que para turcomanos y azeríes, se han obtenido los haplotipos HLA extendidos más frecuentes en población kurda, recogidos en la Tabla R7. Estos doce haplotipos representan el 26.8% del total de haplotipos encontrados en esta población. Observamos que el haplotipo de clase II DRB1*11:01-DQB1*03:01 está presente en asociación con cuatro haplotipos de clase I: A*24:02-B*18:09, A*01:01-B*35:01, A*03:01-B*44:02 y A*02:01-B*35:01 (este a su vez se encuentra también asociado con DRB1*01:01-DQB1*05:01); mientras que el haplotipo DRB1*03:01-DQB1*02:01 lo está con tres haplotipos de clase I distintos: A*24:02-B*58:01, A*01:01-B*08:01 y A*26:01-B*08:01. Solamente dos de estos haplotipos extendidos han sido descritos previamente en otras poblaciones con los alelos de clase I tanto en alta como en baja resolución:

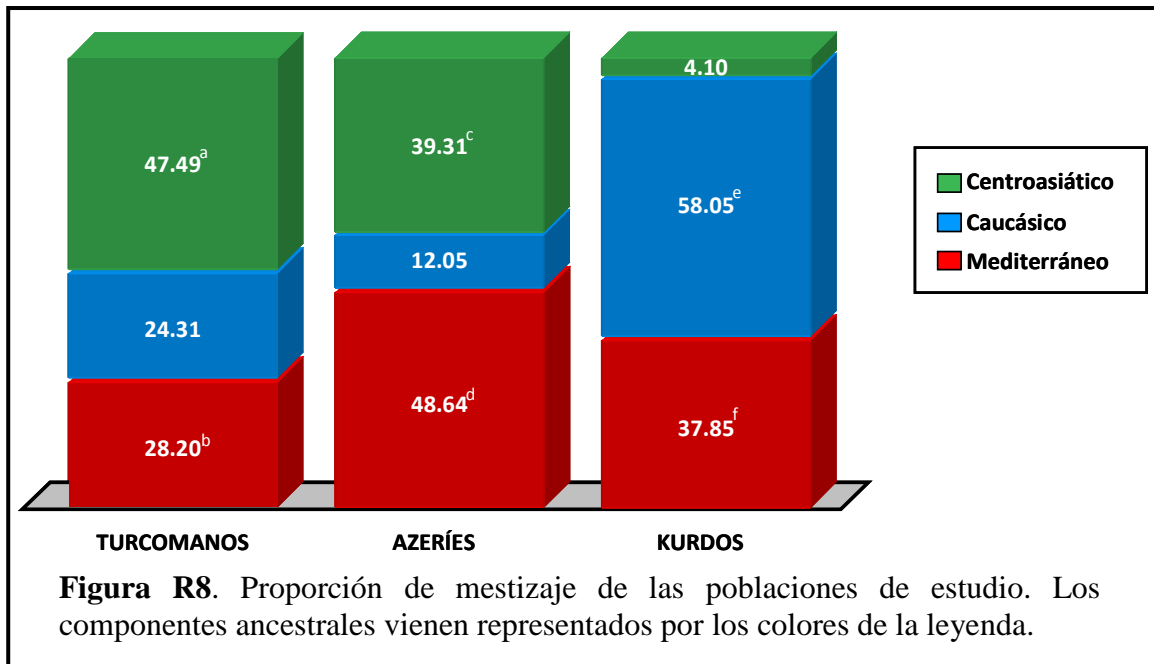
- A*01:01-B*35:01-DRB1*11:01-DQB1*03:01: encontrado únicamente en palestinos (FH: 1.2%) con la clase I en baja resolución.
- A*01:01-B*08:01-DRB1*03:01-DQB1*02:01: encontrado en irlandeses del Sur (FH: 11.5) e ingleses del noroeste (FH: 9.5) y con HLA de clase I en baja resolución en chuvasios (FH: 1.2).

El haplotipo de clase I A*01:01-B*08:01 es frecuente en poblaciones originarias de Asia Central, como los sami y los aleutianos. El resto de haplotipos no han sido descritos en ninguna población, siendo en kurdos la primera vez que se reportan, salvo los haplotipos A*02:01-B*35:01-DRB1*11:01-DQB1*03:01 y A*02:01-B*50:01-DRB1*07:01-DQB1*02:01, ambos hallados en población azerí en este mismo estudio. Sin embargo, la mayoría de los haplotipos de dos loci que componen los haplotipos extendidos coinciden con los obtenidos en los análisis de desequilibrio de ligamiento, y, al igual que en el caso de la muestra de turcomanos y azeríes, algunos son característicos de poblaciones procedentes de Siberia y Asia Central (A*01:01-B*08:01, DRB1*01:01-DQB1*05:01, DRB1*07:01-DQB1*02:01, DRB1*15:01-DQB1*06:01) (Grahovac et al., 1998; Martinez-Laso et al., 2001; Uinuk-Ool et al., 2002; Arnaiz-Villena et al., 2003; Moscoso et al., 2008). No obstante, otros también se encuentra en una frecuencia considerablemente alta en poblaciones del Mediterráneo y de Oriente Próximo (A*02:01-B*50:01, A*26:01-B*08:01, DRB1*03:01-DQB1*02:01) (Izaabel et al., 1998; Gonzalez-Galarza et al., 2011; Tabla R7).

5. ESTIMACIÓN DE LA PROPORCIÓN DE MESTIZAJE EN LAS POBLACIONES DE ESTUDIO

Mediante el programa LEADMIX v1.0 se ha estimado el componente ancestral en las poblaciones de estudio a partir de las poblaciones parentales seleccionadas. Se observa que tanto en turcomanos como en kurdos y azeríes, especialmente en esta última, existe un sustrato mediterráneo relevante en la conformación genética en cuanto al sistema HLA se refiere, con pocas variaciones entre estas poblaciones, siendo los primeros los que presentan una contribución menor de este componente (28.20%). Sin embargo, en lo que respecta al componente centroasiático, se observa un aumento de oeste a este de dicho componente que varía del 47.49% en turcomanos (la mayor contribución ancestral en esta población) al 4.10% en kurdos, una diferencia bastante significativa. Por último, la proporción de ascendencia caucásica es más o menos

similar en las dos poblaciones de lengua túrquica, mientras que en la población kurda es el principal componente ancestral, con un 58.05% (Figura R8).



DISCUSIÓN

1. GENÉTICA HLA DE LAS POBLACIONES DEL SUR DEL MAR CASPIO

Las poblaciones del sur del Mar Caspio (Irán), turcomanos, azeríes y kurdos, presentan un número reducido de alelos HLA de clase I con una frecuencia significativa (por encima del 5%), especialmente los primeros. Así en el caso de los turcomanos iraníes hemos observado únicamente siete alelos para el locus HLA-A y cuatro para el locus HLA-B, en el caso de los kurdos nueve para HLA-A y cuatro para HLA-B, y en el caso de los azeríes diez alelos para HLA-A y seis para HLA-B. Así los alelos más representativos en estas poblaciones se agrupan fundamentalmente en cinco grupos serológicos (superfamilias) para el locus HLA-A: A1, A2, A3, A9 y A11; y en cinco para el locus HLA-B: B5, B8, B13, B18 y B35, siendo este último y el A2 los grupos de clase I más representados.

Otras poblaciones del Suroeste Asiático próximas a las poblaciones de estudio han sido analizadas en los Talleres Internacionales de Histocompatibilidad. Se observa en la mayoría de los casos un reducido número de alelos de clase I representativos que se agrupan en las mismas superfamilias: A1, A2, A9, B5, B18 y B35 en poblaciones judías de Oriente Próximo; A2, A3, B5 y B35 en poblaciones del Cáucaso; y A1, A3, A9, A11, B5, B8 y B53 en poblaciones de India y Pakistán (Clayton & Lonjou, 1997).

El bajo número de alelos presentes en estas poblaciones radica en el entendimiento del origen del polimorfismo MHC que está relacionado con los eventos naturales de especiación. Estos se pueden dividir en dos escenarios: 1.- Las especies (con el sistema MHC) surgieron de un pequeño grupo de poblaciones fundadoras y poseían un limitado espectro de todas las posibles variantes (alelos) presentes en la población ancestral. 2.- Las especies provienen de un gran número de poblaciones fundadoras que recibieron la mayoría del “pool” alélico presente en sus especies ancestrales (proceso trans-específico generador del polimorfismo; Klein, 1987; Figueroa et al., 1988).

Las tres poblaciones presentan las frecuencias más elevadas para los mismos alelos HLA-A: A*02:01, A*03:01 y A*24:02. Estos alelos presentan una distribución poblacional bastante amplia, especialmente el A*02:01, encontrándose en el total de las diez regiones macrogeográficas diferentes con frecuencias significativas (Mack & Erlich, 2007; Solberg et al., 2008). Lo mismo sucede para el locus HLA-B, siendo los alelos más frecuentes para las tres poblaciones los mismos: B*08:01, B*35:01 y B*51:01. En este caso, los alelos también presentan una distribución poblacional amplia, encontrándose en la mayoría de las regiones macrogeográficas, sin embargo el

B*08:01 y el B35:01 están circunscritos a regiones más concretas, Norte de África y Suroeste Asiático, y Norteamérica, Europa y Suroeste Asiático, respectivamente (Solberg et al., 2008; Figura D1).

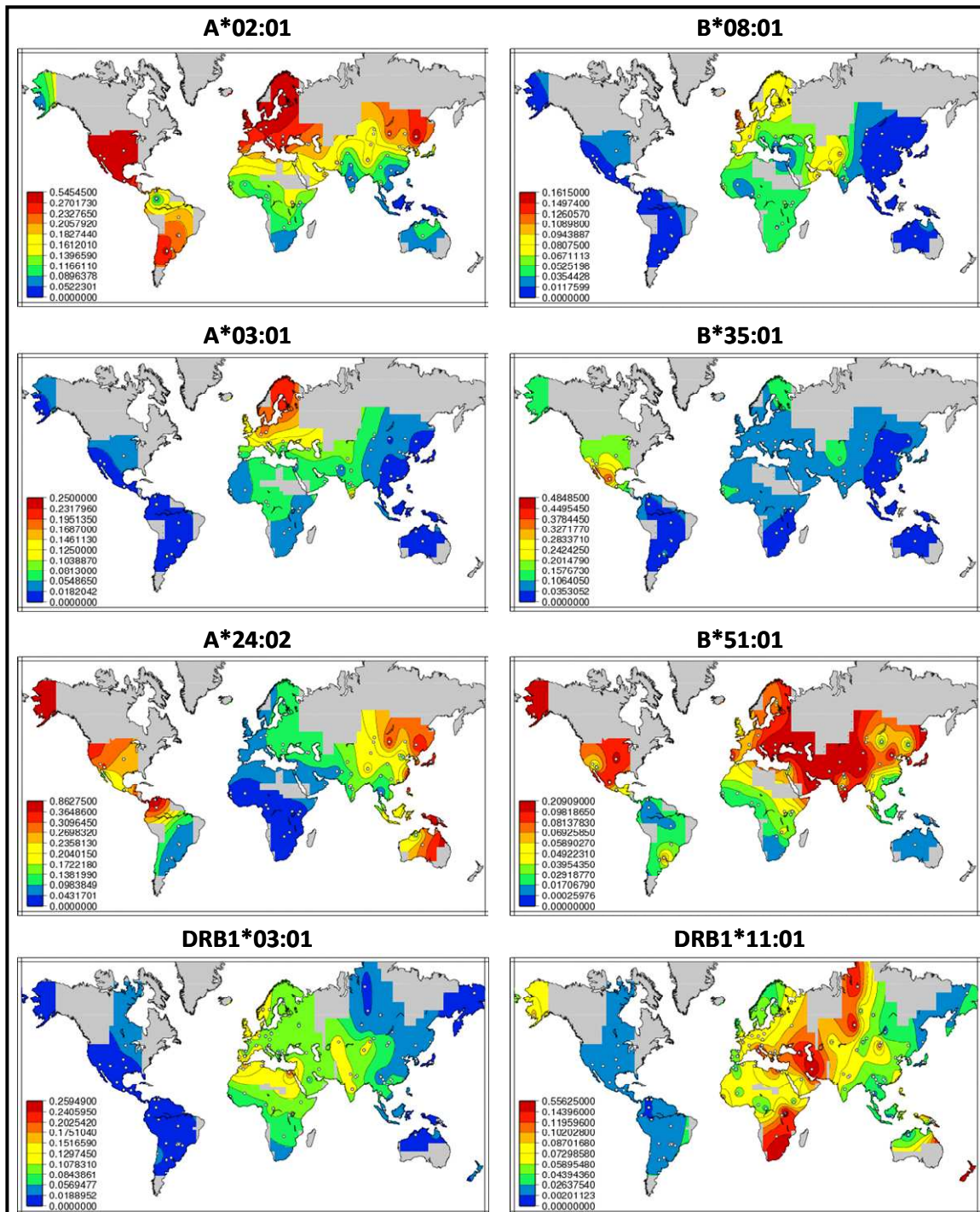


Figura D1. Representación geográfica de las frecuencias de alelos mayoritarios HLA de clase I y clase II presentes en las poblaciones de estudio. Las frecuencias alélicas se identifican mediante los colores de la leyenda (imágenes precedentes de Solberg et al., 2008 - ver www.pypop.org/popdata para más información).

En cuanto a los alelos HLA de clase II se observa la misma situación que en la clase I: hay poco alelos con una frecuencia relativamente importante, de tal forma que en las tres poblaciones para el locus HLA-DRB1 se han descritos seis alelos con frecuencias por encima del 5%. Entre ellos, los alelos más frecuentes coinciden en turcomanos, azeríes y kurdos: DRB1*03:01, DRB1*04:01, DRB1*11:01 y DRB1*15:01. Estos alelos están presentes en poblaciones de las diez regiones macrogeográficas, algunos presentando frecuencias muy elevadas en regiones concretas, como el DRB1*03:01 en poblaciones del Suroeste Asiático o el DRB1*04:01 en poblaciones del Noreste Asiático (Mack & Erlich, 2007; Solberg et al., 2008).

2. RELACIONES GENÉTICAS HLA E HISTÓRICAS ENTRE LAS POBLACIONES DEL SUR DEL MAR CASPIO Y OTRAS POBLACIONES

2.1 Relación con poblaciones iraníes

Los resultados de los análisis moleculares de los genes HLA en turcomanos, azeríes y kurdos del sur del Mar Caspio (Irán), junto con los estudios comparativos basados en el cálculo de las distancias genéticas (dendrogramas NJ y análisis de correspondencias) con otras poblaciones, muestran a las tres poblaciones agrupadas en un mismo nodo, y asociadas a las otras poblaciones iraníes incluidas en los análisis (persas de Shiraz y de Yazd, y baluchis), como se puede observar en las figuras R1, R2, R3, R4 y R5 (particularmente en las dos últimas).

Al comparar estas poblaciones en base a las distancias genéticas a partir de los genes HLA de clase II se observa que las menores distancias son entre los persas de Shiraz con cualquiera de las tres poblaciones en estudio, además de con las otras poblaciones iranianas, siendo este grupo persa y los kurdos los que se sitúan más próximos. Por el contrario, los baluchis son los más distantes a las poblaciones de estudio y a los persas de Yazd (globalmente son los que presentan mayores distancias genéticas), al tratarse de un grupo aislado y alejado geográficamente (se localizan en su mayoría en el sudeste de Irán, Figura I14).

El origen de los pueblos iranios en una hipotética invasión indo-irania procedente del interior de Asia ha suscitado mucho debate. La hipótesis de que tribus pertenecientes a la cultura de Andrónovo y/u otras relacionadas a esta colonizaron la meseta iraní en al menos dos oleadas, introduciéndose hasta Anatolia, (Hennerbichler, 2012) se sustenta únicamente en base a las teorías lingüísticas más asentadas.

Diversos estudios genéticos basados en marcadores clásicos establecen que las distintos grupos iraníes son bastante homogéneos, incluidas las poblaciones del sur del Mar Caspio, con pequeñas diferencias entre poblaciones del norte y del sur, y a su vez, relacionados con poblaciones del Cáucaso y del Mediterráneo (Quintana-Murci et al., 2004; Nasidze et al., 2004; Nasidze et al., 2009). Los estudios HLA realizados hasta la fecha corroboran la homogeneidad entre los iraníes modernos, a pesar de sus diferencias lingüísticas, como ocurre con los pueblos de lenguas túrquicas, así como su relación con pueblos del Mediterráneo y Oriente Próximo (Arnaiz-Villena et al., 2002; Farjadian et al., 2009; Farjadian & Safi, 2013), lo que sugiere la presencia de un sustrato genético común a pesar de las incursiones de pueblos foráneos. Por lo tanto, el hecho de que la invasión antes mencionada se desarrollara como un proceso de “iranización” por una minoría gobernante, a través del cual se impuso una cultura y una lengua, es probable.

Sin embargo, otros estudios basados en ADNmt y cromosoma Y muestran diferencias en cuanto a la distribución de determinados haplogrupos entre poblaciones del norte y del sur de Irán, debido a la presencia de barreras geográficas (desierto de Lut) que hubieran dificultado el flujo génico y favorecido el aislamiento, quizás en épocas más recientes (Regueiro et al., 2006; Terreros et al., 2011). Esto podría explicar las diferencias en cuanto a las frecuencias alélicas HLA de clase II y distancias genéticas que observamos en los baluchis con respecto al resto de poblaciones iraníes. Además también se ha observado en estas poblaciones la presencia de haplogrupos de estos marcadores de linaje con origen en Asia Central (haplogrupos descendientes del macrohaplogrupo M de ADNmt y haplogrupos P-M45, R-M173 y R-M17 de cromosoma Y) que si podrían estar relacionados con la expansión de la cultura de Andrónovo desde la región altaica hacia el “corredor iraní” y su posible huella genética (Wells et al., 2001; Maliarchuk et al., 2002; Schonberg et al., 2011).

2.2 Relación con poblaciones del Mediterráneo y del Cáucaso

El análisis de las distancias genéticas basadas en las frecuencias alélicas no muestra un patrón de agrupamiento geográfico de las poblaciones incluidas en la comparación, sin embargo las tablas R3 y R4, concretamente esta última, asocian las poblaciones del sur del Mar Caspio a poblaciones de su entorno, tanto de Oriente Próximo como, especialmente, del Cáucaso, al presentar valores D_A inferiores a 30×10^{-2} con respecto a las poblaciones de estudio. Con respecto a las distancias genéticas

basadas en las frecuencias haplotípicas HLA de clase II (DRB1-DQB1) sí que se observa una mayor asociación con estas poblaciones (Tabla R6).

Los dendrogramas contruidos a partir de las distancias basadas en las frecuencias alélicas (Figuras R1, R2, R4 y R5) agrupan a las poblaciones de estudio con las poblaciones mediterráneas y del Cáucaso, aunque en un nodo externo. Sin embargo, los análisis de correspondencias con la poblaciones euroasiáticas (Figuras R1 y R4), aunque sitúa a las poblaciones de estudio próximas (respecto a la segunda dimensión), son los kurdos los que se agrupan con las poblaciones de Oriente Próximo y Cáucaso, mientras que turcomanos y azeríes se sitúan próximos a poblaciones del este y centro de Asia. La agrupación con poblaciones del mediterráneo y del Cáucaso se hace patente en los análisis de emparentamiento a partir de las frecuencias haplotípicas de clase II (Figura R7).

2.2.1 Poblaciones del Mediterráneo Oriental: turcos y judíos

Anatolia, que supone la mayor parte de la actual Turquía, ha sido habitada desde el Paleolítico, con la presencia de muchos de los primeros signos de la civilización occidental por parte de los pueblos pre-indoeuropeos e indoeuropeos: desde Catal Huyuk, cerca de la ciudad de Konya, núcleo urbano construido durante el periodo neolítico (7000 a.C.) y la cultura Hacilar, que floreció en el sudeste de la región (5000-2400 a.C.) al contar con los recursos necesarios para extraer y trabajar el bronce, hasta el surgir del Imperio Hitita en la Anatolia central, considerado un pueblo indoeuropeo (3000-2000 a.C.), del reino de Arzawa en la costa egea (1400 a.C.) y del reino de Urartu situado en lo que hoy es Armenia (900 a.C.) (Macqueen, 1999). Después del gobierno impuesto por los Imperios Medo y Persa (siglo VI a.C. – siglo I d.C.), los romanos, y más tarde, los musulmanes en menor grado, tomaron el control de Anatolia hasta la invasión por parte de los pueblos túrquicos oghuz procedentes de Asia Central en el 1055 d.C. que finalmente tomaron Estambul en el 1453 d.C.

Estudios genéticos previos muestran una asociación estrecha de los turcos con otras poblaciones mediterráneas. Uyar *et al.* al analizar las frecuencias haplotípicas HLA de clase I confirman esta asociación, y argumentan que puede ser debida no solamente a la historia compartida entre estas poblaciones, sino a las similitudes derivadas de circunstancias ecológicas comunes (Uyar et al., 2004). Otros estudios basados en HLA y en otros marcadores genéticos apuntan a que, a pesar de esta larga sucesión de invasiones por parte de pueblos procedentes del exterior, el sustrato

genético del pueblo turco actual es mediterráneo (Calafell et al., 1996; Brega et al., 1998; Rosser et al., 2000; Arnaiz-Villena et al., 2002). Nuestros análisis parecen reflejar lo mismo, al situar la muestra de turcos no lejos de cretenses, macedonios y sardos (Figuras R1 y R4).

Los judíos junto con los fenicios son considerados la rama occidental de los pueblos nómadas semitas. Es ampliamente aceptado que alrededor del 1850 a.C. un grupo semita emigró de la ciudad de Ur, próxima al río Éufrates, al Levante Mediterráneo (Longanesi, 1986). Estos pueblos, conocidos como cananeos, se instalaron en la franja Sirio-Palestina fundando numerosas ciudades estado. Debido a las duras condiciones climáticas algunos decidieron dedicarse a la navegación y al comercio, mientras que los de la zona sur (Palestina) se hicieron sedentarios, adoptando la monarquía como forma de gobierno. De esta forma se instauró el reino de Judea gobernado por los judíos. Este estado fue anexionando a distintos imperios en distintas épocas (Egipto, Babilonia, Persia) hasta la conquista por parte del Imperio Romano en el año 70 d.C., (Longanesi, 1986) cuando empezó la conocida y debatida Diáspora.

Estudio genéticos basados en HLA clasifican a las poblaciones judías en asquenazí (los que se asentaron en Europa Central y del Este) y no asquenazí (los que se asentaron en el Mediterráneo y Asia) al mostrar más heterogeneidad entre ellos en cuanto a HLA (Bonne-Tamir et al., 1978). Arnaiz-Villena *et al.* sostienen que ambas poblaciones judías tienen características HLA mediterráneas, aunque de forma coherente, son los no asquenazí los que presentan una mayor asociación con otros pueblos de la región, incluidos los palestinos, al haber preservado el “pool” original de alelos HLA mediterráneos (Martínez-Laso et al., 1996; Arnaiz-Villena et al., 2001b).

Kurdos, azeríes y turcos presentan de forma significativa haplogrupos J1 y J2 del cromosoma Y cuyo origen se supone que está en el Mediterráneo Oriental, ya que son dominantes en poblaciones semitas como judíos y libaneses (Nasidze et al., 2005; Andonian et al., 2011). Concretamente en el caso de los kurdos, el conocido como “haplotipo modal kurdo” difiere solamente en una posición del “haplotipo modal Cohen” judío asociado a la casta sacerdotal Cohanim (Cohen en singular), lo que para varios autores supone que tanto los kurdos como los judíos pertenecieron a las poblaciones indígenas que fundaron las civilizaciones de Oriente Próximo (Hennerbichler, 2012). Sin embargo, los kurdos y los azeríes y turcos hablan lenguas que pertenecen a familias distintas, que a su vez son de origen diferente a las lenguas semíticas habladas por las poblaciones de Oriente Próximo (hebreo y árabe). Estos

hechos apoyarían la tesis de la imposición cultural y lingüística por parte de las élites de los distintos pueblos invasores del Suroeste Asiático, que dieron lugar a procesos de “iranización” y “turquificación” de los antiguos habitantes de Mesopotamia y Anatolia, respectivamente.

Las distancias genéticas a partir de las frecuencias alélicas presentan a los turcos y a otras poblaciones del Mediterráneo oriental (macedonios y cretenses) muy próximos a los azeríes y a los kurdos, especialmente a estos últimos, mientras que el valor aumenta al comparar con los turcomanos, a pesar de hablar estos y los turcos lenguas de origen común (Tablas R3 y R4). Las poblaciones judías son las poblaciones mediterráneas orientales más alejadas a las poblaciones de estudio, junto con libaneses, aunque igual que en el caso anterior, los kurdos son los más próximos a judíos, especialmente a los no asquenazí-marroquíes (Tablas R4). Se puede deducir que, aunque en las representaciones de los análisis de emparentamiento aparecen siempre agrupadas en el mismo conjunto (Figuras R1-R6), la proximidad geográfica influye en las relaciones genéticas HLA entre estas poblaciones. Los resultados en lo que respecta a la población kurda de Irán son coherentes con las tesis de la “dominancia de élite” propuesta anteriormente.

2.2.2 Poblaciones del Cáucaso: georgianos, esvanos y armenios

El Cáucaso es una región natural formada por la cordillera del mismo nombre y las zonas bajas circundantes en la que coexisten en la actualidad varios estados: Georgia, Armenia y Azerbaiyán al sur, y diversas repúblicas de la Federación Rusa al norte. El pueblo georgiano está relacionado con los pueblos de Mesopotamia y Anatolia desde tiempos antiguos cuando los hititas y los asirios se asentaron en la región en el 1200 a.C. y se mezclaron con las poblaciones autóctonas del Cáucaso (Rosen, 1992). En el 730 a.C. pueblos procedentes del norte, los escitas y los cimérios, invadieron Georgia, y muchas de sus pobladores se refugiaron en las zonas altas. Después de esto, las principales tribus del reino de Iberia (también conocido como reino de Kartli que ocupaba el centro y este de la actual Georgia), los tubal y los moschi, formaron el núcleo de la población georgiana, sobreviviendo a las guerras que mantuvieron con el reino de Urartu del sur. Los romanos conquistaron el reino en el siglo I d.C. que a partir del siglo V d.C. pasa a estar bajo la protección de Bizancio, adoptando la rama ortodoxa del cristianismo como religión (Sellier & Sellier, 1993). Los árabes invaden Georgia en el 645 d.C. convirtiéndolo en un estado tributario con cierta independencia, al no ser

islamizados. A finales del siglo IX d.C., bajo la nueva dinastía de gobernadores Bagrationi, se instaura de nuevo un reino georgiano. Los esvanos conforman un grupo étnico que habitan la región de Svanetia, en el noroeste de Georgia. Debido al entorno montañoso que les confirió aislamiento geográfico, los esvanos permanecieron ajenos a las invasiones externas. Tanto la lengua georgiana como la esvana son miembros de la familia lingüística kartveliana (caucásica meridional), muy próxima a las lenguas dene-caucásicas (Ruhlen, 1987; Figura D2).



Los armenios tienen la creencia de que provienen de los antiguos urartitas, del reino de Urartu (véase apartado anterior), pueblo perteneciente a la cultura hurrita, de lengua pre-no-indoeuropea, y que tuvo su máximo esplendor en el 900 a.C. (Cotterel, 1980). Su núcleo de poder se asentó en el lago Van (al este de la actual Turquía), pero los conflictos con los imperios vecinos, asirios al oeste y medos al este, y las invasiones procedentes de los pueblos del norte (cimerios y escitas), lo debilitaron llevándolo a su fin en el 600 a.C. Esto dio paso a la instauración del reino de Armenia, primero como satrapía de los Imperios Medo y Aqueménida. Este reino atravesó a lo largo de su historia periodos de autonomía e independencia, como por ejemplo bajo la dinastía Bagrátida (siglos IX-XI d.C.), pero debido a su situación geográfica entre dos

continentes ha sufrido invasiones de multitud de pueblos e imperios: persas, romanos, bizantinos, árabes, turcos selyúcidas, mongoles, turcos otomanos y rusos. Una de estas invasiones, la llevada a cabo por los selyúcidas (1064 d.C.), dio origen al llamado reino armenio de Cilicia, entre las actuales Turquía y Siria, fundado por la dinastía real de los Rupénidas que huyeron de dicha invasión. Los turcos lo invadieron a finales del siglo XIV d.C., poniendo fin a este reino (Sellier & Sellier, 1993). Esto supuso el inicio de la diáspora Armenia, que continuó durante la dominación otomana, llevando a la mayoría de la población principalmente a países de Oriente Próximo, Europa y América. Armenia fue el primer país en adoptar el cristianismo como religión oficial del estado, en el 301 d.C., doce años antes de que lo hiciese el Imperio Romano, lo que podría haber llevado a un mayor aislamiento con respecto a sus vecinos no cristianos (Hovanissian, 1997). El pueblo armenio habla una lengua de origen indoeuropeo con un alfabeto propio, probablemente impuesta por los imperios persas que lo dominaron, aunque su lengua original pudo pertenecer al grupo de las lenguas dene-caucásicas (Ruhlen, 1987; Arnaiz-Villena et al., 2002; Figura D2).

Estudios basados en marcadores genéticos clásicos presentan a las poblaciones del Cáucaso como portadoras de haplogrupos, tanto de ADNmt (H, J, T y U5) como de cromosoma Y (J1 y J2, en el caso de los armenios), que los relacionan con las poblaciones de Anatolia, Oriente Próximo y con azeríes y kurdos (Quintana-Murci et al., 2004; Hennerbichler, 2012). Esto se sitúa en la línea de los trabajos basados en HLA que establecen una mínima diferenciación entre poblaciones mediterráneas orientales y occidentales en lo que se refiere a estos genes, conformando un sustrato HLA mediterráneo, en el que se incluyen las poblaciones del Cáucaso y los kurdos (Sanchez-Velasco & Leyva-Cobian, 2001; Arnaiz-Villena et al., 2002; Rey et al., 2013). Concretamente los armenios parecen estar muy relacionados con los mediterráneos orientales: turcos, judíos y cretenses (Arnaiz-Villena et al., 1999; Arnaiz-Villena et al., 2001c; Arnaiz-Villena et al., 2002; Matevosyan et al., 2011).

Las distancias genéticas a partir de las frecuencias alélicas sitúan a georgianos y esvanos entre las poblaciones más próximas a kurdos (Tablas R3 y R4). Azeríes y turcomanos también presentan valores de distancia bajos, aunque no tanto como la muestra kurda. Los armenios, aunque se sitúan muy próximos a georgianos y esvanos en cuanto a sus distancias D_A (10.76×10^{-2} y 20.43×10^{-2} , respectivamente), no se encuentran entre los grupos más próximos a las poblaciones de estudio, mostrando mayor relación con los kurdos (Tabla R3). Los dendrogramas y los análisis de

correspondencias muestran a las poblaciones del Cáucaso en el conjunto de poblaciones del mediterráneo, situando en el mismo nodo a las tres poblaciones caucásicas, junto con los turcos (Figura R1). Esto coincide con la tesis del origen común de las poblaciones de Oriente Próximo y Cáucaso, defendida por múltiples estudios previos. Los kurdos, seguidos de los azeríes, son los que guardan una mayor asociación con estas poblaciones.

2.3 Relación con poblaciones de Asia Central

Asia Central es la región asiática comprendida entre el Mar Caspio y China occidental y entre Siberia meridional y norte de la India. Esta región ha servido geográfica e históricamente como vía de comunicación de personas y mercancías, así como crisol de pueblos y culturas, al estar ligada a la Gran Ruta de la Seda desde su origen. Entre las regiones pertenecientes a Asia Central se encuentran las Repúblicas de Tuvá y de Buriatia integradas en la Federación Rusa, zonas de origen de dos poblaciones sugeridas como muy próximas a la población ancestral de gran parte de los grupos euroasiáticos y amerindios: los tuvanos y los buriatos (Uinuk-Ool et al., 2002).

Los primeros son descritos como una mezcla de grupos mongoloides y túrquicos, que habitan un área que se extiende desde el norte de Mongolia hasta la costa norte del lago Baikal, próxima al macizo de Altái. Se les supone un pueblo aislado en relación a otras poblaciones de la región, ya que se diferencian en base a diversos parámetros antropomórficos. Se trata de un pueblo mayoritariamente nómada dedicado al ganado de cabras, ovejas, caballos y camellos. Hablan una lengua túrquica de la familia altaica con la presencia de muchos vocablos de origen mongol (Martinez-Laso et al., 2001; Uinuk-Ool et al., 2002). Los buriatos son un pueblo mongoloide que habitan en los alrededores del lago Baikal, y que al igual que lo tuvanos, practican el nomadismo y la ganadería ovina, caprina, equina y de camellos. Hablan una lengua mongólica de la familia altaica muy próxima al mongol (Uinuk-Ool et al., 2002). Los todja y los tofalar son grupos étnicos que habitan también en la República de Tuvá y zonas colindantes, con una marcada similitud antropomórfica, pero diferentes a los tuvanos (Uinuk-Ool et al., 2002). Se considera que tienen un origen mixto de población túrquica con tribus del norte de Siberia, samoyedos y keti (los tofalar hablaban una lengua samoyeda que ha sido reemplazada recientemente por una lengua de origen túrquico).

Otros pueblos que se supone son originarios de Asia Central aunque localizados en Siberia occidental y este de Europa, y que guardan relación con las poblaciones

descritas anteriormente, son los mansi y los chuvasios. Los primeros son un pueblo ugro que viven en el distrito autónomo ruso de Janty-Mansi y que tradicionalmente subsisten gracias a la pesca, a la caza y a la cría de reno. Hablan la lengua mansi, perteneciente a la familia fino-úgrica (Uinuk-Ool et al., 2002). A los segundos se les considera un origen mixto de población túrquica oghur protobúlgara (pueblo altaico) con los pueblos ugricos que habitaban la región en la que se localizan mayoritariamente, la república de Chuvasia, en la Rusia europea. Hablan chuvasio, una lengua de la rama túrquica oghur (Lewis et al., 2014).

Trabajos previos basados en ADNmt y cromosoma Y constatan un alto nivel de diversidad de haplogrupos en Asia Central, lo que sugiere que las poblaciones centroasiáticas están entre las más antiguas del continente, de las cuales han surgido los linajes ancestrales de Europa, América y el subcontinente indio, pero con escasa contribución a las poblaciones del Suroeste Asiático, entre ellas las poblaciones iraníes (Wells et al., 2001; Quintana-Murci et al., 2001). Estas conclusiones pueden ser sorprendentes si atendemos al gran número de migraciones documentadas por parte de grupos altaicos, en distintas épocas (indo-iranios, mongoles y turcos), pero sirven para apoyar la hipótesis de que las poblaciones ancestrales asentadas en esta región eran lo suficientemente densas para absorber ese flujo génico, y que los procesos de “iranización” y “turquificación” se dieron fundamentalmente en el ámbito cultural y lingüístico a través de la élite gobernante. Derenko *et al.* en un estudio posterior determinan que existe una diferenciación genética clara entre las poblaciones del lago Baikal (buriatos) y las poblaciones altaicas actuales (tuvanos, todja y tofalar), que puede ser explicada por una contribución desigual de los linajes paternos de origen centroasiático y occidental al acervo genético de las poblaciones siberianas actuales (Derenko et al., 2006). También determinan que las poblaciones altaicas actuales son las que tienen una mayor presencia del haplogrupo R-M17, haplogrupo asociado con la expansión de la cultura Andrónovo (véase Introducción apartado 12.3 y Figura D3), por lo que podrían estar muy próximas a los grupos ancestrales indo-iranios. Por otro lado, estudios llevados a cabo en chuvasios detectan divergencias entre los linajes paterno y materno: mientras que el ADNmt evidencia una presencia mayoritaria de haplogrupos representativos de poblaciones europeas (H, U y K) y ausentes en poblaciones altaicas (Graf et al., 2010), el cromosoma Y pone de manifiesto que el haplogrupo mayoritario es el R-M17, de origen centroasiático-altaico, aunque los siguientes haplogrupos en frecuencia son el N3 y el J, representativos de los pueblos ugricos y caucásico-

mesopotámicos respectivamente (Tambets et al., 2004). Esto sugiere que aunque la base de la población es ugría-cuacásica, los varones de los pueblos protobulgaros que colonizaron la región contribuyeron a los linajes paternos.

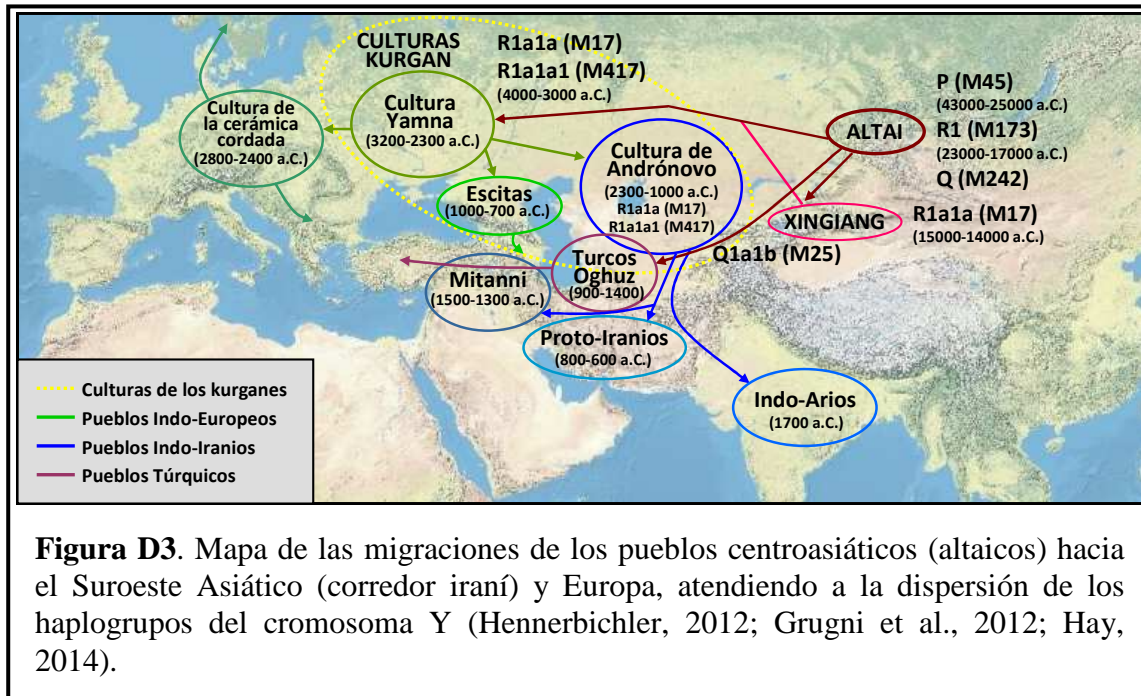


Figura D3. Mapa de las migraciones de los pueblos centroasiáticos (altaicos) hacia el Suroeste Asiático (corredor iraní) y Europa, atendiendo a la dispersión de los haplogrupos del cromosoma Y (Hennerbichler, 2012; Grugni et al., 2012; Hay, 2014).

Estudios basados en HLA en estas poblaciones confirman en gran medida lo expuesto por los trabajos mencionados previamente. Uinok-Ool *et al.* establecen que las poblaciones centroasiáticas y siberianas forman un grupo monofilético en el que los buriatos y mongoles, y las poblaciones altaicas, están bien diferenciados (Uinok-Ool et al., 2002). Por otra parte, los tuvianos son los que presentan un mayor polimorfismo HLA, lo que sugiere que es un grupo mixto a partir de pueblos mongoles, túrquicos y caucasoides, como resultado de la región en la que se localizan, históricamente documentada como crisol demográfico, algo que es confirmado por otros autores (Martinez-Laso et al., 2001). En cuanto a los chuvasios, los genes HLA los relacionan con poblaciones centroeuropeas y caucásicas, con escasa influencia de poblaciones asiáticas y siberianas (Arnaiz-Villena et al., 2003).

En el presente estudio, los dendrogramas presentan a las poblaciones de origen centroasiático en dos grupos bien diferenciados: los buriatos y los mansi, próximos a chuvasios y aleutianos (cuyo origen se presume que está en las poblaciones del lago Baikal, Moscoso et al., 2008) en un grupo más próximo a las poblaciones europeas y mediterráneas; y las poblaciones altaicas (tuvianos, todja y tofalar) próximas a

poblaciones siberianas (Figuras R2 y R5). El dendrograma basado en las frecuencias haplotípicas HLA de clase II, que tiende a homogeneizar las poblaciones, las presenta también separadas pero en el mismo conjunto, y alejadas de chuvasios y aleutianos (Figura R7). Al analizar las distancias genéticas, observamos que son los turcomanos y los azeríes los más próximos a chuvasios, mansi y poblaciones altaicas, con valores D_A entre 15 y 24×10^{-2} . Los buriatos, sin embargo, son el grupo centroasiático más alejado de las poblaciones de estudio (Tablas R3 y R4). El origen caucásico de gran parte del acervo genético de los chuvasios puede explicar la proximidad genética a estas poblaciones, ya que de entre todas éstas es la más próxima a los kurdos (Arnaiz-Villena et al., 2003; Graf et al., 2010). La relación con los mansi, cuyo origen se supone que está en la región Altái-Baikal, podría deberse a una mezcla posterior con genes caucásicos (Uinuk-Ool et al., 2002). En cuanto a la relación entre turcomanos, azeríes y poblaciones altaicas, tanto las distancias genéticas como los análisis de correspondencias parecen indicar que poblaciones ancestrales próximas a tuvanos, todja y tofalar pudieron contribuir en parte a la genética de estas poblaciones. Los turcomanos de Irán proceden de grupos iraníes o pre-iraníes que habitaban al este del Mar Caspio, que se mezclaron con las tribus túrquicas oghuz, procedentes de Asia Central (Altái) entre los siglos X y XV d.C. (Luther, 2001; Johanson, 2006) y que más tarde se establecieron en el noreste de Irán (Golestán y Jorósán), lo que puede justificar esta relación, confirmada por estudios con otros marcadores (Grugni et al., 2012). Sin embargo, a pesar de que existe una fuerte asociación entre los azeríes y poblaciones del Cáucaso y del Mediterráneo oriental, mayor que entre estos y los turcomanos, estos resultados podrían sugerir que el proceso de “turquificación” causado por las tribus altaicas oghuz también dejó su huella en la constitución genética de esta población, conclusión defendida por otros autores (Schonberg et al., 2011) y por datos históricos y lingüísticos (Yarshater, 1988). No obstante, esta relación pudo deberse a un contacto anterior con las tribus de origen altaico (indo-iránicas), como sugieren estudios basados en otros marcadores (Yepiskopasian et al., 2011; Grugni et al., 2012).

2.4 Análisis de haplotipos

El estudio del desequilibrio de ligamiento entre dos loci (HLA-A-B, HLA-B-DRB1 y HLA-DRB1-DQB1) en las poblaciones iraníes del sur del Mar Caspio nos indica, en primer lugar, la frecuencia significativa de un número considerable de haplotipos bi-locus de clase II que contrasta con los pocos haplotipos frecuentes

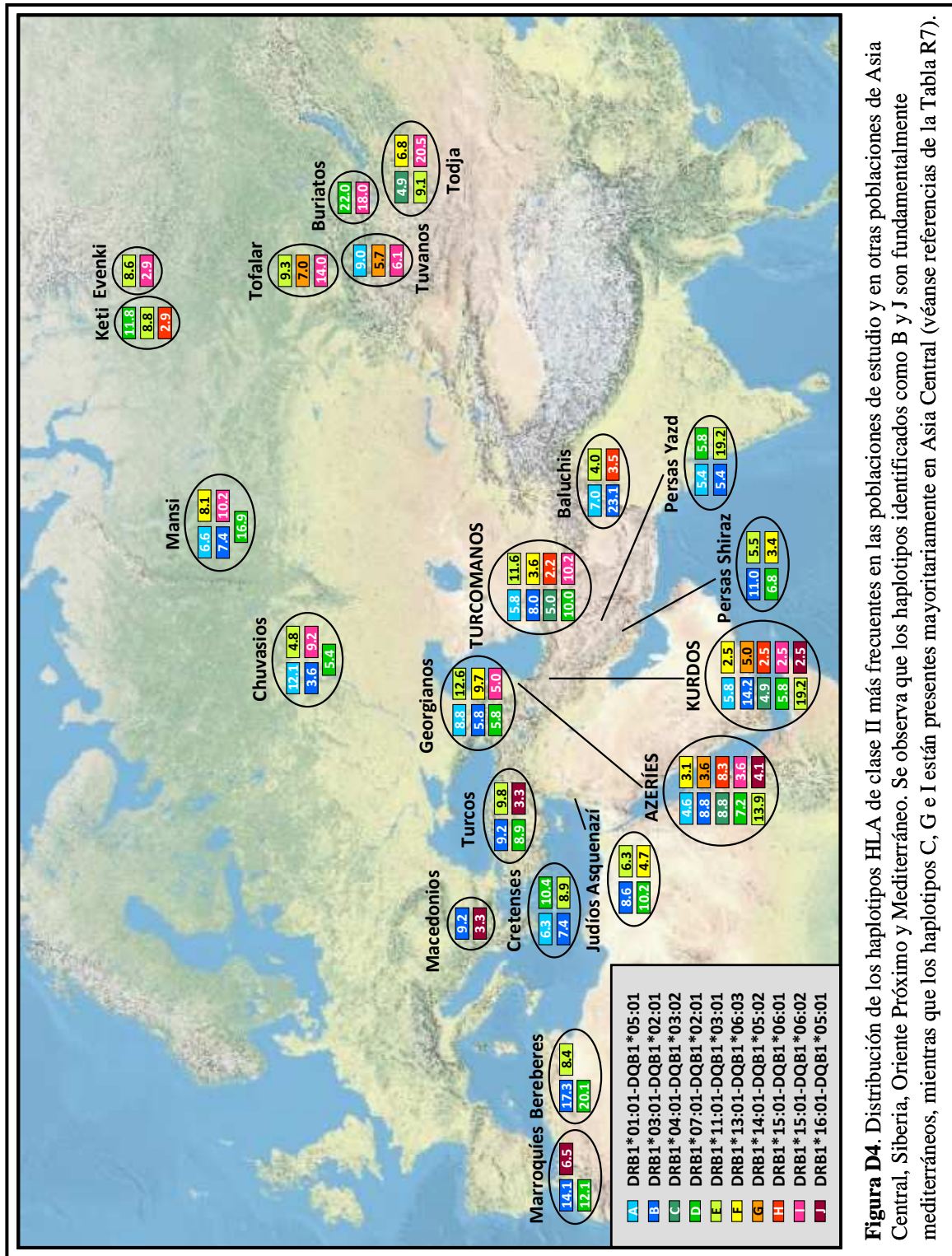
detectados para las otras dos combinaciones analizadas. Esto se debe en parte al elevado polimorfismo de los loci HLA-A y HLA-B tanto a nivel general como a nivel de las muestras estudiadas, y en parte a la mayor distancia en pb que separa a estos loci entre sí y con respecto al locus HLA-DRB1, lo que supone un desequilibrio de ligamiento menos significativo. Por otro lado se observa un número considerable de asociaciones características de poblaciones a priori no asociadas genéticamente con las poblaciones de estudio, o que no han sido identificadas en ninguna población analizada hasta la fecha.

De los 44 haplotipos HLA-A-B encontrados estadísticamente significativos en las poblaciones de estudio (14 en turcomanos, 20 en azeríes y 15 en kurdos, Tabla R5) solamente 19 han sido descritos en otras poblaciones con una frecuencia significativa, y de estos únicamente 10 forman parte de los haplotipos HLA extendidos descritos (Tabla R7). Así, los haplotipos bi-locus A*02:01-B*13:01 (presente en turcomanos), A*02:01-B*35:01 (en kurdos), A*02:01-B*50:01 (en azeríes), A*26:01-B*08:01 (en kurdos), A*30:01-B*13:01 (en turcomanos y azeríes) y A*33:01-B*14:01 (en azeríes) son característicos de poblaciones del Mediterráneo, Cáucaso y de Oriente Próximo (Gómez-Casado, 2002; Gonzalez-Galarza et al., 2011). Por otro lado, los haplotipos A*03:01-B*07:02 (presente en azeríes), A*03:01-B*44:02 (en turcomanos), A*24:02-B*35:01 (en turcomanos y azeríes), A*24:02-B*52:01 (en kurdos) y A*31:01-B*35:01 (en turcomanos) se han descrito previamente en poblaciones cuyo origen se supone que está en Asia Central y del norte de Europa (Johansson et al., 2008; Gonzalez-Galarza et al., 2011). El resto de asociaciones HLA-A-B descritas, con frecuencias bajas en las poblaciones de estudio, como A*01:01-B*37:01, A*02:01-B*07:02, A*68:01-B*44:02, A*68:01-B*51:01 (todos presentes en azeríes) y A*02:01-B*51:01 (en turcomanos) también están presentes en poblaciones de origen centroasiático, como sami y aleutianos (Moscoso et al., 2008; Johansson et al., 2008).

Para los haplotipos HLA-B-DRB1, las poblaciones de estudio comparten con otras poblaciones sólo 11 de los 38 haplotipos encontrados estadísticamente significativos, de los cuales únicamente cinco forman parte de los haplotipos HLA extendidos (Tablas R5 y R7). Estos son: B*07:02-DRB1*15:01, que formando el haplotipo extendido A*03:01-B*07:02-DRB1*15:01-DQB1*06:02, y B*44:02-DRB1*04:01, que asociado al haplotipo extendido A*03:01-B*44:02-DRB1*04:01-DQB1*03:02, se encuentran en poblaciones del norte de Europa, irlandeses y sami (véase párrafo anterior), y que también están presentes en azeríes; B*08:01-DRB1*03:01 (presente en las tres

poblaciones de estudio), que asociado al haplotipo extendido A*01:01-B*08:01-DRB1*03:01-DQB1*02:01 es característico de europeos, y que en este caso está presente en kurdos (Sanchez-Velasco et al., 2003); B*50:01-DRB1*07:01, que formando parte del haplotipo extendido A*02:01-B*50:01-DRB1*07:01-DQB1*02:01 es característico de poblaciones mediterráneas (véase párrafo anterior), y que está presente en azeríes y kurdos; B*58:01-DRB1*03:01 que asociado a A*33:03 están presente en baluchis y asociado a A*33:01 aparece en las tres poblaciones estudiadas. Otras asociaciones presentes en las muestras de estudio se han descrito en sami y aleutianos (grupos cuyo origen se supone centroasiático), como B*37:01-DRB1*10:01 y B*39:01-DRB1*12:01. No ha sido posible comparar los desequilibrios de ligamiento entre dos loci HLA-A-B y HLA-B-DRB1 de muchas de las poblaciones centroasiáticas y siberianas analizadas para HLA de clase II ya que carecen de los perfiles HLA de clase I.

En cuanto a los haplotipos bi-locus HLA de clase II, de los 23 observados, 18 se han descrito en otras poblaciones con frecuencias significativas, y de estos la mitad forman parte de haplotipos extendidos (Tablas R5 y R7). De este modo, los haplotipos DRB1*03:01-DQB1*02:01 (presente en las tres poblaciones pero con elevada frecuencia en kurdos), y DRB1*16:01-DQB1*05:01 (en azeríes y kurdos) son característicos de mediterráneos (Gómez-Casado, 2002; Gonzalez-Galarza et al., 2011; Figura D4); mientras que los haplotipos DRB1*04:01-DQB1*03:02, DRB1*13:01-DQB1*06:03, DRB1*14:01-DQB1*05:02 y DRB1*15:01-DQB1*06:02 (todos descritos en las tres poblaciones, el último con elevada frecuencia en turcomanos) son mayoritarios en poblaciones de Asia Central o cuyo origen se sitúa en esta región (sami y chuvasios) y poblaciones siberianas (Uinuk-Ool et al., 2002; Johansson et al., 2008; Figura D4). El DRB1*07:01-DQB1*02:01 y DRB1*11:01-DQB1*03:01 presentan una distribución global, estando presente en la mayor parte de regiones macrogeográficas, fundamentalmente en poblaciones de Europa atlántica, Mediterráneo, Cáucaso y Asia. El resto de asociaciones HLA-DRB1-DQB1 descritas también están presentes en poblaciones centroasiáticas (DRB1*01:01-DQB1*05:01, DRB1*07:01-DQB1*03:03, DRB1*12:01-DQB1*03:01), siberianas (DRB1*09:01-DQB1*03:03, DRB1*14:02-DQB1*03:01, DRB1*15:01-DQB1*06:01) y mediterráneos (DRB1*13:02-DQB1*06:04, DRB1*10:01-DQB1*05:01, aunque este último también está presente en buriatos con una frecuencia significativa).



Cabe destacar el haplotipo DRB1*13:01-DQB1*06:02, descrito en azeríes y kurdos, ya que esta presente en las poblaciones persas con frecuencias significativas (Gonzalez-Galarza et al., 2011).

Los haplotipos extendidos estimados (HLA-A-B-DRB1-DQB1, Tabla R7 y sus referencias) presentan frecuencias relativamente bajas debido a que los haplotipos parciales también los encontramos en baja frecuencia, salvo algunas excepciones en la clase II. Esto indica que las poblaciones iraníes del sur del Mar Caspio han recibido influencia genética externa y no se trata de aislados genéticos. En el caso de los kurdos la posible mezcla con otras poblaciones no ha diluido con el tiempo el sustrato genético que lo relaciona con otras poblaciones del mediterráneo, principalmente de la región oriental, compartiendo haplotipos extendidos y/o parciales tales como A*02:01-B*35:01-DRB1*01:01-DQB1*05:01 (el de mayor frecuencia, FH: 3.3%), A*01:01-B*35:01-DRB1*11:01-DQB1*03:01, A*24:02-B*58:01-DRB1*03:01-DQB1*02:01, A*02:01-B*50:01-DRB1*07:01-DQB1*02:01 y A*26:01-B*08:01-DRB1*03:01-DQB1*02:01.

Sin embargo, los azeríes y los turcomanos aunque presentan haplotipos que indican ese sustrato mediterráneo como A*01:02-B*08:01-DRB1*03:01-DQB1*02:01, A*02:01-B*50:01-DRB1*07:01-DQB1*02:01, A*33:01-B*14:01-DRB1*01:01-DQB1*05:01 (presentes en azeríes), A*01:01-B*35:01-DRB1*03:01-DQB1*02:01 (en turcomanos), la influencia en la carga genética por parte de las poblaciones centroasiáticas que migraron a la región en diferentes épocas se observa en haplotipos tales como A*24:02-B*35:01-DRB1*11:01-DQB1*03:01 (el de mayor frecuencia en ambas poblaciones), A*03:01-B*44:02-DRB1*04:01-DQB1*03:02 (en azeríes), A*31:01-B*35:01-DRB1*15:01-DQB1*03:01, A*02:50-B*07:03-DRB1*13:01-DQB1*06:03 y A*24:02-B*35:01-DRB1*15:01-DQB1*06:02 (en turcomanos). El haplotipo kurdo A*01:01-B*08:01-DRB1*03:01-DQB1*02:01 y el haplotipo azerí A*03:01-B*07:02-DRB1*15:01-DQB1*06:02 compartidos con poblaciones irlandesa e inglesa, podrían surgir de un efecto fundador procedente de poblaciones mediterráneas ancestrales (tales como los íberos) que colonizaron las Islas Británicas tras la última glaciación (en el caso del primero), o de poblaciones escandinavas con origen en grupos centroasiáticos (como los sami), como defienden diversos autores (Oppenheimer, 2006).

A pesar de que no se han encontrado coincidencias entre los haplotipos HLA extendidos de las poblaciones de estudio con las otras poblaciones iraníes, haplotipos bi-locus de clase II con fuerte desequilibrio de ligamiento y frecuencias significativas en

las tres poblaciones, están muy representados en las poblaciones persas: DRB1*03:01-DQB1*02:01, DRB1*07:01-DQB1*02:01 y DRB1*11:01-DQB1*03:01.

Este estudio comparativo sirve para confirmar lo que nos indican las distancias genéticas y los análisis de emparentamiento: las tres poblaciones tienen un sustrato genético común que las emparenta con las poblaciones de la cuenca mediterránea y del Cáucaso. Este sustrato es mayoritario en kurdos, ya que presenta la mayor parte de los haplotipos, tanto bi-locus como extendidos, definidos como mediterráneos, y que están presentes también en las poblaciones persas. Sin embargo los azeríes y especialmente los turcomanos, presentan con frecuencias considerables haplotipos definidos como siberianos de origen centroasiático, lo que sugiere que las tribus altaicas (con origen en Asia Central) pudieron contribuir al acervo genético de estas poblaciones, en mucha mayor medida que en los kurdos. Poblaciones como los chuvasios tendrían una historia genética similar, ya que aunque muchos de sus haplotipos bi-locus son de origen europeo-caucásico, también presentan haplotipos de origen centroasiático, lo que indica que las tribus túrquicas invasoras dejaron su huella genética.

2.5 Estimación del mestizaje

Los resultados del estudio de la estimación del mestizaje se encuentran en la misma línea de los análisis previamente discutidos, ya que las tres poblaciones de estudio tienen un sustrato mediterráneo relevante y común, que en el caso de los azeríes se aproxima al 50% (Figura R8). Sin embargo, en lo que respecta a los turcomanos se observa que el componente ancestral mayoritario es de origen centroasiático, próximo al 50%. Esto concuerda con lo reportado en estudios previos con otros marcadores. Grugni *et al.*, mediante el análisis de haplogrupos del cromosoma Y, observan en una muestra similar de turcomanos del Golestán la presencia mayoritaria del haplogrupo Q-M25 originario de la región altaica (Grugni *et al.*, 2012; Figura D3). Otros autores identifican la presencia relevante de otros haplogrupos con el mismo origen, como son el P-M45 y su descendiente el R-M173, en población turcomana de Turkmenistán (Wells *et al.*, 2001; Figura D3). Los estudios basados en ADNmt observan hechos similares al identificar con una frecuencia importante en esta población haplogrupos descendientes del macrohaplogrupo M de origen centroasiático (Maliarchuk *et al.*, 2002; Quintana-Murci *et al.*, 2004). De igual modo, los azeríes presentan una proporción ancestral de origen centroasiática que, aunque sin ser la mayoritaria, es bastante relevante. Esto se encuentra en la línea con estudios basados en ADNmt que observan con una frecuencia

significativa haplogrupos descendientes del macrohaplogrupo M (Schonberg et al., 2011). Todo esto sugiere un aporte genético consistente tanto en turcomanos como en azeríes por parte de poblaciones altaicas en distintas épocas y a través de varias migraciones posibles.

En cuanto al componente ancestral caucásico, son los kurdos los que presentan una mayor proporción estando próxima al 60% (Figura R8). Estudios basados en ADNmt identifican con una frecuencia mayoritaria haplogrupos de origen caucásico o mesopotámico tales como H, J o T (Richards et al., 2000; Quintana-Murci et al., 2004). De hecho, los estudios de emparentamiento genético a partir de este marcador demuestran que las poblaciones kurdas procedentes de distintas regiones (Irán, Georgia, Turquía y Turkmenistán) se agrupan con poblaciones del sur del Cáucaso (Nasidze et al., 2005). La presencia dominante de haplogrupos J1 y J2 del cromosoma Y en poblaciones kurdas de distinta procedencia también sugiere que el origen de estas se situaría en torno al Cáucaso y Mesopotamia, al relacionarlo con poblaciones de esta región (armenios y georgianos; Wells et al., 2001; Nasidze et al., 2004).

3. PROGRAMA DE TRASPLANTES Y ELABORACIÓN DE LISTAS DE ESPERA VIRTUALES EN POBLACIONES DEL SUR DEL MAR CASPIO

La compatibilidad HLA entre donante y receptor en caso de trasplante de órganos sólidos y tejidos es necesaria si se quiere lograr la supervivencia del injerto (Gorer, 1936; Snell, 1948). Particularmente, en el caso de trasplante de médula ósea la compatibilidad debe ser estricta. Por esta razón los perfiles genéticos del donante y del receptor son un dato clave en la elaboración de las listas de espera que se encargan de priorizar en la distribución de los injertos susceptibles de ser trasplantados. En España, las listas de espera para la selección del receptor se confeccionan para los donantes cadáver y contienen los datos necesarios para decidir el receptor más adecuado ante un órgano concreto disponible. De los receptores potenciales se dispone de la siguiente información: datos clínicos (historia clínica, complicaciones actuales o pasadas, etc.), ciudad de referencia (para el criterio geográfico), edad, la antigüedad en el programa de trasplante, grupo sanguíneo, datos inmunológicos (HLA-A, -B y -DR) y la presencia o ausencia de anticuerpos preformados anti-HLA, su tipo y su cuantificación. Una vez establecida la compatibilidad HLA, para garantizar la igualdad y la equidad, los criterios se establecen atendiendo a dos aspectos fundamentales: territoriales y clínicos. Los primeros permiten que los órganos disponibles en una determinada área o región sean

trasplantados en esa misma región con prioridad para reducir al máximo el tiempo de isquemia, y los segundos contemplan la gravedad de los posibles receptores. Existen circunstancias adicionales a considerar que pueden estar por encima de los criterios territoriales: la urgencia absoluta o “urgencia 0”, circunstancia en la que el paciente necesita un trasplante en el menor tiempo posible para sobrevivir y por ello se trasplantará el primer órgano compatible por grupo sanguíneo y con prueba cruzada negativa (no existen anticuerpos circulantes citotóxicos que pueden acabar con las células del donante desencadenando el rechazo) sin atender a la compatibilidad HLA; y sujetos hiperinmunizados con anticuerpos circulantes contra una o varias proteínas del sistema HLA, de forma que es muy difícil encontrar un órgano aceptable y se les asignará el primero para el que no tenga los anticuerpos preformados frente a los antígenos del donante.

Los estudios sobre perfiles HLA de poblaciones concretas, como es nuestro caso, sirven para elaborar listas de espera virtuales para trasplantes en las que se definen los alelos y haplotipos HLA (A, B y DR, aunque también DQ) más y menos frecuentes para cada población/grupo étnico. En el caso de Irán, que cuenta con una gran cantidad de grupos étnicos distintos con fuerte presencia poblacional y con características genéticas propias, también en cuanto al sistema HLA, puede ser de gran ayuda, más cuando estas etnias también se distribuyen en otros países, tanto vecinos como lejanos. Nuestras poblaciones de estudio se ubican en un área urbana y no se trata de poblaciones aisladas, ya que comparten varios alelos entre sí con frecuencias significativas que a su vez presentan una amplia distribución global, como A*02:01, A*03:01, A*24:02, B*08:01, B*35:01, B*51:01, DRB1*03:01, DRB1*11:01 y DRB1*15:01. Sin embargo, son frecuentes alelos que en otras poblaciones son raros (A*02:50, B*35:27, DRB1*03:13) y muchos de los haplotipos observados son nuevos. Por esto, en Irán, donde no se analizan los perfiles genéticos HLA para comprobar la compatibilidad donante-receptor de manera rutinaria (Griffin, 2007), las listas virtuales pueden ser de gran utilidad si se obtuvieran los perfiles únicamente de los receptores. En estas poblaciones se aprecia la presencia del haplotipo bi-locus HLA-B8-DR3 (B*08:01-DRB1*03:01, especialmente en kurdos), que predispone a fallo renal en pacientes con nefropatía por IgA durante los trasplantes renales, lo cual se puede convertir en un factor de riesgo para el rechazo del trasplante, que en un hipotético programa de trasplantes debería tenerse en cuenta (Andresdottir et al., 2009).

En resumen, es imprescindible definir el perfil HLA de turcomanos, azeríes y kurdos y de los grupos étnicos vecinos para poder establecer programas de trasplantes y programas preventivos y epidemiológicos para enfermedades ligadas a HLA. También, es importante que exista mayor difusión e interrelación entre los programas de trasplantes existente en los países que albergan a estas poblaciones para establecer un buen programa de intercambio y crear una base de datos HLA accesible para todos que responda a las características genéticas particulares de cada población. Uno de los obstáculos para la realización de estos programas cooperativos es la legalidad de la venta de riñón “in vivo” para trasplante renal en Irán que, además de los conflictos éticos que puede suponer, permite únicamente que el donante y el receptor sean de la misma nacionalidad con el fin de evitar el turismo sanitario (Griffin, 2007). Este programa surgió a finales del siglo XX como respuesta a las largas listas de espera para este tipo de trasplante, a la ausencia de un programa de donación cadavérica y a la precariedad de los sistemas de diálisis (Griffin, 2007; Ghods, 2014).

4. ENFERMEDADES LIGADAS A HLA Y FARMACOGENÉTICA EN LAS POBLACIONES DEL SUR DEL MAR CASPIO

Desde que se observó por primera vez la relación entre determinados alelos del sistema HLA y patologías (determinados alelos de HLA-B y linfoma de Hodgkin; Aimel, 1967), numerosas variantes alélicas de estos genes han sido asociadas con enfermedades autoinmunes (principalmente), inflamatorias e infecciosas. El estudio de estas enfermedades en grupos étnicos distintos ha mostrado con frecuencia asociaciones diferentes entre distintos alelos del mismo grupo y una enfermedad concreta.

Entre las enfermedades autoinmunes ligadas a HLA se encuentra la diabetes mellitus tipo I (DM1, enfermedad que ocasiona la destrucción selectiva de células beta del páncreas causando deficiencia absoluta de insulina) que en población caucasoide se asocia su presencia a los alelos de clase II DRB1*03:01 y DRB1*04:01 (Tabla I4; Vicario et al., 1992), ambos con una amplia presencia en nuestras poblaciones de estudio (Tabla R2). Un estudio sobre la relación entre distintas variantes HLA de clase II y DM1 en población azerí de Azerbaiyán concluye también que distintos subtipos de DR4 se encuentran asociados a esta enfermedad (Ahmedov et al., 2006). Por lo tanto, tanto turcomanos, como azeríes y kurdos presentarían alelos HLA con susceptibilidad diabetogénica. No obstante un estudio sobre la incidencia de esta enfermedad en población total iraní determina una tasa bastante inferior a la de los países de su entorno

(Pishdad, 2005). El síndrome de Behcet, enfermedad reumática crónica que origina inflamación de los vasos sanguíneos, es conocido como enfermedad de “la ruta de la seda”, ya que este síndrome es muy común en Oriente Próximo y Asia, en países que se sitúan a lo largo de la antigua ruta comercial, con una gran incidencia en Irán. Esta distribución coincide con la del grupo HLA-B*51, especialmente del alelo B*51:01, que se supone es uno de los principales factores de riesgo de esta patología (Ohno et al., 1982; Wallace, 2014; Tabla I4). La sintomatología de esta enfermedad parece ser más leve si el individuo presenta en heterocigosis con B*51 alelos del grupo B*27 (Ahn & Park, 2007). Tanto turcomanos como azeríes y kurdos muestran frecuencias significativas de alelos B*51 y muy bajas para alelos B*27 (Tabla R1). La artritis reumatoide, entre otros factores genéticos, está asociada a distintas variantes de HLA-DRB1 que varían en función del grupo étnico. Estudios en población iraní y kurda de Irak demuestran que los principales alelos de susceptibilidad en estos grupos con DRB1*01 y DRB1*03, además del DQB1*06 en kurdos, mientras que DRB1*11 es mucho más frecuente en población sana (Al Timimi et al., 2014; Saghafi et al., 2014). Los dos primeros grupos de alelos son frecuentes en las tres poblaciones de estudio, sin embargo el DRB1*11:01 es el más frecuente en población kurda (Tabla R2) mientras que DQB1*06 tiene una frecuencia inferior, lo que podría suponer una menor incidencia de esta enfermedad en esta población.

En cuanto a las enfermedades infecciosas ligadas a HLA, determinadas variantes alélicas HLA están asociadas a un efecto protector frente a la malaria: B*35, B*53:01 y DRB1*13:02 para la causada por *Plasmodium falciparum* (véase Introducción apartado 8.2) y A*02:01 para la causada por *P. vivax* (Arevalo-Herrera et al., 2002). Estos alelos están presentes en las tres poblaciones de estudio, sin embargo el A*02:01 es el alelo de su clase más frecuente en población azerí (Tabla R1). La malaria causada por ambos *Plasmodium* es endémica fundamentalmente en las áreas rurales de la región sur-sureste de Irán (zona tropical). Sin embargo, hay pequeños focos de malaria por *P. vivax* al norte de los montes Zagros (provincias de Azerbaiyán Oriental y Ardabil) en determinadas épocas del año (Hemami et al., 2013). Esto podría sugerir una relación entre esta patología y la frecuencia significativa del alelo protector en el principal grupo étnico de esta región. Una de las asociaciones entre HLA y enfermedad infecciosa mejor documentada es la del SIDA. La infección por el virus VIH-1 es tratada y pronosticada hoy en día con base en el perfil HLA. El grupo de alelos B*57 está asociada con un control efectivo de la infección por este virus y un mejor pronóstico, incluso con poco o

ningún tratamiento (Crawford et al., 2009). Este grupo de alelos tienen poca representación en las poblaciones de estudio, estando totalmente ausente en kurdos. Sin embargo, el grupo B*58 (B*58:01) con fuertes similitudes en los péptidos de unión al antígeno a los alelos del grupo B*57, y que en algunas poblaciones se le asocia con un papel protector frente a la progresión del VIH-2 (Trachtenberg & Erlich, 2001), se encuentra en una frecuencia significativa en kurdos. Esto permite pensar que sea este grupo de alelos el mejor candidato para el pronóstico y tratamiento de esta enfermedad en esta población, en el caso de que futuros estudios demuestren su asociación. Los alelos del grupo A*32, bien representados en las tres poblaciones de estudio, también parecen estar asociados con una lenta progresión de la enfermedad en poblaciones caucásicas (Trachtenberg & Erlich, 2001). Por otra parte el grupo B*35, cuya ausencia está relacionada con un mejor pronóstico de la enfermedad, también está bien representado, siendo el mayoritario del locus HLA-B. Sin embargo la tasa de prevalencia en adultos de esta enfermedad en Irán es baja (0.2%) y muy similar a la de los países de su entorno (Central Intelligence Agency, 2013).

En lo que respecta a la farmacogenética HLA, la mayor parte de los estudios se han realizado en población caucásica de origen europeo y asiática, pero no se tiene constancia de ningún trabajo concreto en poblaciones de Irán. Sin embargo, diversos alelos HLA relacionados con efectos adversos a diversos fármacos están presentes, algunos con frecuencias considerables, en las poblaciones de estudio. Cabe destacar la elevada frecuencia del alelo A*24:02 tanto en turcomanos como en azeríes y kurdos, que está asociado a la aparición de rash cutáneo en tratamiento con el fármaco anticonvulsivo carbamazepina; la presencia del alelo DRB1*15:01 en las tres poblaciones, siendo el más frecuente de este locus en turcomanos, que está asociado con daño hepático en tratamiento con co-amoxiclav (combinación de antibióticos); y la frecuencia significativa del alelo DRB1*01:01, especialmente en kurdos, asociado con síndrome de hipersensibilidad en tratamiento con el inhibidor de la transcriptasa para el virus VIH-1 nevirapina (Becquemont, 2010; Gonzalez-Galarza et al., 2011; Tabla I5).

Los perfiles genéticos HLA determinados en el presente trabajo pueden ser de gran utilidad para establecer programas de epidemiología preventiva en estas poblaciones para aquellas patologías con una demostrada asociación a HLA, así como para minimizar los posibles efectos adversos generados por la acción de determinadas variantes alélicas ante determinados fármacos.

CONCLUSIONES

A partir del estudio genético HLA realizado en las poblaciones iraníes del sur del Mar Caspio -turcomanos, azeríes y kurdos- y su comparación con 13000 cromosomas de otras poblaciones, podemos concluir:

1. Los estudios comparativos de las poblaciones, a través de los dendrogramas y análisis de correspondencia, permiten observar que las tres poblaciones estudiadas tienen un sustrato genético común que las emparenta con las poblaciones del Mediterráneo oriental y del Cáucaso, muy evidente en el caso de kurdos y azeríes.
2. Estos estudios y los análisis de estimación de mestizaje sugieren que poblaciones ancestrales próximas a los pueblos altaicos actuales (Asia Central) pudieron contribuir, en parte, a la genética HLA de los turcomanos, y en menor medida a la de los azeríes.
3. El estudio de los haplotipos HLA de dos loci y extendidos sitúan a las poblaciones de estudio en el contexto HLA de las poblaciones del Mediterráneo oriental y del Cáucaso, como se deduce de la presencia de los siguientes haplotipos entre los más frecuentes: A*01:01-B*35:01-DRB1*03:01-DQB1*02:01, A*02:01-B*50:01-DRB1*07:01-DQB1*02:01, A*01:02-B*08:01-DRB1*03:01-DQB1*02:01 y A*24:02-B*58:01-DRB1*03:01-DQB1*02:01.
4. La presencia de haplotipos HLA, A*24:02-B*35:01-DRB1*15:01-DQB1*06:02, A*03:01-B*44:02-DRB1*04:01-DQB1*03:02 y A*03:01-B*07:02-DRB1*15:01-DQB1*06:02 en turcomanos y azeríes apoya que la contribución genética de poblaciones centroasiáticas que migraron a la región en diferentes épocas es importante. Esto se corrobora por otros datos genéticos y arqueológicos.
5. Al estudiar las relaciones entre estas poblaciones comprobamos que hay una fuerte correlación entre la genética HLA y la geografía, pero no entre genética y la lingüística: estas tres poblaciones próximas en su situación geográfica están estrechamente emparentadas entre sí según sus genes HLA y sin embargo hablan lenguas pertenecientes a familias lingüísticas distintas.
6. El hecho de que estas poblaciones estén genéticamente emparentadas pero tengan rasgos lingüísticos y culturales que las diferencian, sustenta la hipótesis de la “dominancia de élite” por la cual se produce una imposición cultural y lingüística por parte de una minoría gobernante de invasores, que da lugar en

nuestro estudio a una “iranización” cultural en el caso de los kurdos y “turquificación” cultural en el caso de los azeríes.

7. Los perfiles HLA obtenidos en nuestras poblaciones iraníes tienen una aplicación directa en el campo de los trasplantes de tejidos y órganos, permitiendo la elaboración de una lista de espera virtual de trasplante para un futuro terapéutico.
8. Los perfiles HLA obtenidos también se pueden aplicar al diagnóstico y terapéutica de enfermedades ligadas a HLA, como el SIDA y su asociación con alelos B*57 y B*58 (serotipo B17) y a la farmacogenética, ya que diversos alelos, como A*24:02 y DRB1*15:01, están asociados a reacciones adversas a determinados fármacos.

BIBLIOGRAFÍA

Abdi K (2012) The Iranian Plateau from Paleolithic Times of the Rise of the Achaemenid Empire. In: Daryaei T ed. *The Oxford Handbook of Iranian History*. Oxford: Oxford University Press, pp. 13-36.

Abele R, Tampe R (2004) The ABCs of immunology: structure and function of TAP, the transporter associated with antigen processing. *Physiology (Bethesda)* **19**: 216-224.

Ahmed AR, Wagner R, Khatri K, Notani G, Awdeh Z, Alper CA, Yunis EJ (1991) Major histocompatibility complex haplotypes and class II genes in non-Jewish patients with pemphigus vulgaris. *Proc Natl Acad Sci U S A* **88** (11): 5056-5060.

Ahmed AR, Yunis EJ, Khatri K, Wagner R, Notani G, Awdeh Z, Alper CA (1990) Major histocompatibility complex haplotype studies in Ashkenazi Jewish patients with pemphigus vulgaris. *Proc Natl Acad Sci U S A* **87** (19): 7658-7662.

Ahmedov G, Ahmedova L, Sedlakova P, Cinek O (2006) Genetic association of type 1 diabetes in an Azerbaijanian population: the HLA-DQ, -DRB1*04, the insulin gene, and CTLA4. *Pediatr Diabetes* **7** (2): 88-93.

Ahn JK, Park YG (2007) Human leukocyte antigen B27 and B51 double-positive Behcet uveitis. *Arch Ophthalmol* **125** (10): 1375-1380.

Aimel JL (1967) Study of the leukocyte phenotypes in Hodgkin's disease. In: Curtioni ES, Mattiuz PL, Tosi RM eds. *Histocompatibility Testing 1967*. Copenhagen: Munksgaard, pp. 79-81.

Ajam M (2013) *Geography of Iran: names of Caspian Sea*. Iran Chamber Society. http://www.iranchamber.com/geography/articles/names_of_caspian_sea.php. Accedido el 15-4-2014.

Al Timimi DJ, Rasool MT, Sulaiman DM (2014) HLA-DR/DQ Genotypes in Kurd Patients with Rheumatoid Arthritis: Relation to Disease Activity. *J Clin Diagn Res* **8** (5): CC01-CC04.

Alexeev L, Khaitov R, Boldyreva M, Trofimov D, Guskova I, Akopian A, Fedoseeva T, Undritsov I, Pospelov L, Bendukidze N et al (1997) HLA in some ethnic groups of the former Soviet Union (FSU). In: Charron D ed. *Genetic diversity of HLA. Functional and Medical Implication. Proceedings of the Twelfth International Histocompatibility Workshop and Conference*. Paris: EDK, pp. 364-373.

Alexeev LP, Petranyi G (1992) HLA in ethnic groups from the former USSR. In: Tsuji K, Aizawa M, Sasazuki T eds. *Oxford: Oxford University Press, HLA 1991*, pp. 666-673.

Amirzargar A, Mytilineos J, Farjadian S, Doroudchi M, Scherer S, Opelz G, Ghaderi A (2001) Human leukocyte antigen class II allele frequencies and haplotype association in Iranian normal population. *Hum Immunol* **62** (11): 1234-1238.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F et al (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290** (5806): 457-465.

Andonian L, Rezaie S, Margaryan A, Farhud D, Mohammad K, Naieni KH, Khorramizadeh M, Sanati MH, Jamali M, Bayatian P et al (2011) Iranian Azeri's Y-Chromosomal Diversity in the Context of Turkish-Speaking Populations of the Middle East. *Iran J Public Health* **40** (1): 119-123.

Andresdottir MB, Haasnoot GW, Persijn GG, Claas FH (2009) HLA-B8, DR3: a new risk factor for graft failure after renal transplantation in patients with underlying immunoglobulin A nephropathy. *Clin Transplant* **23** (5): 660-665.

Arevalo-Herrera M, Valencia AZ, Vergara J, Bonelo A, Fleischhauer K, Gonzalez JM, Restrepo JC, Lopez JA, Valmori D, Corradin G et al (2002) Identification of HLA-A2 restricted CD8(+) T-lymphocyte responses to Plasmodium vivax circumsporozoite protein in individuals naturally exposed to malaria. *Parasite Immunol* **24** (3): 161-169.

Arnaiz-Villena A, Alonso-Garcia J (2001) *Caucasicos, Turcos, Mesopotamicos y Vascos*. Madrid: Editorial Complutense SA.

Arnaiz-Villena A, Benmamar D, Alvarez M, Diaz-Campos N, Varela P, Gomez-Casado E, Martinez-Laso J (1995) HLA allele and haplotype frequencies in Algerians. Relatedness to Spaniards and Basques. *Hum Immunol* **43** (4): 259-268.

Arnaiz-Villena A, Dimitroski K, Pacho A, Moscoso J, Gomez-Casado E, Silvera-Redondo C, Varela P, Blagoevska M, Zdravkovska V, Martinez-Laso J (2001a) HLA genes in Macedonians and the sub-Saharan origin of the Greeks. *Tissue Antigens* **57** (2): 118-127.

Arnaiz-Villena A, Elaiwa N, Silvera C, Rostom A, Moscoso J, Gomez-Casado E, Allende L, Varela P, Martinez-Laso J (2001b) The origin of Palestinians and their genetic relatedness with other Mediterranean populations. stml. <http://www.stml.net/text/Populations.pdf>. Accedido el 5-4-2013b.

Arnaiz-Villena A, Gomez-Casado E, Martinez-Laso J (2002) Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens* **60** (2): 111-121.

Arnaiz-Villena A, Iliakis P, Gonzalez-Hevilla M, Longas J, Gomez-Casado E, Sfyridaki K, Trapaga J, Silvera-Redondo C, Matsouka C, Martinez-Laso J (1999) The origin of Cretan populations as determined by characterization of HLA alleles. *Tissue Antigens* **53** (3): 213-226.

Arnaiz-Villena A, Karin M, Bendikuze N, Gomez-Casado E, Moscoso J, Silvera C, Oguz FS, Sarper DA, De Pacho A, Allende L et al (2001c) HLA alleles and haplotypes in the Turkish population: relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens* **57** (4): 308-317.

Arnaiz-Villena A, Martinez-Laso J, Moscoso J, Livshits G, Zamora J, Gomez-Casado E, Silvera-Redondo C, Melvin K, Crawford MH (2003) HLA genes in the Chuvashian population from European Russia: admixture of Central European and Mediterranean populations. *Hum Biol* **75** (3): 375-392.

Asatrian G (2009) Prolegomena to the study of the Kurds. *Iran and the Caucasus* **13**: 1-58.

Asatrian G, Borjian H (2005) Talish and the Talishis (the state of research). *Iran and the Caucasus* **9** (1): 43-72.

Askarikhaneha A, Sharifkamali M (1995) *Turkmen Iranians: Anthropological and Demographic Study*. Teheran: Asatir Publications.

Ayala FJ (1995) The myth of Eve: molecular biology and human origins. *Science* **270** (5244): 1930-1936.

- Bannai M, Tokunaga K, Imanishi T, Harihara S, Fujisawa K, Juji T, Omoto K (1996) HLA class II alleles in Ainu living in Hidaka District, Hokkaido, northern Japan. *Am J Phys Anthropol* **101** (1): 1-9.
- Barber LD, Parham P (1993) Peptide binding to major histocompatibility complex molecules. *Annu Rev Cell Biol* **9**: 163-206.
- Beck S, Abdulla S, Alderton RP, Glynne RJ, Gut IG, Hosking LK, Jackson A, Kelly A, Newell WR, Sanseau P et al (1996) Evolutionary dynamics of non-coding sequences within the class II region of the human MHC. *J Mol Biol* **255** (1): 1-13.
- Becquemont L (2010) HLA: a pharmacogenomics success story. *Pharmacogenomics* **11** (3): 277-281.
- Belich MP, Glynne RJ, Senger G, Sheer D, Trowsdale J (1994) Proteasome components with reciprocal expression to that of the MHC-encoded LMP proteins. *Curr Biol* **4** (9): 769-776.
- Berger R, Bernheim A, Sasportes M, Hauptmann G, Hors J, Legrand L, Fellous M (1979) Regional mapping of the HLA on the short arm of chromosome 6. *Clin Genet* **15** (3): 245-251.
- Bermejo J (2013) Historia completa de Irán. Los pueblos prearios. La cultura de Gorgán. Lengua Persa. <http://www.lenguapersa.com/Articulos/Hair1.htm#gorgan>. Accedido el 15-3-2014.
- Bjorkman PJ, Saper MA, Samraoui B, Bennett WS (1987a) Structure of the human class I histocompatibility antigen, HLA-A 2. *Nature* **329**: 506-512.
- Bjorkman PJ, Saper MA, Samraoui B, Bennett W, Strominger J, Wiley DC (1987b) The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329** (6139): 512-518.
- Bodmer JG, Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Charron D, Dupont B, Erlich HA, Fauchet R, Mach B et al (1997) Nomenclature for factors of the HLA system, 1996. *Eur J Immunogenet* **24** (2): 105-151.
- Bodmer WF, Bodmer JG (1978) Evolution and function of the HLA system. *Br Med Bull* **34** (3): 309-316.
- Bonne-Tamir B, Bodmer JG, Bodmer WF, Pickbourne P, Brautbar C, Gazit E, Nevo S, Zamir R (1978) HLA polymorphism in Israel. 9. An overall comparative analysis. *Tissue Antigens* **11** (3): 235-250.
- Bosworth CE (2002) Gorgan VI. History from the rise of Islam to the beginning of the Safavid period. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge & Kegan Paul, vol. XI, pp. 153-154. <http://www.iranicaonline.org/articles/gorgan-vi>
- Braud VM, Allan DS, Wilson D, McMichael AJ (1998) TAP- and tapasin-dependent HLA-E surface expression correlates with the binding of an MHC class I leader peptide. *Curr Biol* **8** (1): 1-10.
- Brega A, Scacchi R, Cuccia M, Kirdar B, Peloso G, Corbo RM (1998) Study of 15 protein polymorphisms in a sample of the Turkish population. *Hum Biol* **70** (4): 715-728.
- Brocke P, Garbi N, Momburg F, Hammerling GJ (2002) HLA-DM, HLA-DO and tapasin: functional similarities and differences. *Curr Opin Immunol* **14** (1): 22-29.

Brown WM, Wright JW (1979) Mitochondrial DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). *Science* **203** (4386): 1247-1249.

Calafell F, Underhill P, Tolun A, Angelicheva D, Kalaydjieva L (1996) From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann Hum Genet* **60** (Pt 1): 35-49.

Cammarota G, Scheirle A, Takacs B, Doran DM, Knorr R, Bannwarth W, Guardiola J, Sinigaglia F (1992) Identification of a CD4 binding site on the beta 2 domain of HLA-DR molecules. *Nature* **356** (6372): 799-801.

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* **325** (6099): 31-36.

Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K, O'Brien SJ (1999) HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283** (5408): 1748-1752.

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton: Princeton University Press.

Central Intelligence Agency (2013) The World Factbook. Washington DC: Central Intelligence Agency.

Cepellini R, Curtini ES, Mattiuz PL, Miggiano V, Scudeller G, Serra A (1967) Genetics of leukocyte antigens. A family study of segregation and linkage. In: Curtini ES, Mattiuz PL, Tosi RM eds. *Histocompatibility Testing 1967*. Copenhagen: Munksgaard, pp. 149-187.

Cereb N, Kong Y, Lee S, Maye P, Yang SY (1996) Nucleotide sequences of MHC class I introns 1, 2, and 3 in humans and intron 2 in nonhuman primates. *Tissue Antigens* **47** (6): 498-511.

Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, Vignali DA, Strominger JL (1992) Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358** (6389): 764-768.

Clayton J, Lonjou C (1997) Allele and Haplotype frequencies for HLA loci in various ethnic groups. In: Charron D ed. *Genetic diversity of HLA. Functional and medical implications*. Paris: EDK, pp. 665-820.

Collins EJ, Garboczi DN, Karpusas MN, Wiley DC (1995) The three-dimensional structure of a class I major histocompatibility complex molecule missing the alpha 3 domain of the heavy chain. *Proc Natl Acad Sci U S A* **92** (4): 1218-1221.

Comas D, Calafell F, Bendukidze N, Fananas L, Bertranpetit J (2000) 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.

Connolly JM, Hansen TH, Ingold AL, Potter TA (1990) Recognition by CD8 on cytotoxic T lymphocytes is ablated by several substitutions in the class I alpha 3 domain: CD8 and the T-cell receptor recognize the same class I molecule. *Proc Natl Acad Sci U S A* **87** (6): 2137-2141.

Coon CS (1952) Excavations in Huto Cave, Iran, 1951: A Preliminary Report. *Proceedings of the American Philosophical Society* **96**: 231-269.

Cotterel A (1980) The encyclopedia of ancient civilizations. London: Rainbird.

- Crawford H, Lumm W, Leslie A, Schaefer M, Boeras D, Prado JG, Tang J, Farmer P, Ndung'u T, Lakhi S et al (2009) Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. *J Exp Med* **206** (4): 909-921.
- Danilovs JA, Ayoub G, Terasaki PI (1980) B lymphocyte isolation by thrombin-nylon wool. In: Terasaki PI ed. *Histocompatibility Testing 1980*. Los Angeles: UCLA Tissue Typing Laboratory, pp. 287-288.
- Dausset J (1958) Iso-leuko-antibodies. *Acta Haematol* **20** (1-4): 156-166.
- Dausset J, Ivanyi P, Ivanyi D (1965) Tissue alloantigens in humans: identification of a complex system (Hu-1). In: Dausset J, Rappaport FJ, Ivanyi P, Colombani J eds. *Histocompatibility Testing 1965*. Copenhagen: Munksgaard, pp. 51-62.
- de Pablo R, Garcia-Pacheco JM, Vilches C, Moreno ME, Sanz L, Rementeria MC, Puente S, Kreisler M (1997) HLA class I and class II allele distribution in the Bubi population from the island of Bioko (Equatorial Guinea). *Tissue Antigens* **50** (6): 593-601.
- Degos L, Dausset J (1974) Human migrations and linkage disequilibrium of HLA system. *Immunogenetics* **1**: 195-210.
- Denzin LK, Sant'Angelo DB, Hammond C, Surman MJ, Cresswell P (1997) Negative regulation by HLA-DO of MHC class II-restricted antigen processing. *Science* **278** (5335): 106-109.
- Derenko M, Malyarchuk B, Denisova GA, Wozniak M, Dambueva I, Dorzhu C, Luzina F, Miscicka-Sliwka D, Zakharov I (2006) Contrasting patterns of Y-chromosome variation in South Siberian populations from Baikal and Altai-Sayan regions. *Hum Genet* **118** (5): 591-604.
- Diakonoff I, Starostin S (1986) Hurro-Urartian as an Eastern Caucasian Language. *Münchener Studien zur Sprachwissenschaft* **12**: 164-207.
- Diodoro Sículo (2001) *Biblioteca Histórica*, volumen VI: libro XVIII, capítulo 3, 3. Madrid: Gredos.
- Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moreau P (2011) Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. *Cell Mol Life Sci* **68** (3): 369-395.
- Dunbar SA, Vander Zee CA, Oliver KG, Karem KL, Jacobson JW (2003) Quantitative, multiplexed detection of bacterial pathogens: DNA and protein applications of the Luminex LabMAP system. *J Microbiol Methods* **53** (2): 245-252.
- Encyclopaedia Britannica (2014) Mannai. Encyclopaedia Britannica, Inc. <http://global.britannica.com/EBchecked/topic/362506/Mannai>. Accedido el 15-4-2014.
- Erlich HA, Mack SJ, Bergstrom T, Gyllensten UB (1997) HLA class II alleles in Amerindian populations: implications for the evolution of HLA polymorphism and the colonization of the Americas. *Hereditas* **127** (1-2): 19-24.
- Erlich HA, Zeidler A, Chang J, Shaw S, Raffel LJ, Klitz W, Beshkov Y, Costin G, Pressman S, Bugawan T et al (1993) HLA class II alleles and susceptibility and resistance to insulin dependent diabetes mellitus in Mexican-American families. *Nat Genet* **3** (4): 358-364.
- Estrabón (1991a) *Geografía*, libro II, capítulo 5, sección 12. Madrid: Gredos.
- Estrabón (1991b) *Geografía*, libro XI, capítulo 13, sección 1. Madrid: Gredos.

- Estrabón (1991c) Geografía, libro XI, capítulo 4, sección 6. Madrid: Gredos.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47-50.
- Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351** (6324): 290-296.
- Fan QR, Wiley DC (1999) Structure of human histocompatibility leukocyte antigen (HLA)-Cw4, a ligand for the KIR2D natural killer cell inhibitory receptor. *J Exp Med* **190** (1): 113-123.
- Farjadian S, Ghaderi A (2007) HLA class II similarities in Iranian Kurds and Azeris. *Int J Immunogenet* **34** (6): 457-463.
- Farjadian S, Naruse T, Kawata H, Ghaderi A, Bahram S, Inoko H (2004) Molecular analysis of HLA allele frequencies and haplotypes in Baloch of Iran compared with related populations of Pakistan. *Tissue Antigens* **64** (5): 581-587.
- Farjadian S, Ota M, Inoko H, Ghaderi A (2009) The genetic relationship among Iranian ethnic groups: an anthropological view based on HLA class II gene polymorphism. *Mol Biol Rep* **36** (7): 1943-1950.
- Farjadian S, Safi S (2013) Genetic connections among Turkic-speaking Iranian ethnic groups based on HLA class II gene diversity. *Int J Immunogenet* **40** (6): 509-514.
- Fernandez-Vina MA, Lazaro AM, Marcos CY, Nulf C, Raimondi E, Haas EJ, Stastny P (1997) Dissimilar evolution of B-locus versus A-locus and class II loci of the HLA region in South American Indian tribes. *Tissue Antigens* **50** (3): 233-250.
- Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, Vyse TJ, Rioux JD (2008) Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet* **4** (4): e1000024.
- Figueroa F, Gunther E, Klein J (1988) MHC polymorphism pre-dating speciation. *Nature* **335** (6187): 265-267.
- Frey RN (2004) Iran V. Peoples of Iran (1) A general survey. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge and Kegan Paul, vol. XIII, pp. 321-326.
<http://www.iranicaonline.org/articles/iran-v1-peoples-survey#1>
- Fruh K, Gossen M, Wang K, Bujard H, Peterson PA, Yang Y (1994) Displacement of housekeeping proteasome subunits by MHC-encoded LMPs: a newly discovered mechanism for modulating the multicatalytic proteinase complex. *EMBO J* **13** (14): 3236-3244.
- Fujii T, Ishitani A, Geraghty DE (1994) A soluble form of the HLA-G antigen is encoded by a messenger ribonucleic acid containing intron 4. *J Immunol* **153** (12): 5516-5524.
- Fulton RJ, McDade RL, Smith PL, Kienker LJ, Kettman JR, Jr. (1997) Advanced multiplexed analysis with the FlowMetrix system. *Clin Chem* **43** (9): 1749-1756.
- Gabunia L, Vekua A, Lordkipanidze D, Swisher CC, III, Ferring R, Justus A, Nioradze M, Tvalchrelidze M, Anton SC, Bosinski G et al (2000) Earliest Pleistocene hominid cranial remains from Dmanisi, Republic of Georgia: taxonomy, geological setting, and age. *Science* **288** (5468): 1019-1025.

- Gao X, Bhatia K, Trent RJ, Serjeantson SW (1992a) HLA-DR,DQ nucleotide sequence polymorphisms in five Melanesian populations. *Tissue Antigens* **40** (1): 31-37.
- Gao X, Veale A, Serjeantson SW (1992b) HLA class II diversity in Australian aborigines: unusual HLA-DRB1 alleles. *Immunogenetics* **36** (5): 333-337.
- Gao X, Zimmet P, Serjeantson SW (1992c) HLA-DR,DQ sequence polymorphisms in Polynesians, Micronesians, and Javanese. *Hum Immunol* **34** (3): 153-161.
- Garrett TP, Saper MA, Bjorkman PJ, Strominger JL, Wiley DC (1989) Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature* **342** (6250): 692-696.
- Ghods AJ (2014) The history of organ donation and transplantation in Iran. *Exp Clin Transplant* **12 Suppl 1**: 38-41.
- Ghosh P, Amaya M, Mellins E, Wiley DC (1995) The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature* **378** (6556): 457-462.
- Gilbert SC, Plebanski M, Gupta S, Morris J, Cox M, Aidoo M, Kwiatkowski D, Greenwood BM, Whittle HC, Hill AV (1998) Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* **279** (5354): 1173-1177.
- Gimbutas M (1977) The first wave of Eurasian steppe pastoralists into Copper Age Europe. *Journal of Indo-European Studies* **5**: 277-338.
- Gimbutas M (1980) The Kurgan wave #2 (c.3400-3200 BC) into Europe and the following transformation of culture. *Journal of Indo-European Studies* **8**: 273-315.
- Glynne R, Powis SH, Beck S, Kelly A, Kerr LA, Trowsdale J (1991) A proteasome-related gene between the two ABC transporter loci in the class II region of the human MHC. *Nature* **353** (6342): 357-360.
- Gómez-Casado E (2002) Relaciones entre cretenses y otros pueblos mediterráneos según los genes HLA y las culturas. Tesis doctoral dirigida por Dr. A. Arnaiz-Villena y Dr. J. Martínez-Laso. Universidad Complutense de Madrid, Facultad de Ciencias Químicas.
- Gomez-Casado E, del Moral P, Martinez-Laso J, Garcia-Gomez A, Allende L, Silvera-Redondo C, Longas J, Gonzalez-Hevilla M, Kandil M, Zamora J et al (2000) HLA genes in Arabic-speaking Moroccans: close relatedness to Berbers and Iberians. *Tissue Antigens* **55** (3): 239-249.
- Gomez-Casado E, Vargas-Alarcon G, Martinez-Laso J, Granados J, Varela P, Alegre R, Longas J, Gonzalez-Hevilla M, Martin-Villa JM, Arnaiz-Villena A (1999) Evolutionary relationships between HLA-B alleles as indicated by an analysis of intron sequences. *Tissue Antigens* **53** (2): 153-160.
- Gomez-Casado E, Vargas-Alarcon G, Martinez-Laso J, Perez-Blas M, Granados J, Layrisse Z, Montoya F, Varela P, Arnaiz-Villena A (1997) Generation of the HLA-B35, -B5, -B16, and B15 groups of alleles studied by intron 1 and 2 sequence analysis. *Immunogenetics* **46** (6): 469-476.
- Gomez-Prieto P, Parga-Lozano C, Rey D, Moreno E, Arnaiz-Villena A (2010) HLA-G, -F and -E: Polymorphism, Function, and Evolution. In: Mehra NK ed. *The HLA Complex in Biology and Medicine. A Resource Book*. New Delhi: Jaypee Brothers Medical Publishers, pp. 159-174.

- Gonem A (1996) The Encyclopedia of the Peoples of the World. Jerusalem, Israel: Jerusalem Publishing House Ltd.
- Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR (2011) Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res* **39** (Database issue): D913-D919.
- Gorer PA (1936) The detection of a hereditary antigenic difference in the blood of mice by means of human group A serum. *Journal of Genetics* **32**: 17-31.
- Graf OM, Mitchell J, Wilcox S, Livshits G, Crawford MH (2010) Chuvash origins: evidence from mtDNA markers. Abstracts of AAPA poster and podium presentations. *American Journal of Physical Anthropology* **141** (S50): 52-252.
- Grahovac B, Sukernik RI, O'hUigin C, Zaleska-Rutczynska Z, Blagitko N, Raldugina O, Kosutic T, Satta Y, Figueroa F, Takahata N et al (1998) Polymorphism of the HLA class II loci in Siberian populations. *Hum Genet* **102** (1): 27-43.
- Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA (2013) Genenames.org: the HGNC resources in 2013. *Nucleic Acids Res* **41** (Database issue): D545-D552.
- Grey HM, Kubo RT, Colon SM, Poulik MD, Cresswell P, Springer T, Turner M, Strominger JL (1973) The small subunit of HL-A antigens is beta 2-microglobulin. *J Exp Med* **138** (6): 1608-1612.
- Griffin A (2007) Iranian organ donation: kidneys on demand. *British Medical Journal* **334** (7592): 502-505.
- Grugni V, Battaglia V, Hooshiar KB, Parolo S, Al Zahery N, Achilli A, Olivieri A, Gandini F, Houshmand M, Sanati MH et al (2012) Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One* **7** (7): e41252.
- Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K et al (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*.
- Hay M (2014) Haplogroup R1a (Y-DNA). Eupedia. http://www.eupedia.com/europe/Haplogroup_R1a_Y-DNA.shtml. Accedido el 30-9-2014.
- Hegland J (1997) Transplant immunology. HLA and issues of stem cell donation. In: Whedon MB, Wujcik D eds. Blood and marrow stem cell transplantation: principles, practice and nursing insights. Burlington, Massachusetts: Jones & Barlett Learning, pp. 43-65.
- Hemami MR, Sari AA, Raeisi A, Vatandoost H, Majdzadeh R (2013) Malaria elimination in iran, importance and challenges. *Int J Prev Med* **4** (1): 88-94.
- Hennerbichler F (2011) The origin of Kurds. Edition Winterwork: Borsdorf, Germany.
- Hennerbichler F (2012) The origin of Kurds. *Advances in Anthropology* **2** (2): 64-79.
- Henning WB (1954) The ancient language of Azerbaijan. *Transactions of the Philological Society* **53** (1): 157-177.
- Herodoto (1989) Historia, libro III. Madrid: Gredos.

- Hetherington S, Hughes AR, Mosteller M, Shortino D, Baker KL, Spreen W, Lai E, Davies K, Handley A, Dow DJ et al (2002) Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. *Lancet* **359** (9312): 1121-1122.
- Hill AV (2001) Immunogenetics and genomics. *Lancet* **357** (9273): 2037-2041.
- Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, Brewster D, McMichael AJ, Greenwood BM (1991) Common west African HLA antigens are associated with protection from severe malaria. *Nature* **352** (6336): 595-600.
- Hiller C, Bischoff M, Schmidt A, Bender K (1978) Analysis of the HLA-ABC linkage disequilibrium: decreasing strength of gametic association with increasing map distance. *Hum Genet* **41** (3): 301-312.
- Horton R, Gibson R, Coggill P, Miretti M, Allcock RJ, Almeida J, Forbes S, Gilbert JG, Halls K, Harrow JL et al (2008) Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics* **60** (1): 1-18.
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC, Jr., Wright MW et al (2004) Gene map of the extended human MHC. *Nat Rev Genet* **5** (12): 889-899.
- Hovanissian RG (1997) The Armenian People from ancient to modern times. New York: St. Martin's Press.
- Humphreys T, Reinherz EL (1994) Invertebrate immune recognition, natural immunity and the evolution of positive selection. *Immunol Today* **15** (7): 316-320.
- Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T (1992a) Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T eds. HLA 1991. Oxford: Oxford University Press, pp. 1065-1220.
- Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T (1992b) Estimation of allele and haplotype frequencies for HLA and complement loci. In: Tsuji K, Aizawa M, Sasazuki T eds. HLA 1991. Oxford: Oxford University Press, pp. 76-79.
- Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T (1992c) Genetic relationships among various human populations indicated by MHC polymorphisms. In: Tsuji K, Aizawa M, Sasazuki T eds. HLA 1991. Oxford: Oxford University Press, pp. 627-632.
- International Business Publication (2009) Iran: Country Study Guide. Washington DC: International Business Publication USA.
- International HIV Controllers Study (2010) The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **330** (6010): 1551-1557.
- Ishitani A, Geraghty DE (1992) Alternative splicing of HLA-G transcripts yields proteins with primary structures resembling both class I and class II antigens. *Proc Natl Acad Sci U S A* **89** (9): 3947-3951.
- Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, Kashiwase K, Kikkawa E, Kulski JK, Satake M, Inoko H (2005) High-throughput DNA typing of HLA-A, -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population. *Immunogenetics* **57** (10): 717-729.

Izaabel H, Garchon HJ, Caillat-Zucman S, Beaurain G, Akhayat O, Bach JF, Sanchez-Mazas A (1998) HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* **51** (1): 106-110.

Izady M (1992) *The Kurds: a concise handbook*. Washington: Crane Russak.

Jensen CO (1903) Experimentelle untersuchungen iiber krebs bei mausen. *Zentralbl Bakteriol Parasitenkd Infektkr* **34**: 28.

Johanson L (2006) Turkmen. In: Brown K ed. *Encyclopedia of Language and Linguistic*. Amsterdam: Elsevier, vol XIII, pp. 172.

Johansson A, Ingman M, Mack SJ, Erlich H, Gyllensten U (2008) Genetic origin of the Swedish Sami inferred from HLA class I and class II allele frequencies. *Eur J Hum Genet* **16** (11): 1341-1349.

Jordan BR, Caillol D, Damotte M, Delovitch T, Ferrier P, Kahn-Perles B, Kourilsky F, Layet C, Le Bouteiller P, Lemonnier FA et al (1985) HLA class I genes: from structure to expression, serology and function. *Immunol Rev* **84**: 73-92.

Kapustin S, Lyshchov A, Alexandrova J, Imyanitov E, Blinov M (1999) HLA class II molecular polymorphisms in healthy Slavic individuals from North-Western Russia. *Tissue Antigens* **54** (5): 517-520.

Kasimova RM (2001) Anthropological research of Azykh man osseous remains. *Human Evolution* **16** (1): 37-44.

Katunuma N, Matsunaga Y, Saibara T (1994) Mechanism and regulation of antigen processing by cathepsin B. *Adv Enzyme Regul* **34**: 145-158.

Kaufman JF, Auffray C, Korman AJ, Shackelford DA, Strominger J (1984) The class II molecules of the human and murine major histocompatibility complex. *Cell* **36** (1): 1-13.

Keyser C, Bouakaze C, Crubezy E, Nikolaev VG, Montagnon D, Reis T, Ludes B (2009) Ancient DNA provides new insights into the history of south Siberian Kurgan people. *Hum Genet* **126** (3): 395-410.

Khan MA (2013) Polymorphism of HLA-B27: 105 subtypes currently known. *Curr Rheumatol Rep* **15** (10): 362.

Kiani MY (2002) Gorgan IV. Archeology. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge & Kegan Paul, vol. XI, pp. 148-151. <http://www.iranicaonline.org/articles/gorgan-iv>

Kirszenbaum M, Moreau P, Gluckman E, Dausset J, Carosella E (1994) An alternatively spliced form of HLA-G mRNA in human trophoblasts and evidence for the presence of HLA-G transcript in adult lymphocytes. *Proc Natl Acad Sci U S A* **91** (10): 4209-4213.

Klein J (1990) *The major histocompatibility complex Immunology*. Massachusetts: Blackwell Scientific Publications.

Klein J (1986) *Natural history of the Major Histocompatibility Complex*. New York, USA: J. Willey and Sons.

Klein J (1987) Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Human Immunology* **19**: 155-162.

Klein J, O'hUigin C, Kasahara M, Vincek V, Klein D, Figueroa F (1991) Frozen haplotypes in MHC evolution. In: Klein J, Klein D eds. *Molecular Evolution of the Major Histocompatibility Complex*. Heidelberg: Springer-Verlag, pp. 261-281.

Klein J, Satta Y, O'hUigin C, Takahata N (1993) The molecular descent of the major histocompatibility complex. *Annu Rev Immunol* **11**: 269-295.

Konig R, Huang LY, Germain RN (1992) MHC class II interaction with CD4 mediated by a region analogous to the MHC class I binding site for CD8. *Nature* **356** (6372): 796-798.

Koopmann JO, Hammerling GJ, Momburg F (1997) Generation, intracellular transport and loading of peptides associated with MHC class I molecules. *Curr Opin Immunol* **9** (1): 80-88.

Krausz C, Quintana-Murci L, Forti G (2004) Y chromosome polymorphisms in medicine. *Ann Med* **36** (8): 573-583.

Lajoie J, Hargrove J, Zijenah LS, Humphrey JH, Ward BJ, Roger M (2006) Genetic variants in nonclassical major histocompatibility complex class I human leukocyte antigen (HLA)-E and HLA-G molecules are associated with susceptibility to heterosexual acquisition of HIV-1. *J Infect Dis* **193** (2): 298-301.

Lee N, Goodlett DR, Ishitani A, Marquardt H, Geraghty DE (1998) HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. *J Immunol* **160** (10): 4951-4960.

Lester S, Cassidy S, Humphreys I, Bennett G, Hurley CK, Boettcher B, McCluskey J (1995) 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. *Hum Immunol* **42** (2): 154-160.

Lewis MP, Simons GF, Fenning CDe (2014) *Ethnologue: languages of the world*, seventeenth edition. Dallas, TX: SIL International. <http://www.ethnologue.com/>

Li L, Bouvier M (2004) Structures of HLA-A*1101 complexed with immunodominant nonamer and decamer HIV-1 epitopes clearly reveal the presence of a middle, secondary anchor residue. *J Immunol* **172** (10): 6175-6184.

Little CC, Johnson BW (1922) The inheritance of susceptibility to implants of splenic tissue in mice. *Proceedings of the Society for Experimental Biology and Medicine* **19**: 163.

Loeb L (1908) Ueber Entstehung eines Sarkoms nach Transplantation eines Adenocarcinoms einer japanischen Maus. *Journal of Cancer Research and Clinical Oncology* **7**: 80.

Longanesi C (1986) *Cività e Imperi del Medio Oriente*. Milán: Jaca Book s.p.a.

Luther KA (2001) *The history of the Seljuk Turku from the Jamil Al-tawarik: an ilkhanid adaptation of the saljuk-nama of Zahir Al-din Nishapuri*. Richmon, Surrey: Cruzon Press.

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonne-Tamir B, Sykes B, Torroni A (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* **64** (1): 232-249.

Mack SJ, Erlich HA (2007) 13th International Histocompatibility Workshop Anthropology/Human Genetic Diversity Joint Report. In: Hansen JA ed. *Immunology of the Human MHC. Proceedings of the 13th International Histocompatibility Workshop and Conference*. Seattle, USA: IHWG Press, vol. 1, pp. 557-766.

MacKenzie DN (1964) The origins of Kurdish. *Transactions of the Philological Society* **60** (1): 68-86.

Macqueen JC (1999) *The hittites*. London: Thames and Hudson.

Madden DR (1995) The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol* **13**: 587-622.

Madden DR, Gorga JC, Strominger JL, Wiley DC (1991) The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature* **353** (6342): 321-325.

Madden DR, Gorga JC, Strominger JL, Wiley DC (1992) The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* **70** (6): 1035-1048.

Maliarchuk BA, Derenko MV, Denisova GA, Nassiri MR, Rogaev EI (2002) Mitochondrial DNA polymorphism in populations of the Caspian region and southeastern Europe. *Genetika* **38** (4): 534-538.

Malissen M, Malissen B, Jordan BR (1982) Exon/intron organization and complete nucleotide sequence of an HLA gene. *Proc Natl Acad Sci U S A* **79** (3): 893-897.

Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, Sayer D, Castley A, Mamotte C, Maxwell D et al (2002) Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet* **359** (9308): 727-732.

Mallory JP (1989) *In search of the Indo-Europeans: language, archaeology, and myth*. London: Thames & Hudson.

Mallory JP (1997a) Andronovo culture. In: Mallory JP, Adams DQ eds. *Encyclopedia of Indo-European Culture*. Londres: Fitzroy Dearborn, pp. 20-21.

Mallory JP (1997b) BMAC. In: Mallory JP, Adams DQ eds. *Encyclopedia of Indo-European Culture*. Londres: Fitzroy Dearborn, pp. 73.

Mallory JP (1997c) Yamna culture. In: Mallory JP, Adams DQ eds. *Encyclopedia of Indo-European Culture*. Londres: Fitzroy Dearborn.

Marks MS, Germain RN, Bonifacino JS (1995) Transient aggregation of major histocompatibility complex class II chains during assembly in normal spleen cells. *J Biol Chem* **270** (18): 10475-10481.

Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernandez-Vina M, Geraghty DE, Holdsworth R, Hurley CK et al (2010) Nomenclature for factors of the HLA system, 2010. *Tissue Antigens* **75** (4): 291-455.

Martinez-Laso J, de Juan D, Martinez-Quiles N, Gomez-Casado E, Cuadrado E, Arnaiz-Villena A (1995) The contribution of the HLA-A, -B, -C and -DR, -DQ DNA typing to the study of the origins of Spaniards and Basques. *Tissue Antigens* **45** (4): 237-245.

Martinez-Laso J, Gazit E, Gomez-Casado E, Morales P, Martinez-Quiles N, Alvarez M, Martin-Villa JM, Fernandez V, Arnaiz-Villena A (1996) HLA DR and DQ polymorphism in Ashkenazi and non-Ashkenazi Jews: comparison with other Mediterraneans. *Tissue Antigens* **47** (1): 63-71.

Martinez-Laso J, Montoya F, Areces C, Moscoso J, Silvera C, Rey D, Parga-Lozano C, Gomez-Prieto P, Enriquez dS, Arnaiz-Villena A (2011) HLA in Jaidukama: an Amerindian secluded Colombian population with new haplotypes and Asian and Pacific-shared alleles. *Mol Biol Rep* **38** (6): 3689-3701.

Martinez-Laso J, Sartakova M, Allende L, Konenkov V, Moscoso J, Silvera-Redondo C, Pacho A, Trapaga J, Gomez-Casado E, Arnaiz-Villena A (2001) HLA molecular markers in Tuvinians: a population with both Oriental and Caucasoid characteristics. *Ann Hum Genet* **65** (Pt 3): 245-261.

Matevosyan L, Chattopadhyay S, Madelian V, Avagyan S, Nazaretyan M, Hyussian A, Vardapetyan E, Arutunyan R, Jordan F (2011) HLA-A, HLA-B, and HLA-DRB1 allele distribution in a large Armenian population sample. *Tissue Antigens* **78** (1): 21-30.

Mattiuz PL, Ihde D, Piazza A, Ceppellini R, Bodmer WF (1971) New approaches to the population genetics and segregation analysis of the HLA system. In: Terasaki PI ed. *Histocompatibility testing 1970*. Copenhagen: Munksgaard, pp. 193-206.

McKusick VA (2014) Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). www.omim.org. Accedido el 30-5-2014.

McMahon RM, Friis L, Siebold C, Friese MA, Fugger L, Jones EY (2011) Structure of HLA-A*0301 in complex with a peptide of proteolipid protein: insights into the role of HLA-A alleles in susceptibility to multiple sclerosis. *Acta Crystallogr D Biol Crystallogr* **67** (Pt 5): 447-454.

Mehra NK, Rajalingam R, Kanga U, McEneny L, Cullen C, Agarwal S, Middleton D, Pollack MS, Amirzargar A, Singal DP (1997) Genetic diversity of HLA in the populations of India, Sri Lanka and Iran. In: Charron D ed. *Genetic diversity of HLA. Functional and Medical Implications. Proceedings of the Twelfth International Histocompatibility Workshop and Conference*. Paris: EDK, pp. 314-320.

Merriwether DA, Hall WW, Vahlne A, Ferrell RE (1996) mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. *Am J Hum Genet* **59** (1): 204-212.

Minorsky V (1957) *Studies in caucasian history*. Cambridge: Cambridge University Press.

Mobini N, Yunis EJ, Alper CA, Yunis JJ, Delgado JC, Yunis DE, Firooz A, Dowlati Y, Bahar K, Gregersen PK et al (1997) Identical MHC markers in non-Jewish Iranian and Ashkenazi Jewish patients with pemphigus vulgaris: possible common central Asian ancestral origin. *Hum Immunol* **57** (1): 62-67.

Mohyuddin A, Ayub Q, Khaliq S, Mansoor A, Mazhar K, Rehman S, Mehdi SQ (2002) HLA polymorphism in six ethnic groups from Pakistan. *Tissue Antigens* **59** (6): 492-501.

Monsalve MV, Edin G, Devine DV (1998) Analysis of HLA class I and class II in Na-Dene and Amerindian populations from British Columbia, Canada. *Hum Immunol* **59** (1): 48-55.

Moscoso J, Crawford MH, Vicario JL, Zlojutro M, Serrano-Vela JI, Reguera R, Arnaiz-Villena A (2008) HLA genes of Aleutian Islanders living between Alaska (USA) and Kamchatka (Russia) suggest a possible southern Siberia origin. *Mol Immunol* **45** (4): 1018-1026.

Mungall AJ, Palmer SA, Sims SK, Edwards CA, Ashurst JL, Wilming L, Jones MC, Horton R, Hunt SE, Scott CE et al (2003) The DNA sequence and analysis of human chromosome 6. *Nature* **425** (6960): 805-811.

- Munkhbat B, Sato T, Hagihara M, Sato K, Kimura A, Munkhtuvshin N, Tsuji K (1997) Molecular analysis of HLA polymorphism in Khoton-Mongolians. *Tissue Antigens* **50** (2): 124-134.
- Nasidze I, Ling EY, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA et al (2004) Mitochondrial DNA and Y-chromosome variation in the caucasus. *Ann Hum Genet* **68** (Pt 3): 205-221.
- Nasidze I, Quinque D, Ozturk M, Bendukidze N, Stoneking M (2005) MtDNA and Y-chromosome variation in Kurdish groups. *Ann Hum Genet* **69** (Pt 4): 401-412.
- Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Asadova P, Zhukova O, Stoneking M (2009) mtDNA and Y-chromosome variation in the Talysh of Iran and Azerbaijan. *Am J Phys Anthropol* **138** (1): 82-89.
- Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Stoneking M (2006) Concomitant replacement of language and mtDNA in South Caspian populations of Iran. *Curr Biol* **16** (7): 668-673.
- Nebel A, Filon D, Oppenheim A, Faerman M (2007) The genetic history of populations in the Southern Levant as revealed by Y chromosome polymorphisms. In: Faerman M, Horwitz LK, Kahana T, Zilberman U eds. *Faces from the Past: Diachronic Patterns in the Biology and Health Status of Human Populations in the Eastern Mediterranean*. Papers in honour of Patricia Smith. Oxford: Archeopress, BAR International Series 1603, pp. 257-270.
- Nei M (1972) Genetic distances between populations. *Am Nat* **106**: 283.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* **70** (12): 3321-3323.
- Nejestani and Elr J (2002) Gorgan VII. History from the Safavids to the end of the Pahlavi era. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge & Kegan Paul, vol. XI, pp. 154-162. <http://www.iranicaonline.org/articles/gorgan-vii>
- Nepom GT, Erlich H (1991) MHC class-II molecules and autoimmunity. *Annu Rev Immunol* **9**: 493-525.
- O'Callaghan CA, Tormo J, Willcox BE, Blundell CD, Jakobsen BK, Stuart DI, McMichael AJ, Bell JI, Jones EY (1998) Production, crystallization, and preliminary X-ray analysis of the human MHC class Ib molecule HLA-E. *Protein Sci* **7** (5): 1264-1266.
- Ohno S, Ohguchi M, Hirose S, Matsuda H, Wakisaka A, Aizawa M (1982) Close association of HLA-Bw51 with Behcet's disease. *Arch Ophthalmol* **100** (9): 1455-1458.
- Ohta T (1990) How gene families evolve. *Theor Popul Biol* **37** (1): 213-219.
- Ohta T (1991) Multigene families and the evolution of complexity. *J Mol Evol* **33** (1): 34-41.
- Oppenheimer S (2006) *The Origin of the British - A Genetic Detective Story*. Londres: Constable and Robinson.
- Panneerchelvam S, Norazmi MN (2012) HLA Polymorphism in Anthropology. In: Abdel-Salam B ed. *Histocompatibility*. InTech. <http://www.intechopen.com/books/histocompatibility/hla-polymorphisms-in-phylogenetics>

- Parham P, Adams EJ, Arnett KL (1995) The origins of HLA-A,B,C polymorphism. *Immunol Rev* **143**: 141-180.
- Parham P, Lomen CE, Lawlor DA, Ways JP, Holmes N, Coppin H, Salter RD, Wan AM, Ennis PD (1988) Nature of polymorphism in HLA-A, -B, and -C molecules. *PNAS* **85** (11): 4005-4009.
- Parham P, Ohta T (1996) Population biology of antigen presentation by MHC class I molecules. *Science* **272** (5258): 67-74.
- Paul L (1998) The position of Zazaki among West Iranian languages. In: Sims-Williams N ed. Proceedings of the 3rd European Conference of Iranian Studies. Wiesbaden: Reichert, pp. 163-176.
- Paul P, Cabestre FA, Ibrahim EC, Lefebvre S, Khalil-Daher I, Vazeux G, Quiles RM, Bermond F, Dausset J, Carosella ED (2000) Identification of HLA-G7 as a new splice variant of the HLA-G mRNA and expression of soluble HLA-G5, -G6, and -G7 transcripts in human transfected cells. *Hum Immunol* **61** (11): 1138-1149.
- Pavlos R, Mallal S, Phillips E (2012) HLA and pharmacogenetics of drug hypersensitivity. *Pharmacogenomics* **13** (11): 1285-1306.
- Payne R, Rolfs MR (1958) Fetomaternal leukocyte incompatibility. *The Journal of Clinical Investigation* **37** (12): 1756-1763.
- Peterson PA, Rask L, Lindblom JB (1974) Highly purified papain-solubilized HL-A antigens contain beta2-microglobulin. *Proc Natl Acad Sci U S A* **71** (1): 35-39.
- Pieters J (1997) MHC class II restricted antigen presentation. *Curr Opin Immunol* **9** (1): 89-96.
- Pishdad GR (2005) Low incidence of type 1 diabetes in Iran. *Diabetes Care* **28** (4): 927-928.
- Planhol X (2004) Iran I. Lands of Iran. In: Yarshater E ed. Encyclopaedia Iranica. Londres: Routledge and Kegan Paul, vol. XIII, pp. 204-212. <http://www.iranicaonline.org/articles/iran-i-lands-of-iran>
- Ploegh HL, Orr HT, Strominger JL (1981) Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell* **24** (2): 287-299.
- Profaizer T, Eckels D (2012) HLA alleles and drug hypersensitivity reactions. *Int J Immunogenet* **39** (2): 99-105.
- Prugnolle F, Manica A, Charpentier M, Guegan JF, Guernier V, Balloux F (2005) Pathogen-driven selection and worldwide HLA class I diversity. *Curr Biol* **15** (11): 1022-1027.
- Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al Zahery N, Semino O, Santachiara-Benerecetti AS et al (2004) Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet* **74** (5): 827-845.
- Quintana-Murci L, Krausz C, Zerjal T, Sayar SH, Hammer MF, Mehdi SQ, Ayub Q, Qamar R, Mohyuddin A, Radhakrishna U et al (2001) Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *Am J Hum Genet* **68** (2): 537-542.

- Radley E, Alderton RP, Kelly A, Trowsdale J, Beck S (1994) Genomic organization of HLA-DMA and HLA-DMB. Comparison of the gene organization of all six class II families in the human major histocompatibility complex. *J Biol Chem* **269** (29): 18834-18838.
- Rammensee HG (1995) Chemistry of peptides associated with MHC class I and class II molecules. *Curr Opin Immunol* **7** (1): 85-96.
- Rammensee HG, Falk K, Rotzschke O (1993) Peptides naturally presented by MHC class I molecules. *Annu Rev Immunol* **11**: 213-244.
- Reed CA (1959) Animal Domestication in the Prehistoric Near East: The origins and history of domestication are beginning to emerge from archeological excavations. *Science* **130** (3389): 1629-1639.
- Regueiro M, Cadenas AM, Gayden T, Underhill PA, Herrera RJ (2006) Iran: tricontinental nexus for Y-chromosome driven migration. *Hum Hered* **61** (3): 132-143.
- Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R et al (2001) Linkage disequilibrium in the human genome. *Nature* **411** (6834): 199-204.
- Rey D, Areces C, Alonso-Rubio J, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Bendikuzé N, Fernandez-Honrado M, Barbolla L, Martín-Villa JM, Arnaiz-Villena A (2013) HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans. *Mol Biol Rep.*
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T et al (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67** (5): 1251-1276.
- Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG (2015) The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* **43** (Database issue): D423-D431.
- Roche PA, Marks MS, Cresswell P (1991) Formation of a nine-subunit complex by HLA class II glycoproteins and the invariant chain. *Nature* **354** (6352): 392-394.
- Roder G, Blicher T, Justesen S, Johannesen B, Kristensen O, Kastrup J, Buus S, Gajhede M (2006) Crystal structures of two peptide-HLA-B*1501 complexes; structural characterization of the HLA-B62 supertype. *Acta Crystallogr D Biol Crystallogr* **62** (Pt 11): 1300-1310.
- Roitberg-Tambur A, Witt CS, Friedmann A, Safirman C, Sherman L, Battat S, Nelken D, Brautbar C (1995) Comparative analysis of HLA polymorphism at the serologic and molecular level in Moroccan and Ashkenazi Jews. *Tissue Antigens* **46** (2): 104-110.
- Rosen R (1992) Georgia: a sovereign country of the Caucasus. Hong-Kong: Odyssey Publications.
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G et al (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67** (6): 1526-1543.
- Rudensky AY, Preston-Hurlburt P, al Ramadi BK, Rothbard J, Janeway CA, Jr. (1992) Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature* **359** (6394): 429-431.

- Ruhlen M (1987) *A Guide to the World's Languages. Volume 1: Classification*. Stanford (USA): Stanford University Press.
- Ruhlen M (1994) *The Origin of Language: Tracing the Evolution of the Mother Tongue*. New York: John Wiley and Sons.
- Sabery R (2008) Ethnicity and ethnic relations in Golestan province of Iran with special reference to Turkmen. Tesis doctoral dirigida por Gambhir RD. Universidad de Pune, Facultad de Ciencias Social, Moral y Mental.
- Saghafi M, Nohesara N, Rafatpanah H, Shariati J, Shakeri MT (2014) HLA-DRB1 frequency in patients with familial and sporadic rheumatoid arthritis in north east of Iran. *Clin Rheumatol* **33** (10): 1397-1402.
- Saiki R, Scharf N, Faloona F, Mullis K, Horn G, Erlich H, Arnheim N (1985) Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4** (4): 406-425.
- Salter RD, Benjamin RJ, Wesley PK, Buxton SE, Garrett TP, Clayberger C, Krensky AM, Norment AM, Littman DR, Parham P (1990) A binding site for the T-cell co-receptor CD8 on the alpha 3 domain of HLA-A2. *Nature* **345** (6270): 41-46.
- Salvador Figueras M (2003) Análisis de Correspondencias. 5campus.com, Estadística. <http://www.5campus.com/leccion/correspondencias>. Accedido el 15-1-2015.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Sanchez-Velasco P, Gomez-Casado E, Martinez-Laso J, Moscoso J, Zamora J, Lowy E, Silvera C, Cemborain A, Leyva-Cobian F, Arnaiz-Villena A (2003) HLA alleles in isolated populations from North Spain: Origin of the Basques and the ancient Iberians. *Tissue Antigens* **61** (5): 384-392.
- Sanchez-Velasco P, Leyva-Cobian F (2001) The HLA class I and class II allele frequencies studied at the DNA level in the Svanetian population (Upper Caucasus) and their relationships to Western European populations. *Tissue Antigens* **58** (4): 223-233.
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74** (12): 5463-5467.
- Saper MA, Bjorkman PJ, Wiley DC (1991) Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol* **219** (2): 277-319.
- Satta Y, O'hUigin C, Takahata N, Klein J (1994) Intensity of natural selection at the major histocompatibility complex loci. *Proc Natl Acad Sci U S A* **91** (15): 7184-7188.
- Schippman K (1987) Azerbaijan III. Pre-islamic history. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge & Kegan Paul, vol. III, pp. 221-224. <http://www.iranicaonline.org/articles/azerbaijan-iii>
- Schmidt M, Zantopf D, Kraft R, Kostka S, Preissner R, Kloetzel PM (1999) Sequence information within proteasomal prosequences mediates efficient integration of beta-subunits into the 20 S proteasome complex. *J Mol Biol* **288** (1): 117-128.

- Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: a software for population genetics database analysis 2.0.
- Schonberg A, Theunert C, Li M, Stoneking M, Nasidze I (2011) High-throughput sequencing of complete human mtDNA genomes from the Caucasus and West Asia: high diversity and demographic inferences. *Eur J Hum Genet* **19** (9): 988-994.
- Sellier J, Sellier A (1993) Atlas des peuples d'Orient. Paris: Editions La Decouverte.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S et al (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* **290** (5494): 1155-1159.
- Shiina T, Hosomichi K, Inoko H, Kulski JK (2009) The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet* **54** (1): 15-39.
- Shiroishi M, Kuroki K, Ose T, Rasubala L, Shiratori I, Arase H, Tsumoto K, Kumagai I, Kohda D, Maenaka K (2006) Efficient leukocyte Ig-like receptor signaling and crystal structure of disulfide-linked HLA-G dimer. *J Biol Chem* **281** (15): 10439-10447.
- Singh R, Kaul R, Kaul A, Khan K (2007) A comparative review of HLA associations with hepatitis B and C viral infections across global populations. *World J Gastroenterol* **13** (12): 1770-1787.
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T et al (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423** (6942): 825-837.
- Slade RW, McCallum HI (1992) Overdominant vs. frequency-dependent selection at MHC loci. *Genetics* **132** (3): 861-864.
- Sloan VS, Cameron P, Porter G, Gammon M, Amaya M, Mellins E, Zaller DM (1995) Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* **375** (6534): 802-806.
- Snell GD (1948) Methods for the study of histocompatibility genes. *Journal of Genetics* **49**: 87-108.
- Solberg OD, Mack SJ, Lancaster AK, Single RM, Tsai Y, Sanchez-Mazas A, Thomson G (2008) Balancing selection and heterogeneity across the classical human leukocyte antigen loci: a meta-analytic review of 497 population studies. *Hum Immunol* **69** (7): 443-464.
- Spies T, Bresnahan M, Bahram S, Arnold D, Blanck G, Mellins E, Pious D, DeMars R (1990) A gene in the human major histocompatibility complex class II region controlling the class I antigen presentation pathway. *Nature* **348** (6303): 744-747.
- Srivatsan A, Han Y, Peng J, Tehrani AK, Gibbs R, Wang JD, Chen R (2008) High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. *PLoS Genet* **4** (8): e1000139.
- Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, Wiley DC (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* **368** (6468): 215-221.

- Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coggill P, Dunham I, Forbes S, Halls K, Howson JM et al (2004) Complete MHC haplotype sequencing for common disease gene mapping. *Genome Res* **14** (6): 1176-1187.
- Suarez-Alvarez B, Lopez-Vazquez A, Gonzalez MZ, Fdez-Morera JL, Diaz-Molina B, Blanco-Gelaz MA, Pascual D, Martinez-Borra J, Muro M, Alvarez-Lopez MR et al (2007) The relationship of anti-MICA antibodies and MICA expression with heart allograft rejection. *Am J Transplant* **7** (7): 1842-1848.
- Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124** (4): 967-978.
- Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogvali EL, Tolk HV, Reidla M, Metspalu E, Pliss L et al (2004) The western and eastern roots of the Saami--the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet* **74** (4): 661-682.
- Tapper R (1988) Azerbaijan VI. Population and its occupations and culture. In: Yarshater E ed. *Encyclopaedia Iranica*. Londres: Routledge and Kegan Paul, vol. III, pp. 234-238.
<http://www.iranicaonline.org/articles/azerbaijan-vi>
- Taxman DJ, Cressman DE, Ting JP (2000) Identification of class II transcriptional activator-induced genes by representational difference analysis: discoordinate regulation of the DN alpha/DO beta heterodimer. *J Immunol* **165** (3): 1410-1416.
- Terasaki PI, McClelland JD (1964) Microdroplet assay of human serum cytotoxins. *Nature* **204**: 998-1000.
- Terreros MC, Rowold DJ, Mirabal S, Herrera RJ (2011) Mitochondrial DNA and Y-chromosomal stratification in Iran: relationship between Iran and the Arabian Peninsula. *J Hum Genet* **56** (3): 235-246.
- The International Society of Genetic Genealogy (2014) Y-DNA Haplogroup Tree 2014. International Society of Genetic Genealogy. Accedido el 15-6-2014.
- The MHC sequencing consortium (1999) Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* **401** (6756): 921-923.
- Tonnelle C, DeMars R, Long EO (1985) DO beta: a new beta chain gene in HLA-D with a distinct regulation of expression. *EMBO J* **4** (11): 2839-2847.
- Trachtenberg EA, Erlich HA (2001) A review of the role of the Human Leukocyte Antigen (HLA) system as a host immunogenetic factor influencing HIV transmission and progression to AIDS. In: Korber BTM, Koup R, Walker BD, Brander C, Kuiken C, Watkins DI, Haynes BF, Moore JP eds. *HIV Molecular Immunology 2001*. Los Alamos: Theoretical Biology and Biophysics, pp. I43-I60. <http://hiv-web.lanl.gov/immunology>
- Tripathi P, Agrawal S (2007) The role of human leukocyte antigen E and G in HIV infection. *AIDS* **21** (11): 1395-1404.
- Trowsdale J, Hanson I, Mockridge I, Beck S, Townsend A, Kelly A (1990) Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters. *Nature* **348** (6303): 741-744.
- Trowsdale J, Kelly A (1985) The human HLA class II alpha chain gene DZ alpha is distinct from genes in the DP, DQ and DR subregions. *EMBO J* **4** (9): 2231-2237.

- Trowsdale J, Ragoussis J, Campbell RD (1991) Map of the human MHC. *Immunol Today* **12** (12): 443-446.
- Tulp A, Verwoerd D, Dobberstein B, Ploegh HL, Pieters J (1994) Isolation and characterization of the intracellular MHC class II compartment. *Nature* **369** (6476): 120-126.
- Uinuk-Ool TS, Takezaki N, Sukernik RI, Nagl S, Klein J (2002) Origin and affinities of indigenous Siberian populations as revealed by HLA class II gene frequencies. *Hum Genet* **110** (3): 209-226.
- Uyar FA, Dorak MT, Saruhan-Direskeneli G (2004) Human leukocyte antigen-A, -B and -C alleles and human leukocyte antigen haplotypes in Turkey: relationship to other populations. *Tissue Antigens* **64** (2): 180-187.
- van Rood J, van Leeuwen, Eernisse JG (1959) Leucocyte antibodies in sera of pregnant women. *Vox sanguinis* **4**: 427-444.
- Vicario JL, Martinez-Laso J, Corell A, Martin-Villa JM, Morales P, Lledo G, Segurado OG, de Juan D, Arnaiz-Villena A (1992) Comparison between HLA-DRB and DQ DNA sequences and classic serological markers as type 1 (insulin-dependent) diabetes mellitus predictive risk markers in the Spanish population. *Diabetologia* **35** (5): 475-481.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science* **253** (5027): 1503-1507.
- Wallace DC, Brown MD, Lott MT (1999) Mitochondrial DNA variation in human evolution and disease. *Gene* **238** (1): 211-230.
- Wallace GR (2014) HLA-B*51 the primary risk in Behcet disease. *Proc Natl Acad Sci U S A* **111** (24): 8706-8707.
- Walsh JB (1987) Sequence-dependent gene conversion: can duplicated genes diverge fast enough to escape conversion? *Genetics* **117** (3): 543-557.
- Wang J (2003) Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics* **164** (2): 747-765.
- Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S et al (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci U S A* **98** (18): 10244-10249.
- WHO Nomenclature Committee (1968) Nomenclature for factors of the HL-a system. *Bull World Health Organ* **39** (3): 483-486.
- WHO Nomenclature Committee (1975) Nomenclature for factors of the HLA system. *Bull World Health Organ* **52** (3): 261-265.
- Williams F, Meenagh A, Darke C, Acosta A, Daar AS, Gorodezky C, Hammond M, Nascimento E, Middleton D (2001) Analysis of the distribution of HLA-B alleles in populations from five continents. *Hum Immunol* **62** (6): 645-650.
- Wixman R (1984) The peoples of the USSR. An ethnographic handbook. Armonk, NY: M. E. Sharpe.
- Wu C, Macleod I, Su AI (2013) BioGPS and MyGene.info: organizing online, gene-centric information. *Nucleic Acids Res* **41** (Database issue): D561-D565.

- Yarshater E (1988) Azerbaijan VII. The Iranian language of Azerb. In: Yarshater E ed. Encyclopaedia Iranica. Londres: Routledge and Kegan Paul, vol. III, pp. 238-245. <http://www.iranicaonline.org/articles/azerbaijan-vii>
- Yepiskopasian L, Margarian S, Andonian L, Rashidvash V (2011) The location of Azaris on the patrilineal genetic landscape of the Middle East (a preliminary report). *Iran and the Caucasus* **15** (1-2): 73-78.
- Young FW, Bann CM (1996) A visual statistics system. In: Stine RA, Fox J eds. Statistical Computing Environments for Social Researches. London: Sage Publications, pp. 207-236.
- Young K, Frodsham A, Doumbo OK, Gupta S, Dolo A, Hu JT, Robson KJ, Crisanti A, Hill AV, Gilbert SC (2005) Inverse associations of human leukocyte antigen and malaria parasite types in two West African populations. *Infect Immun* **73** (2): 953-955.
- Zadok R (2006) Mannea. In: Yarshater E ed. Encyclopaedia Iranica. Londres: Routledge and Kegan Paul. <http://www.iranicaonline.org/articles/mannea>
- Zanjani HA (2002) Gorgan III. Population. In: Yarshater E ed. Encyclopaedia Iranica. Londres: Routledge & Kegan Paul, vol. XI, pp. 145-148. <http://www.iranicaonline.org/articles/gorgan-iii>
- Zerjal T, Wells RS, Yuldasheva N, Ruzibakiev R, Tyler-Smith C (2002) A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *Am J Hum Genet* **71** (3): 466-482.
- Zimdahl H, Schiefenhovel W, Kayser M, Roewer L, Nagy M (1999) Towards understanding the origin and dispersal of Austronesians in the Solomon Sea: HLA class II polymorphism in eight distinct populations of Asia-Oceania. *Eur J Immunogenet* **26** (6): 405-416.
- Zwirner NW, Dole K, Stastny P (1999) Differential surface expression of MICA by endothelial cells, fibroblasts, keratinocytes, and monocytes. *Hum Immunol* **60** (4): 323-330.
- Zwirner NW, Marcos CY, Mirbaha F, Zou Y, Stastny P (2000) Identification of MICA as a new polymorphic alloantigen recognized by antibodies in sera of organ transplant recipients. *Hum Immunol* **61** (9): 917-924.

ANEXOS

ANEXO 1**ABREVIATURAS MÁS FRECUENTES EMPLEADAS EN EL TEXTO**

(ordenadas por orden alfabético).

aa Aminoácidos
a. C. Antes de Cristo
ADN Ácido desoxirribonucleico
ADNmt Ácido desoxirribonucleico mitocondrial
ADR Reacción adversa a fármaco
APC Células presentadoras de antígenos
ARN Ácido ribonucleico
ARNm Ácido ribonucleico mensajero
ARNr Ácido ribonucleico ribosómico
ARNt Ácido ribonucleico transferente
ATP Adenosín trifosfato
BiP Proteína de unión de la cadena pesada de la inmunoglobulina
BTN Butirofilinas
β2M β2-microglobulina
CD Marcadores de diferenciación
CLIP Péptido de la cadena invariante de clase II
D Desequilibrio de ligamiento
D_A Distancia genética de Nei
DC Células dendríticas
d. C. Después de Cristo
DM1 Diabetes mellitus tipo 1
dNTPs Deoxinucleótidos trifosfato
FH Frecuencia haplotípica
HLA Antígenos leucocitarios humanos
HSP Proteínas de choque térmico
Ii Cadena invariante
Kd Kilodalton
LMP Proteínas de bajo peso molecular
LY6 Antígeno leucocitario 6
Mb Megabase
MHC Complejo principal de Histocompatibilidad
MIC Genes relacionados con la cadena de MHC de clase I
MIIC Compartimiento de MHC de clase II
MSY Región específico del cromosoma Y
NJ Neighbor-joining
NK Células *Natural Killer*
p Probabilidad
PAR Región pseudoautosómica
pb Pares de bases
PCR Reacción en cadena de la Polimerasa
PSMB Proteosoma β
PV Pénfigo vulgar
RE Retículo endoplasmático
RFLP Polimorfismo de longitud de fragmentos de restricción
SNP Polimorfismo de nucleótido simple

SSOP Oligosondas específicas de secuencia
TAP Transportadores asociados a la presentación antigénica
TCR Receptor de la célula T
TNF Factor de necrosis tumoral
UT Región no traducida
VIH Virus de inmunodeficiencia humana

ANEXO 2

TABLA A1. Perfiles HLA de todos los individuos incluidos en el estudio. El número de identificación de la muestra va precedido del prefijo “Go” en el caso de tratarse de turcomanos, “T” en el caso de tratarse de azeríes y “K” en el caso de tratarse de kurdos.

Muestra	HLA-A		HLA-B		HLA-DRB1		HLA-DQB1	
	Alelo 1	Alelo 2	Alelo 1	Alelo 2	Alelo 1	Alelo 2	Alelo 1	Alelo 2
Go2	01:01	24:02	07:02	35:01	11:01	14:01	03:01	03:04
Go3	02:01	03:01	38:01	39:01	08:01	12:09	03:04	04:01
Go4	02:01	03:01	13:01	44:02	04:01	15:01	03:05	06:02
Go5	24:02	29:01	08:01	35:20	04:01	15:01	03:02	06:02
Go6	29:01	68:01	35:01	54:01	11:01	13:08	03:01	05:03
Go9	11:01	24:02	13:04	40:01	11:01	15:01	03:01	06:02
Go13	01:02	29:01	13:01	51:06	04:01	07:01	02:01	03:02
Go14	03:01	63:03	08:01	35:27	03:01	15:01	02:01	06:02
Go16	02:01	03:01	13:01	35:01	04:01	07:01	02:01	04:01
Go17	02:01	30:01	07:03	40:46	04:01	08:20	03:07	04:01
Go18	24:02	33:01	58:01	58:06	03:01	15:01	02:01	06:01
Go21	02:01	11:01	01:01	35:01	10:01	15:01	03:07	04:01
Go22	03:01	26:01	35:10	53:07	03:01	15:01	03:01	05:03
Go23	30:01	68:01	35:01	40:46	10:01	15:01	05:01	06:02
Go24	26:13	32:04	08:01	35:01	03:01	15:01	02:01	05:01
Go29	01:01	02:34	37:01	40:39	08:20	10:01	04:01	05:02
Go30	03:02	24:02	27:01	35:01	01:01	11:01	03:01	05:01
Go31	02:50	31:07	07:03	51:06	07:01	13:01	03:01	06:03
Go32	02:50	32:03	07:03	14:01	04:01	07:01	05:01	06:03
Go33	02:34	30:01	13:01	38:01	04:01	15:01	02:01	06:03
Go34	01:02	23:01	37:01	44:02	03:01	08:01	02:01	03:02
Go36	02:01	30:01	07:03	08:01	04:01	07:01	02:01	06:03
Go37	24:02	31:02	35:01	51:06	04:01	15:01	03:02	06:02
Go38	02:01	02:50	27:01	51:01	03:01	08:01	02:01	04:01
Go39	11:01	26:13	35:27	40:02	03:01	12:09	02:01	03:01
Go40	02:01	26:01	08:01	35:27	03:01	15:01	02:01	06:01
Go41	11:01	30:01	13:01	55:01	11:17	15:01	03:01	06:02
Go43	11:01	33:01	35:01	53:01	07:01	11:28	02:01	03:01
Go45	01:02	24:02	15:17	57:01	01:01	13:01	05:01	06:03
Go46	02:34	24:02	15:46	51:06	07:01	11:01	02:01	03:01
Go50	02:01	24:02	15:24	35:20	11:01	15:01	03:01	05:01
Go51	02:01	03:01	08:01	35:01	03:13	13:01	02:03	06:03
Go54	02:34	24:02	35:01	40:01	04:01	15:01	03:02	06:02
Go56	02:01	24:02	15:24	35:01	11:01	15:01	03:01	05:01
Go57	23:01	32:04	35:01	49:01	11:01	15:01	03:01	06:02
Go58	30:01	32:01	13:01	49:01	11:01	15:01	03:01	06:02
Go61	02:50	29:01	07:03	15:10	04:01	13:01	03:02	06:03
Go79	29:02	69:01	40:39	51:06	11:01	15:01	03:01	06:02
Go80	02:52	29:01	13:01	35:01	03:13	14:05	03:01	05:01
Go81	31:01	32:01	35:01	51:01	03:01	15:01	03:01	05:02
Go82	24:02	26:01	07:02	49:01	11:01	15:01	02:01	03:01

TABLA A1. Continuación

Go85	02:01	02:01	13:01	48:01	07:01	11:01	02:01	03:01
Go86	02:50	24:02	35:01	49:01	01:01	13:02	05:01	06:04
Go87	01:02	24:02	35:01	57:01	03:03	07:01	03:02	06:03
Go90	02:52	03:01	13:01	55:01	04:01	07:01	02:01	03:02
Go91	02:48	24:08	35:01	51:01	11:01	15:01	03:02	06:03
Go92	02:34	33:01	18:01	27:01	01:01	04:01	03:02	05:01
Go93	02:52	11:01	13:01	35:01	03:08	08:20	02:01	03:01
Go94	01:01	02:61	35:01	57:01	03:01	07:01	02:01	03:03
Go95	02:50	31:01	35:01	39:01	12:01	15:01	03:01	06:02
Go97	11:01	29:01	07:02	15:02	07:01	15:01	02:01	05:01
Go98	24:02	26:01	38:01	51:01	01:01	11:01	03:01	05:01
Go99	02:01	33:01	14:02	15:02	01:01	15:01	05:01	06:01
Go100	26:01	31:01	35:01	40:01	11:01	15:01	03:01	06:02
Go101	02:50	24:02	35:01	48:01	03:01	04:01	02:01	03:01
Go102	01:01	03:02	08:01	35:01	01:01	03:01	02:01	05:01
Go103	02:50	03:01	08:01	51:01	03:13	14:24	02:01	06:04
Go104	03:01	25:01	18:09	35:10	03:01	10:01	02:01	05:01
Go105	01:02	25:01	07:02	39:01	08:01	15:01	04:01	06:04
Go106	31:01	32:01	18:01	35:27	11:01	15:01	03:01	06:02
Go108	29:01	31:01	35:01	39:01	07:01	12:01	02:01	03:01
Go109	11:01	26:01	38:01	51:01	10:01	13:01	05:01	06:03
Go110	02:50	30:01	14:02	35:38	01:01	07:01	02:01	05:01
Go112	03:01	26:01	07:02	50:01	07:01	11:01	02:01	03:01
Go115	29:01	68:01	08:01	13:01	03:01	11:01	02:01	03:01
Go120	23:01	69:01	35:01	50:01	07:01	13:06	05:01	06:02
Go123	02:01	02:52	51:01	51:08	07:01	15:01	06:02	06:02
Go129	01:02	02:01	35:01	51:01	11:20	13:02	03:07	03:07
T1	32:01	33:01	38:01	49:01	11:01	14:01	03:01	05:02
T2	24:03	26:01	18:01	18:01	01:01	04:01	03:02	05:01
T3	01:02	03:01	08:01	18:01	03:01	04:01	02:01	03:01
T4	01:01	33:01	52:01	58:01	15:01	15:01	06:01	06:01
T5	24:02	33:01	14:01	52:01	01:01	04:01	03:02	05:01
T6	01:01	26:01	35:01	56:01	07:01	10:01	02:01	05:01
T7	11:01	68:01	35:01	51:01	11:01	15:01	03:01	06:01
T8	24:02	24:02	35:01	35:01	07:01	11:01	02:01	03:01
T9	02:01	29:01	51:08	58:01	03:01	08:01	02:01	04:01
T10	02:50	11:01	35:01	52:01	04:01	11:01	03:01	04:01
T11	02:50	24:02	35:01	57:01	07:01	11:01	03:01	03:03
T12	24:02	33:01	35:01	55:01	13:01	14:01	05:02	06:03
T13	02:50	24:02	35:01	44:02	04:01	11:01	03:01	04:01
T14	02:01	02:01	35:01	52:01	14:01	14:01	05:02	05:02
T15	02:50	30:01	38:01	49:01	14:01	15:01	03:05	05:01
T16	01:01	01:01	27:01	27:01	04:01	04:01	03:02	03:02
T17	02:01	26:19	07:02	40:01	04:01	15:01	03:02	05:01
T18	01:02	30:01	14:01	55:01	04:01	13:01	04:01	06:03
T19	29:01	68:01	51:01	58:01	08:01	13:01	04:01	06:02
T20	02:01	24:02	35:01	51:01	04:01	14:01	03:02	05:01

TABLA A1. Continuación

T21	01:02	01:02	52:01	52:01	15:01	15:01	06:01	06:01
T22	03:01	24:02	35:01	78:01	07:01	11:01	02:01	03:01
T23	01:02	32:01	35:01	51:01	04:01	14:02	03:01	03:01
T24	24:03	68:01	18:01	51:01	12:01	12:01	03:01	03:01
T25	33:01	68:01	39:01	58:01	03:01	16:01	02:01	05:01
T26	11:01	11:01	13:01	51:01	07:01	16:01	02:01	05:01
T27	01:01	02:01	07:02	37:01	10:01	16:01	05:01	05:01
T28	03:01	11:01	35:01	44:02	04:01	13:01	03:02	06:03
T29	01:02	68:01	41:01	51:01	07:01	13:01	03:03	06:03
T30	24:02	30:01	38:01	41:01	03:01	03:01	02:01	02:01
T31	01:01	03:01	07:02	35:01	10:01	13:02	05:01	06:04
T32	03:01	33:01	35:01	52:01	07:01	15:01	03:03	06:01
T33	29:01	33:01	14:01	15:10	01:01	03:01	02:01	05:01
T34	01:01	24:02	18:01	35:01	11:01	11:01	03:01	03:01
T35	02:01	11:01	07:02	51:08	11:01	15:01	03:01	05:01
T36	24:02	31:01	35:01	55:01	08:01	14:01	04:01	05:01
T37	03:01	11:01	18:01	44:02	04:01	15:01	03:02	06:01
T38	03:02	24:02	35:01	52:01	04:01	11:01	03:01	03:02
T39	02:01	30:01	18:01	35:01	04:01	15:01	03:02	06:01
T40	01:02	29:01	35:01	37:01	10:01	11:01	03:01	05:01
T41	01:02	68:01	08:01	51:01	03:01	15:01	02:01	06:01
T42	01:01	24:02	51:01	52:01	11:01	16:01	03:01	05:01
T43	03:01	29:01	07:02	51:01	10:01	15:01	05:01	06:02
T44	01:01	03:01	14:01	18:01	07:01	13:01	03:03	06:03
T45	11:01	29:01	07:02	55:01	04:01	10:01	03:05	05:01
T46	24:02	30:01	27:01	35:01	13:03	16:01	03:01	05:01
T47	02:50	24:02	46:01	51:01	08:01	09:01	03:03	04:01
T48	02:01	33:01	50:01	58:01	03:01	07:01	02:01	02:01
T49	02:50	24:02	35:01	39:01	08:01	16:01	04:01	05:01
T50	01:01	24:02	37:01	50:01	10:01	11:01	03:01	05:01
T51	02:01	24:02	15:01	41:01	07:01	09:01	02:01	03:03
T52	02:01	11:01	51:07	52:01	04:01	15:01	03:02	06:01
T53	02:01	68:01	15:01	44:02	01:01	13:02	05:01	06:04
T54	03:01	24:02	38:01	44:02	04:01	14:01	03:02	05:01
T55	03:01	33:01	52:01	53:01	07:01	14:01	02:01	05:02
T56	24:03	30:01	13:01	40:01	07:01	15:01	02:01	06:01
T57	03:01	32:01	52:01	58:08	04:01	13:08	03:01	06:04
T58	02:01	03:02	44:02	51:01	04:01	13:01	03:02	06:03
T59	26:01	68:01	08:01	49:01	03:01	11:01	02:01	03:01
T60	01:02	24:02	08:01	15:01	03:01	11:01	02:01	03:03
T61	01:02	30:01	13:01	35:01	01:01	07:01	02:01	05:01
T62	01:01	03:02	08:01	50:01	03:01	07:01	02:01	03:12
T63	24:02	66:01	41:01	51:01	04:01	07:01	02:01	03:01
T64	01:01	03:01	35:01	49:01	01:01	13:01	05:01	06:03
T65	02:01	23:01	49:01	50:01	15:01	15:01	06:02	06:02
T66	02:01	30:01	35:01	38:01	14:01	15:01	05:02	06:01
T68	26:01	34:02	35:01	44:02	04:01	07:01	02:01	03:01

TABLA A1. Continuación

T69	02:01	03:01	13:01	18:01	07:01	11:01	02:01	03:01
T70	11:01	68:01	44:02	55:01	11:01	15:01	03:01	06:03
T71	03:01	11:01	07:02	51:01	04:01	15:01	03:01	06:02
T72	01:01	32:01	40:01	52:01	15:01	15:01	06:01	06:02
T73	02:01	32:01	41:01	50:01	03:02	14:01	04:01	05:01
T74	01:02	33:01	14:01	51:01	01:01	11:01	03:01	05:01
T75	02:01	30:01	35:01	35:01	11:01	11:01	03:01	03:01
T76	01:01	03:01	38:01	50:01	11:01	15:01	03:01	06:01
T77	01:01	32:01	41:01	44:02	03:01	04:01	02:01	03:02
T78	02:01	31:01	14:01	35:01	01:01	11:01	03:01	05:01
T81	02:01	03:01	07:02	50:01	07:01	15:01	02:01	06:02
T82	03:01	32:01	18:09	35:10	01:01	11:01	03:01	05:01
T86	01:01	03:01	18:01	49:01	04:01	11:01	03:01	03:02
T87	02:01	26:01	50:01	52:01	03:01	11:01	02:01	03:01
T89	01:02	02:01	08:01	35:19	03:01	11:01	02:01	03:01
T90	32:01	69:01	55:01	55:01	03:01	15:01	02:01	05:01
T91	01:01	32:01	57:01	57:01	07:01	15:01	03:03	06:01
T92	01:01	02:01	49:01	51:06	09:01	15:01	03:03	06:02
T93	02:01	24:02	15:01	41:01	04:01	14:02	03:01	03:02
T94	11:01	24:02	50:01	52:01	04:01	07:01	02:01	03:01
T95	01:01	02:01	41:01	51:01	16:01	16:01	05:01	05:01
T97	01:01	03:01	13:01	18:01	11:01	14:01	03:01	05:02
T98	02:01	31:01	52:01	54:01	04:01	15:01	04:01	06:01
T99	24:02	33:01	38:01	44:02	13:01	13:03	03:07	06:02
T100	01:01	32:01	13:01	38:01	03:01	03:01	02:01	02:01
T102	02:50	02:50	15:24	41:01	12:01	13:05	03:07	03:07
T105	01:01	02:52	07:02	35:05	03:08	14:01	03:07	03:07
T110	01:01	68:01	18:01	41:01	08:01	11:01	02:01	04:01
T111	29:01	68:01	44:18	51:06	04:01	13:02	03:05	03:05
T112	30:01	32:01	13:01	15:17	04:01	07:01	03:02	06:02
K1	02:52	03:01	53:01	44:02	03:01	11:01	02:01	03:01
K2	01:01	02:52	08:01	13:01	03:01	04:01	02:01	03:05
K3	02:01	02:52	49:01	51:01	10:01	13:01	05:01	06:03
K4	02:52	24:02	18:09	35:01	03:01	11:01	02:01	03:01
K5	03:01	24:02	35:01	58:01	03:01	04:01	02:01	03:02
K6	02:52	11:01	35:01	50:01	07:01	11:01	02:01	03:01
K7	13:01	23:01	07:24	18:01	11:01	15:01	03:01	06:02
K8	02:01	24:02	55:01	58:01	03:01	15:01	02:01	06:03
K9	02:50	26:01	08:01	51:01	03:01	16:01	02:02	05:01
K10	02:50	24:02	27:01	44:02	10:01	14:01	05:02	05:02
K11	02:61	30:01	07:02	35:01	11:01	13:01	03:01	06:03
K13	03:02	24:02	51:01	51:01	04:01	08:01	03:02	04:01
K14	02:01	11:01	35:01	44:10	01:01	07:01	02:01	05:01
K15	03:01	23:01	35:01	49:01	01:03	15:01	03:05	06:03
K16	02:50	30:01	13:01	44:02	04:01	07:01	03:01	03:05
K17	02:50	11:01	13:01	35:01	01:03	16:01	06:01	06:02
K18	02:52	11:01	35:27	40:01	01:02	15:01	06:02	06:02

TABLA A1. Continuación

K19	03:01	32:01	35:01	44:02	11:01	11:01	03:01	03:01
K22	02:01	24:02	18:09	35:01	11:01	11:01	03:01	03:01
K24	01:01	01:02	35:01	35:27	03:01	13:02	02:01	06:04
K26	02:01	02:48	49:01	51:06	04:01	15:01	03:05	05:01
K29	02:48	24:02	18:01	58:01	03:01	07:01	02:01	02:01
K31	26:01	32:01	38:01	55:01	03:01	14:01	03:07	06:03
K33	24:02	33:01	18:01	58:01	03:02	04:01	03:02	06:04
K34	24:02	31:01	08:01	35:01	01:01	03:01	02:01	05:01
K35	01:02	02:01	50:01	50:01	07:01	13:05	02:01	03:01
K36	01:02	11:01	35:01	38:01	13:01	14:01	05:02	06:03
K37	26:01	32:01	49:01	55:01	11:01	14:01	03:01	05:02
K38	11:01	32:01	35:01	52:01	03:01	15:01	02:01	06:01
K39	02:01	33:01	50:01	58:01	03:01	14:01	02:01	05:02
K40	01:01	24:02	18:01	52:01	15:01	16:01	05:01	06:01
K41	24:02	26:01	38:01	55:01	01:01	04:01	03:07	05:02
K42	01:01	26:01	08:01	35:01	03:01	11:01	02:01	03:01
K44	01:01	02:52	35:01	51:01	11:01	14:01	03:01	05:02
K47	02:50	24:02	35:01	35:01	03:01	14:01	02:01	05:02
K49	03:01	23:01	44:02	49:01	03:01	11:01	02:01	03:01
K50	01:01	68:01	53:01	55:01	13:01	13:02	03:01	06:02
K51	02:50	24:03	35:01	51:01	04:01	13:01	03:02	06:02
K53	03:01	24:02	18:09	35:01	11:01	11:01	03:01	03:01
K54	03:01	24:02	50:01	51:01	04:01	08:01	04:01	04:01
K56	03:01	24:02	44:02	51:07	04:01	15:01	05:01	06:04
K57	24:02	26:01	38:01	52:01	04:01	15:01	03:05	06:01
K58	02:52	24:02	18:01	51:01	11:01	11:01	03:01	03:01
K59	02:01	26:01	35:01	38:01	11:01	13:01	03:01	06:02
K60	11:01	24:02	35:20	52:01	01:01	15:01	05:01	06:01
K62	02:01	03:01	35:01	51:01	01:01	04:15	03:01	05:01
K63	02:50	03:01	18:01	40:01	11:01	16:01	03:01	03:02
K64	03:01	32:01	07:02	18:01	11:01	15:01	03:01	06:02
K65	24:02	69:01	18:09	35:01	07:01	11:01	02:01	03:01
K66	02:01	24:02	40:15	49:01	13:02	16:01	03:02	06:01
K67	01:01	24:02	15:24	52:01	01:01	12:01	03:01	05:01
K68	01:01	31:01	08:01	54:01	03:01	04:01	02:01	04:01
K69	02:01	11:01	35:01	50:01	07:01	15:01	02:01	05:02
K70	02:01	26:01	08:01	41:01	03:01	04:01	02:01	02:01
K72	01:01	02:01	14:01	35:01	01:01	04:01	03:02	05:01
K73	01:01	03:01	35:01	53:01	03:01	11:01	02:01	03:01
K74	03:01	11:01	50:01	55:01	07:01	11:01	02:01	03:01
K75	02:01	02:61	15:17	51:01	04:01	13:02	03:02	06:04
K76	26:01	32:01	13:01	58:01	03:01	14:01	02:01	05:02
K78	02:01	11:01	35:01	51:01	01:01	11:01	03:01	05:01

ANEXO 3**CURRICULUM VITAE****Formación académica**

Nombre: Diego Federico Rey Medrano
 DNI: 72069990L
 Fecha de Nacimiento: 28/06/1984
 Nacionalidad: España
 Dirección de contacto: C/ Virtudes 22 2ºB 28010 Madrid
 Teléfono: 660 992 478
 Correo electrónico: dreymedrano@gmail.com

Situación profesional actual

Organismo: Universidad Complutense – Centro de Transfusión de la Comunidad de Madrid

Facultad, Escuela o Instituto: Medicina

Depto./Secc./Unidad estr.: Microbiología I (Inmunología)

Dirección postal: Avda. Complutense s/n 28040

Teléfono (indicar prefijo, número y extensión): 91 301 7356

Fax: 91 301 7229

Correo electrónico: diegoreymedrano@med.ucm.es

Especialización (Códigos UNESCO): 2412.99 (Inmunogenética) 2409.03 (Genética de las Poblaciones)

Categoría profesional: Becario predoctoral Fecha de inicio: octubre 2009

Situación administrativa

Plantilla Contratado Interino Becario
 Otras situaciones especificar:

Dedicación A tiempo completo
 A tiempo parcial

Líneas de investigación

Estudio de genética comparada de poblaciones mediante el uso del sistema HLA.

Descripción y caracterización del polimorfismo y la estructura genética de los alelos HLA.

Estudio evolutivo del sistema MHC (Complejo Principal de Histocompatibilidad) mediante el estudio del polimorfismo de los locus MHC-G y MHC-DMB en diferentes especies de primates (chimpancé, gorila, orangután, macaco).

Relaciones filogenéticas entre Paseriformes; evolución de su sistema MHC (principal de histocompatibilidad).

Formación académica

Titulación Superior	Centro	Fecha
Licenciado en CC. Biológicas	Universidad Complutense de Madrid	2002-2008

Postgrado	Centro	Fecha
Postgrado en Ciencias Biomédicas	Universidad Complutense de Madrid	2009-2011

Actividades anteriores de carácter científico profesional

Puesto	Institución	Fechas
Técnico de laboratorio hematológico y bioquímico Becario en prácticas (Beca ARGO)	Ematos Vetlab, SRL	2008-2009
Becario de Colaboración en Departamentos	Universidad Complutense de Madrid	2007-2008

Idiomas (R = regular, B = bien, C = correctamente)

Idioma	Habla	Lee	Escribe
Inglés	B	B	B
Italiano	C	C	C
Francés	B	B	R

Participación en proyectos de I+D+i

- Estudio epidemiológico de los genes de obesidad e histocompatibilidad y farmacogenómico HLA en Amerindios en España: genes PC-1, PPARG2, apM1, y HLA inmunosupresores y clásicos. Entidad financiadora: Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III, 2015-2017. Investigador Principal: Antonio Arnaiz Villena.
- Genes de trasplante no clásicos (HLA-E, -F, -G, -DMB y MIC), y de obesidad en Amerindios inmigrantes a España: medicina preventiva, trasplante y farmacocinética. Entidad financiadora: Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III, 2012-2014. Investigador Principal: Antonio Arnaiz Villena.
- Genes de obesidad e histocompatibilidad no clásicos (HLA-E, -F, -G) MIC y HLA-DMB en grupos Amerindios de América inmigrantes a Madrid: medicina preventiva, trasplante y farmacocinética. Entidad financiadora: Fundación Mutua Madrileña Automovilística, 2012-2014. Investigador Principal: Antonio Arnaiz Villena.

Publicaciones o Documentos Científico-Técnicos

A.- Artículos publicados en Revistas Internacionales y Nacionales

- 1.- Arnaiz-Villena A, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Marco J, Muñiz E, Fernández-Honrado M, Martín-Villa M, Rey D. *Pacific Islanders and Amerindian relatedness according to HLA autosomal genes*. *International Journal of Modern Anthropology*, 7:44-67, 2014.
- 2.- Arnaiz-Villena A, Ruiz-del-Valle V, Gomez-Prieto P, Rey D, Enríquez-de-Salamanca M, Marco J, Muñiz E, Martín-Villa M, Areces C. *Carduelini New Systematics: Crimson-winged Finch (*Rodhoptechys sanguineus*) is included in "Arid-Zone" Carduelini Finches by Mitochondrial DNA Phylogeny*. *The Open Ornithology Journal*, 7:55-62, 2014.
- 3.- Arnaiz-Villena A, Vargas-Alarcón G, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Fernández-Honrado M, Marco J, Martín-Villa JM, Rey D. *Mixtec Mexican Amerindians: an HLA alleles study for America peopling, pharmacogenomics and transplantation*. *Immunological Investigations*, 43(8):738-755, 2014.
- 4.- Rey D, Amirzargar A, Areces C, Enríquez-de-Salamanca M, Marco J, Abd-El-Fatah-Khalil S, Fernández-Honrado M, Muñiz E, Martín-Villa JM, Arnaiz-Villena A. *Gorgan (Turkmen in Iran) HLA genetics: transplantation, pharmacogenomics and anthropology*. *Immunological Investigations*, 44(1): 88-100, 2014.
- 5.- Rey D, Areces C, Alonso-Rubio J, Enríquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Bendikuze N, Fernández-Honrado M, Barbolla L, Martín-Villa JM, Arnaiz-Villena A. *HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans*. *Molecular Biology Reports*, 40(10):5523-5530, 2013.
- 6.- Rey D, Parga-Lozano C, Moscoso J, Areces C, Enríquez-de-Salamanca M, Fernández-Honrado M, Abd-El-Fatah-Khalil S, Alonso-Rubio J, Arnaiz-Villena A. *HLA genetic profile Mapuche (Araucanian) Amerindians from Chile*. *Molecular Biology Reports*, 40(7): 4257-4267, 2013.
- 7.- Arnaiz-Villena A, Enríquez-de-Salamanca M, Areces C, Abd-El-Fatah Khalil S, Fernández-Honrado M, Rey D. *HLA-G*01:05N null allele in Mayans (Guatemala) and Uros (Titikaka Lake, Peru): evolution and population genetics*. *Human Immunology*, 74(4): 478-482, 2013.
- 8.- Arnaiz-Villena A, Rey D, Enríquez-de-Salamanca M, Abd-El-Fatah Khalil S, Arribas I, Coca C, Fernández-Honrado M, Algora M, Areces C. *Amerindians show association of obesity with adiponectin gene SNP45 and SNP276: population genetics of a food intake and "thrifty" gene*. *Molecular Biology Reports*, 40(2): 1819-1826, 2013.
- 9.- Arnaiz-Villena A, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah Khalil S, Arribas I, Coca C, Fernández-Honrado M, Algora M, Rey D. *Amerindians show no association of PPAR- γ 2 gene Ala12 allele and obesity: an "unthrifty" variant population genetics*. *Molecular Biology Reports*, 40(2): 1767-1774, 2013.
- 10.- Arnaiz-Villena A, Areces C, Rey D, Enríquez-de-Salamanca M, Alonso-Rubio J, Ruiz-del-Valle V. *Three different North American Siskin/Goldfinch evolutionary radiations (Genus *Carduelis*): Pine Siskin green morphs and European Siskin in America*. *The Open Ornithology Journal*, 5:73-81, 2012.
- 11.- Rey D, Areces C, Enríquez-de-Salamanca M, Parga-Lozano C, Abd-El-Fatah S, Fernández M, Arnaiz-Villena A. *Los primeros pobladores de América y sus relaciones con las poblaciones del Océano Pacífico según los genes HLA*. *Inmunología*, 31:83-91, 2012.
- 12.- Arnaiz-Villena A, Rey D. *Iberian-Tartessian scripts/graffiti in Iruna-Veleia (Basque Country, North Spain): findings in both Iberia and Canary Islands-Africa*. *International Journal of Modern Anthropology*, 1(5): 24-38, 2012.
- 13.- Rey D, Fernández-Honrado M, Areces C, Algora M, Abd-El-Fatah S, Enríquez-de-Salamanca M, Coca C, Arribas I, Arnaiz-Villena A. *Ameridians show no association of PC-1 gene Gln121 allele and obesity: a thrifty gene population genetics*. *Molecular Biology Reports*, 39(7): 7687-7693, 2012.
- 14.- Arnaiz-Villena A, Fernández-Honrado M, Areces C, Arribas I, Coca C, Enríquez-de-Salamanca M, Parga-Lozano C, Abd-El-Fatah S, Rey D. *Amerindians normalized waist circumference and obesity*

diagnosis standardized by biochemical and HLA data. Molecular Biology Reports, 39(4): 4875-4878, 2012.

15.-Longás J, Martínez-Laso J, Rey D, Areces C, Casado EG, Parga-Lozano C, Luna F, Enríquez-de-Salamanca M, Moral P, Arnaiz-Villena A. *Las Alpujarras region (South East Spain) HLA genes study: evidence of a probable success of 17th century repopulation from North Spain.* Molecular Biology Reports, 39:1387-1394, 2012.

16.-Arnaiz-Villena A, Abd-El-Fatah S, Granados-Silvestre MA, Parga-Lozano C, Gómez-Prieto P, Rey D, Areces C, Peñaranda P, Menjivar M, Rodríguez-Perez JM, Granados J, Vargas-Alarcón G. *Human Leukocyte Antigen-DRB1 class II genes in Mexican Amerindian Mazahuas: genes and language do not correlate.* Genetic Testing and Molecular Biomarkers, 15 (1-2): 97-102, 2011.

17.-Martínez-Laso J, Montoya F, Areces C, Moscoso J, Silvera C, Rey D, Parga-Lozano C, Gómez-Prieto P, Enríquez-de-Salamanca M, Arnaiz-Villena A. *HLA in Jaidukama: Amerindian secluded Colombian population with new haplotypes and Pacific-shared alleles.* Molecular Biology Reports, 38(6): 3689-3701, 2011.

18.-Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moureau P. *Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association.* Cell & Molecular Life Sciences, 68(3): 369-395, 2011.

19.-Parga-Lozano C, Rey D, Gómez-Prieto P, Areces C, Moscoso J, Abd-El-Fatah-Khalil S, Arnaiz-Villena A. *HLA genes in Amerindian immigrants to Madrid: epidemiology and a virtual transplantation waiting list.* Molecular Biology Reports, 38 (4): 2263-2271, 2011.

20.-Silvera C, Vargas-Alarcón G, Areces C, Rey D, Parga-Lozano C, Gómez-Prieto P, Barbolla L, Arnaiz-Villena A. *HLA genes in Wayu Amerindian from Colombia.* Immunological Investigations, 40: 92-100, 2011.

21.-Rey D, Areces C, Parga-Lozano C, Gómez-Prieto P, Crawford MH, Arnaiz-Villena A. *HLA genes variation in populations of the Aleutian Islands.* Human Biology, 82 (5-6): 737-744, 2010.

22.-Arnaiz-Villena A, Ruiz-del-Valle V, Reche P, Gómez-Prieto P, Lowry E, Zamora J, Areces C, Rey D, Parga C, Serrano-Vela JI. *Songbirds conserved sites and introns size of MHC class I molecules reveal a unique evolution in vertebrates.* The Open Ornithology Journal, 3: 156-165, 2010.

23.-Arnaiz-Villena A, Areces C, Gómez-Prieto P, Parga-Lozano C, Moreno E, Abd-El-Fatah-Khalil S, Rey D. *The peopling of the Americas: a complex issue for Amerindian, Na-Dene, Aleut and Eskimo first inhabitants.* International Journal of Modern Anthropology, 3:65-79, 2010.

24.-Arnaiz-Villena A, Parga-Lozano C, Moreno E, Areces C, Rey D, Gómez-Prieto P. *The Origin of Amerindians and the Peopling of the Americas According to HLA Genes: Admixture with Asian and Pacific People.* Current Genomics, 11 (2): 104-113, 2010.

25.-Vargas-Alarcón G, Granados J, Rodríguez-Pérez JM, Parga-Lozano C, Pérez-Hernández N, Rey D, Zuñiga J, Arnaiz-Villena A. *Distribution of HLA Class II Alleles and Haplotypes in Mexican Mestizo Population: Comparison with Other Populations.* Immunological Investigations, 39 (3): 268-283, 2010.

26.-Mesa MS, López-Parra AM, Rey D, Baeza C, Alfaro E, Castilla F, Marrodán MD, Tirado M, Iturbe P, Gamba C, Fernández E, Dipierri JE, Arroyo-Pardo E. *La aplicación de los Y-STRs a la investigación de los linajes familiares.* Revista Española de Antropología Física, Vol. 29: 68, 2009.

B.- Artículos Pendientes de Publicación en Revistas Internacionales

27.-Rey D, Amirzargar A, Areces C, Enríquez-de-Salamanca M, Marco J, Abd-El-Fatah-Khalil S, Fernández-Honrado M, Muñiz E, Martín-Villa JM, Arnaiz-Villena A. *Kurds HLA genetics: transplantation, pharmacogenomics and anthropology.* Immunological Investigations, En Prensa, 2015.

28.-Rey D, Amirzargar A, Areces C, Enríquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Fernández-Honrado M, Marco J, Muñiz E, Martín-Villa JM, Arnaiz-Villena A. *Azeris HLA genetics: anthropology, transplantation and pharmacogenomics.* Immunological Investigations, En Prensa, 2015.

C.- Capítulos de Libros

29.-Gómez-Prieto P, Reguera R, Parga-Lozano C, Moreno E, Rey D, Areces C, Arnaiz Vileña A. *HLA-G, -F y -E: Polymorphism, Function and Evolution*. Capítulo 9 en *HLA complex in Biology and Medicine: A resource book*. Editor: N. Mehra. Jaypee Brothers Medical Publishers Ltd, New Delhi, India, 2010.

Contribuciones a Congresos

71 contribuciones a los siguientes congresos.

Internacionales:

- 28th European Immunogenetics and Histocompatibility Conference "From genes to therapy". Stockholm, Sweden, June 25th-28th, 2014.
- 27th European Immunogenetics and Histocompatibility Conference. Maastricht, the Netherlands, May 11th-14th, 2013.
- 5th North American Ornithological Conference (NAOC-V). Vancouver, Canada, August 14th-18th, 2012.
- 16th International HLA and Immunogenetics Workshop Report Conference, 26th European Immunogenetics and Histocompatibility Conference and 23rd British Society for Histocompatibility and Immunogenetics Conference. Liverpool, United Kingdom, May 31st-June 3rd, 2012.
- 129th Stated Meeting of the American Ornithologists' Union. Jacksonville, Florida, USA, July 24th-29th, 2011.
- 25th European Immunogenetics and Histocompatibility Conference. Prague, Czech Republic, May 4th-7th, 2011.
- Convención Internacional de Antropología ANTHROPOS 2011. La Habana, Cuba, March 14th-18th, 2011.
- 25th International Ornithological Congress. Campos de Jordão, Brazil, August 22nd-28th, 2010.
- 24th European Immunogenetics and Histocompatibility Conference. Florence, Italy, May 15th-18th, 2010.
- 35th Human Biology Association Annual Meeting. Albuquerque, New Mexico, USA, April 14th-15th, 2010.
- COS/AOU/SCO 2010 Joint Meeting (80th annual meeting of the Cooper Society, 128th stated meeting of the American Ornithologists' Union, and the 28th annual meeting of the Society of Canadian Ornithologists/Société des Ornithologistes du Canada). San Diego, California, USA, February 7th-11th, 2010.
- V Congreso de la Asociación Latinoamericana de Antropología Forense. Buenos Aires, Argentina, October 26th-28th, 2009.
- 51st Symposium and Proffered Papers Meeting "Human Dispersals". Rome, Italy, June 18th-20th, 2009
- X Congreso de la Asociación Latinoamericana de Antropología Biológica (ALAB). La Plata, Argentina, October 20th-23rd, 2008.

Nacionales:

- XXII Congreso Español de Ornitología. Madrid, 6-9 Diciembre 2015.
- 56 Congreso de la Sociedad Española de Endocrinología y Nutrición (SEEN). Valencia, 21-23 Mayo 2014.
- XXXVIII Congreso Nacional de Inmunología. Badajoz, 8-10 Mayo 2014.
- IV Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Barcelona, 27-29 Noviembre 2013.
- XXXVII Congreso Nacional de Inmunología. Salamanca, 29-31 Mayo 2013.
- XXI Congreso Español y V Ibérico de Ornitología. Vitoria, 6-9 Diciembre 2012.

- I Congreso Ibérico de Sistemática Animal. Madrid, 17-19 Enero 2012.
- III Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Madrid, 21-25 Noviembre 2011.
- XXXVI Congreso Nacional de la Sociedad Española de Inmunología. Pamplona, 8-11 Junio 2011.
- XX Congreso Español de Ornitología. Tremp (Lérida), 4-8 Diciembre 2010.
- XXXV Congreso de la Sociedad Española de Inmunología. San Sebastián, 23-26 Junio 2010.
- II Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Valencia, 29 Noviembre-2 Diciembre 2009.
- XVI Congreso de la Sociedad Española de Antropología Física. Alcalá de Henares, 6-9 Julio 2009.

Formación complementaria

Mayo 2013	Curso. EndNote Web. Biblioteca de la Facultad de Medicina, Universidad Complutense de Madrid. 3 horas
Mayo 2013	Curso. Pubmed. Biblioteca de la Facultad de Medicina, Universidad Complutense de Madrid. 3 horas
Octubre 2009	Seminario. Origen, Diversidad y Mestizaje de las Poblaciones Humanas desde el punto de vista molecular. Facultad de Biología, Universidad Complutense de Madrid. 5 horas
Agosto 2008	Curso. La antropología física del siglo XXI, una ciencia sin fronteras. Universidad Complutense de Madrid. 30 horas
Julio 2008	Seminario. La Policía Científica del siglo XXI en el marco europeo. Universidad Internacional de Menéndez Pelayo. 30 horas
Febrero 2008	Curso. Introducción a la Genealogía Molecular. Laboratorio de Genética Forense y Genética de las Poblaciones de la Universidad Complutense de Madrid y Servicio de Genealogía Molecular SIGENlab. Madrid. 6 horas
Julio 2007	Seminario. Los primeros humanos en el viejo continente y la revolución faunística del Plioceno al Pleistoceno. Universidad Internacional Menéndez Pelayo. Santander. 15 horas

Experiencia en organización de actividades I+D+i

Agosto 2010	Organización del Seminario titulado “El Origen de los Vascos” en el ámbito de los Cursos de Verano de la Universidad Complutense de Madrid
-------------	--

Otros méritos o aclaraciones que se desee hacer constar

BECAS RECIBIDAS

2011-2015	Beca predoctoral UCM-FPI
2009-2011	Beca de la Fundación Mutua Madrileña Automovilística con cargo a proyecto "Historia poblacional de los genes anti-stress y ahorradores de energía: su importancia en la patogenia de la diabetes y del síndrome metabólico"
2008-2009	Beca ARGO concedida por el Ministerio de Educación y el Programa Europeo de Movilidad Leonardo da Vinci
2007-2008	Beca de Colaboración en Departamentos concedida por la Universidad

2005-2006	Complutense de Madrid Beca ERASMUS concedida por la Unión Europea y el Ministerio de Educación
2000-2001	Beca de la Fundación Marcelino Botín para la realización del Curso de Creación Literaria

EXPERIENCIA DOCENTE

2012-2015	Colaborador en tareas docentes del Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid en docencia práctica y seminarios prestando servicio a diferentes asignaturas de inmunología de las licenciaturas y grados de Biología, Farmacia y Medicina y del máster de Inmunología (con acreditación)
2010-2012	Colaborador Honorífico del Departamento de Microbiología I (Inmunología) de la Facultad de Medicina de la Universidad Complutense de Madrid en docencia práctica prestando servicio a diferentes asignaturas de inmunología de las licenciaturas de Biología, Farmacia, Medicina y Bioquímica (con acreditaciones)

ESTANCIAS EN EL EXTRANJERO

2005-2006	Estancia en la Università degli Studi di Roma Tor Vergata (programa ERASMUS)
-----------	--

ASISTENCIA A CONGRESOS

- I Congreso Ibérico de Sistemática Animal. Madrid, 17-19 Enero 2012.
- III Congreso de la Sociedad Española de Biología Evolutiva (SESBE). Madrid, 21-25 Noviembre 2011.
- XXXVI Congreso Nacional de la Sociedad Española de Inmunología. Pamplona, 8-11 Junio 2011.
- 25th European Immunogenetics and Histocompatibility Conference. Prague, Czech Republic, May 4th-7th, 2011.
- XXXV Congreso de la Sociedad Española de Inmunología. San Sebastián, 23-26 Junio 2010.

ANEXO 4**TRABAJOS PUBLICADOS CON RESULTADOS OBTENIDOS EN ESTA TESIS**

1.- **Rey D**, Areces C, Alonso-Rubio J, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Bendikuze N, Fernandez-Honrado M, Barbolla L, Martin-Villa JM, Arnaiz-Villena A. HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans. *Molecular Biology Reports*, 40(10):5523-5530, 2013. A continuación se adjunta una copia del artículo.

2.- **Rey D**, Amirzargar A, Areces C, Enriquez-de-Salamanca M, Marco J, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Muñiz E, Martin-Villa JM, Arnaiz-Villena A. Gorgan (Turkmen in Iran) HLA genetics: transplantation, pharmacogenomics and anthropology. *Immunological Investigations*, 44(1): 88-100, 2014. A continuación se adjunta una copia del artículo.

TRABAJOS ENVIADOS PARA SU PUBLICACIÓN CON RESULTADOS OBTENIDOS EN ESTA TESIS

3.- **Rey D**, Amirzargar A, Areces C, Enriquez-de-Salamanca M, Marco J, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Muñiz E, Martin-Villa JM, Arnaiz-Villena A. Kurds HLA genetics: transplantation, pharmacogenomics and anthropology. *Immunological Investigations*, En Prensa, 2015.

4.- **Rey D**, Amirzargar A, Areces C, Enriquez-de-Salamanca M, Abd-El-Fatah-Khalil S, Fernandez-Honrado M, Marco J, Muñiz E, Martin-Villa JM, Arnaiz-Villena A. Azeris HLA genetics: anthropology, transplantation and pharmacogenomics. *Immunological Investigations*, En Prensa, 2015.

HLA in Georgians (Caucasus) and their relationship with Eastern Mediterraneans

Diego Rey · Cristina Areces · Javier Alonso-Rubio · Mercedes Enríquez-de-Salamanca · Sedeka Abd-El-Fatah-Khalil · Nina Bendikuze · Mercedes Fernández-Honrado · Luz Barbolla · José Manuel Martín-Villa · Antonio Arnaiz-Villena

Received: 12 April 2013 / Accepted: 7 August 2013 / Published online: 20 August 2013
© Springer Science+Business Media Dordrecht 2013

Abstract HLA-A, -B, -DQB1, and -DRB1 typing has been performed in a sample of Georgian population (South Caucasus). Allele frequencies, neighbour joining and correspondence relatedness analyses and extended HLA haplotypes have been obtained with comparison with other Middle East and Mediterranean populations. Our Georgian sample tends to be genetically related in these analyses with Eastern Mediterraneans and Middle East people. This is important for future regional transplant programs, and Georgian HLA and disease epidemiology and pharmacogenomics.

Keywords Caucasians · Georgians · Greeks · HLA · Pharmacogenomics · Population genetics

Introduction

Southern Caucasus Georgian Republic includes South Ossetia and the main Caucasus Pass is placed between

Diego Rey and Cristina Areces contributed equally for this work and the order of authorship is arbitrary.

D. Rey · C. Areces · J. Alonso-Rubio · M. Enríquez-de-Salamanca · J. M. Martín-Villa · A. Arnaiz-Villena (✉)
Department of Immunology, School of Medicine, Madrid Regional Blood Center, University Complutense, 28020 Madrid, Spain
e-mail: aarnaiz@med.ucm.es
URL: <http://chopo.pntic.mec.es/biolmol/>

S. Abd-El-Fatah-Khalil · M. Fernández-Honrado · L. Barbolla
Department of Haematology, Madrid Regional Blood Center, Madrid, Spain

N. Bendikuze
Institute for Clinical and Experimental Medicine, Prague, Czech Republic

North and South Ossetia. In addition, inside Georgia two Republics there exist: Abkhazia and Adjara. Georgian capital (Tbilisi) has a big number of Chechen and Kurd inhabitants. Eastern Georgia has been named Iberia since antiquity and Black Sea part was identified as Cholchid (The Golden Fleece and Medea mythological placement) [1–4].

Georgia is related with Mesopotamian (Anatolian) history in ancient times [1]. It is believed that Assyrians and Hittites took shelter in Georgia by 1200 BC and mixed with autochthonous Caucasus population [4]. After 730 AD Scythians and Cimmerians coming from North went through Georgia and many people fled to highlands [1]. Later the main Iberian tribes Tibal and Mushki formed Georgian population nucleus, in spite of war with southern tribes coming from ancient Urartu (Armenians). Alexander the Great did not invade Georgia, but Romans did it in 1st century AD after defeating Iberians (Fig. 1) [1, 2]. Romans stayed for 300 years and St. Nino christianized Eastern part, Kartli/Iberia, but not the western Georgian part. Arabs invaded Georgia in 645 AD but left it going with its culture and Christian Orthodox religion. Later Bragations dynasty held power of Eastern Georgia until Russia annexed Georgia in 1801 AD. Several Turks and Iranians invasions have run over Georgia in different times. North and South Ossetians are descents of Alans [4] and speak an Iranian language (Fig. 1). Georgia was independent from Russia between 1918 and 1921. Stalin was born in Gora (Georgia) and brought Georgia under Russia rule [1–4].

The Georgian language is close to the Dene-Caucasian and Basque languages [5–8]. In fact, the Georgians name their country, Georgia, Sakartvelo (in Basque language saka = hollow, art = between and velo = summit), which describes perfectly Georgian orography.

Gorgan (Turkmen in Iran) HLA genetics: transplantation, pharmacogenomics and anthropology

Diego Rey,¹ Ali Amirzargar,² Cristina Areces,¹
Mercedes Enríquez-de-Salamanca,¹ Javier Marco,¹
Sedeka Abd-El-Fatah-Khalil,³ Mercedes Fernández-Honrado,³
Ester Muñiz,¹ José Manuel Martín-Villa,¹ and
Antonio Arnaiz-Villena¹

¹Departament of Immunology, University Complutense, School of Medicine, Madrid Regional Blood Center, Madrid, Spain,

²Molecular Immunology Research Center, Medical School, Tehran University of Medical Sciences, Tehran, Iran, and

³Departament of Haematology, Madrid Regional Blood Center, Madrid, Spain

HLA class I and II alleles have been studied in a population from Gorgan (North East Iranian city bordering Turkmenistan). This population is composed of mainly Turkmen who speak Oghuz Turkish language. Comparison of Gorgan people HLA profile has been carried out with about 7984 HLA chromosomes from other worldwide populations; extended haplotypes and three dimension genetic distances have been calculated by using neighbor-joining and correspondence relatedness analyses. Most frequent extended HLA haplotypes show a Siberian/Mediterranean admixture and closest populations are Chuvashians (North Caspian Sea, Russia) and other geographically close populations like Siberian Mansi, Buryats and other Iranians. New extended HLA haplotypes have been found, such as: A*31:01-B*35:01-DRB1*15:01-DQB1*03:01, A*01:01-B*35:01-DRB1*03:01-DQB1*02:01. Relationships of Turkmen with Kurgan (Gorgan) archaeological mounds, Scythians and Sarmatians are discussed. This study is also useful for a future transplantation Gorgan waiting list, Gorgan HLA and disease epidemiology and HLA pharmacogenomics.

Keywords Anthropology, Asian, Gorgan, HLA, Iran, Kurgan, Mediterranean, pharmacogenomics, Turkmen

INTRODUCTION

The HLA system is the most polymorphic genetic system described in humans and consists of several closely linked loci. The strength of the multiallelic HLA system to single out individuals for paternity testing has been shown to be greater than the additive strength of many other polymorphic enzymes and blood group systems. The discovery of new loci and the presently available DNA typing and sequencing of new alleles have dramatically increased the