



**UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO**

---

**FACULTAD DE QUÍMICA**

*EVOLUCIÓN DEL VIRUS DE LA INMUNODEFICIENCIA  
HUMANA (VIH) EN MÉXICO: IMPACTO DE LOS GENES HLA  
DE CLASE I EN DIFERENTES REGIONES DE LA REPÚBLICA  
MEXICANA.*

**T E S I S**

**PARA OBTENER EL TÍTULO DE:  
QUIMICA FARMACÉUTICA BIÓLOGA**

**PRESENTA:**

**MARIBEL SOTO NAVA**



**MÉXICO, D.F.**

**AÑO 2010**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**JURADO ASIGNADO:**

**PRESIDENTE:**       **Profesor: Rosa Elena Sarmiento Silva**

**VOCAL:**             **Profesor: Laura Cecilia Bonifaz Alfonzo**

**SECRETARIO:**     **Profesor: Santiago Avila Ríos**

**1er. SUPLENTE:**   **Profesor: Sonia Mayra Pérez Tapia**

**2° SUPLENTE:**     **Profesor: León Patricio Martínez Castilla**

**SITIO DONDE SE DESARROLLÓ EL TEMA:**

**CENTRO DE INVESTIGACIÓN EN ENFERMEDADES INFECCIOSAS (CIENI)**

**ASESOR DEL TEMA:**

**M EN C. SANTIAGO AVILA RÍOS**

(Nombre y firma)

**SUSTENTANTE:**

**MARIBEL SOTO NAVA**

(Nombre y firma)

## AGRADECIMIENTOS

Hay tantas personas a las que debo agradecer por todo el apoyo y cariño que me han brindado a lo largo de mi carrera profesional y también a lo largo de toda mi vida. En este momento mencionare aquellas personas que hicieron posible el término de esta meta tan importante en mi vida.

Primero quiero agradecer a mis Padres: Micaela y Guillermo, que me han dado la vida, que gracias a todo el apoyo, cariño y comprensión que he recibido a lo largo de esta carrera, me ha ayudado a cumplir con cualquiera que sea mi sueño. Mil gracias a mi hermana Aracely, por todos los momentos tan divertidos que hemos pasado juntas, eres la mejor compañía que podría tener.

A mis tíos Carolina y Ernesto, que han estado conmigo en los momentos buenos y malos que he tenido, no solo en mi carrera profesional, sino también a lo largo de mi vida. GRACIAS por cuidar de mí y hacerme la persona que soy ahora, que termina con una meta más en su vida. SIN USTEDES SERÍA IMPOSIBLE.

A Alfonso Sánchez, quien ha estado conmigo durante toda mi carrera y que no sólo es parte de mi vida, sino que es el pilar de ella. GRACIAS por todo el cariño y apoyo incondicional que me brindas; por todas las palabras de ánimo que tanto me han ayudado a seguir adelante; por escucharme tantas y tantas veces que lo he necesitado y por supuesto, por acompañarme todo este tiempo. TE AMO.

A mi amiga, Eréndira Rojas que siempre me brinda una amistad sincera e incondicional. Gracias por todos los consejos y los buenos momentos que hemos compartido, que espero no sean los últimos.

Por supuesto a mi tutor Santiago Ávila, por darme la oportunidad de integrarme al grupo de investigación. Gracias por el apoyo, consejos y enseñanzas que hicieron posibles el término de este proyecto.

Gracias a los que colaboraron conmigo en este trabajo, Christopher E. Ormsby, por los análisis estadísticos, Daniela Garrido y Claudia García por apoyarme con las técnicas y por supuesto por darme consejos sobre ellas.

Y finalmente al Dr. Gustavo Reyes Terán, director del CIENI. Por el apoyo y por procurar siempre que no nos falte nada para seguir adelante en nuestro trabajo.

## ÍNDICE

RESUMEN.....	8
1. INTRODUCCIÓN.....	10
SITUACIÓN EPIDEMIOLÓGICA DEL VIH/SIDA EN EL MUNDO.....	10
EPIDEMIOLOGÍA DE LA INFECCIÓN POR VIH EN MÉXICO.....	14
VIRUS DE LA INMUNODEFICIENCIA HUMANA (VIH).....	16
Historia.....	16
Clasificación.....	17
Estructura y organización del genoma del VIH.....	19
Ciclo de replicación del VIH.....	23
VIAS DE TRANSMISIÓN.....	25
HISTORIA NATURAL DE LA INFECCIÓN POR VIH.....	26
Infección Primaria.....	26
Fase Crónica.....	29
Avance de la enfermedad, Síndrome de la Inmunodeficiencia Humana (SIDA).....	30
ANTÍGENO LEUCOCITARIO HUMANO (HLA) Y RESPUESTA DE LOS LINFOCITOS T CITOTÓXICOS (CTLs).....	30
VARIABILIDAD DEL HLA EN EL MUNDO.....	33
HLA Y VIH.....	34
IMPACTO DEL HLA EN LA EVOLUCIÓN DEL VIH.....	38
Nivel intra-paciente.....	38
Nivel poblacional.....	39

2. PLANTEAMIENTO DEL PROBLEMA.....	40
3. OBJETIVOS.....	42
GENERALES.....	42
PARTICULARES.....	42
4. HIPÓTESIS.....	42
5. METODOLOGÍA.....	43
PACIENTES.....	43
SEPARACIÓN DE PBMC's.....	44
EXTRACCIÓN DE DNA.....	45
TIPIFICACIÓN DE HLA DE CLASE I A ALTA RESOLUCIÓN.....	45
TIPIFICACIÓN DE HLA DE CLASE I A BAJA RESOLUCIÓN.....	47
ANÁLISIS DE FRECUENCIAS DE LOS DIFERENTES ALELOS HLA....	49
Análisis de frecuencias HLA en las distintas regiones en estudio (Norte, Centro y Sur).....	49
Análisis de frecuencias de los distintos alelos HLA con relación al estado y/o zona de procedencia.....	49
Desequilibrio de ligamiento.....	49
Análisis de correlación entre valores significativamente más altos o bajos de carga viral y la expresión de los alelos HLA específicos.....	50
ANÁLISIS DE SECUENCIAS DEL VIH.....	50
Alineamiento de secuencias.....	50
Análisis filogenéticos.....	51
ANÁLISIS EVOLUTIVOS.....	51

Análisis de selección positiva con modelo de codones.....	51
Análisis de correlación de alelo HLA – polimorfismo del VIH.....	52
6. RESULTADOS.....	53
CARACTERÍSTICAS DE LA COHORTE ESTUDIADA.....	53
FRECUENCIA HLA EN LA COHORTE EN ESTUDIO.....	58
Frecuencias HLA en la cohorte mexicana.....	58
Frecuencias HLA de la cohorte de la región Norte del país.....	61
Frecuencias HLA de la cohorte de la región Centro del país.....	62
Frecuencias HLA en la región Sur del país.....	64
Comparación de frecuencias alélicas entre la región Norte y Sur del país.....	66
Comparación de las frecuencias alélicas HLA entre las regiones Norte, Centro y Sur del país.....	70
Comparación de las frecuencias alélicas HLA entre los estados que se incluyeron en el estudio.....	76
COMPARACIÓN DE FRECUENCIAS ALÉLICAS DE LA COHORTE DE INDIVIDUOS VIH POSITIVOS CON UNA COHORTE DE INDIVIDUOS VIH NEGATIVOS.....	82
Comparación de frecuencias HLA de los individuos VIH (+) de la región Norte con las observadas en una población VIH (-).....	84
Comparación de las frecuencias HLA de los individuos VIH (+) de la región Centro con las observadas en una población VIH (-).....	84
Comparación de las frecuencias HLA de los individuos VIH (+) de la	86

región Sur con las observadas en una población VIH (-).....	
COMPARACIÓN DE LAS FRECUENCIAS ALÉLICAS HLA DE LA	
POBLACIÓN MEXICANA VIH (+) EN ESTUDIO CON UNA	89
POBLACIÓN VIH (+) CAUCÁSICA.....	
ANÁLISIS DE DESEQUILIBRIO DE LIGAMIENTO.....	91
ANÁLISIS DE LA EXPRESIÓN DE ALELOS ESPECÍFICOS HLAS	
CON LOS VALORES DE CARGA VIRAL (CV).....	93
ANÁLISIS EVOLUTIVOS.....	95
Análisis de selección positiva por modelo de codones.....	95
Análisis de asociaciones alelos HLA – polimorfismo del VIH.....	98
Análisis de asociación entre la presencia de un polimorfismo en el	
genoma viral y la procedencia de los individuos.....	101
Confirmación de las asociaciones HLA – polimorfismos de VIH	
mediante mapeo de epítopes de CTLs.....	102
7. DISCUSIÓN.....	105
CARÁCTERÍSTICAS DE LA COHORTE ESTUDIADA.....	105
COMPARACIÓN DE FRECUENCIAS HLA DE CLASE I.....	107
FACTORES DE RIESGO O PROTECCIÓN DENTRO DE LA	
COHORTE EN ESTUDIO.....	112
EVOLUCIÓN DEL VIH.....	113
8. CONCLUSIONES.....	120
9. PERSPECTIVAS.....	123
10. REFERENCIAS.....	124

## RESUMEN

Este trabajo representa sólo una parte de un estudio evolutivo poblacional para obtener información sobre el papel de los genes HLA de clase I como presión selectiva sobre el VIH, así como el impacto de la diversidad inmunogenética humana en la adaptación del virus a nivel poblacional.

El análisis de frecuencias HLA a alta y baja resolución de 139 individuos incluidos en el estudio y la comparación de esta diversidad de frecuencias entre las regiones Norte, Centro y Sur del país, muestra que la población mexicana caracterizada por un claro mestizaje tiene una gran diversidad inmunogenética dentro del territorio nacional [5-7]. La diversidad de expresión de frecuencias HLA tiene características únicas en comparación con cohortes enriquecidas con individuos caucásicos [5]. Esta evidencia sugiere que la evolución del VIH dentro del territorio nacional mediada por alelos HLA, podría seguir un camino diferente para cada una de las regiones del país y ser única para la población mexicana.

Las frecuencias alélicas de las diferentes regiones incluidas en el estudio, mostraron la diversidad de alelos HLA de origen amerindio expresados mayoritariamente en la región Centro y Sur del país. En la comparación de las frecuencias alélicas entre el Norte y Sur del país se encontraron alelos como A\*0206, que mostró estar sobre expresado en la región Sur en comparación a la baja frecuencia en la región Norte del país, principalmente en el estado de Nuevo León. Mientras que los alelos sobre expresados en la región Norte del país fueron: A\*2402, Cw\*0401 y Cw\*0501 los dos primeros son alelos amerindios [7] y el tercero es un alelo predominante en población caucásica [5]. Estas diferencias en la expresión de alelos HLA, nos sugieren las diferencias en las características inmunogenéticas existentes en población mexicana en regiones distintas del país. En esta cohorte en estudio, se encontraron alelos que están funcionando como alelos protectores como son los alelos: A\*3101, Cw\*0202 y Cw\*05 asociados con baja carga viral (CV). Se encontró también al alelo B\*1801 que ya se encuentra

reportado como un alelo de riesgo, por estar asociado a CV alta y a una progresión rápida de la infección por el VIH [8].

Aunque existen importantes limitaciones para el análisis de asociaciones HLA – codón del VIH, debido al número limitado de individuos incluidos en el estudio y al bajo poder de detección de asociaciones HLA – polimorfismo del VIH. El presente trabajo muestra evidencias que sugieren características de evolución del VIH mediada por HLA, algunas de estas consistentes con las ya reportadas en trabajos previos [5]. Usando un fragmento de 1302pb el gen *pol* del VIH, que incluye 99 posiciones de la Proteasa (PR) y 335 posiciones de la retrotranscriptasa (RT), de 123 individuos incluidos en el estudio, se encontraron en total 24 asociaciones HLA – codón de VIH a alta resolución y 33 a baja resolución. Sabiendo que muchas de estas asociaciones son falsos positivos y cuentan con un bajo poder de detección, algunas de estas asociaciones pudieron comprobarse con epítopes óptimos ya reportados o epítopes predichos para los alelos considerados importantes en el contexto de este trabajo. De esta forma la posición 41 de la RT asociado con el alelo B\*44 coincidió con el epítope de los alelos B\*4402 y B\*4403. De la misma forma las posiciones asociadas a los alelos B\*1801 y Cw\*05 que mostraron tener una asociación con CV alta, coincidieron con epítopes predichos. Se comprobaron algunas posiciones con desequilibrio de ligamiento que se encuentran muy frecuentes en la cohorte mexicana, como son lo alelos B\*40/Cw\*07 y B\*07/Cw\*07 ya reportados en Avila – Ríos y colaboradores [5].

El estudio a profundidad de la dinámica evolución del VIH mediada por la diversidad única de frecuencias HLA a nivel poblacional dentro del territorio nacional, lleva al diseño de tratamientos específicos para cada región del país para así tener un control y una mejora de la calidad de vida de la población que se encuentra infectada por el VIH.

# 1. INTRODUCCIÓN

## ***SITUACIÓN EPIDEMIOLÓGICA DEL VIH/SIDA EN EL MUNDO***

Algunas epidemias en la historia se han detenido sólo al sucumbir los individuos susceptibles (es el caso de la peste); otras han sido superadas cuando el avance en el conocimiento permite la disponibilidad de técnicas para prevenirlas o combatirlas (viruela); otras más persisten después de siglos (lepra y tuberculosis); incluso otras, cuyo extenso control llevó a pensar que estaban en camino de eliminación, han resurgido (paludismo). Pocas enfermedades emergen como condiciones nuevas, para las que aún no se está lo suficientemente preparado; entre ellas se pueden contar las fiebres hemorrágicas y el síndrome de inmunodeficiencia adquirida (SIDA) [9].

El SIDA ya tiene un lugar en la historia de la humanidad por su gran impacto no solo en el ámbito de la salud sino también en el ámbito social, al igual que la lepra, la peste o la viruela. El SIDA se convirtió en epidemia y rápidamente en pandemia, y su naturaleza modificó por igual la conciencia de grupos progresistas y liberales, conservadores y religiosos [3].

Desde 1984, cuando se aisló por primera vez el agente etiológico del SIDA, conocido posteriormente como VIH (Virus de la Inmunodeficiencia Humana), este se ha confirmado a través del aislamiento de los retrovirus de personas infectadas y la detección de anticuerpos. Es claro que existe un grupo particular de riesgo, como son los usuarios de drogas intravenosas; personas que reciben transfusiones sanguíneas, como los hemofílicos y aquellas personas que practican relaciones sexuales de riesgo [10].

Para el 2007 existían 33.2 millones de personas infectadas por el virus de la inmunodeficiencia humana (VIH) en el mundo, un 16% menos que la cifra estimada publicada en 2006 (39.5 millones); de las cuales 30.8 millones eran personas adultas y el resto eran niños menores de 15 años. La principal razón individual de esta reducción fue el ejercicio intensivo de evaluar la epidemia del VIH algunos países como son la India y África subsahariana [9].

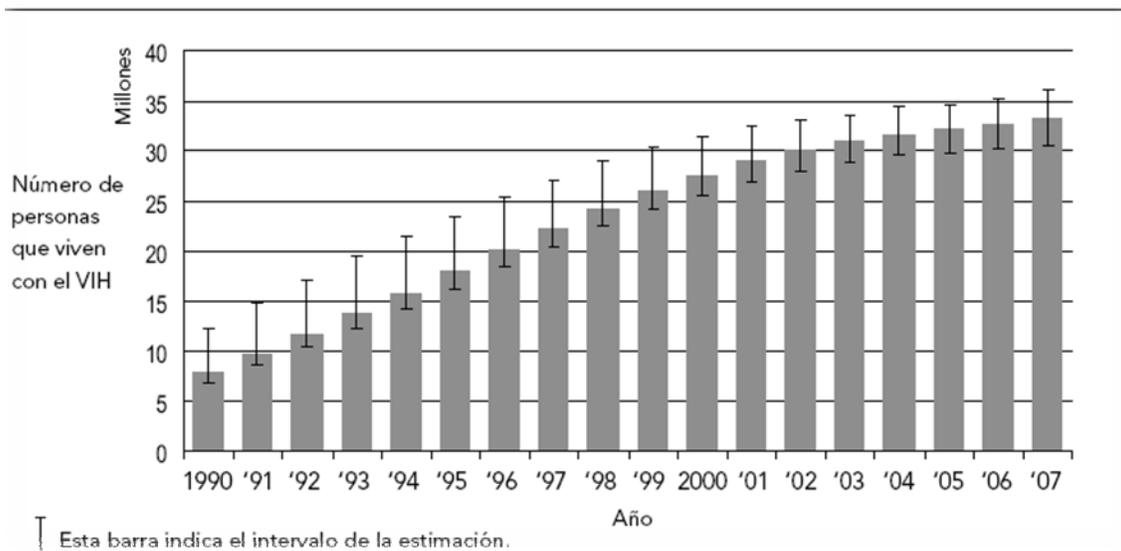
Cada día, más de 6800 personas contraen infección por VIH y más de 5700 individuos fallecen a diario a causa del SIDA, en la mayoría de los casos debido

a un acceso inadecuado a los servicios de prevención y tratamiento del VIH (Figura 1.1). La estadística muestra que para el 2007 se observaron 2.5 millones de casos nuevos por VIH en todo el mundo [9].

El tratamiento antirretroviral es la mejor opción para suprimir la viremia plasmática de manera duradera y subsecuentemente reducir la morbilidad y mortalidad que es generalmente asociada con la infección del VIH-1.

El alto grado de la replicación viral, la baja fidelidad de la transcriptasa inversa y la habilidad para recombinarse son las características virales que permiten la diversidad de las especies de VIH-1 (cuasiespecies) en los individuos con infección crónica. La terapia antirretroviral altamente activa (TARAA ó HAART por sus siglas en inglés), es una combinación de varios agentes antirretrovirales potentes, con los que se suprime la replicación viral a niveles indetectables con las técnicas de laboratorio. Con este tratamiento se incrementa el número de CD4+, disminuye la morbilidad y mortalidad, aumentando la esperanza de vida, transformando la infección de VIH-1 en una infección crónica tratable [11].

Número estimado de personas que viven con el VIH a nivel mundial, 1990–2007



**Figura 1.1 Número estimado de personas que viven con el VIH a nivel mundial, datos de 1990 al 2007.** En los últimos años se ha observado un aumento sustancial en el número de personas infectadas con el VIH en el mundo. A pesar de que existe una disminución en el número total de casos nuevos por año, se estima que el uso extensivo de los fármacos antirretrovirales, confiriendo una sobrevida más larga a los individuos infectados, contribuye a que el número total de casos se mantenga elevado. Tomado de [3].

La pandemia del VIH sigue constituyendo uno de los desafíos más importantes en enfermedades infecciosas para la salud pública. No obstante, la evaluación epidemiológica actual presenta elementos alentadores, ya que sugiere que:

- La prevalencia mundial de la infección por el VIH (porcentaje de personas infectadas por el virus) se mantiene en el mismo nivel, aunque el número general de personas que viven con el VIH está aumentando debido a la acumulación continua de nuevas infecciones con períodos más prolongados de supervivencia, medidos en una población general en constante crecimiento.
- Existen reducciones localizadas en la prevalencia en países específicos.
- Se observa una reducción en la mortalidad asociada al VIH, en parte atribuible al reciente aumento de acceso al tratamiento antirretroviral.
- Se ha logrado una reducción en el número total de nuevas infecciones anuales por el VIH a nivel mundial.

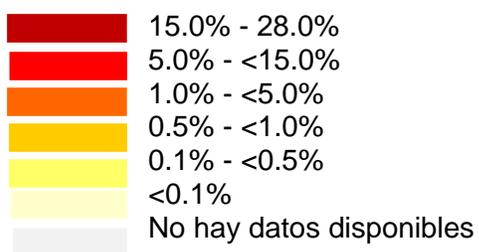
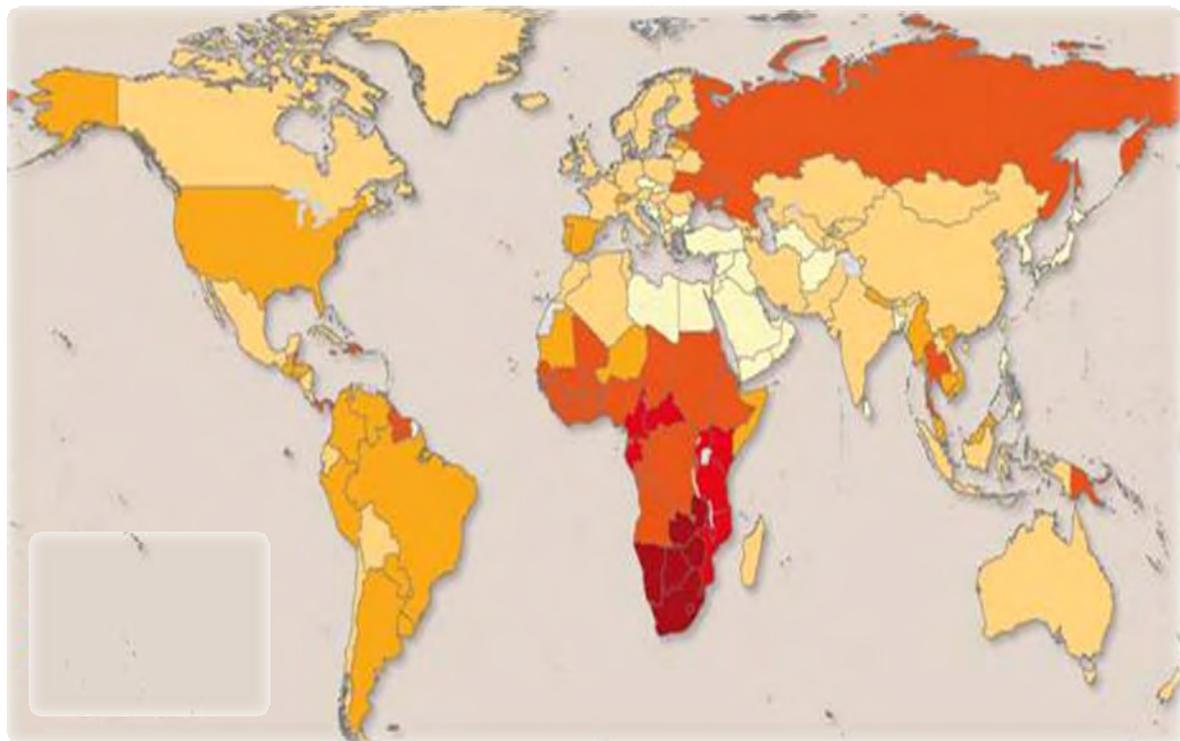
La evaluación de las tendencias mundiales y regionales sugiere que la pandemia ha tomado dos patrones generales:

- Epidemias generalizadas en las poblaciones totales de muchos países de África subsahariana, en especial en la parte meridional del continente; y
- Epidemias en el resto del mundo que se concentran principalmente entre las poblaciones de mayor riesgo como los hombres que tienen relaciones sexuales con hombres, usuarios de drogas inyectables, profesionales del sexo y sus parejas sexuales.

África subsahariana continúa siendo la región más gravemente afectada por la epidemia, siendo el SIDA la causa principal de mortalidad [9]. La prevalencia del VIH tiende a reducirse lentamente con el paso del tiempo, debido a una reducción en el número de nuevas infecciones totales y debido a la defunción de las personas infectadas por el VIH. Por otra parte, la prevalencia de infección por el virus podría aumentar al incrementarse la sobrevivencia de las personas infectadas por el VIH que reciben tratamiento antirretroviral.

En todo el mundo, el número estimado de fallecimientos a causa del SIDA en 2007 fue de 2.1 millones, de los cuales 76% se produjo en África Subsahariana [9]. Las disminuciones registradas en los dos últimos años se deben en parte a la ampliación de los servicios de tratamiento antirretroviral.

La prevalencia heterogénea de adultos infectados con el VIH en las diversas partes del mundo se muestra en la Figura 1.2.



**Figura 1.2 Prevalencia de infección por VIH en adultos.**

Se muestran prevalencias de infección por VIH en individuos mayores de 15 años en el mundo. La zona más afectada es África subsahariana. La mayoría de las regiones del continente americano tienen < 1.0% de prevalencia, con excepción de algunos países del Caribe. Modificada de: [1]

## ***EPIDEMIOLOGÍA DE LA INFECCIÓN POR VIH EN MÉXICO***

La epidemia mexicana por VIH/SIDA está claramente definida por periodos: una etapa inicial que comenzó en 1983, caracterizada por un crecimiento lento, seguida de un crecimiento exponencial a mitad de los años ochentas y nuevamente un crecimiento lento que comenzó en 1994 [3].

Hasta el 14 de noviembre del 2008 se habían acumulado 124 505 casos de SIDA en el país, de los cuales el 82% son hombres y 17.5% para las mujeres [12]. El mayor número de casos se presentan en el Distrito Federal (17.2%), Estado de México (11.0%), Veracruz (9%) y Jalisco (8.3%). Un 48.2% de los casos documentados están entre los 30 y 44 años de edad. Para esta fecha los casos de mortalidad por cada 100 000 habitantes fue de 4.8. En este año, se diagnosticaron 26 200 casos de infecciones por VIH. La mayoría de los casos se dieron en las edades comprendidas de 15 a 44 años de edad tanto para mujeres como para varones. Para los jóvenes de 15 a 29 años de edad la transmisión sexual es responsable de 95.6% de los casos para hombres y 98.2% de los casos para mujeres, mientras que la transmisión sanguínea constituye el 1.3% y 1.8% de los casos respectivamente. En menores de 15 años, la gran mayoría de las infecciones corresponden a transmisión perinatal (90.9% en hombres y 94.3% en mujeres) [12].

Para todos estos datos debe considerarse tanto el subregistro como el retraso en la notificación ya que son problemas que afectan, en mayor o menor medida, la calidad de los datos en todos los países del mundo incluyendo nuestro país. Estos fenómenos son el resultado de múltiples mecanismos, entre los que se destacan el desconocimiento de la normatividad, la falta de diagnóstico oportuno, el no envío u ocultamiento de la notificación y los mecanismos burocráticos que estimulan el retraso [13].

En el 2007 se observó que la prevalencia de VIH/SIDA en individuos de 15 a 44 años es de 0.3%, ya que se ha mantenido estable en los últimos años. La prevalencia es mayor a 5% en grupos específicos con una alta frecuencia de actividades de riesgo para contraer la infección. En hombres que tienen sexo con hombres (HSH) es de 12.6%, en hombres que trabajadores del sexo (HTS)

comercial es de 6.1%, mientras que en usuarios de drogas inyectables, mujeres trabajadoras del sexo (MTS), hombres reclusos y mujeres embarazadas la prevalencia es inferior a 2 por ciento [3].

La emigración a gran escala de mexicanos hacia Estados Unidos ha sido considerada por numerosos estudios como un importante factor de diseminación del VIH/SIDA. También se ha relacionado con la difusión de la infección en las áreas rurales del país, y con una mayor prevalencia de MTS, HSH y usuarios de drogas intravenosas (UDI) de ambos sexos en las regiones fronterizas [13].

Para los migrantes mexicanos, aun sin incrementar las prácticas de riesgo, las posibilidades de haberse infectado en Estados Unidos eran 5.5 veces mayores. Para el 31 de Diciembre de 2007, el 12.7% de los casos acumulados en México se había producido en personas que habían radicado previamente en Estados Unidos, con una incidencia mayor (27.8%) entre los casos que se habían presentado en localidades menores a cinco mil habitantes. Los estados de Michoacán y Zacatecas poseen a más de 20% de los casos acumulados con historial de residencia en Estados Unidos [13].

Una de las primeras medidas que se tomaron al inicio de la epidemia fue la descripción epidemiológica del padecimiento, mediante vigilancia y aplicación de encuestas seroepidemiológicas y de comportamiento. Debido a la importancia global de la epidemia y la aparición cada vez más frecuente de casos nuevos en el ámbito nacional, el gobierno mexicano creó en 1986 el Comité Nacional para la Prevención del SIDA (Conasida), cuyas actividades se limitaban a la coordinación de esfuerzos en la lucha contra el SIDA. Para 1988, dicho comité fue elevado a la categoría de Consejo Nacional. Conasida se consolidó como la entidad encargada de la formulación y difusión de las políticas y estrategias nacionales en materia de atención control de VIH /SIDA y otras enfermedades de transmisión sexual. Recientemente Conasida pasó a ser el Centro Nacional para la Prevención y Control del VIH/SIDA (Censida), conservando su estructura organizacional y sus funciones.

El Instituto Mexicano del Seguro Social (IMSS) ha realizado esfuerzos para que el paciente con infección por VIH/SIDA pueda ser recibido en todas las unidades

médicas de la institución y para que todo el personal de salud pueda participar en su atención. Desde 1994 el IMSS integró guías clínicas para orientar el proceso diagnóstico-terapéutico, en las que se establecen en un formato flexible, lineamientos para el empleo de medicamentos antirretrovirales (ARV), quimioprofilaxis e inmunizaciones, así como para la atención de las complicaciones asociadas con la infección, las cuales permiten orientar todo el proceso institucional [3].

## ***VIRUS DE LA INMUNODEFICIENCIA HUMANA (VIH)***

### **Historia**

Al final de 1970 y principios de 1980, algunos individuos antes sanos con síntomas de una disfunción inmunológica buscaron consejo y tratamiento con médicos de los Estados Unidos y Europa. Este síndrome nuevo e inusual, fue caracterizado por una linfadenopatía generalizada, infecciones oportunistas (típicamente *Pneumocystis jirovecii*, neumonía, retinitis asociada a Citomegalovirus, meningitis cryptococcal y una variedad inusual de cánceres (linfoma no Hodgkin y sarcoma de Kaposi). Los individuos afectados tenían una marcada depleción de los linfocitos T CD4+ de la sangre periférica. La enfermedad comenzó a tener atención cuando en junio de 1981 el CDC (Center for Disease Control) describió a 5 hombres con severa inmunodeficiencia en el *Morbidity and Mortality Weekly Report*. Este síndrome de inmunodeficiencia se observó también en otros grupos que incluían hemofílicos, receptores de transfusión de sangre y más significativamente parejas sexuales o niños de miembros de varios grupos de riesgo [10].

Entre 1981 y 1983, numerosos microorganismos fueron propuestos como posibles agentes etiológicos del síndrome de inmunodeficiencia adquirida (SIDA), como fue llamado. En la primavera de 1983, Montagnier y colegas en el Instituto Pasteur aislaron al agente de nodos linfáticos de un individuo asintomático que presentaba la linfadenopatía generalizada de origen desconocido. Durante este cultivo celular, el virus asociado a linfadenopatía, o LAV como fue llamado, producía viriones que

contenían una transcriptasa inversa dependiente de magnesio que es parte de los retrovirus [10].

Gallo y colaboradores en los Institutos Nacionales de la Salud en Estados Unidos, reportaron el aislamiento del retrovirus de los pacientes con SIDA, el cual llamaron virus de leucemia de células T humano tipo III (HTLV-III) para distinguirlo del virus no citopático HTLV-1 [10].

Al mismo tiempo Levy y colaboradores recuperaron un virus similar de pacientes con SIDA y de individuos saludables de varios grupos de riesgo, el cual llamaron retrovirus asociado al SIDA (ARV). Este reporte sugirió que el nuevo patógeno humano descubierto podía inducir una infección sintomática y asintomática. El nuevo retrovirus, asociado con el SIDA en los Estados Unidos, Europa y África central que tenía características morfológicas y genéticas típicas del género de los lentivirus fue nombrado Virus de la Inmunodeficiencia Humana o VIH (y subsecuentemente llamado VIH-1) [10].

En 1986 un relacionado, pero distinto retrovirus humano (ahora llamado VIH-2) fue recuperado de individuos que residían en algunos países del Oeste de África como Senegal, Costa de Marfil y Guinea – Bissau. Comparado con el VIH-1, el VIH-2 causa una lenta depleción en los niveles de linfocitos T CD4+, un periodo más largo de infección asintomática y un rango menor de mortalidad.

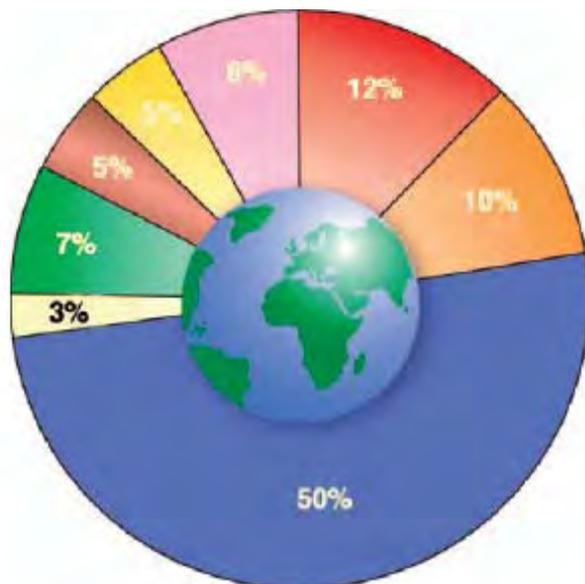
### **Clasificación**

La alta variabilidad genética y la rápida evolución del HIV-1 son factores importantes en su extensión alrededor del mundo. La heterogeneidad genética del HIV-1 es originada por: (a) la síntesis errónea del DNA por la transcriptasa inversa ( $3 \times 10^{-5}$  mutaciones /nucleótido/ ciclo de replicación), (b) la alta frecuencia de recombinación que acompaña a la transcriptasa inversa, (c) los altos niveles de producción de virus in vivo ( $10^9$  partículas/día; 150 a 300 ciclos de replicación/año), y (d) el número tan grande de individuos infectados [10]. Estos factores originan una alta diversidad genética en el virus dentro de un individuo infectado, el cual se caracteriza por presentar diversas poblaciones de cuasiespecies de virus que pueden diferir hasta un 10% [14].

La diversidad del VIH también se refleja a nivel poblacional. El análisis filogenético de aislados de VIH de diversas partes del mundo permite la clasificación del HIV-1 dentro de tres grupos genéticos: M (major), O (outlier) y N (non-M, non-O) [10, 11, 14]. La mayoría de las infecciones a nivel global es causada por los virus del grupo M. Los grupos O y N han sido aislados de individuos que viven en Cameroon, Gabon y Guinea Ecuatorial. Dentro del grupo M, se han organizado nueve subgrupos designados con las letras A – D, F – H, J y K [14].

El análisis de múltiples regiones genómicas del virus ha revelado recombinaciones inter-subtipo, las cuales se producen a través de los individuos con múltiples infecciones. Existen alrededor de 40 CRFs (formas recombinantes circulares. Algunas de ellas como CRF01\_AE y CRF02\_AG, circulan de manera extensa en regiones específicas del mundo (Figura 1.3) [14].

El VIH-1 subgrupo C es el predominante en el mundo por ser el más frecuente en África Subsahariana y en la India, mientras que el VIH-1 del subgrupo B es el predominante en América [14], incluyendo a México.



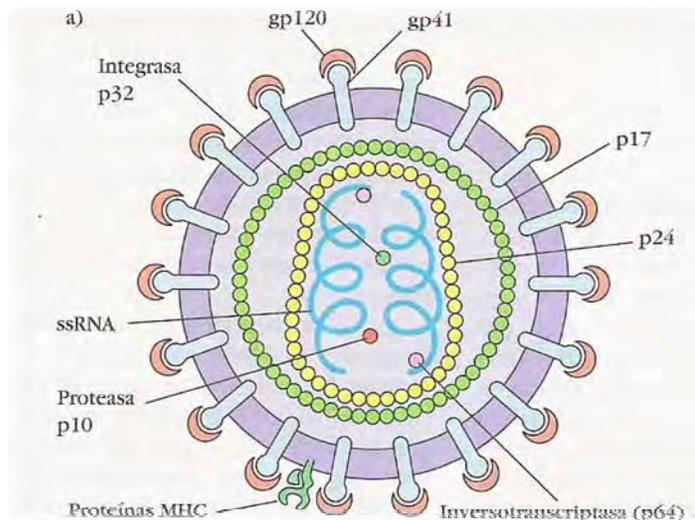
Se muestra la proporción de infecciones causadas por los diferentes subtipos y formas recombinantes del VIH-1 en el total de individuos infectados. Se observa que el 50% de las infecciones por VIH-1 son por el subtipo C. Los colores representan diferentes subtipos y formas recombinantes de VIH-1. ■ A; ■ B; ■ C; ■ D; ■ F,G,H,J,K; ■ CRF01\_AE; ■ CRF02\_AG; ■ CRF03\_AB; ■ Otras recombinantes. Tomado de [14].

**Figura 1.3 Prevalencia global de los subtipos y formas recombinantes de HIV-1 en 2004.**

## Estructura y organización del genoma del VIH

El VIH pertenece al género de los lentivirus, dentro de la familia de los retrovirus. Como otros retrovirus prototípicos, el VIH es un virus envuelto con un genoma de RNA de cadena sencilla positiva que tiene la capacidad de transcribir su genoma a DNA en la célula hospedera mediante la enzima llamada transcriptasa inversa [3].

El VIH es aproximadamente esférico, de un diámetro comprendido entre 80 y 120 nm (Figura 1.4). La envoltura contiene glicoproteínas virales y se adquiere por gemación a través de la membrana plasmática. La envoltura rodea una cápside que contiene dos copias



idénticas del genoma de RNA dentro de un centro viral denso a los electrones. El

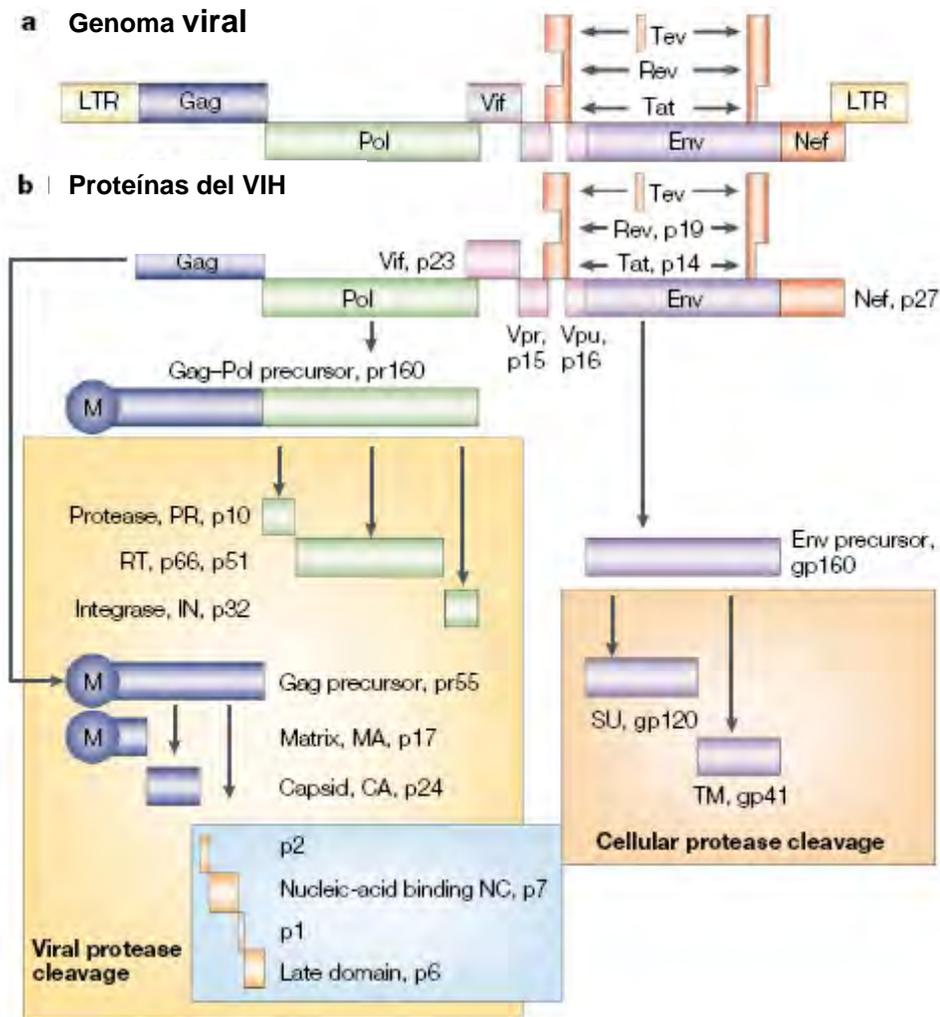
virión también contiene entre 10 y 50 copias de las enzimas transcriptasa inversa e integrasa y dos RNA celulares de transferencia (RNAt).

**Figura 1.4 Corte transversal del virión del VIH.** Los viriones con envoltura contienen dos cadenas de RNA idénticas, las cuales se relacionan con moléculas de transcriptasa inversa (p64) y proteínas de nucleótido p10, una proteasa y p32, una integrasa. Se observan las proteínas de membrana como gp120 unida a gp41 además de proteínas de la célula hospedera como el MHC [2].

Estos RNAt emparejan sus bases con cada copia del genoma para ser usados como cebadores por la transcriptasa inversa. La morfología de la cápside viral es semejante a un cono truncado [15].

Cada virión expresa un promedio de 72 espículas compuestas por trímeros de las glicoproteínas virales gp120 y gp41 (Figura 1.4). La gp41 es una molécula transmembranal que cruza la bicapa lipídica de la cubierta viral. La molécula gp120 se relaciona directamente con la molécula gp41 y funciona como receptor para la molécula CD4 de la célula hospedadora. La envoltura viral, al derivarse de la célula hospedadora, contiene algunas proteínas de la membrana de ésta, entre las que destacan moléculas HLA de clase I y II.

El VIH comparte tres genes característicos que codifican para proteínas estructurales y enzimáticas requeridas para producir infecciones productivas en todos los retrovirus: *gag* (antígeno grupo-específico), *pol* (polimerasa) y *env* (envoltura) [2]. Su genoma es de 10kb de largo y codifica para 16 proteínas distintas (Figura 1.5).



**Figura 1.5 Organización genómica y proteínas del VIH. Modificada de: [4].** a. El genoma viral de 10 kb, tiene una cabeza en el extremo 5' y una cola poliadenilada en el extremo 3'. En cada extremo del genoma existen unas extensas secuencias de repeticiones terminales (LTR). Las secuencias LTR contienen secuencias promotoras, potenciadoras y otras secuencias genéticas utilizadas para generar distintos factores de transcripción celular [15]. b. Los precursores de antígeno de grupo-específico (Gag) y la poliproteína Gag-Pol (Polimerasa) son procesados por la proteasa viral en nueve subunidades: proteasa (PR), transcriptasa inversa (RT), la cual contiene RNAasa H, integrasa (IN), matriz (MA), cápside (CA), p2, nucleocápside (NC), p1 y p6. Env es cortada por proteasas celulares en superficie (SU) gp120 y transmembranal (TM) y gp41[4].

La cápside y la matriz proteica del virus están conformadas por las proteínas codificadas por gag (las proteínas virales por convención se nombran con un número que corresponde a su tamaño): p24 (CA), que constituye la cápside viral; p17 (MA), que forma la matriz proteica asociada a la cara interna de la envoltura viral; p7 (NC), que constituye la nucleocápside asociada al RNA viral y p6, que es importante para la liberación de los viriones nuevos al interactuar con proteínas asociadas al tránsito de vesículas en la célula hospedera [3].

Las enzimas codificadas por *pol* se encuentran también formando parte de los viriones. Las enzimas virales se producen por autocatálisis del precursor Gag-Pol (Pr 160<sup>Gag-Pol</sup>), resultando en la producción de la proteasa (PR, p10), la transcriptasa inversa (RT, p66/p51), y la integrasa (IN, p32) [3].

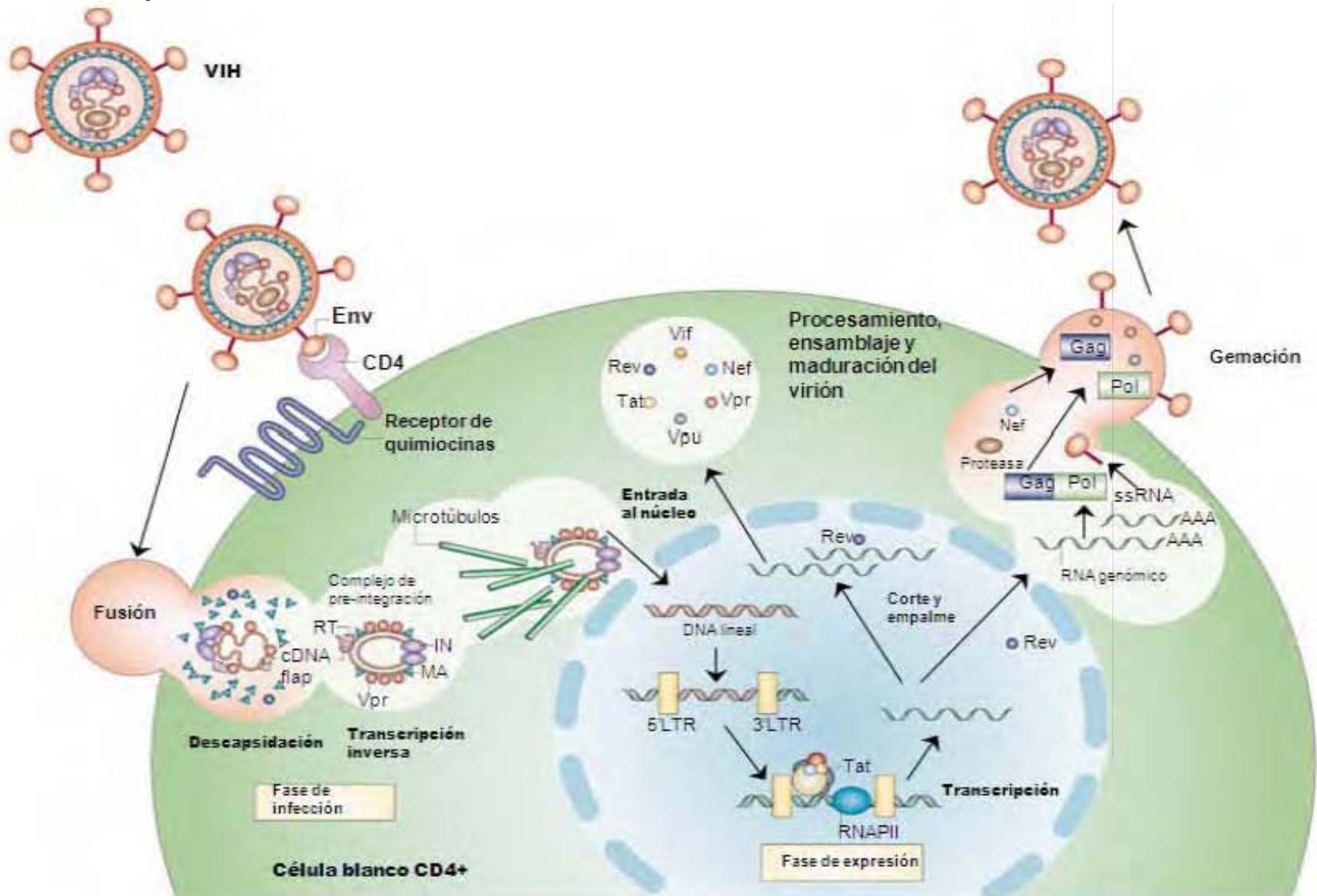
Adicionalmente, el VIH codifica dos proteínas reguladoras, el transactivador transcripcional (Tat) y el regulador de la expresión de genes del virión (Rev). Finalmente el virus contiene cuatro genes que codifican para las proteínas accesorias: el 'efector negativo' (Nef), el factor de infectividad viral (Vif), y las proteínas virales r (Vpr) y u (Vpu). Las proteínas virales son sintetizadas a partir de más de 30 especies de RNA mensajeros, los cuales derivan de la transcripción primaria. Mientras que los transcritos tempranos, que codifican para las proteínas Tat, Rev y Nef, sufren varios eventos de corte y empalme, los transcritos tardíos, que codifican para las proteínas estructurales y componentes enzimáticos del virión, sufren pocos o nulos eventos de corte y empalme [4].

El VIH se une a la célula blanco del hospedero a través de la molécula CD4, la cual funciona como receptor principal del virus. Así, el VIH puede infectar a linfocitos T ayudadores CD4+, células CD4+ del linaje de los macrófagos y a algunas poblaciones de células dendríticas. El CD4 se une a la glicoproteína de superficie viral gp120, lo que induce cambios conformacionales en la proteína viral que exponen los dominios de unión a un grupo de receptores de quimiocinas, los cuales funcionan como correceptores del virus (Figura 1.6). Los receptores de quimiocinas más usados son CCR5 y CXCR4. Según el uso de estos correceptores los virus se clasifican en: R5 o macrófagos-trópicos que usan el

receptor CCR5; X4 o linfotrópicos que utilizan el receptor CXCR4; y R5/X4 que pueden usar ambos receptores [3].

La unión del correceptor causa otro cambio conformacional que expone el dominio de fusión en gp41, a través del cual se unen las membranas celular y viral, permitiendo la entrada del virus a la célula. Una vez dentro de la célula se descapsida, llevándose a cabo la transcripción inversa del genoma del virus. Al ocurrir la transcripción inversa del genoma y formarse el ala central de DNA, una triple cadena de DNA que se forma por la interacción en cis de dos secuencias lentivirales: el tracto central de polipurina (cPPT) y la secuencia central de terminación (CTS), se forma el complejo de preintegración (PIC), incluyendo la proteína de matriz (MA), la transcriptasa inversa o retrotranscriptasa (RT), la integrasa (IN) y Vpr [3, 4]. El PIC viaja al núcleo en un principio por microtúbulos y posteriormente por filamentos de actina en las proximidades del núcleo. Gracias a las señales de localización nuclear en MA, IN y Vpr, el PIC es transportado al interior del núcleo. Una vez en el núcleo, el DNA viral puede integrarse en la cromatina de la célula por la actividad de la integrasa. Alternativamente, el DNA viral puede permanecer en forma episomal como círculos de dos LTR, círculos de un LTR y formas circulares reorganizadas. Estas formas episomales no generan interacciones productivas [3].

## Ciclo de replicación del VIH



**Figura 1.6 El ciclo de replicación del VIH.** La proteína de la envoltura (Env) gp120 se une al receptor CD4+ en la célula hospedera, esta unión causa un cambio conformacional en gp41 y se exponen los dominios de unión al correceptor (receptor de quimiocinas, CCR5 o CXCR4). La envoltura del virus se fusiona con la membrana plasmática, lo que permite la introducción de la cápside y posteriormente la desencapsidación, Esta favorece la formación del complejo de pre-integración (PIC). El PIC viaja hacia el núcleo, una vez en el poro nuclear entra al núcleo por las señales de localización nuclear en MA, IN y Vpr. El DNA se integra a la cromatina de la célula, Tat activa la transcripción de genes virales. Los primeros transcritos sufren múltiples eventos de corte y empalme produciéndose así las proteínas accesorias/reguladoras como Rev, Vif, Vpr y Vpu, En la transcripción tardía, los transcritos que no sufren corte y empalme, gracias a que Rev los transporta fuera del núcleo, se usan para producir los precursores de las proteínas estructurales como Gag, Env y Pol. Antes de la salida del nuevo virión se agrupan los precursores Gag-Pol, la proteasa y el RNA genómico dentro de la precapside. Estos nuevos viriones deben madurar para ser infecciosos. La proteasa viral lleva a cabo este proceso cortando las poliproteínas Gag y Gag-Pol para general las proteínas maduras [3]. Imagen modificada de:[4].

Una vez integrado, el genoma viral es transcrito, produciéndose RNA con múltiples eventos de corte y empalme (splicing) en etapas tempranas. Estos RNA tempranos originan a las proteínas reguladoras y accesorias Tat, Rev, Vif, Vpr y Vpu. Tat funciona como un transactivador de la transcripción de los genes virales, provocando que la RNA polimerasa II (RNAPII) produzca mensajeros largos. En etapas tardías de la transcripción se forman RNA con uno o sin ningún evento de corte y empalme, los cuales originan a las proteínas estructurales, Env, Pol y Gag. Los RNA tardíos son transportados al citoplasma por Rev, evitando así los eventos de corte y empalme. Los precursores de las proteínas estructurales y los genomas virales se ensamblan en balsas lipídicas en membranas de endosomas tardíos o en la membrana plasmática [3], MA se dirige a las membranas por sus características de miristilación. Para la salida de nuevos viriones de la célula, el virus usurpa a los sistemas ESCRT, ordenadores del tránsito intracelular de vesículas [3]. Gracias a interacciones con proteínas celulares ESCRT y a la contribución de la proteína viral Vpu, los nuevos viriones pueden gemar hacia el exterior a partir de la membrana plasmática o gemar hacia el lumen de cuerpos multivesiculares (MVB), para después ser liberados por fusión de estas vesículas con la membrana plasmática. Los nuevos viriones deben madurar para ser infecciosos. La proteasa viral lleva a cabo este proceso cortando la poliproteínas Gag y Gag-Pol para genera proteínas maduras [3].

## **VÍAS DE TRANSMISIÓN**

La infección por VIH puede transmitirse mediante:

- ✓ Las relaciones sexuales sin protección con una persona infectada. Esta es la vía de transmisión más importante de infección por VIH en el mundo [16]. La presencia de otras enfermedades de transmisión sexual incrementan notablemente el riesgo de infección por VIH. Es importante resaltar que no siempre hay correlación entre la concentración de la carga viral en la sangre con la de otros fluidos corporales. Por tanto, el riesgo individual aún es difícil de calcular. Además, los pacientes infectados por VIH no están protegidos de la superinfección con nuevas cepas virales. Sin embargo, en general se ha observado que a menor carga viral, el paciente es menos infeccioso y a mayor carga viral el paciente es más infeccioso. Esto es particularmente cierto para los pacientes durante la infección aguda por VIH [16].
- ✓ La inyección o transfusión de sangre contaminada o de productos de la sangre (la infección mediante inseminación artificial, los trasplantes de piel y los trasplantes de órganos también es posible). En la mayoría de los países occidentales la administración de la transfusión de sangre contaminada o de productos de la sangre contaminada con VIH se ha convertido en un evento raro. Con los métodos actuales de prueba, el riesgo de adquirir el VIH es aproximadamente 1:1 000 000.
- ✓ Compartir el equipo de infección no esterilizado que se utilizó previamente por alguien infectado (uso de drogas intravenosas). Esta forma de transmisión es importante en muchos países con alta prevalencia de usuarios de drogas intravenosas.
- ✓ La transmisión maternofetal (durante el embarazo, en el nacimiento y durante la lactancia). En ausencia de intervención, aproximadamente el 15-30% de las madres habrán de transmitir la infección durante el embarazo y el parto. En aproximadamente el 75% de los casos, el VIH se transmite

durante el último período de embarazo o durante el parto. Aproximadamente el 10% de las infecciones verticales por HIV tienen lugar antes de el tercer trimestre y el 10 – 15% son causadas por la lactancia.

## **HISTORIA NATURAL DE LA INFECCIÓN POR VIH**

El curso de la infección por VIH puede dividirse en tres etapas: infección primaria (o aguda), infección crónica (o asintomática) y avance de la enfermedad (SIDA). La duración de cada etapa es variable y puede ser alterada por terapia antirretroviral [10]. El avance de la inmunodeficiencia causada por la progresiva pérdida de los linfocitos T CD4+ es la manifestación clínica más marcada de la infección por el VIH-1. La aparición de las infecciones oportunistas y las características malignas del SIDA son consecuencias de los profundos defectos en la inmunidad celular. Adicionalmente, un número de síndromes clínicos son atribuidos directamente a la infección de los órganos específicos por el VIH-1.

La Organización Mundial de la Salud clasifica la infección de acuerdo a datos clínicos y de laboratorio de los pacientes (Tabla 1.1).

**Tabla 1.1 Sistema de clasificación para la infección por VIH [10]**

Categorías de las células CD4+	Categorías clínicas		
	A Asintomática, Linfadenopatía o Infección aguda	B Sintomática♣, No categorizada en A o C.	C Clínica SIDA♣
>500/ml (>29%)	A1	B1	C1
200 – 499 /ml (14% a 28%)	A2	B2	C2
<200 /ml (<14%)	A3	B3	C3

♣ Se refiere a la aparición de síntomas que incluyen candidosis vulvovaginal, displasia cervical o carcinoma in situ; síntomas como fiebre o diarrea por más de un mes de duración, herpes zoster trombocitopenia púrpura, enfermedad inflamatoria pelviana.

♣ Candidosis de tracto respiratorio, cáncer cervical invasivo, coccidioidomicosis extrapulmonar, criptosporidiosis, infección por Citomegalovirus extralinfática, herpes simplex con ulcera mucocutánea por más de un mes de duración.

### **Infección Primaria**

La adquisición de la infección por VIH-1 es acompañada de síntomas no específicos de una enfermedad viral en aproximadamente 50% a 70% de las infecciones individuales [10]. Los síntomas, que comienzan 2 semanas después

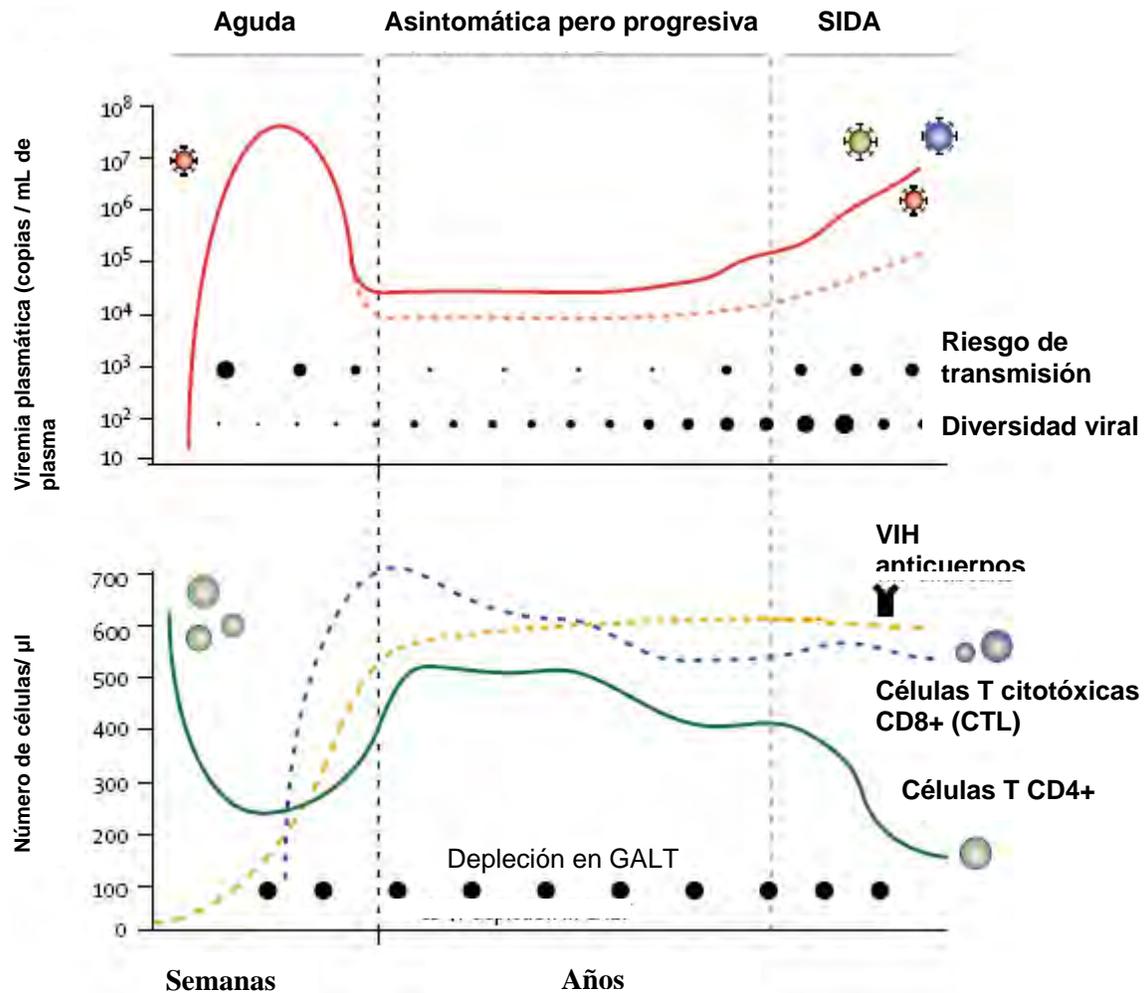
de la exposición, frecuentemente incluyen fiebre, faringitis, dolor de cabeza, artralgias, mialgias y pérdida de peso y normalmente se resuelven dentro de 3 a 5 semanas en la mayoría de los pacientes. Existe una linfadenopatía generalizada que es frecuentemente encontrada durante la etapa primaria de la enfermedad. Ulceraciones mucocutáneas y pérdida de peso ayudan a distinguir la infección primaria por VIH-1 de otras infecciones virales. La meningoencefalitis aséptica es la manifestación neurológica más común de la infección primaria del VIH-1. Algunos síntomas persisten más de 8 a 12 semanas, junto con una severa depresión de la cuenta de linfocitos T CD4+ y niveles altos de RNA del VIH-1 en plasma que pueden predecir la progresión rápida de la enfermedad [10].

La replicación inicial toma lugar en los órganos linfoides regionales (ejemplo: nodos linfáticos) y se caracteriza por la presencia de pocas variantes virales y una modesta amplificación primaria. Con la migración de los linfocitos T infectados o viriones dentro de la sangre periférica, la amplificación secundaria en el tracto gastrointestinal, el bazo y la médula ósea resultando una infección masiva de las células susceptibles, resultando en un pico de viremia (alrededor de  $10^6$  y  $10^7$  por ml de plasma) y la aparición de los síntomas de la infección primaria por el VIH-1 [11] (Figura 1.7).

Existe una pronunciada depleción de células T CD4+ de memoria localizadas en los tejidos linfoides asociados al intestino (GALT) que se ha identificado tempranamente después de la infección. Esta depleción de las células CD4+ en los tejidos linfoides de las mucosas permanece por años a pesar del uso del tratamiento antirretroviral [11, 17], a pesar de que el número de linfocitos T CD4+ en sangre periférica pueda reestablecerse parcialmente [11].

Las características de laboratorio de la infección primaria incluyen linfopenia y un decremento en la cuenta de linfocito T CD4+ absolutos, usualmente acompañado de un incremento en circulación de células T CD8+ activadas. Puede presentarse también trombocitosis pero es inusual. Se encuentra aumentada la aspartato transaminasa (AST) en suero y la fosfatasa alcalina puede estar presente. Después del pico de viremia y coincidiendo con la aparición de la respuesta citotóxica contra el virus, se observa un control parcial de la viremia la cual se

establece en un punto de equilibrio conocido como “set-point” viral (entre  $10^3$  a  $10^5$  copias/ml) dos meses después de la adquisición de la infección [10].



**Figura 1.2** Curso de la infección por VIH-1 definido por los niveles de replicación viral y los cambios en el número de células T CD4+. Se muestran los cambios típicos en la viremia plasmática (arriba) y los cambios dinámicos de los linfocitos T CD4+ (abajo). La infección primaria se caracteriza por una alta viremia y bajos linfocitos T CD4+ (línea verde), en ausencia de anticuerpos específicos para el VIH-1 (línea punteada amarilla). La caída de la viremia se desarrolla cuando aparecen los linfocitos T citotóxicos (CTLs) (línea punteada azul) y el establecimiento de un set-point de carga viral durante la infección crónica. El set-point difiere en cada individuo y predice la progresión de la enfermedad. La diversidad viral se incrementa a través del curso de la enfermedad. El riesgo de transmisión es muy alto dentro de la primera semana cuando existe el pico de la viremia. Se observa la depleción de los linfocitos T de memoria en tejidos linfoides asociados al intestino (GALT). Imagen modificada de [11].

## **Fase crónica**

Siguiendo la resolución de la infección primaria y el establecimiento de un estado cuasiestable virológico, se inicia la etapa de infección asintomática crónica. Aunque la mayoría de los pacientes permanecen asintomáticos durante esta fase, la replicación viral sigue en curso y la depleción de los linfocitos CD4+ marca el término de una latencia clínica aparente. A pesar de que existe una pérdida de las células T CD4+ en un rango medio de 30 a 60 células / $\mu$ l /año, el conteo de las células CD4+ puede permanecer estable por muchos años antes del periodo de rápido decremento que caracteriza a la etapa de SIDA. Una porción pequeña de pacientes (<1%) experimentan una progresión a SIDA dentro de 1 a 2 años [10].

Existen también algunos individuos, como una cohorte de individuos infectados por recibir una transfusión sanguínea entre los años 1981 y 1984, que han permanecido como asintomáticos por 27 años sin tratamiento antirretroviral, con cuentas de células T CD4+ estables y niveles de copias de RNA viral por/ml de sangre por debajo de niveles detectables. Estos individuos son conocidos como no progresores a largo plazo o controladores de la infección por VIH-1 [18]. La razón de este control sobresaliente de la replicación y patogénesis viral es multifactorial, ya que depende de características intrínsecas de la respuesta inmune del individuo, como el grado de activación del sistema inmune; factores genéticos, como la presencia de alelos HLA específicos o la presencia de polimorfismos en TLRs importantes en los mecanismos de la respuesta inmune innata; factores virales como la presencia de mutaciones que causan costos en la aptitud replicativa del virus o deleciones en el gen *nef* viral que disminuyen la patogenicidad viral, etc. [18].

La replicación del VIH-1 es dinámica a través de la enfermedad. La vida media de un virión es corta y la mitad de la población viral que se encuentra en plasma es reemplazada en menos de 30 minutos [11]. El total de los viriones producidos en las personas infectadas crónicamente puede llegar hasta  $10^{10}$  partículas por día [11].

### **Avance de la enfermedad, síndrome de la inmunodeficiencia humana (SIDA)**

Las infecciones oportunistas y malignas son raras en las personas infectadas por VIH con cuentas de linfocitos T CD4+ alrededor de 500 células/ $\mu$ l, pero se incrementan cuando la cuenta de linfocitos T CD4+ declina por debajo de esta marca. Candidosis oral, infección por pneumococco, tuberculosis, y varicela zoster llegan a ser comunes. El riesgo de tener complicaciones, incluyendo neumonía por *Pneumocystis jiroveci* (antes *P. carinii*), *Candida esophagitis*, histoplasmosis diseminada y otras infecciones fúngicas sistémicas, encefalitis toxoplásmica, y meningitis cryptococcal, se incrementa substancialmente una vez que los linfocitos T CD4+ caen a valores cercanos a 200 células/ $\mu$ l. Infecciones oportunistas, como una diseminación de una infección compleja de *Mycobacterium avium*; reactivación de infección por citomegalovirus (CMV), criptosporidiosis y microsporidiosis; además de una leucoencefalopatía progresiva multifocal causada por la reactivación del virus JC son indicativos de un defecto profundo en la inmunidad celular que usualmente ocurre con una cuenta de CD4 por debajo de 50 células/ $\mu$ l [10].

Las enfermedades malignas asociadas con SIDA generalmente están relacionadas con la progresión de la infección viral, estas incluyen sarcoma de Kaposi causada por una infección con el virus herpes tipo 8; linfomas asociados con infecciones virales por Epstein-Barr; y carcinoma cervical y anal asociado con la infección del virus del papiloma humano. A través de una terapia antirretroviral potente se ha disminuido el riesgo del sarcoma de Kaposi y el linfoma no Hodgkin [10].

### **ANTÍGENO LEUCOCITARIO HUMANO (HLA) Y RESPUESTA DE LOS LINFOCITOS T CITOTÓXICOS (CTLs)**

El antígeno leucocitario humano (HLA), también llamado complejo mayor de histocompatibilidad (MHC) es una región de genes muy polimórficos que se encuentran en el brazo corto del cromosoma 6 en los seres humanos [2, 8, 19], cuyos productos se expresan en la superficie de varias células. Los genes del HLA desempeñan un papel central en la respuesta inmunitaria frente a antígenos

proteicos, ya que los linfocitos T específicos para un antígeno no reconocen a estos antígenos en su forma libre ni en forma soluble, sino que reconocen porciones de los antígenos proteicos (es decir péptidos), unidos de forma no covalente a productos génicos del HLA [19].

Los genes del HLA se organizan en regiones que codifican para tres clases de moléculas [2]:

- Genes HLA de clase I codifican para glicoproteínas que se expresan en la superficie de casi todas las células nucleadas; la principal función de los productos génicos de clase I es la presentación de antígenos peptídicos a los linfocitos T citotóxicos (CTLs).
- Genes HLA de clase II: codifican para glicoproteínas que se expresan sobre todo en células presentadoras de antígeno (macrófagos, células dendríticas y células B), donde presentan péptidos antigénicos procesados a células T ayudadoras (Th).
- Genes HLA de clase III: codifican, entre otros productos genéticos para varias proteínas secretadas que desempeñan funciones inmunitarias, incluyendo componentes del sistema de complemento y moléculas relacionadas con la inflamación.

Las moléculas del HLA de clase I contienen una cadena  $\alpha$  vinculada de manera no covalente con una molécula de microglobulina  $\beta_2$ . La cadena  $\alpha$  es una glicoproteína transmembranal codificada por los genes polimórficos dentro de las regiones A, B y C del complejo HLA. El dominio  $\alpha$  forma una hendidura de unión al péptido, que es lo bastante grande para unir péptidos de ocho a 10 aminoácidos [2].

Un aspecto fundamental del reconocimiento del antígeno por las células Th y los CTLs es que los linfocito reconocen antígenos peptídicos sólo cuando éstos están unidos a una forma alélica de una molécula del HLA. A este fenómeno se le denomina restricción por HLA [19]. Una molécula HLA particular puede unir muchos péptidos diferentes. Como los alelos del HLA se expresan de manera codominante, un individuo heterocigoto expresa en sus células los productos génicos codificados por ambos alelos en cada locus del HLA; lo que significa que

expresa alelos A, B y C de cada padre (seis HLA de clase I diferentes) en la membrana de cada célula nucleada. La expresión de tantas moléculas HLA de clase I permite que cada célula tenga la capacidad de presentar un gran número de péptidos en las hendiduras de unión al péptido de sus moléculas HLA [2]. Cuando un individuo es homocigoto tiene una expresión limitada de péptidos y por lo tanto se reduce el repertorio de antígenos que pueden ser presentados a linfocitos T.

Existen dos vías de procesamiento y presentación de antígeno, la vía citosólica y la vía endocítica. Como regla general, los antígenos endógenos (los que se general dentro de la célula) se procesan en la vía citosólica y se presentan en el contexto de moléculas HLA de clase I; los antígenos exógenos (los que se captan por endocitosis) se procesan en la vía endocítica y se presentan en la membrana con moléculas HLA de clase II [2].

Para la vía citosólica o endógena, el complejo HLA de clase I se une al péptido que deriva de las proteínas virales, bacterianas y propias (que derivan del recambio normal de proteínas propias [2]). Esta vía comienza con el marcaje de las proteínas para la proteólisis con una proteína pequeña, llamada ubiquitina. Esta marca de las proteínas es reconocida por un sistema proteolítico presente en todas las células llamado proteasoma. El proteasoma rompe enlaces peptídicos en un proceso dependiente de ATP, formando así péptidos pequeños para la presentación con moléculas HLA clase I [2, 19]. Estos péptidos son transportados al retículo endoplásmico rugoso por una proteína llamada TAP (transportador relacionado con procesamiento de antígenos), para su unión con el HLA de clase I. Una vez unido el péptido, la molécula de clase I presenta mayor estabilidad y puede salir del retículo endoplásmico y avanzar a la superficie celular vía el complejo de Golgi [2, 19].

La expresión de las moléculas HLA de clase I en las células nucleadas y la presencia de la vía de carga de péptidos en dichas moléculas, a través de un mecanismo ligado a la ubiquitina para la degradación de proteínas celulares, garantiza que los péptidos de prácticamente cualquier proteína intracelular puedan ser presentados para ser reconocidos por los CTLs.

De esta manera los CTLs pueden lisar cualquier célula nucleada que produzca un antígeno extraño. Así, los CTLs específicos reconocen células infectadas, mediante la unión de sus receptores de células T (TCR) a los complejos HLA – péptido expresados en la superficie [2]. Esto permite la activación de los CTLs, con la consecuente liberación del contenido de sus gránulos citotóxicos que incluye citotoxinas como perforina, que forman poros en las membranas celulares y serín proteasas llamadas granzimas, que inducen la apoptosis de la célula infectada. De esta manera, la respuesta de CTLs está definida y restringida por los genes HLA de clase I.

### ***VARIABILIDAD DEL HLA EN EL MUNDO***

Como se ha mencionado arriba, la región del HLA en los humanos que se localiza en el cromosoma 6, es muy polimórfica. De los tres loci del HLA (HLA-A, HLA-B y HLA-C), el HLA-B es el más polimórfico. Se han descrito 1249 alelos diferentes para HLA-B, comparados con los 853 alelos para HLA-A y los 463 alelos distintos HLA-C (ver IMG/HLA database [20]). Además de que el HLA-B es el locus más polimórfico en el cromosoma 6 es el más polimórfico dentro del genoma humano. Este alto polimorfismo de los genes HLA trae como consecuencia que los diferentes alelos HLA se distribuyan con frecuencias altamente variables en las diferentes poblaciones del mundo [8].

Los números mencionados antes para las formas alélicas de HLA-A, B y C humanas permite calcular la cifra teórica de combinaciones al multiplicar  $1246 \times 853 \times 463$ , con lo que se obtiene más de 492 millones de diferentes haplotipos clase I posibles en la población [2]. Se sabe que la diversidad real es menor porque ciertas combinaciones alélicas ocurren con mayor frecuencia en los haplotipos HLA que lo predicho por la combinación aleatoria, un estado que se denomina desequilibrio de ligamiento o de enlace. Brevemente, el desequilibrio de ligamiento es la diferencia entre la frecuencia observada para una combinación particular de los alelos y la esperada con base en las frecuencias de los dos alelos si estos se distribuyeran al azar. Por ejemplo si HLA-A1 ocurre en 16% de los individuos en una población (frecuencia = 0.16) y HLA-B8 en 9% de ese grupo

(frecuencia = 0.09), cabe esperar que alrededor de 1.4% del grupo tenga ambos alelos ( $0.16 \times 0.09 = 0.014$ ). Sin embargo, los datos muestran que HLA-A1 y HLA-B8 se encuentran juntos en 8.8% de los individuos de la población estudiada. Esta diferencia es una medida del desequilibrio de ligamiento entre estos alelos de genes HLA clase I [2].

Cabe mencionar que aún considerando el fenómeno de desequilibrio de ligamiento, se mantiene un enorme polimorfismo en los haplotipos HLA humanos. Este polimorfismo ha sido generado por recombinación, mutación puntual y conversión génica, todo lo cual contribuye a la diversidad de genes HLA dentro de la población [2].

El extenso polimorfismo del HLA clase I se limita a los residuos que se encuentran en el sitio de unión al péptido de estas moléculas [8, 21]. De esta forma se definen los péptidos que pueden unirse a cada molécula HLA. El significado evolutivo del notable polimorfismo de las moléculas HLA de clase I radica en que pequeñas diferencias entre las múltiples moléculas de HLA y los péptidos que estas presentan conllevan a diferencias funcionales importantes en las respuestas inmunes restringidas por diferentes moléculas HLA hacia patógenos específicos. Estas diferencias funcionales son blanco de selección natural Darwiniana, es decir, definen a individuos dentro de una población con mayor o menor susceptibilidad al avance de la enfermedad [8].

## ***HLA Y VIH***

Como se ha mencionado arriba, los genes HLA de clase I son un componente fundamental del sistema inmune, ya que definen y restringen respuestas necesarias para controlar infecciones intracelulares por medio de la presentación de péptidos provenientes de organismos patógenos a CTLs. La unión del péptido a la molécula HLA es crítica, ya que el HLA actúa como un filtro que restringe los epítopes específicos dentro de un patógeno que serán reconocidos por el sistema inmune, algunos de los cuales son más eficaces que otros para controlar la replicación viral, jugando un papel importante en la velocidad de avance de la enfermedad por VIH [22].

El papel de los CTLs es eliminar las células infectadas por el virus a través del reconocimiento de péptidos lineales y cortos procesados intracelularmente y presentados en la superficie celular por moléculas HLA de clase I. De esta forma todas las proteínas virales tienen la capacidad de unirse y ser presentadas por moléculas HLA de clase I. Las respuestas de CTLs restringidas por los genes HLA de clase I seleccionan mutaciones de escape inmune. Este hecho, aunado a la observación de que los CTLs probablemente contribuyen más con el control inmune en la infección por el VIH-1 que los anticuerpos, destaca el papel de los CTLs como una importante fuerza de selección que conduce la evolución del virus en términos de su genoma [23].

Existe una correlación entre la aparición de las respuestas de CTLs específicas contra VIH *in vivo* y la disminución de la viremia en la fase aguda [8, 23]. Aunado a esto, se ha observado una fuerte correlación epidemiológica entre la presencia de alelos específicos HLA de clase I y diferentes velocidades de progresión de la enfermedad por VIH-1, lo cual sugiere que la calidad de las respuestas de CTLs y/o las características de los epítopes blanco influyen fuertemente en la efectividad del control viral [23].

Se sabe que la mayoría de los epítopes que están relacionados con valores bajos de carga viral, son aquellos específicos del gen Gag [23]. La razón es porque las proteínas de Gag son altamente inmunogénicas y sus secuencias son altamente conservadas, por lo que una mutación de escape dentro de ellas implica frecuentemente una reducción en la capacidad de replicación viral. De esta manera, muchos estudios de control inmune se han realizado siguiendo la aparición de mutaciones de escape tempranas en Gag [23]. Adicionalmente, el número de polimorfismos de aminoácidos específicos de Gag asociados con cada alelo HLA-B tiene una fuerte correlación con la media de carga viral de los individuos que presentan estos alelos y con eso en la progresión de la infección por VIH.

Por lo tanto, no sólo el número de epítopes de Gag que son presentados por cada alelo es importante sino también la presencia de respuestas de CTLs específicos de Gag que conducen a la selección de mutantes de escape que tienen un

impacto en la capacidad replicativa viral. Cabe destacar que esta relación aplica predominantemente a los alelos HLA-B y las repuestas específicas de Gag [8].

Es evidente que en la infección por VIH, existe un mayor impacto de las moléculas HLA-B en el establecimiento del set-point viral, en comparación con las moléculas HLA-A y HLA-C. Una explicación a esto es que debido a que el HLA-B es el locus más polimórfico, tiene la capacidad de presentar una variedad más amplia de péptidos [8].

El cambio de un aminoácido, aparentemente insignificante, entre alelos HLA cercanos puede ser crucial para la resolución de una infección. En el caso de la infección por VIH, un ejemplo muy claro es la asociación del HLA-B\*3502 y HLA-B\*3503 con una progresión rápida de la enfermedad, lo que no sucede con el alelo HLA-B\*3501, el cual difiere de HLA-B\*3502 y HLA-B\*3503 por solo tres y un aminoácido respectivamente [8].

En el caso del gen HLA-B\*35, los pequeños cambios observados en los aminoácidos que conforman la hendidura de unión al péptido de los diferentes alelos pueden afectar la unión del aminoácido carboxilo-terminal del péptido antigénico: el residuo preferencial para la unión del péptido en la hendidura del HLA-B\*3501 es tirosina, mientras que un residuo hidrofóbico es preferido para la unión de péptidos a la hendidura del HLA-B\*3502 y HLA-B\*3503. Así, los péptidos capaces de unirse a la hendidura de la molécula HLA-B\*3501 que incluyen epítopes de Gag p24 (residuos de Gag 254 – 262), no son reconocidos por los alelos HLA-B\*3502 y HLA-B\*3503. Estas diferencias traen como consecuencia una ausencia de respuestas de CTLs específicas para Gag en pacientes con los alelos HLA-B\*3502 y HLA-B\*3503, las cuales son cruciales para el control viral [8]. Otro ejemplo más extremo, es la asociación del HLA-B\*5801 y valores de carga viral bajos, mientras que el alelo HLA-B\*5802, el cual difiere del HLA-B\*5801 por solo tres aminoácidos, es asociado con valores de carga viral altos. Estas pequeñas diferencias en las moléculas del HLA de clase I, se trasladan a una diferencia substancial en la unión del péptido específico [8].

Adicionalmente, existen ejemplos de alelos HLA relacionados muy cercanamente por el tipo de péptidos que pueden unir y presentar. Este es el caso de los alelos

del supertipo HLA-B7 (HLA-B\*0702, HLA-B\*3910, HLA-B\*4201, HLA-B\*4202 y HLA-B\*8101), los cuales presentan péptidos idénticos en el contexto de moléculas HLA diferentes [8, 21].

Con la diversidad de HLA en diferentes poblaciones y la correlación encontrada entre alelos HLA y polimorfismos del VIH, se han identificado alelos HLA de clase I que tiene un impacto significativo en la progresión de la enfermedad o en el set-point viral (el cual es fuertemente predictivo del grado de progresión de la enfermedad). Estos HLA relacionados con la progresión de la enfermedad se muestran en la Tabla 1.2.

Un ejemplo de esto es el alelo B\*57 o B\*5801 asociados al epítipo en Gag TSTLQEQIAW (el cual se abrevia como TW10 y corresponde al aminoácido 240 – 249), el cual es el principal blanco de la respuesta de CTLs en la infección con VIH-1 en sujetos con dichos alelos y está asociado al control de la infección.

**Tabla 1.2 Fuerte asociación de HLA-1 de clase I con particular desenlace de la infección por VIH-1. Modificada de [8].**

<b>Alelo HLA de clase I</b>	<b>de Progresión</b>	<b>Set-point viral</b>	<b>Epítipo Implicado</b>	<b>Proteína asociada del VIH</b>
<b>HLA-B*1302</b>	Lenta	Bajo	RQANFLGKI	Gag 429 – 437
<b>HLA-B*1801</b>	Rápida	Alto	-	-
<b>HLA-B*2705</b>	Lenta	Bajo	KRWIILGLNK	Gag 263 – 272
<b>HLA-B*3501</b>	Lenta	Bajo	PPIPVGGEIY	Gag 254 – 262
<b>HLA-B*3502 y HLA-B*3503</b>	Rápida	Alto	-	-
<b>HLA-B*5101</b>	Lenta	Bajo	TAFTIPSI	RT 128 – 135
<b>HLA-B*5701, HLA-B*5702 y HLA-B*5703</b>	Lenta	Bajo	TSTLQEQIAW KSGDPRBIPMF ISPRTLNAW IVLPEKDSW	Gag 240 – 249 Gag 162 -172 Gag 147 -155 RT 235 – 242
<b>HLA-B*5801</b>	Lenta	Bajo	TSTLQEQIAW ISPRTLNAW	Gag 240 – 249 Gag 147 – 155
<b>HLA-B*5802</b>	Rápida	Alto	QTRVLAIERYL	gp41 577 – 587
<b>HLA-B*8101</b>	Lenta	Bajo	-	-

## ***IMPACTO DEL HLA EN LA EVOLUCIÓN DEL VIH***

### **Nivel intra-paciente.**

La extraordinaria capacidad de mutación del VIH-1 constituye una de las principales fuentes de variación del VIH, introduciendo la retrotranscriptasa en promedio una mutación por ciclo de replicación [23]. Adicionalmente, la recombinación es un mecanismo adicional para generar variabilidad en el virus, con 2 a 3 eventos de recombinación por cada ciclo de replicación [23]. Dentro de un individuo infectado, existe una expansión progresiva de la diversidad viral en el transcurso de la enfermedad, con múltiples variantes co-existiendo como una población heterogénea o “cuasiespecies” que son únicas en cada individuo [23].

En el curso natural de la infección por VIH, la respuesta inmune del hospedero actúa como una presión de selección, que dirige la evolución del VIH-1 en un proceso dinámico y continuo conocido como “escape inmune” [23]. La respuesta de los CTLs restringida por los genes HLA, es el principal mediador del control antiviral del hospedero durante la infección por el VIH-1 [23] y su aparición está relacionada con la disminución de la viremia en la fase aguda [8, 23]. La respuesta específica de CTLs puede ser evadida por el virus, mediante la selección de mutaciones en sitios cruciales para la unión de los péptidos virales a la molécula HLA de clase I, al encontrarse dentro de los epítopes banco de los CTLs [24, 25]. La sustitución de un aminoácido puede eliminar o reducir la unión del péptido a la molécula del HLA, interferir con el procesamiento de antígenos virales por el proteasoma o reducir el reconocimiento por los CTLs específicos, lo que finalmente provoca una pérdida del control inmune del virus y la progresión de la enfermedad [24, 26]. Sin embargo, en algunos casos, el sistema inmune tiene la habilidad de generar respuestas específicas de CTLs para las nuevas variantes de escape en el contexto de la infección crónica por VIH-1 [24].

Se ha observado que un gran número de mutaciones de escape a CTLs son reproducibles en el contexto de una molécula HLA selectora específica. Esto tiene como consecuencia la importante observación de que la evolución del VIH sigue generalmente patrones reproducibles cuando se aplica una presión inmune

específica. Un ejemplo de ello es para las personas que expresan el alelo protector HLA-B\*57, el cual selecciona un cambio de T por N en la posición 3 del epítotope TW10 en la proteína p24 de Gag, dentro de las primeras semanas de infección. Para los individuos que expresan HLA-B\*27, la primera mutación que se selecciona es un cambio de L por M en la posición 6 del epítotope KK10 en Gag, seguida de un cambio de R por K en la posición 2 después de un año de seguimiento [23].

### **Nivel poblacional**

El hecho de que la respuesta de CTLs, definida por los genes HLA de clase I, dirige y restringe de manera predecible la evolución del VIH en un individuo infectado, trae la consecuencia inmediata de que la evolución poblacional del virus y su adaptación a diferentes poblaciones humanas podría también ser un fenómeno predecible [23].

Las mutaciones de escape pueden tener un costo sustancial en la capacidad replicativa del virus (*fitness*) por lo que al ser transmitidas a un individuo que carece del HLA selector revierten a la secuencia original o *wild-type* [8]. Sin embargo, se ha demostrado también la existencia de selección de mutaciones de escape que no tienen costos elevados para la aptitud replicativa viral y que no revierten después de la transmisión a un nuevo hospedero incluso en ausencia del alelo HLA selector [23, 27]. En el contexto de alelos HLA altamente frecuentes en una población, este tipo de mutaciones podrían permitir la fijación de una forma “escapada” de epítotope específicos de CTLs en las cepas circulantes en una población provocando la extinción de respuestas inmunes efectivas para la población en cuestión [23, 27]. Así, la selección por HLA está dejando una “huella” en la secuencia del virus al funcionar como una presión inmune que dirige la evolución del VIH-1 que se refleja en la adaptación viral a un alelo HLA comúnmente expresado en una población en particular [23, 28, 29].

Moore y colaboradores [25] fueron los primeros en identificar la existencia de esta huella del HLA sobre el VIH a nivel poblacional al identificar asociaciones entre la presencia de polimorfismos en la secuencia del virus y la expresión de alelos HLA específicos a nivel poblacional en un estudio con una cohorte de más de 400

pacientes mayoritariamente caucásicos incluyendo los codones 20 – 227 de la retrotranscriptasa [27].

Aunque se sabe que algunas variantes de escape a CTLs pueden ser transmitidas y generar una respuesta inmunodominante durante la infección aguda en un nuevo hospedero, el grado en el cual las mutaciones de escape dentro de los epítopes bajo presión selectiva, pueden acumularse en la población depende de múltiples factores incluyendo la frecuencia del alelo HLA selector en la población, el impacto de la mutación en el *fitness* viral y la estabilidad genética de la mutación [28].

## **2. PLANTEAMIENTO DEL PROBLEMA**

Debido al notable polimorfismo de los genes HLA de clase I, y que las distribuciones de frecuencias alélicas en diferentes poblaciones alrededor del mundo son altamente variables [8], pudiendo impactar de manera distinta la evolución del VIH a nivel poblacional. Esta variación en la evolución del virus podría estarse dando incluso en regiones distintas dentro del territorio nacional, ya que se conoce que las características inmunogenéticas de los individuos del norte y del sur de México son diferentes, presentándose un enriquecimiento en alelos de origen caucásico en el norte y de origen amerindio en el sur [7]. Recientemente, nuestro grupo de investigación demostró que la evolución del VIH mediada por HLA en una cohorte de individuos mexicanos presentaba características únicas al compararse con la evolución del virus en una cohorte de individuos mayoritariamente caucásicos [5]. Dada la diversidad genética de la población mexicana y la heterogeneidad en la distribución de alelos HLA incluso dentro del territorio nacional, existe la posibilidad de que el VIH presente características únicas de adaptación ligadas a la presión selectiva por HLA en diferentes regiones del país. El hecho de que la evolución del VIH presente características particulares en diferentes poblaciones, tendría implicaciones en el diseño de vacunas y tratamientos efectivos contra el virus con efectividad a nivel local. Para abordar la pregunta de si la evolución del VIH presenta características únicas en poblaciones genéticamente distintas de acuerdo a sus frecuencias de alelos HLA de clase I

incluso dentro del territorio nacional, en el presente trabajo se estudió la evolución dependiente de HLA del VIH a nivel poblacional en una cohorte de pacientes del norte y del sur de México. Los datos obtenidos se compararon con información previamente obtenida en el laboratorio en una cohorte de pacientes del centro del país [5], así como con observaciones de la evolución del virus en una cohorte internacional de individuos mayoritariamente caucásicos [30].

Se tipificaron a alta resolución los genes HLA de clase I A, B y C de una cohorte de pacientes VIH positivos, provenientes del norte y del sur del país y se determinaron las frecuencias de los alelos y haplotipos HLA más frecuentes en cada una de las regiones, comparándose las distribuciones de frecuencias de alelos HLA entre las dos regiones del país. Por otro lado, se construyó una base de secuencias del gen *pol* viral, previamente obtenidas en el laboratorio, para cada uno de los pacientes incluidos en el estudio. Se buscaron asociaciones correlativas entre la presencia de alelos HLA específicos y la presencia de polimorfismo en cada una de las posiciones de la secuencia de VIH estudiada mediante pruebas exactas de Fisher corregidas para comparaciones múltiples por métodos estadísticos basados en el concepto de razón de descubrimientos falsos [31]. Se buscaron también alelos HLA que se encuentren relacionados con el control de la viremia plasmática, los cuales puedan considerarse alelos HLA protectores o por el contrario aquellos que no tengan control con la viremia plasmática o alelos de riesgo en la población mexicana, esto se hizo relacionando los alelos HLA y la carga viral con pruebas de Man – Whitney con alta y baja resolución. Este estudio será de gran valor para profundizar en los mecanismos evolutivos del VIH, así como de su capacidad de adaptación a diferentes poblaciones humanas, teniendo implicaciones en el desarrollo de vacunas y otras estrategias de tratamiento efectivas para la población mexicana.

### **3. OBJETIVOS**

#### ***GENERALES***

Estudiar y comparar la evolución del VIH a nivel poblacional y la posible existencia de factores de riesgo o de protección hacia el avance de la enfermedad por el virus en tres cohortes de pacientes mexicanos con características inmunogenéticas diferentes.

#### ***PARTICULARES***

- Tipificar los genes HLA de clase I A, B y C de una cohorte de individuos VIH positivos provenientes de las regiones Norte, Sur y Centro de la República Mexicana.
- Obtener las frecuencias alélicas HLA de pacientes del Norte Sur y Centro de la República Mexicana.
- Generar una base de secuencias de gen *pol* viral para los pacientes incluidos en el estudio.
- Buscar asociaciones HLA – polimorfismos del VIH en las regiones Norte y Sur de México y compararlas con datos previamente obtenidos en el laboratorio de una cohorte de pacientes VIH positivos del Centro del país.
- Buscar asociaciones HLA – Carga Viral (CV) en las regiones Norte, Centro y Sur del país, para los alelos en alta y baja resolución.

#### **4. HIPÓTESIS**

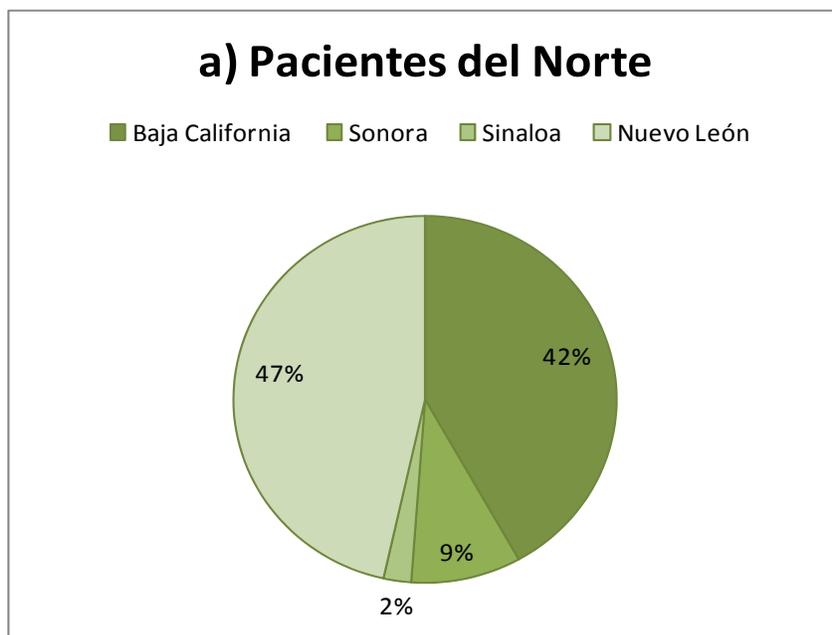
Debido a diferencias en las distribuciones de frecuencias alélicas HLA entre las poblaciones del norte, centro y sur de la República Mexicana, las características de adaptación del VIH a las respuestas restringidas por HLA a nivel poblacional serán diferentes para las tres regiones del país.

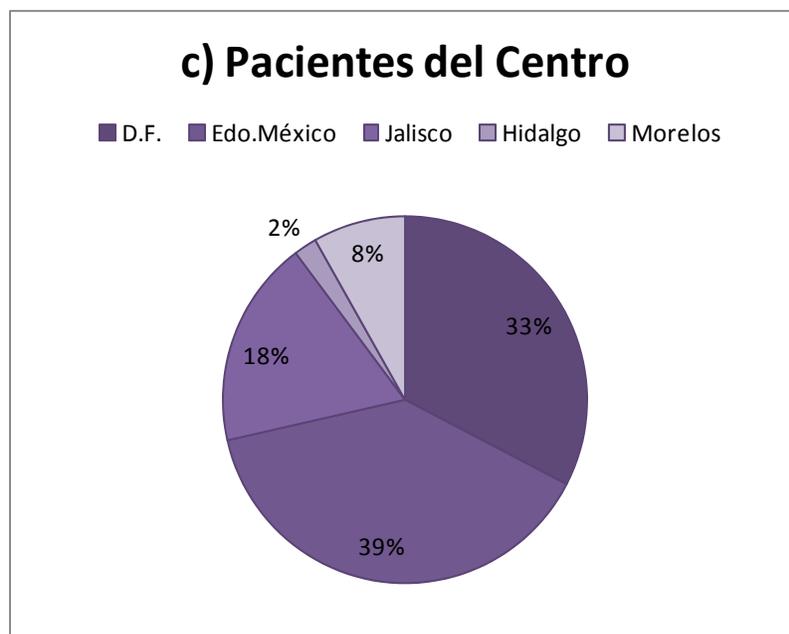
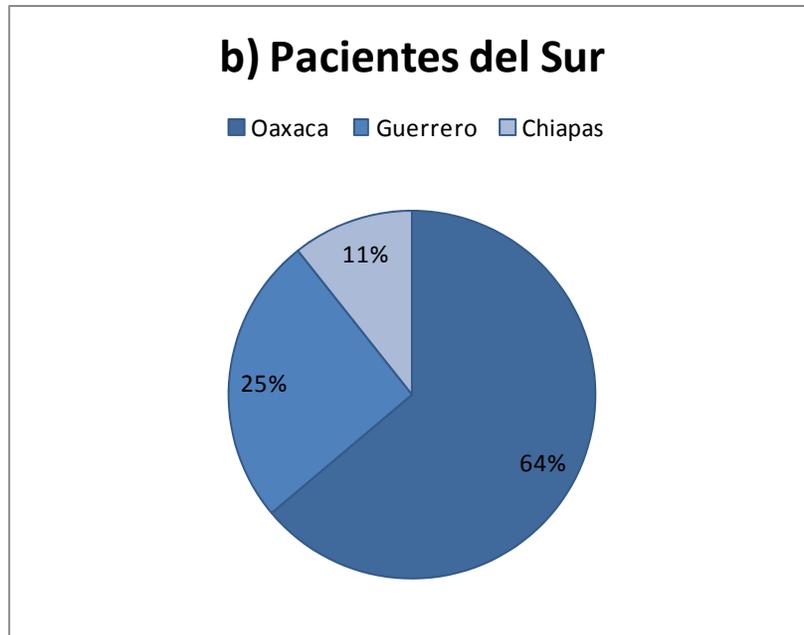
Debido a las distribuciones de frecuencias alélicas HLA únicas de la población mexicana, será posible encontrar asociaciones únicas entre la presencia de alelos HLA específicos y un mayor o menor riesgo de avance de la enfermedad por VIH en nuestro país.

## 5. METODOLOGÍA

### **PACIENTES**

Se seleccionaron tres grupos de individuos VIH positivos, sin tratamiento antirretroviral previo, en etapas tempranas de la infección ( $>500$  células T CD4 $^{+}$ / $\mu$ l), provenientes del Norte, Sur y Centro de la República Mexicana, de la cohorte de Epidemiología Molecular y Evolución del VIH en México del Centro de Investigación en Enfermedades Infecciosas del Instituto Nacional de Enfermedades Respiratorias. Se incluyeron un total de 43 individuos del Norte del país, siendo estos de los siguiente estados: Baja California (18, 42%), Sonora (4, 9%), Sinaloa (1, 2%) y Nuevo León (20, 47%) (Figura 5.1.a); 49 individuos de la zona Sur del país (Figura 5.1.b), provenientes de los siguientes estados: Oaxaca (30, 64%), Guerrero (12, 25%) y Chiapas (5, 11%); y 49 individuos del Centro del país (Figura 5.1.c), provenientes de los siguientes estados: Distrito Federal (16, 33%), Estado de México (19, 39%), Jalisco (9, 18%), Hidalgo (1, 2%) y Morelos (4, 8%). Para cada individuo se descongelaron células mononucleares de sangre periférica (PBMCs), se extrajo DNA genómico y se tipificaron los genes HLA de clase I a alta y/o baja resolución como se indica a continuación.





**Figura 5.1. Distribución de los pacientes en las diferentes zonas de la República Mexicana.** Con un total de 139 pacientes de las tres zonas de la república se muestra una distribución de: a) Individuos del Norte de la República Mexicana (n = 43) que provienen de estados como Baja California, Sonora, Sinaloa y Nuevo León; b) Individuos del Sur del país (n = 47), que provienen de estados como: Oaxaca, Guerrero y Chiapas y c) Individuos de la zona Centro del país (n = 49), provenientes de los estados: Distrito Federal, Estado de México, Jalisco, Hidalgo y Morelos.

### **SEPARACIÓN DE PMBCs**

Se obtuvieron PMBCs (células mononucleares de sangre periférica) de muestras sanguíneas de cada individuo participante, separándolas por gradiente de

densidad usando Ficoll. Brevemente, el paquete globular fue diluido 1:2 en solución salina amortiguada de Hanks (HBSS) y superpuesto sobre una alícuota de Ficoll. Después de obtener el gradiente de densidad por centrifugación a 2000 rpm por 20 min, la capa de PBMCs se aspiró y se lavó tres veces con HBSS. Las células fueron contadas en cámara de Neubauer y conservadas en medio de congelación (suero fetal bovino 90%, DMSO 10%) en nitrógeno líquido a -140°C.

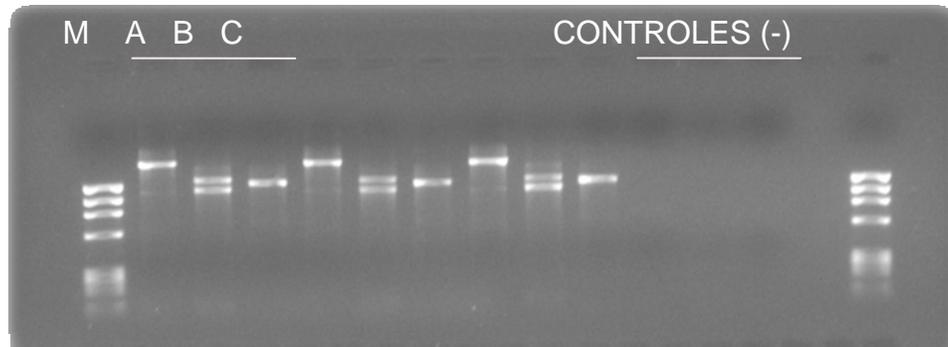
### ***EXTRACCIÓN DE DNA***

Se extrajo DNA genómico a partir de al menos 6 millones de PBMCs para cada individuo participante, utilizando el estuche de reactivos QIAmp DNA Blood Mini Kit (QIAGEN, Valencia, CA) de acuerdo a las instrucciones del fabricante. Brevemente, las células se lavaron con solución amortiguadora de fosfatos (PBS) para retirar el medio de congelación, y se lisaron por lisis alcalina para liberar el DNA genómico. El DNA se unió a una columna de sílice y se lavó dos veces con solución amortiguadora de fuerza iónica alta. El DNA purificado se eluyó posteriormente con solución amortiguadora de fuerza iónica baja. La concentración del DNA eluido se midió espectrofotométricamente con el equipo NanoDrop ND1000 UV-Vis (NanoDrop Technologies, Inc. Wilmington, DE). En el cual, la muestra de DNA sin diluir se coloca directamente en el pedestal y se mide a una longitud de onda de 260nm para conocer así la concentración del DNA. Posteriormente el DNA se almacenó a -20 °C hasta su uso.

### ***TIPIFICACIÓN DE HLA DE CLASE I A ALTA RESOLUCIÓN***

Para la tipificación de los alelos de HLA se usó el kit de alta resolución Allele SEQR- HLA Sequencing Kit (Atria Genetics, San Francisco, CA). Brevemente, se realizó una reacción en cadena de la polimerasa (PCR) con cebadores específicos para amplificar una sección de los exones 2, 3 y 4 de los genes HLA-A y B, y los exones 2 y 3 del gen HLA-C, utilizando DNA genómico de cada uno de los individuos participantes a una concentración de  $20 \text{ ng}/\mu\text{l}$ . Los productos de amplificación se corrieron en un gel de Agarosa al 1% con  $0.1\mu\text{g/ml}$  de bromuro de etidio para evaluar la calidad de la PCR, buscando un patrón de amplificación

similar al mostrado en la Figura 5.2. Los productos amplificados positivos aparecen como bandas aisladas.



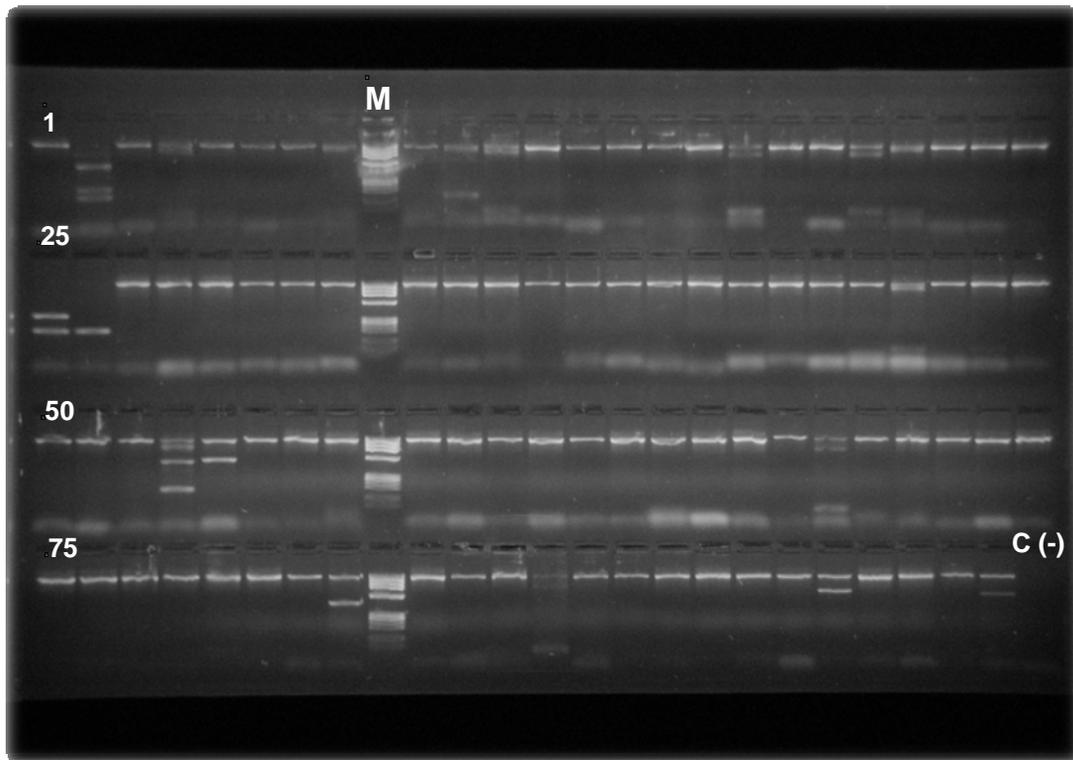
**Figura 5.2. Patrón de bandas para la PCR de amplificación de los exones 2, 3 y 4 de los genes HLA-A y HLA-B y los exones 2 y 3 del gen HLA-C para la tipificación HLA en alta resolución.** Se observa el patrón de bandas de una reacción típica de PCR para tipificación HLA a alta resolución. Por cada paciente se tienen tres PCRs (A, B y C). En el primer y último pozo se encuentra el marcador de peso molecular, después se observan los productos de las tres PCRs para cada paciente (gen A, B y C), observándose una banda para A, dos para B y una banda para C, se muestran también los controles negativos para cada uno de los tres genes.

Los productos de PCR se purificaron con ExoSAP-IT, una mezcla de exonucleasa I y fosfatasa alcalina, para eliminar los cebadores y dNTPs que no se incorporaron durante la PCR. El DNA tratado se utilizó como molde en las reacciones de secuenciación con cebadores específicos de los exones 2, 3 y 4 para los genes HLA-A y HLA-B y los exones 2 y 3 para el gen HLA-C, utilizando ddNTPs marcados con fluorocromos como terminadores, empleando la química de reactivos Big Dye Terminator v3.1 (Applied Biosystems, Foster City, CA). Los productos de PCR de secuenciación se purificaron por precipitación con 1% v/v de acetato de sodio (3 M, pH = 5.2) y 2.5 volúmenes de etanol absoluto. Se lavó con etanol al 80% y finalmente se resuspendió con formamida HiDi. Las secuencias de DNA se obtuvieron por electroforesis capilar en un secuenciador 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, CA).

La tipificación de los alelos HLA-A, HLA-B y HLA-C se hizo por comparación con una base de datos actualizada de HLA en el software Assign SBT (Conexio Genomics, Western Australia), provisto por el fabricante (Figura 5.3). Brevemente las secuencias se importan con el programa Assign SBT para cada uno de los



Figura 5.4. A las bandas que aparecen además del control positivo en cada pozo, se determina el peso molecular con ayuda del marcador que se encuentra en cada nivel de 24 pozos de muestra y se compara con los pesos moleculares propuestos por una hoja de trabajo que se incluye en el kit. Si existe una banda que coincida con el peso molecular propuesto, el PCR de dicho número de pozo se considera positivo. De acuerdo al patrón de amplificación y a los pozos que resultaron positivos, se asignó el alelo HLA correspondiente, utilizando el software UniMatch v3.2 provisto por el fabricante, utilizando bases de datos actualizadas para determinar el haplotipo HLA. Se incluyeron controles de amplificación internos para validar cada una de las reacciones así como un control negativo para detectar posibles contaminaciones por prueba.



**Figura 5.4. Patrón de bandas observado para tipificación de HLA en baja resolución.** Se muestra un gel de agarosa al 2% de una tipificación HLA a baja resolución representativa. Cada pozo corresponde a una PCR con iniciadores secuencia específicos para determinar polimorfismos útiles en la tipificación. Se corrió todo el producto de PCR (96 pozos), incluyendo un carril con marcador de peso molecular en cada nivel de los pozos para poder identificar los pesos de las bandas esperadas comparando con la hoja de trabajo que se incluye en el kit. Con este patrón se puede identificar el HLA A, B y C de un solo paciente. Los primeros 24 pozos corresponden al gen A, de los pozos 25 al 73 corresponden al B y del 74 al 95 corresponden al gen C. Todas las PCRs incluyeron un control de amplificación interno para validar la reacción. En el caso de los pozos 25 y 26, el peso de la banda control de amplificación es diferente al resto.

## ***ANÁLISIS DE FRECUENCIAS DE LOS DIFERENTES ALELOS HLA.***

### **Análisis de frecuencias HLA en las distintas regiones en estudio (Norte, Centro y Sur)**

Las frecuencias alélicas y poblacionales de los genes HLA de clase I A, B y C para las cohortes estudiadas fueron obtenidas utilizando las herramientas de Inmunología de la Base de Datos de VIH de Los Álamos <http://www.hiv.lanl.gov/content/immunology/hla/>. Las comparaciones de distribuciones de frecuencias en diferentes regiones se realizaron por pruebas de Fisher corregidas para múltiples comparaciones mediante técnicas de razón de descubrimientos falsos (false discovery rate, FDR), empleando el estadístico q [32]. Se consideraron diferencias significativas con valores con  $p < 0.05$  y  $q < 0.2$ .

### **Análisis de frecuencia de los distintos alelos HLA con relación al estado y/o zona de procedencia.**

Se buscaron asociaciones entre la expresión de alelos HLA específicos y el estado y/o región de procedencia de los individuos incluidos en el estudio, con el objeto de reconocer patrones característicos de expresión de alelos HLA en la población de diferentes zonas del país. . Estos análisis se llevaron a cabo con el programa estadístico "R" v2.8.1, usando el paquete q value v1.1, comparando distribución de frecuencias por pruebas de Fisher corregidas para múltiples comparaciones mediante técnicas de descubrimientos falsos (FDR), empleando el estadístico q [32]. Se consideraron diferencias significativas con valores con  $p < 0.05$  y  $q < 0.2$ .

### **Desequilibrio de ligamiento.**

Se analizó la posibilidad de existencia de desequilibrio de ligamiento entre los alelos HLA característicos de cada una de las tres regiones en estudio (Norte, Centro y Sur) en alta y baja resolución, con ayuda de las herramientas de Inmunología de la Base de Datos de VIH de Los Álamos <http://www.hiv.lanl.gov/content/immunology/hla/>. El análisis se lleva a cabo calculando la frecuencia observada de tener un HLA (P (HLA1)) y tener otro alelo (P(HLA2)) en la población; se obtiene también la frecuencia observada de los

individuos que tienen ambos alelos en la población ( $P(\text{HLA1} \ \& \ \text{HLA2})$ ), y finalmente se obtiene la frecuencia esperada (E) si HLA1 y HLA2 se heredan de manera independiente uno de otro ( $E(P(\text{HLA1} \ \& \ \text{HLA2}))$ ). Con pruebas exactas de Fisher se obtiene el valor de p, el cual es un estimado de la probabilidad de que  $P(\text{HLA1})$  y  $P(\text{HLA2})$  sean independientes. Un valor pequeño o significativo de p indica que la ocurrencia de tener HLA1 y HLA2 está vinculada en dicha población.

### **Análisis de correlación entre valores significativamente más altos o bajos de carga viral y la expresión de alelos HLA específicos.**

Con el fin de buscar alelos HLA protectores o de riesgo para el avance de la enfermedad por VIH en la población mexicana, se buscaron correlaciones entre la presencia de un alelo HLA específico y niveles altos o bajos de carga viral. Se utilizaron alelos HLA tipificados a alta y baja resolución e incluyendo las tres regiones en estudio (Norte, Centro y Sur). Este análisis se realizó con la prueba de Man – Whitney con ayuda del programa estadístico “R” v2.8.1, usando el paquete q value v1.1 y se consideraron diferencias significativas con  $p < 0.05$  y  $q < 0.2$ .

## **ANÁLISIS DE SECUENCIAS DEL VIH**

### **Alineamiento de secuencias**

De la base de datos de secuencias del proyecto de Epidemiología Molecular y Evolución del VIH en México del CIENI, se tomaron las secuencias de un fragmento del gen *pol* (de 1302 pb, que incluye 99 codones de la proteasa y 335 codones de la RT), de la base de datos del CIENI). Posteriormente estas secuencias se alinearon junto con 9 secuencias de referencia tomadas de la base de datos de los Álamos (<http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>), con ayuda del programa MEGA 4 (Análisis Molecular de Genética Evolutiva) [33]. Cada una de las secuencias se importó dentro del programa MEGA 4, utilizando la herramienta de CLUSTALW. CLUSTALW se usa cuando se quiere un alineamiento de múltiples secuencias, donde el grupo de secuencias más similares se alinea primero y progresivamente se van alineando las secuencias más alejadas hasta

que se obtiene un alineamiento global. El algoritmo comienza con una matriz de distancia aproximada entre cada par de secuencias basadas en las puntuaciones de alineación en pares de secuencias. A continuación el algoritmo utiliza el método de Neighbor – Joining para crear un árbol guía, el cual es usado para general el alineamiento global, es decir, primero se alinean las secuencias con mayor similitud y se van agregando las demás ramas más alejadas. El árbol guía sirve como plantilla para los subtipos que tienden a compartir las inserciones y deleciones. En general, CLUSTALW proporciona un resultado cercano a lo óptimo, especialmente cuando el conjunto de datos contiene secuencias con diversos grados de divergencia, por lo que el árbol guía es menos sensible al ruido [34].

### **Análisis filogenéticos**

Se construyó un árbol filogenético para las secuencias de los individuos en estudio (tomadas de la base de datos del CIENI), a partir de un fragmento de 1302pb del gen *pol* del VIH, abarcando todo el gen de la proteasa (99 codones) y dos terceras partes de la RT (335 codones), en el programa MEGA 4 [33]. Se utilizó el método de Neighbor – Joining [35], calculando las distancias evolutivas con el modelo de dos parámetros de Kimura [36]. La significancia del árbol se probó mediante 1000 réplicas de bootstrap [37]. Se incluyeron 9 secuencias de referencia obtenidas de la Base de Datos de Secuencias de Los Álamos (<http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>), de los 9 subtipos del grupo M: A – D, F – H, J y K.

## **ANÁLISIS EVOLUTIVOS**

### **Análisis de selección positiva con modelo de codones**

Se buscaron posiciones en el genoma viral con evidencia de estar sujetas a selección positiva, utilizando un modelo molecular de evolución de codones basado en razones de sustituciones sinónimas (dS) y no sinónimas (dN), disponible en el servidor Datamonkey ([www.datamonkey.org](http://www.datamonkey.org)), utilizando el análisis SLAC (Single-Likelihood Ancestor Counting) [38]. SLAC ajusta un modelo de nucleótidos al alineamiento de secuencias y a su árbol filogenético estimado

por el método Neighbor-Joining. Posteriormente, se ajusta un modelo de codones para obtener un valor global de  $\omega$ , con la siguiente definición:

$$\omega = \frac{dN}{dS}$$

Con esto se reconstruyen secuencias ancestrales sitio por sitio utilizando máxima verosimilitud (maximum likelihood, ML), utilizando los parámetros determinados en los ajustes de modelos de nucleótidos y de codones anteriores. Finalmente se infiere si existe selección calculando los valores normalizados esperados (ES y EN) y observados (NS y NN) de sustituciones sinónimas y no sinónimas para cada sitio. SLAC estima los valores de:

$$dN = \frac{NN}{EN} \quad y \quad dS = \frac{NS}{ES}$$

y si  $dN < (\text{ó}) > dS$  para cada codón, estimando la significancia mediante valores de  $p$  derivados de una distribución binomial de dos colas. Si  $dN > dS$ , entonces se determinará que la posición se encuentra bajo selección positiva; si  $dN < dS$ , entonces se determinara que la posición se encuentra bajo presión negativa. Se consideraron como significativas posiciones con  $dN \neq dS$  con  $p < 0.1$ .

### **Análisis de correlación de alelo HLA- polimorfismo de VIH**

Se buscó la existencia de alguna correlación entre la expresión de cada uno de los alelos HLA y la presencia de aminoácidos específicos en cada una de las posiciones de la secuencia del gen *pol* del VIH en la cohorte de pacientes mexicanos del estudio. De la misma forma, se buscaron correlaciones entre la presencia de aminoácidos específicos en la secuencia viral y el estado y/o región de procedencia de los pacientes. Todas estas asociaciones se hicieron con pruebas exactas de Fisher, corrigiendo para múltiples comparaciones mediante técnicas de razón de descubrimientos falsos, empleando el estadístico  $q$  [32], con ayuda del programa estadístico "R" v2.8.1, usando el paquete  $q$  value v1.1. Se consideraron diferencias significativas con  $p < 0.01$  y  $q < 0.2$ .

Con los alelos HLA asociados a un polimorfismo en el genoma del VIH resultantes del análisis estadístico se realizó un mapa de epítopes, utilizando la secuencia consenso del fragmento del gen *pol* que contiene 99 posiciones de la PR y 335

posiciones de la RT, tomada de [27]. Los epítomos de CTLs reportados se obtuvieron a partir de la base de datos de los Álamos ([http://www.hiv.lanl.gov/content/immunology/tables/ctl\\_summary.html](http://www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html)).

## 6. RESULTADOS

### ***CARACTERÍSTICAS DE LA COHORTE ESTUDIADA***

Se reunió una cohorte compuesta por 139 individuos infectados por VIH de la zona Norte, Centro y Sur del país, para el estudio de la evolución dependiente de HLA del VIH en diferentes regiones de la República Mexicana. La cohorte de individuos del estudio presentó un promedio de células T CD4+ de 590.842 células/ $\mu$ l (Tabla 6.1) con una desviación estándar de 192.029 células/ $\mu$ l y una mediana de 558 células/ $\mu$ l, además de un promedio de carga viral de 49352.782 copias/ml de sangre, con una desviación estándar de 121481.733 copias/ml de sangre.

**Tabla 6.1 Descripción estadística de la población en estudio •.**

	NORTE	CENTRO	SUR	TOTAL
<b>Total de pacientes</b>	43	49	47	139
<b>CD4 (células/<math>\mu</math>l)</b>				
Promedio	617.067	605.727	551.33	590.842
Mediana	556.78	571	523.02	558
Desviación Estándar	219.474	171.234	183.236	192.029
Error Estándar	33.47	24.462	26.728	16.288
Rango	(304.540 - 1360.810)	(315.000 - 1232.000)	(146.00 - 919.00)	(146.00 - 1360.00)
<b>CV (copias/ml)</b>				
Promedio	24208.272	75473.184	45125.426	49352.782
Mediana	6437	12000	13213	11300
Desviación Estándar	38027.686	186252.61	73612.073	121481.733
Error Estándar	5799.168	26607.516	10737.424	10303.946
Rango	(50.00 - 158552.00)	(173.00 - 1000000.00)	(50.00 - 411859.00)	(50.00 - 1000000.0)

•Estos datos se obtuvieron con el programa StatView, se muestran estadísticos para los parámetros clínicos de cuenta de células T CD4+/ $\mu$ l y Carga Viral plasmática (CV, copias de RNA/ml), para los individuos totales y los individuos por la zona de procedencia.

En general, la población presentó una distribución homogénea tanto para los valores de carga viral (CV) como para las cuentas de células T CD4+. En la Figura 6.1 se incluyen diagramas de caja mostrando la homogeneidad de ambos

parámetros clínicos en los grupos de individuos de las tres zonas del país en estudio. Estos resultados reflejan el hecho de que en el presente estudio se incluyeron preferentemente individuos con un número de células T CD4+/ $\mu$ l de sangre de alrededor de 500, para asegurar que la mayoría de sus niveles de linfocitos T CD4+ se encontraran en valores cercanos al set-point viral. Es sabido que el valor de la carga viral en este punto de equilibrio es un buen predictor de la velocidad de avance de la enfermedad por VIH [8]. De ahí que se haya considerado como criterio de inclusión en el presente estudio el tener cuentas de linfocitos T CD4+ altas, con la idea de acercarse lo más posible a este valor de set-point viral y poder estudiar la existencia de alelos HLA asociados con rápida o lenta progresión de la enfermedad en el contexto de la población mexicana. Se obtuvieron logaritmos de la CV de todos los individuos del estudio para la facilitar el manejo, análisis y presentación de los datos. A diferencia de la cuenta de las cuentas de los linfocitos T CD4+, la CV no se encuentra distribuida de manera simétrica, se observó que los individuos del la zona Norte del país tuvieron un rango más amplio de valores de CV que los individuos de las zonas Centro y Sur de país, mientras que los valores de CV de la zona Sur tuvieron el menor rango (Figura 6.1.b) La homogeneidad del número de linfocitos T CD4+ y la diferencia significativa ( $p < 0.5$ ) que se encontró para los valores de CV de los individuos del estudio de las tres zona geográficas en estudio se comprobó con la prueba de Kruskal-Wallis (Tabla 6.2). Con esta prueba se concluyó que la tres zonas (Norte, Centro y Sur), tienen una distribución homogénea entre ellas de CD4+/ $\mu$ l, y existe una diferencia significativa en los valores de CV (copias RNA/ml).

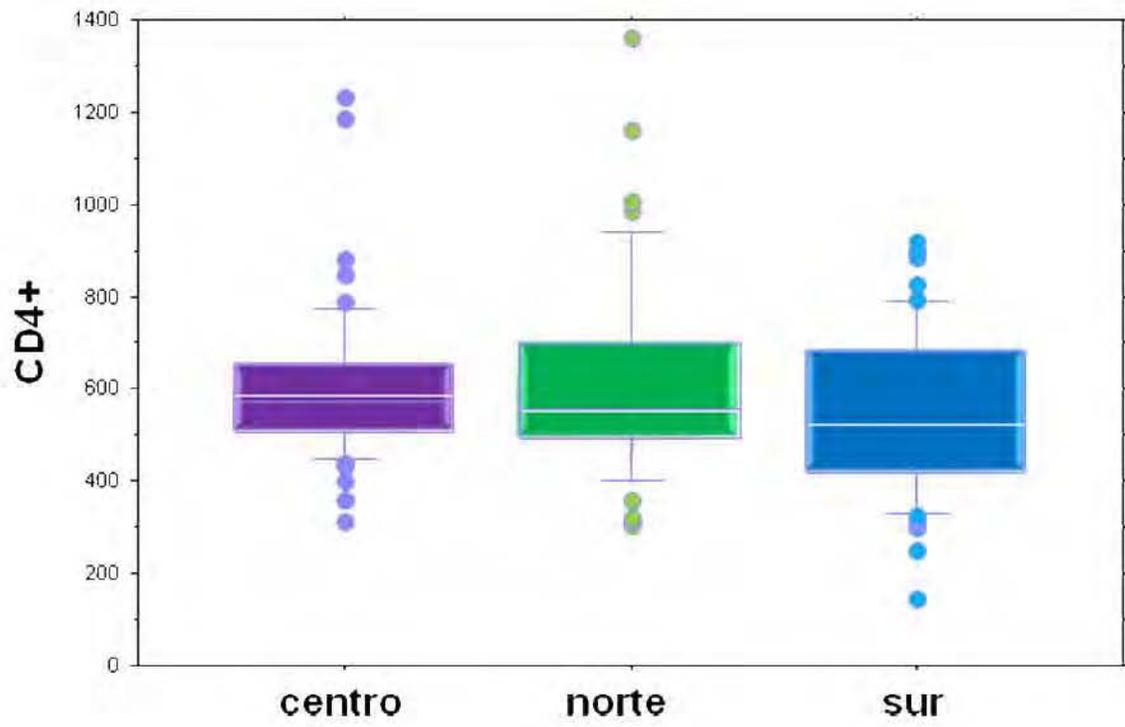
**Tabla 6.2 Prueba de Kruskal - Wallis para los valores de conteo de linfocitos T CD4+ y CV para las tres zonas en estudio ♦.**

<b>Parámetros</b>	<b>CD4+</b>	<b>CV</b>
GL	2	2
Número de Grupos	3	3
Número de Rangos	4	4
Valor de H	2.162	6.476
Valor de p	0.3393	0.0392
Corrección de H	2.162	6.48
Valor de p de Rangos	0.3393	0.0392

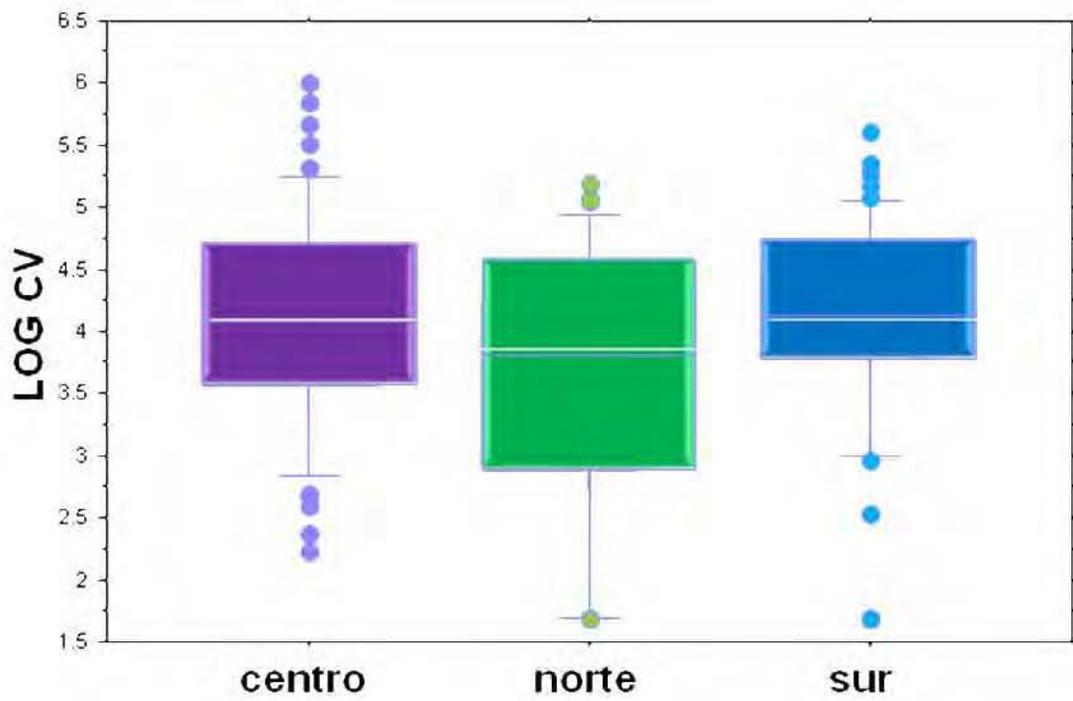
♦ La prueba se realizó con el programa StatView para la comparación de la cuenta de linfocitos T CD4+ y los valores de Carga Viral (CV) de las tres zonas en estudio. GL: grados de libertad.

Se encontró que todos los individuos de la cohorte estudiada estaban infectados con VIH-1 del subtipo B (Figura 6.2), el cual se ha descrito como el subtipo predominante en los países de América [14]. Un árbol filogenético con las secuencias del gen *pol* de los virus de los individuos incluidos en el estudio (tomadas de la base de secuencias de Epidemiología y Evolución del VIH en México del CIENI), mostró que todos los virus estudiados se agrupaban en la rama donde se encontraba la referencia del VIH-1 subtipo B.

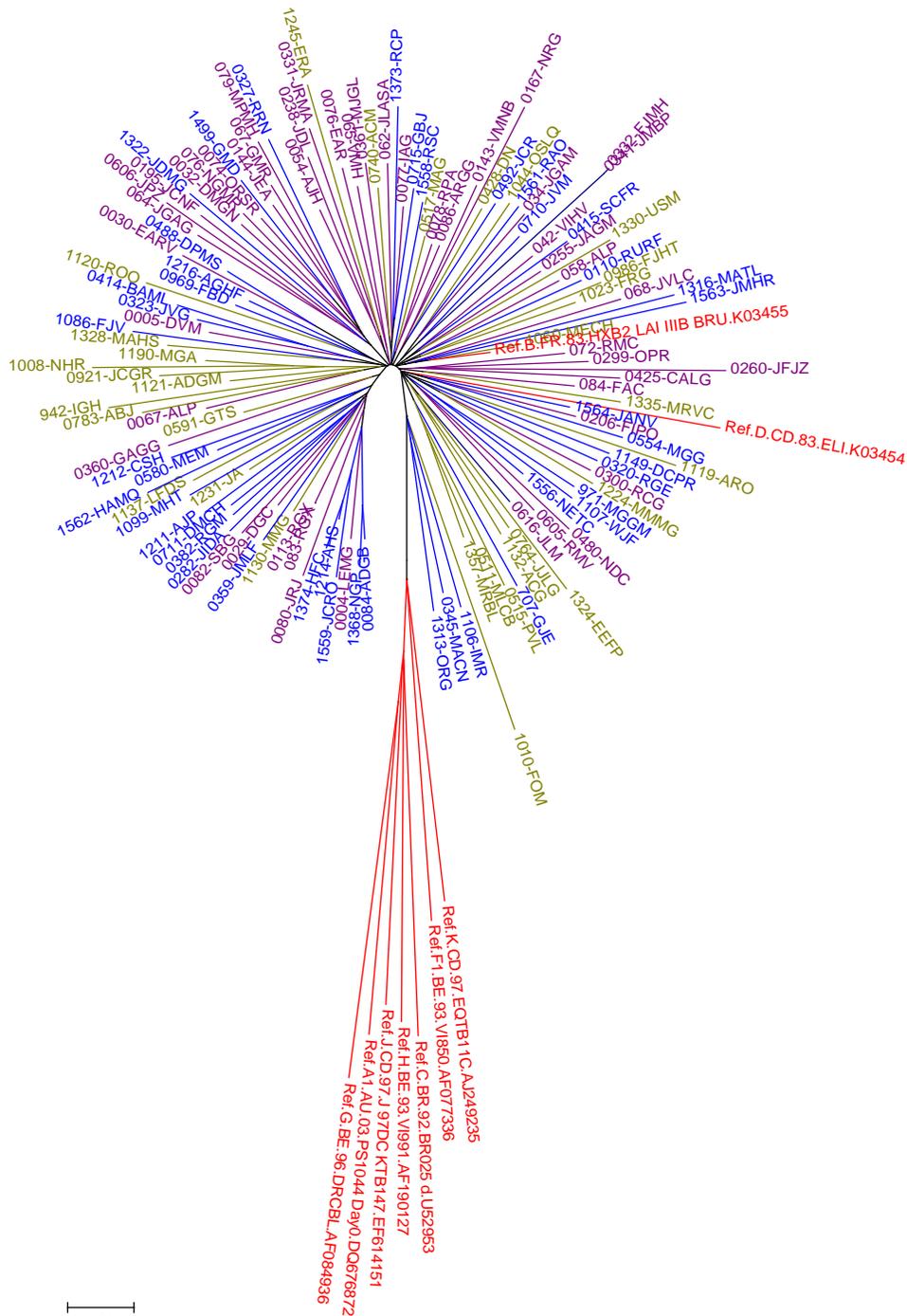
a)



b)



**Figura 6.1 Gráficas de cajas y bigotes para los parámetros clínicos de los individuos incluidos en el estudio. a) Distribución del número de células CD4+/ $\mu$ l de sangre, para las tres zonas del país en estudio. b) Distribución del logaritmo de carga viral (log CV) por cada zona en estudio. Las cajas incluyen los percentiles 25 a 75 (50% de los datos). Las cajas más los bigotes representan los percentiles 10 a 90 (80% de los datos). Se muestran como puntos los valores fuera de los percentiles 10 a 90. La mediana está indicada por una línea horizontal dentro de las cajas. Estos diagramas se obtuvieron con el programa StatView.**



**Figura 6.2** **Árbol filogenético de distancias para 123 secuencias del gen *pol* de los virus que infectan a los individuos estudiados en la cohorte.** La historia de evolución fue inferida usando el método Neighbor – Joining [35]. El árbol consenso de 1000 repeticiones de bootstrap fue tomado como representativo de la historia evolutiva de los datos analizados [37]. Las ramas corresponden a la porción reproducida en al menos 50% de las repeticiones. La distancia de evolución fue calculada usando el método de dos parámetros de Kimura [36] y tiene como unidades el número de sustituciones estimadas por el modelo por sitio. Todas las posiciones se encuentran alineadas y los datos faltantes fueron eliminados solo por comparación de las secuencias en pares (opción de delección en pares). Se consideraron 1302 posiciones del gen *pol* en total. El análisis filogenético fue realizado con MEGA 4 [33]. En color rojo se muestran las secuencias de referencia tomadas de la base de datos de los Álamos <http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>, las secuencias de los individuos del Centro se muestran en color morado, las secuencias de los individuos del Sur en color azul y las secuencias de los individuos del Norte en color verde.

## **FRECUENCIAS HLA EN LA COHORTE EN ESTUDIO**

### **Frecuencias HLA en la cohorte mexicana**

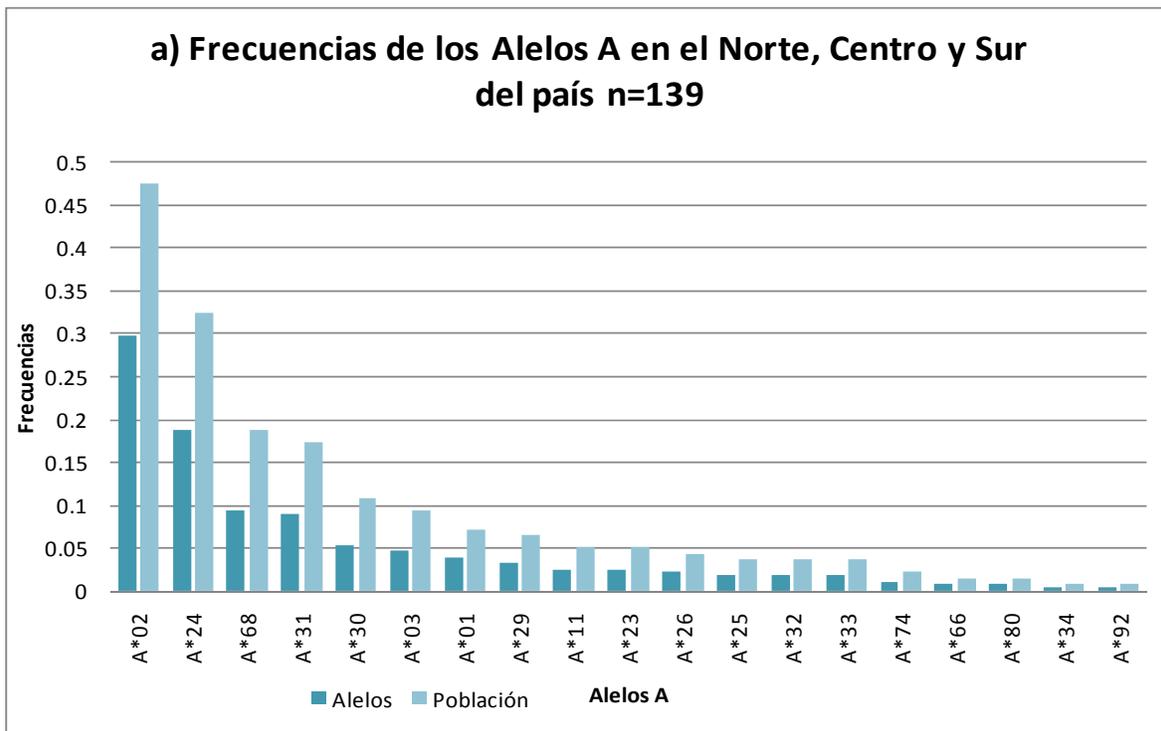
En la Tabla 6.3 se encuentran las frecuencias a baja resolución de los diferentes alelos HLA de clase I en la cohorte en estudio (n=139). En total se encontraron 19 alelos diferentes del HLA-A, 27 alelos diferentes de HLA-B y 14 alelos diferentes para el HLA-C. De estos, los alelos más frecuentes fueron A\*02, A\*24, A\*68, A\*31, para el gen HLA-A (Figura 6.3.a); B\*35, B\*39, B\*40, B\*15, para el gen HLA-B (Figura 6.3.b); y Cw\*07, Cw\*03, Cw\*04, Cw\*08, para el gen HLA-C (Figura 6.3.c).

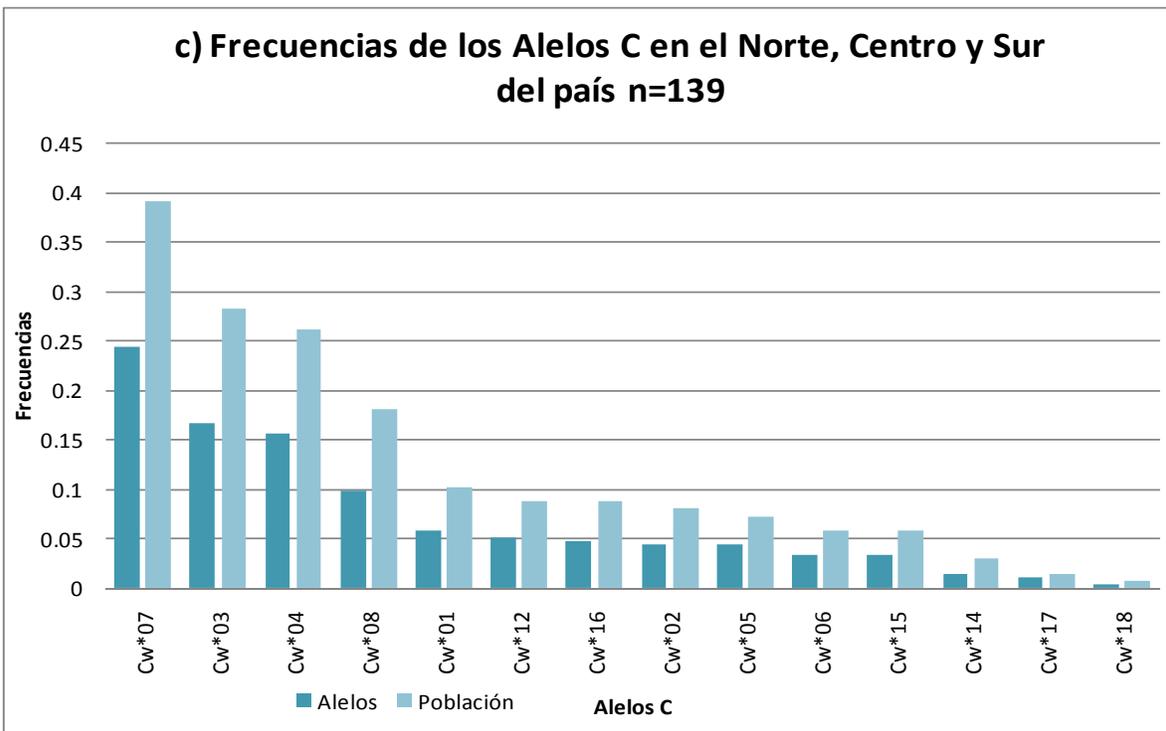
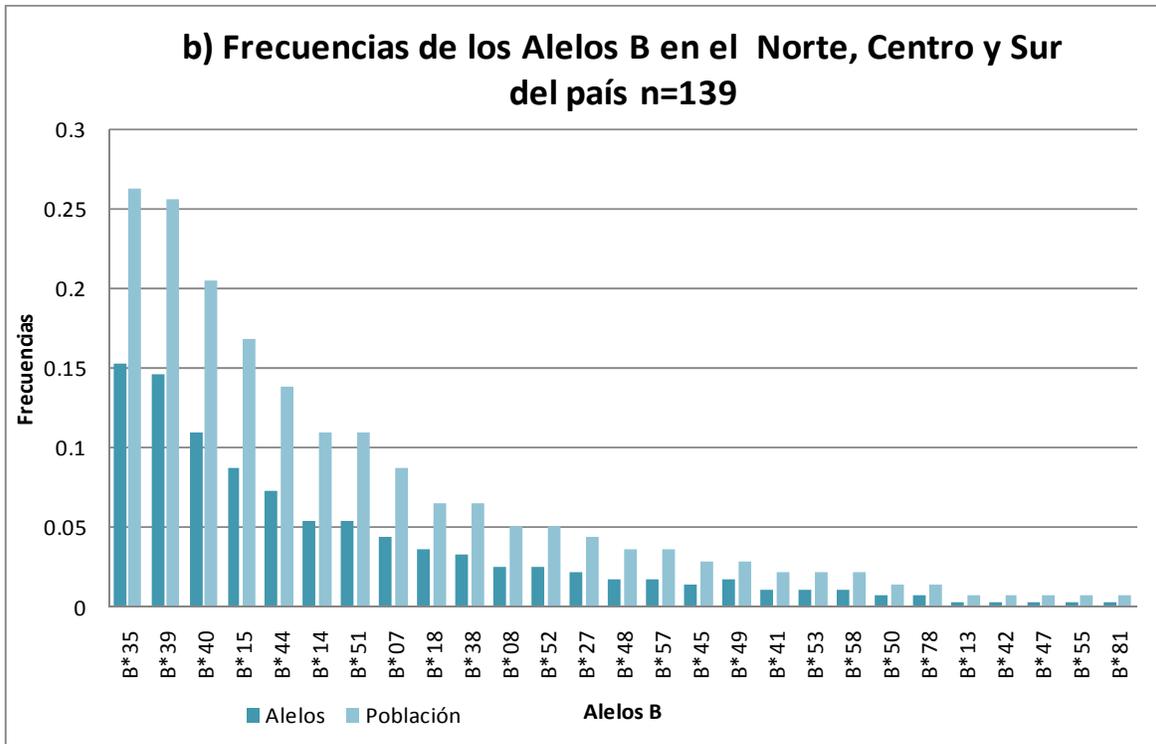
**Tabla 6.3 Frecuencias alélicas de las tres regiones estudiadas ▲.**

<b>HLA A</b>	<b>FRECUENCIA ALÉLICA</b>	<b>HLA B</b>	<b>FRECUENCIA ALÉLICA</b>	<b>HLA C</b>	<b>FRECUENCIA ALÉLICA</b>
<b>A*01</b>	0.0396	<b>B*07</b>	0.0438	<b>Cw*01</b>	0.058
<b>A*02</b>	0.2986	<b>B*08</b>	0.0255	<b>Cw*02</b>	0.0435
<b>A*03</b>	0.0468	<b>B*13</b>	0.0036	<b>Cw*03</b>	0.1667
<b>A*11</b>	0.0252	<b>B*14</b>	0.0547	<b>Cw*04</b>	0.1558
<b>A*23</b>	0.0252	<b>B*15</b>	0.0876	<b>Cw*05</b>	0.0435
<b>A*24</b>	0.1871	<b>B*18</b>	0.0365	<b>Cw*06</b>	0.0326
<b>A*25</b>	0.018	<b>B*27</b>	0.0219	<b>Cw*07</b>	0.2428
<b>A*26</b>	0.0216	<b>B*35</b>	0.1533	<b>Cw*08</b>	0.0978
<b>A*29</b>	0.0324	<b>B*38</b>	0.0328	<b>Cw*12</b>	0.0507
<b>A*30</b>	0.054	<b>B*39</b>	0.146	<b>Cw*14</b>	0.0145
<b>A*31</b>	0.0899	<b>B*40</b>	0.1095	<b>Cw*15</b>	0.0326
<b>A*32</b>	0.018	<b>B*41</b>	0.0109	<b>Cw*16</b>	0.0471

<b>A*33</b>	0.018	<b>B*42</b>	0.0036	<b>Cw*17</b>	0.0109
<b>A*34</b>	0.0036	<b>B*44</b>	0.073	<b>Cw*18</b>	0.0036
<b>A*66</b>	0.0072	<b>B*45</b>	0.0146		
<b>A*68</b>	0.0935	<b>B*47</b>	0.0036		
<b>A*74</b>	0.0108	<b>B*48</b>	0.0182		
<b>A*80</b>	0.0072	<b>B*49</b>	0.0182		
<b>A*92</b>	0.0036	<b>B*50</b>	0.0073		
		<b>B*51</b>	0.0547		
		<b>B*52</b>	0.0255		
		<b>B*53</b>	0.0109		
		<b>B*55</b>	0.0036		
		<b>B*57</b>	0.0182		
		<b>B*58</b>	0.0109		
		<b>B*78</b>	0.0073		
		<b>B*81</b>	0.0036		

▲ Los valores de frecuencia se obtuvieron por medio de las herramientas de la base de datos de Inmunología de Los Álamos. Se estudió una cohorte de n=139, donde se incluyeron individuos de las tres regiones del país estudiadas en este trabajo. En total hubo 248 alelos HLA-A de los cuales fueron 19 alelos diferentes (los cuales se muestran en la tabla); 274 alelos totales HLA-B con 27 alelos diferentes y 276 alelos totales HLA-C con 14 alelos diferentes.





**Figura 6.3 Frecuencias alélicas y poblacionales de HLA para la cohorte de individuos infectados por VIH estudiada.** Se muestran las frecuencias de los alelos encontrados para cada gen HLA para una cohorte de 139 individuos de las tres regiones de la República Mexicana en orden de frecuencia: a) Alelos HLA-A, b) Alelos HLA-B y c) Alelos HLA-C.

## Frecuencias HLA de la cohorte de la región Norte del país

Las frecuencias alélicas que se muestran en la Tabla 6.4 corresponden a los alelos HLA tipificados en alta resolución de los individuos procedentes del Norte del país (n=43). En total se obtuvieron 29 alelos diferentes para HLA-A, 36 alelos diferentes para HLA-B y 20 alelos diferentes para HLA-C.

Los alelos que se encontraron con mayor frecuencia en la región Norte del país son: A\*0201, A\*2402, A\*0101, A\*3101, B\*4002, B\*3517, B\*0801, B\*1402, Cw\*0401, Cw\*0304, Cw\*0702, Cw\*0701 (Figura 6.4).

Tabla 6.4 Frecuencia de alelos de la región Norte del país ►.

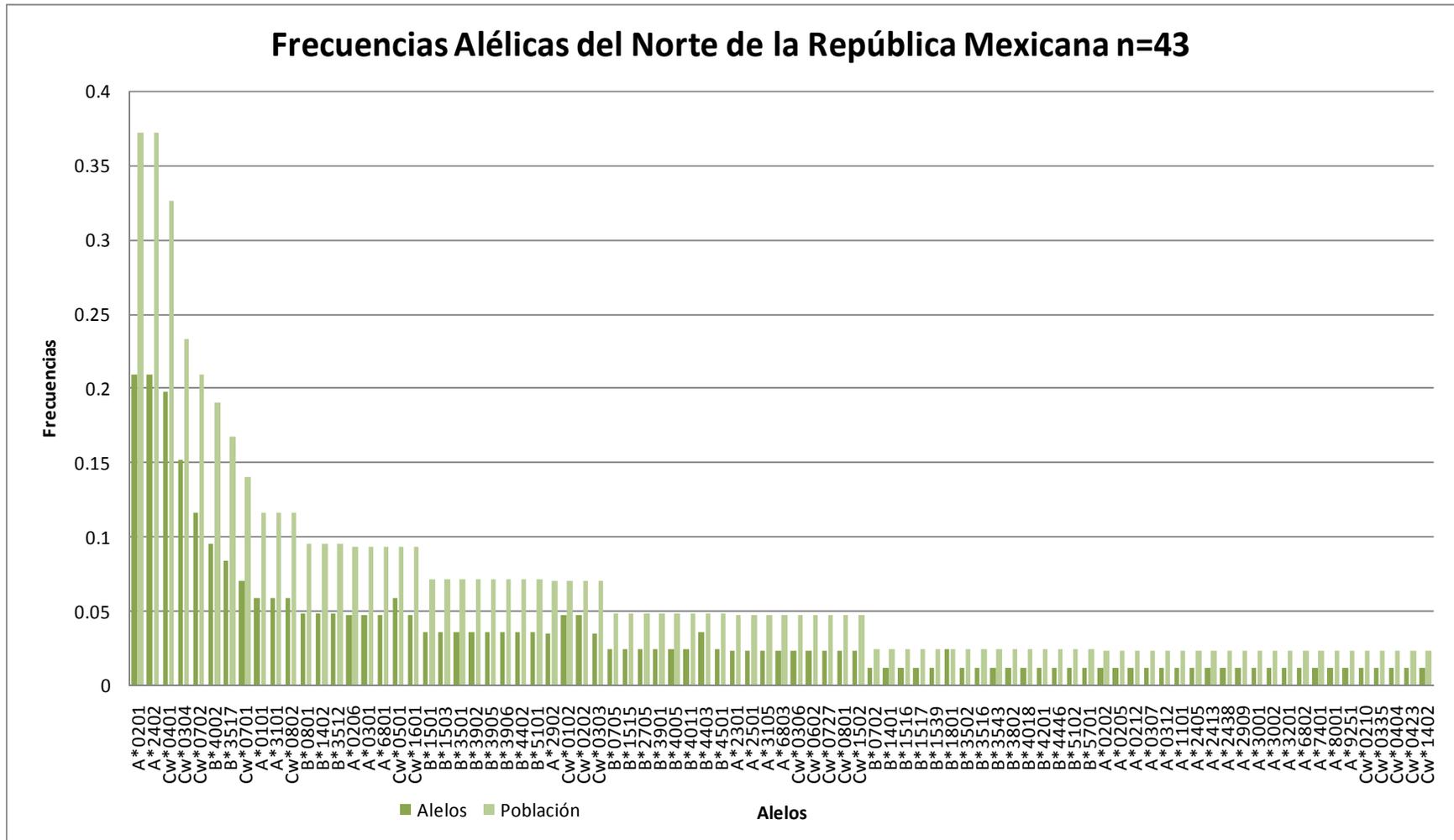
HLA A	FRECUENCIA ALÉLICA	HLA B	FRECUENCIA ALÉLICA	HLA C	FRECUENCIA ALÉLICA
A*0101	0.0581	B*0702	0.0119	Cw*0102	0.0465
A*0201	0.2093	B*0705	0.0238	Cw*0202	0.0465
A*0202	0.0116	B*0801	0.0476	Cw*0210	0.0116
A*0205	0.0116	B*1401	0.0119	Cw*0303	0.0349
A*0206	0.0465	B*1402	0.0476	Cw*0304	0.1512
A*0212	0.0116	B*1501	0.0357	Cw*0306	0.0233
A*0301	0.0465	B*1503	0.0357	Cw*0335	0.0116
A*0307	0.0116	B*1515	0.0238	Cw*0401	0.1977
A*0312	0.0116	B*1516	0.0119	Cw*0404	0.0116
A*1101	0.0116	B*1517	0.0119	Cw*0423	0.0116
A*2301	0.0233	B*1539	0.0119	Cw*0501	0.0581
A*2402	0.2093	B*1801	0.0238	Cw*0602	0.0233
A*2405	0.0116	B*2705	0.0238	Cw*0701	0.0698
A*2413	0.0116	B*3501	0.0357	Cw*0702	0.1163
A*2438	0.0116	B*3502	0.0119	Cw*0727	0.0233
A*2501	0.0233	B*3512	0.0476	Cw*0801	0.0233
A*2902	0.0349	B*3516	0.0119	Cw*0802	0.0581
A*2909	0.0116	B*3517	0.0833	Cw*1402	0.0116
A*3001	0.0116	B*3543	0.0119	Cw*1502	0.0233
A*3002	0.0116	B*3802	0.0119	Cw*1601	0.0465
A*3101	0.0581	B*3901	0.0238		
A*3105	0.0233	B*3902	0.0357		
A*3201	0.0116	B*3905	0.0357		
A*6801	0.0465	B*3906	0.0357		
A*6802	0.0116	B*4002	0.0952		
A*6803	0.0233	B*4005	0.0238		
A*7401	0.0116	B*4011	0.0238		
A*8001	0.0116	B*4018	0.0119		

<b>A*9251</b>	0.0116	<b>B*4201</b>	0.0119
		<b>B*4402</b>	0.0357
		<b>B*4403</b>	0.0357
		<b>B*4446</b>	0.0119
		<b>B*4501</b>	0.0238
		<b>B*5101</b>	0.0357
		<b>B*5102</b>	0.0119
		<b>B*5701</b>	0.0119

► Se muestra las frecuencias alélicas de los genes HLA-A, HLA-B y HLA-C de los individuos de la región norte del país (n=43). Se encontraron 86 alelos HLA-A, de los cuales 29 fueron diferentes; 84 alelos HLA-B con 36 alelos diferentes y 86 alelos HLA-C con 20 alelos diferentes mostrados en la tabla.

### **Frecuencias HLA de la cohorte de la región Centro del país.**

Los alelos HLA diferentes que se encontraron en la región Centro del país fueron 15 para el gen HLA-A, 22 para el gen HLA-B y 13 para el gen HLA-C (Tabla 6.5). Con un total de 50 alelos diferentes entre los tres alelos HLA de clase I, tipificados a baja resolución. Los alelos HLA más frecuentes en la región Centro fueron: A\*02, A\*24, A\*31, A\*30, B\*35, B\*39, B\*44, B\*07, Cw\*07, Cw\*04, Cw\*08, Cw\*03 (Figura 6.5). De manera similar a lo observado en la región Norte, la población del Centro del país mostró una alta frecuencia de los grupos alélicos A\*02, A\*24, A\*31, B\*35, B\*39, Cw\*07, Cw\*04.



**Figura 6.4 Frecuencias alélicas y poblacionales HLA de los individuos provenientes de la región Norte del país.** Se encontraron en total 85 alelos entre HLA-A (29), HLA-B (36) y HLA-C (20), en el grupo de 43 individuos infectados por VIH procedentes del Norte del país, los cuales se muestran en orden del de mayor a menor frecuencia. Todos ellos tipificados en alta resolución (ver metodología). Las barras de color verde oscuro corresponden a las frecuencias alélicas y las barras de color verde claro a las frecuencias poblacionales.

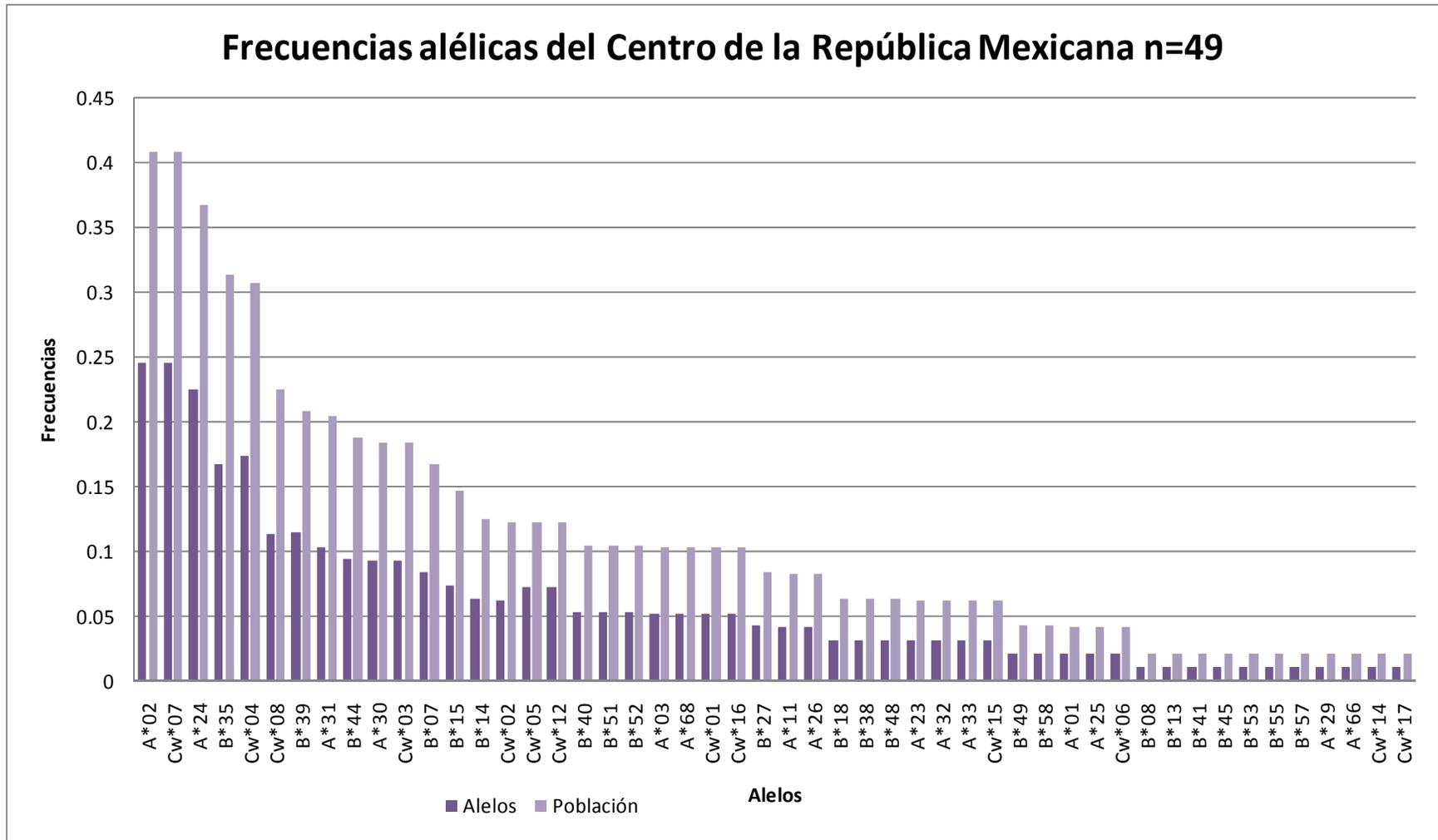
**Tabla 6.5 Frecuencias alélicas de los alelos encontrados en la región Centro del país ▼**

HLA A	FRECUENCIA ALÉLICA	HLA B	FRECUENCIA ALÉLICA	HLA C	FRECUENCIA ALÉLICA
<b>A*01</b>	0.0204	<b>B*07</b>	0.0833	<b>Cw*01</b>	0.051
<b>A*02</b>	0.2449	<b>B*08</b>	0.0104	<b>Cw*02</b>	0.0612
<b>A*03</b>	0.051	<b>B*13</b>	0.0104	<b>Cw*03</b>	0.0918
<b>A*11</b>	0.0408	<b>B*14</b>	0.0625	<b>Cw*04</b>	0.1735
<b>A*23</b>	0.0306	<b>B*15</b>	0.0729	<b>Cw*05</b>	0.0714
<b>A*24</b>	0.2245	<b>B*18</b>	0.0312	<b>Cw*06</b>	0.0204
<b>A*25</b>	0.0204	<b>B*27</b>	0.0417	<b>Cw*07</b>	0.2449
<b>A*26</b>	0.0408	<b>B*35</b>	0.1667	<b>Cw*08</b>	0.1122
<b>A*29</b>	0.0102	<b>B*38</b>	0.0312	<b>Cw*12</b>	0.0714
<b>A*30</b>	0.0918	<b>B*39</b>	0.1146	<b>Cw*14</b>	0.0102
<b>A*31</b>	0.102	<b>B*40</b>	0.0521	<b>Cw*15</b>	0.0306
<b>A*32</b>	0.0306	<b>B*41</b>	0.0104	<b>Cw*16</b>	0.051
<b>A*33</b>	0.0306	<b>B*44</b>	0.0938	<b>Cw*17</b>	0.0102
<b>A*66</b>	0.0102	<b>B*45</b>	0.0104		
<b>A*68</b>	0.051	<b>B*48</b>	0.0312		
		<b>B*49</b>	0.0208		
		<b>B*51</b>	0.0521		
		<b>B*52</b>	0.0521		
		<b>B*53</b>	0.0104		
		<b>B*55</b>	0.0104		
		<b>B*57</b>	0.0104		
		<b>B*58</b>	0.0208		

▼ Se muestran las frecuencias alélicas de los genes HLA-A, HLA-B y HLA-C de los individuos del Centro del país (n=49). Se encontraron 98 alelos para el gen HLA-A, con solo 15 alelos diferentes (datos mostrados en la tabla); 96 alelos totales con 22 alelos diferentes para el gen HLA-B y hubo 98 alelos totales y 13 alelos diferentes para el gen HLA-C.

### **Frecuencias alélicas en la región Sur del país**

En los individuos procedentes de la región Sur del país (n=47), se encontraron 29 alelos diferentes para HLA-A, 43 alelos diferentes para HLA-B y 23 alelos diferentes para HLA-C (Tabla 6.6) todos ellos tipificados a alta resolución. En la Figura 6.6 se encuentran los alelos con mayor frecuencia como son: A\*0206, A\*0201, A\*2402, A\*3101, B\*3906, B\*4002, B\*1801, B\*3905, Cw\*0702, Cw\*0304, Cw\*0102, Cw\* 0401.



**Figura 6.5 Frecuencias alélicas y poblacionales HLA de los individuos procedentes de la región Centro del país.** Se encontraron un total de 15 alelos diferentes HLA-A, 22 alelos diferentes de HLA-B, y 13 alelos diferentes para HLA-C en un total de 49 individuos, los cuales se muestran por orden de frecuencia. En morado se muestran las frecuencias alélicas y en lila las frecuencias poblacionales. La tipificación se llevó a cabo a resolución (ver Metodología).

**Tabla 6.6 Frecuencias alélicas de los alelos encontrados en la región Sur del país ◀.**

HLA A	FREC. ALÉLICA	HLA A	FREC. ALÉLICA	HLA B	FREC. ALÉLICA	HLA B	FREC. ALÉLICA	HLA C	FREC. ALÉLICA
	CA		A		CA		A		A
A*0101	0.0426	A*6801	0.0426	B*0702	0.0106	B*4006	0.0213	Cw*0102	0.0761
A*0201	0.1809	A*6802	0.0426	B*0801	0.0213	B*4008	0.0106	Cw*0210	0.0109
A*0205	0.0106	A*6803	0.0426	B*1401	0.0319	B*4101	0.0213	Cw*0303	0.0326
A*0206	0.1596	A*6805	0.0213	B*1402	0.0106	B*4403	0.0319	Cw*0304	0.1196
A*0236	0.0106	A*7401	0.0213	B*1501	0.0213	B*4427	0.0106	Cw*0305	0.0435
A*0301	0.0213	A*8001	0.0106	B*1510	0.0106	B*4501	0.0106	Cw*0401	0.0761
A*1101	0.0213			B*1530	0.0213	B*4701	0.0106	Cw*0602	0.0543
A*2301	0.0106			B*1548	0.0106	B*4801	0.0213	Cw*0701	0.0761
A*2304	0.0106			B*1801	0.0532	B*4901	0.0319	Cw*0702	0.163
A*2402	0.0745			B*3501	0.0426	B*5001	0.0213	Cw*0704	0.0109
A*2403	0.0106			B*3514	0.0213	B*5101	0.0532	Cw*0727	0.0217
A*2428	0.0106			B*3517	0.0213	B*5102	0.0106	Cw*0801	0.0326
A*2501	0.0106			B*3543	0.0106	B*5201	0.0213	Cw*0802	0.0652
A*2601	0.0213			B*3801	0.0426	B*5301	0.0213	Cw*1202	0.0217
A*2902	0.0426			B*3802	0.0106	B*5701	0.0213	Cw*1203	0.0543
A*3002	0.0319			B*3901	0.0106	B*5703	0.0106	Cw*1402	0.0109
A*3004	0.0106			B*3902	0.0213	B*5802	0.0106	Cw*1403	0.0109
A*3101	0.0745			B*3903	0.0106	B*7802	0.0106	Cw*1502	0.0217
A*3102	0.0106			B*3905	0.0532	B*7805	0.0106	Cw*1503	0.0109
A*3201	0.0106			B*3906	0.0851	B*8101	0.0106	Cw*1509	0.0109
A*3303	0.0213			B*3913	0.0106			Cw*1601	0.0435
A*3402	0.0106			B*4002	0.0638			Cw*1701	0.0217
A*6601	0.0106			B*4005	0.0319			Cw*1801	0.0109

◀Se muestran frecuencias alélicas HLA a alta resolución para los individuos provenientes de la región Sur del país (n=47). Se encontraron 94 alelos HLA-A con 29 alelos diferentes, 94 alelos HLA-B con 43 alelos diferentes y 92 alelos HLA-C totales con 23 alelos diferentes.

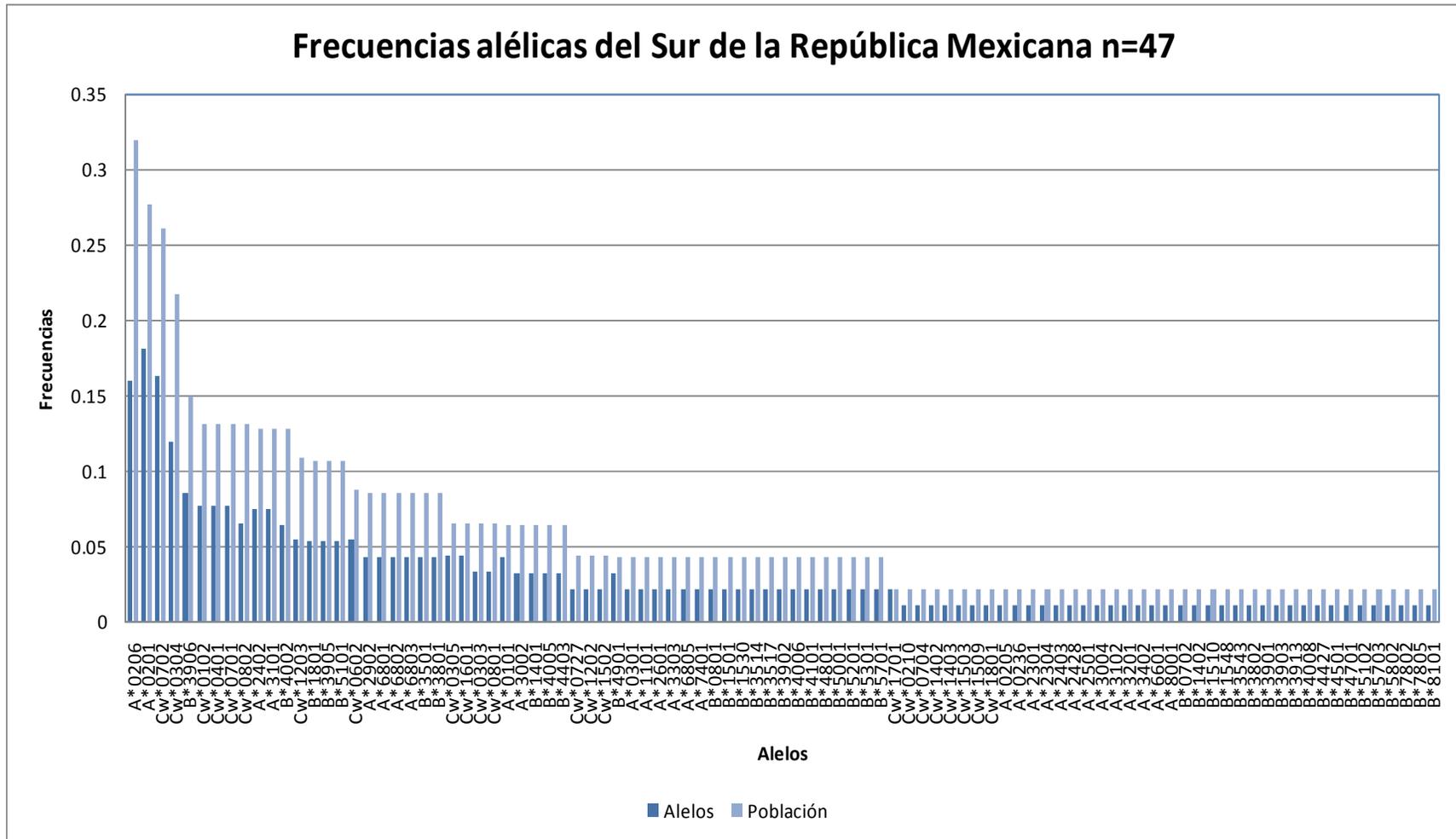
### **Comparación de frecuencias alélicas entre la región Norte y Sur del país**

Al comparar las frecuencias alélicas HLA, con ayuda de las herramientas de inmunología de la base de datos de los Álamos (<http://www.hiv.lanl.gov/content/immunology/hla/>) entre la región Norte (n=43) y Sur (n=47) del país a alta resolución, se encontraron cuatro diferencias significativas (con  $p < 0.01$  y  $q < 0.2$ ) (Figura 6.7): dos para los alelos HLA-A y dos para los alelos HLA-C. El alelo A\*0206 ( $p = 0.0206$ ,  $q = 0.148$ ), se encontró sobre expresado en los individuos procedentes de la región Sur del país mientras que el alelo A\*2402 ( $p = 0.01$ ,  $q = 0.164$ ) se encontró sobre expresado en los individuos

procedentes de la región Norte del país (Figura 6.7.a). Cw\*0401 ( $p= 0.0267$ ,  $q=0.145$ ) y Cw\*0501 ( $p= 0.0248$ ,  $q=0.181$ ), ambos se encontraron sobre expresados en la región Norte del país (Figura 6.7.c). Para los alelos HLA-B no se encontraron diferencias significativas entre las dos regiones comparadas con el número de individuos considerados; sin embargo, existe una tendencia de sobre expresión de B\*3517 en la región Norte sin alcanzarse los valores de significancia establecidos para  $q$  (Figura 6.7.b), lo cual se podría confirmar aumentando el número de individuos en la cohorte.

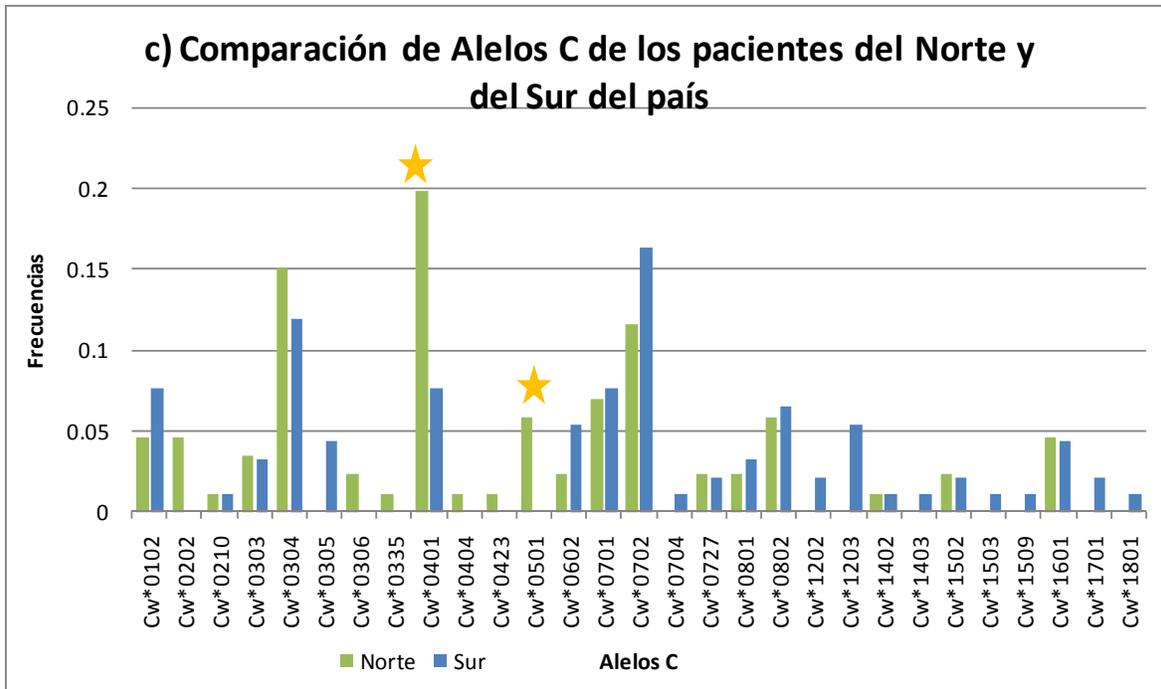
Se observaron también posibles diferencias en las frecuencias de alelos HLA minoritarios en dichas regiones estudiadas que deberán confirmarse al aumentar el número de individuos en la cohorte, ya que hubo alelos como: A\*0202, A\*0212, A\*0307, A\*0312, A\*2405, A\*2413, A\*2438, A\*2909, A\*3001, A\*3105, A\*9251, B\*0705, B\*1503, B\*1515, B\*1516, B\*1517, B\*1539, B\*2705, B\*3502, B\*3512, B\*3516, B\*4011, B\*4018, B\*4201, B\*4402, B\*4446, Cw\*0202, Cw\*0306, Cw\*0335, Cw\*0404, Cw\*0423, Cw\*0501 que sólo se observaron el Norte del país (Figura 6.7.a,b y c; barras color verde) y alelos que se observaron exclusivamente en el sur del país, los cuales fueron más que los exclusivos del norte (Figura 6.7.a,b y c; barras color azul): A\*0236, A\*2304, A\*2403, A\*2428, A\*2601, A\*3004, A\*3102, A\*3303, A\*3402, A\*6601, A\*6805, B\*1510, B\*1530, B\*1548, B\*3514, B\*3801, B\*3903, B\*3913, B\*4006, B\*4008, B\*4101, B\*4427, B\*4701, B\*4801, B\*4901, B\*5001, B\*5201, B\*5301, B\*5703, B\*5802, B\*7802, B\*7805, B\*8101, Cw\*0305, Cw\*0704, Cw\*1202, Cw\*1203, Cw\*1403, Cw\*1503, Cw\*1509, Cw\*1701, Cw\*1801.

## Frecuencias alélicas del Sur de la República Mexicana n=47



**Figura 6.6 Frecuencias alélicas y poblacionales HLA de los individuos procedentes de la región Sur del país.** Se muestran las frecuencias de los 95 alelos diferentes encontrados para los tres alelos HLA de clase I (HLA-A con 29 alelos diferentes ; HLA-B con 43 alelos diferentes y HLA-C con 23 alelos diferentes ) de los individuos provenientes del Sur del país (n=47) ordenados del alelo con mayor frecuencia poblacional al de menor frecuencia. Las barras de color azul marino corresponden a las frecuencias alélicas y las barras azul cielo a las frecuencias poblacionales. La tipificación HLA se realizó a alta resolución (ver Metodología).





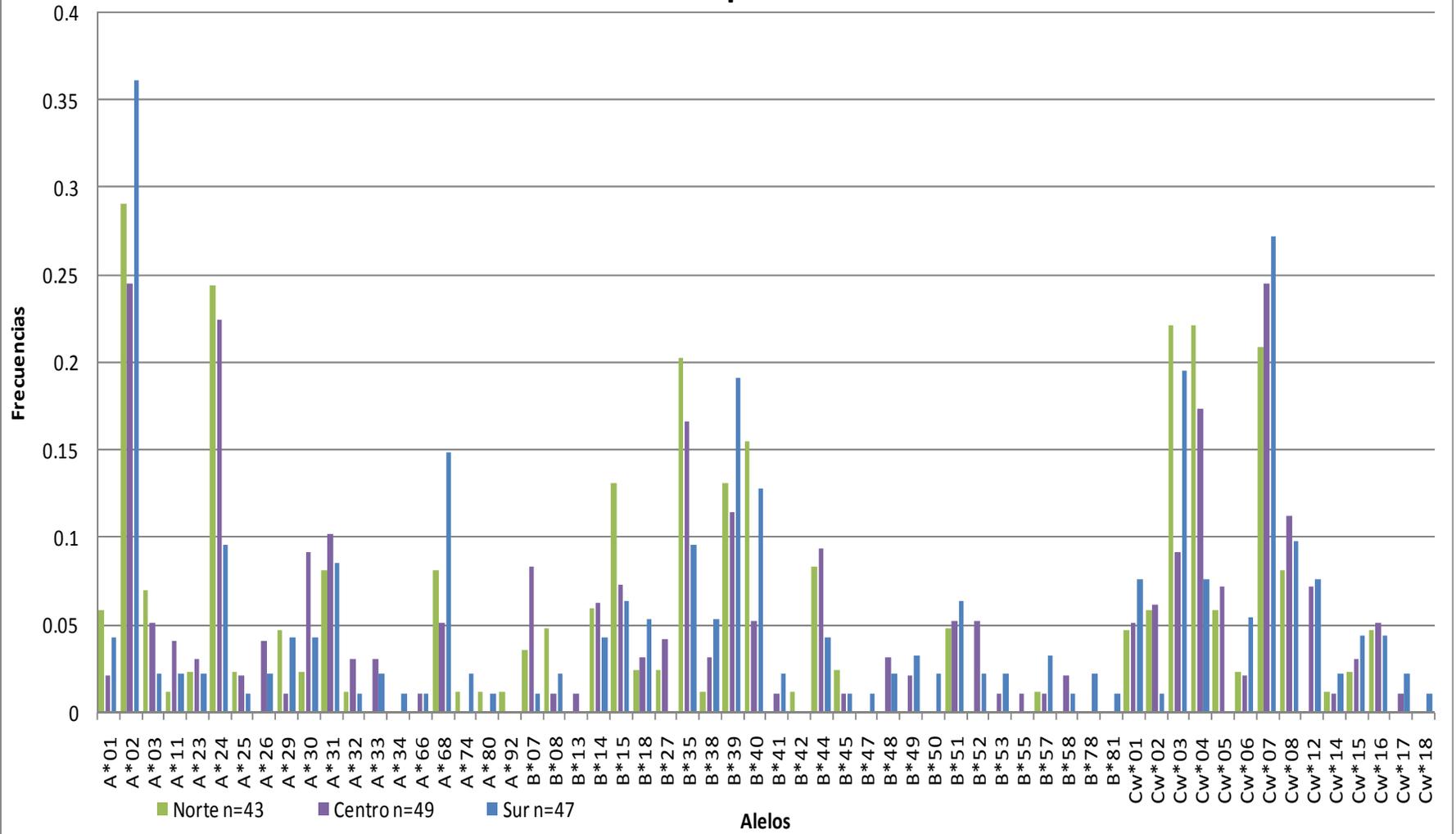
**Figura 6.7 Comparación de frecuencias alélicas de los genes HLA-A, B y C entre el Norte y Sur del país.** Se compararon frecuencias alélicas HLA de los individuos provenientes de la región Norte (n=43) y la región Sur (n=47) del país. En verde se muestran las frecuencias HLA de los individuos del Norte y en azul las de los individuos del Sur. Las tipificaciones HLA se llevaron a cabo a alta resolución como se explica en la sección de Metodología. Las diferencias significativas en las frecuencias HLA de los individuos del Norte y del Sur están marcadas con una estrella. La comparación de frecuencias HLA se hizo con las herramientas de la base de datos de Los Álamos (ver Metodología).

### Comparación de las frecuencias alélicas entre las regiones Norte, Centro y Sur del país

Se compararon las frecuencias alélicas HLA de los individuos de las tres regiones en estudio: Norte vs Centro vs Sur (Figura 6.8). Para poder comparar entre las tres poblaciones, se ocupó el programa estadístico “R” v2.8.1, usando el paquete q value v1.1 (ver Metodología). Se buscó si existían diferencias significativas en la expresión de alelos HLA entre las regiones de procedencia. La comparación se llevó a cabo a baja resolución y a alta resolución.

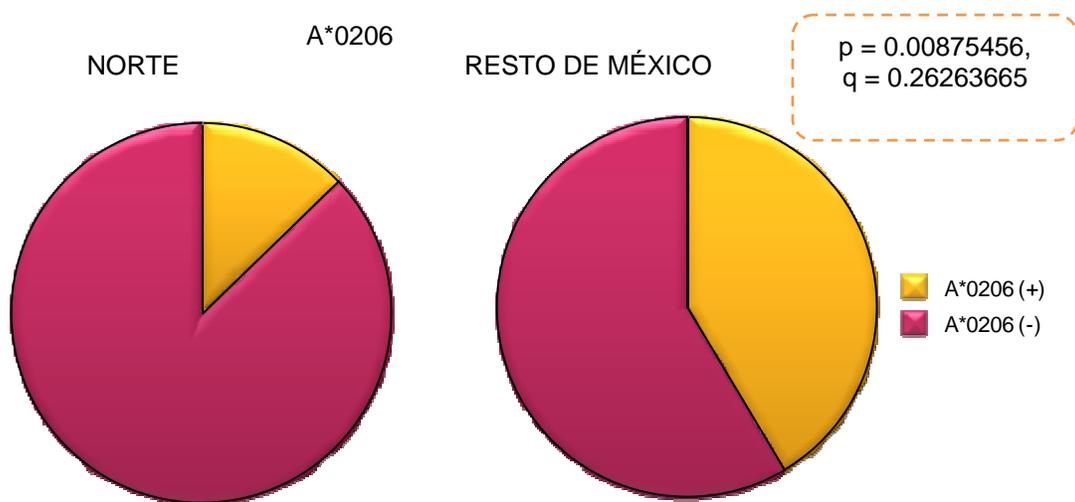
En la comparación de frecuencias HLA a alta resolución sólo se encontró un alelo: A\*0206 que se encontraba con menor frecuencia en la región Norte de la república en comparación con las dos regiones restantes (Figura 6.9).

## Comparación de las frecuencias alélicas de los pacientes del Norte, Centro y Sur de la República Mexicana



**Figura 6.8 Comparación de las frecuencias alélicas HLA-A, B y C de los pacientes de las regiones Norte, Centro y Sur del país.** Se muestra una comparación de frecuencias alélicas observadas en cada una de las tres regiones en estudio para cada alelo HLA a baja resolución. La comparación fue realizada con programa estadístico “R” v2.8.1. Las barras de color verde corresponden a los individuos provenientes de la región Norte del país, las barras moradas a los individuos del Centro del país y las barras azules a los individuos del Sur del país.

Las frecuencias del alelo A\*0206 también mostraron diferencia significativa en la comparación de los alelos HLA-A entre la región Norte y la región Sur en alta resolución al analizarse con las herramientas de análisis de frecuencias HLA de la Base de Datos de Los Álamos (Figura 6.7.a), donde se observó una menor frecuencia del este alelo en el Norte con respecto al Sur.



**Figura 6.9 Alelo HLA-A con frecuencias significativamente diferentes en las tres regiones del país en estudio.** En amarillo se muestra la proporción de individuos que presentaron el alelo A\*0206 y en fucsia se muestra la proporción de individuos carentes del alelo. La gráfica de la izquierda muestra la proporción de individuos que presentaron el alelo exclusivamente en la región Norte, la gráfica de la derecha muestra la proporción de individuos que presentaron el alelo en el resto del país (Regiones Sur y Centro). En el recuadro se muestran los valores de p y q. Este análisis se hizo con el programa estadístico “R” v2.8.1, usando el paquete q value v1.1. Las graficas se obtuvieron con el programa StatView.

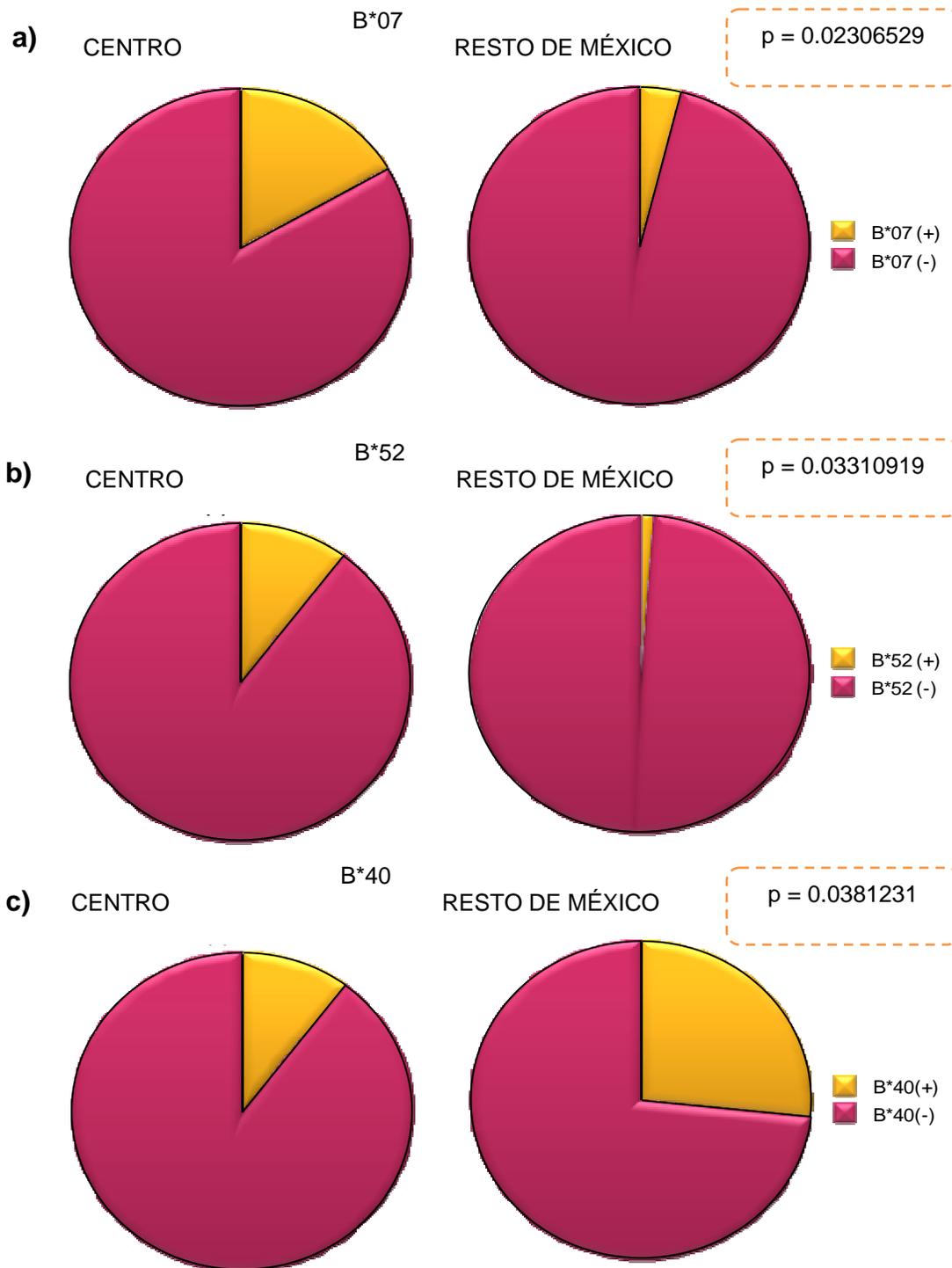


**Figura 6.10 Alelos HLA-A con frecuencias diferentes en las tres regiones del país.** Las comparaciones se realizaron con el programa estadístico “R” v2.8.1, usando el paquete q value v1.1. Las gráficas muestran en color amarillo la porción de individuos que presentaron el alelo HLA-A y en fucsia la porción de individuos que no presentaron el alelo HLA-A. Las gráfica de la izquierda muestran la proporción de individuos que presentaron el alelo exclusivamente en la región con diferencia significativa, las gráficas de la derecha muestran la proporción de individuos que presentaron el alelo en el resto del país. Las graficas se encuentran ordenadas del valor de p más bajo al más alto, en los recuadros se muestra el valor de p.

En la comparación de las frecuencias de los alelos HLA a baja resolución no se encontraron diferencias significativas en la expresión de los alelos en las tres regiones en estudio de acuerdo a los criterios de significancia establecidos ( $p < 0.01$ ,  $q < 0.2$ ); sin embargo, los valores de q ( $q < 0.5$ ) muestran una tendencia a presentar frecuencias mayores o menores en alguna de las tres regiones con respecto al resto del país en la expresión de algunos alelos HLA-A, HLA-B y HLA-C. (Figura 6.10, 6.11 y 6.12 respectivamente). Un aumento en el número de individuos tipificados para las tres regiones podrá confirmar estas observaciones.

La figura 6.10 muestra a los alelos HLA-A que mostraron una tendencia de diferencia en su expresión entre las regiones. El alelo A\*68 se observó con mayor frecuencia en la región Sur del país que en el resto de la República, (Figura 6.10.a); el alelo A\*24 que se encontró menos expresado en la región Sur del país, que en el resto de las regiones (Figura 6.10.b) y el alelo A\*30 se observó con mayor frecuencia en la región Centro (Figura 6.10.c) que en el resto de las regiones .

Dentro de los alelos HLA-B se observaron también tres alelos con  $p < 0.05$  (aunque con valores de  $q > 0.2$ ) que se muestran en la Figura 6.11. Los alelos B\*07 y B\*52, se encontraron con mayor frecuencia en la región Centro del país con respecto al resto de la República (Figura 6.11.a y 6.11.b respectivamente), y el alelo B\*40 se encontró menos expresado en la región Centro que en las regiones restantes (Figura 6.11.b).



**Figura 6.11 Alelos HLA-B con frecuencias diferentes en las tres regiones del país.** Las comparaciones se realizaron con el programa estadístico “R” v2.8.1, usando el paquete q value v1.1. Las gráficas se obtuvieron con el programa StatView. Las gráficas muestran en color amarillo la porción de individuos que presentaron el alelo HLA-B y en fucsia la porción de individuos que no presentaron el alelo HLA-B. Las gráficas de la izquierda muestran la proporción de individuos que

presentaron el alelo exclusivamente en la región con diferencia significativa; las gráficas de la derecha muestran la proporción de individuos que presentaron el alelo en el resto del país. Las gráficas se encuentran ordenadas del valor de p más bajo al más alto, en los recuadros se muestra el valor de p.

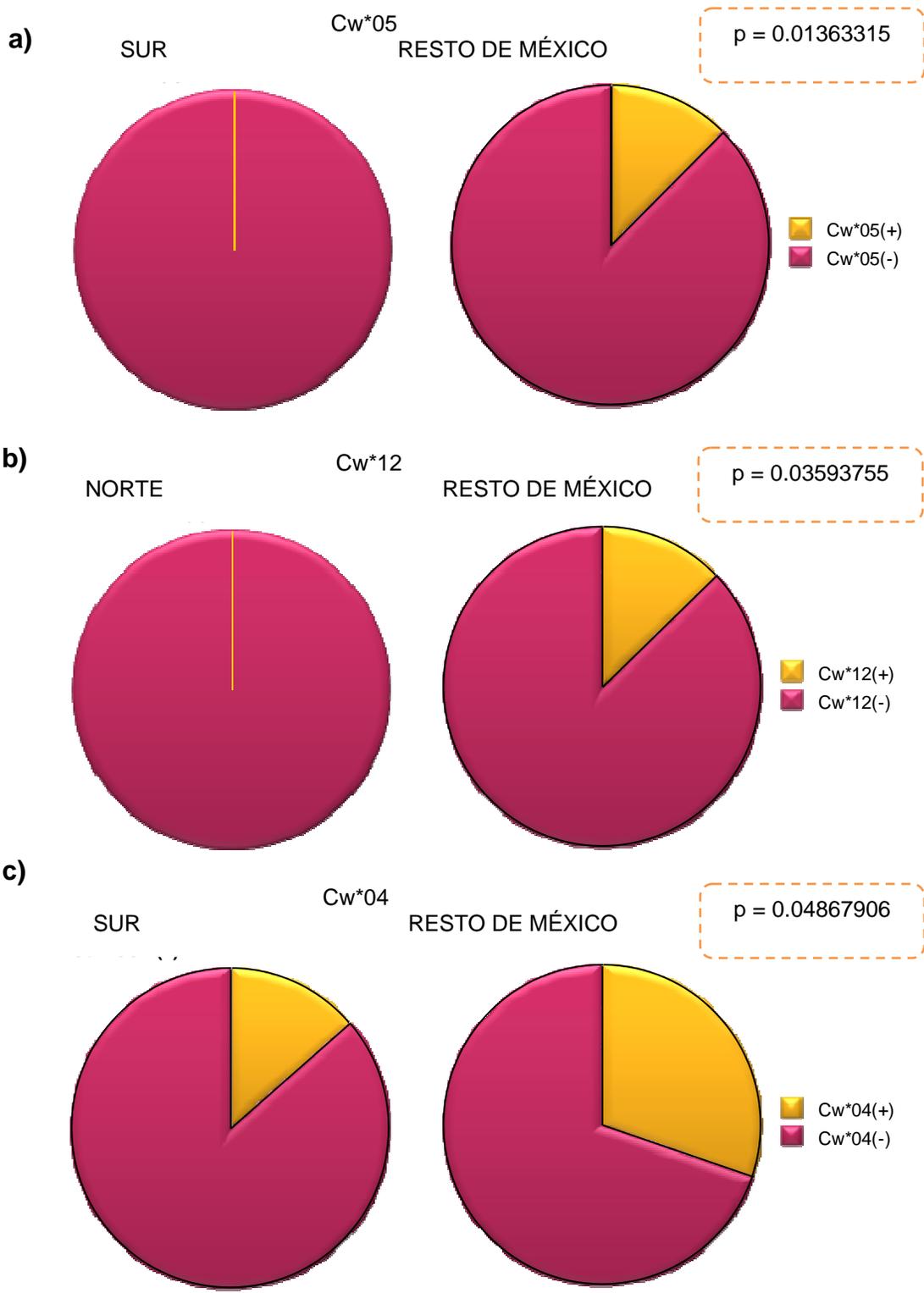
Dentro de los alelos HLA-C (Figura 6.12) los alelos que mostraron diferencias en sus frecuencias en las regiones estudiadas fueron Cw\*05 que no se encontró en ningún individuo de la región Sur del país (Figura 6.12.a), Cw\*12 que no se encontró en la región Norte (Figura 6.12.b) y el alelo Cw\*04 que se encontró con una baja frecuencia en la región Sur en comparación con las regiones restantes (Figura 6.12.c).

### **Comparación de las frecuencias alélicas HLA entre los estados que se incluyeron en el estudio**

La comparación de frecuencias alélicas HLA por estado se hizo con ayuda del programa estadístico "R" v2.8.1, usando el paquete q value v1.1. El análisis se hizo con alelos HLA a alta y baja resolución. No se observaron diferencias significativas en las frecuencias HLA por estado de acuerdo a los criterios de significancia establecidos ( $p < 0.01$ ,  $q < 0.2$ ) aunque los valores de p mostraron una tendencia de expresión diferencial para algunos alelos.

En la comparación de frecuencias de los alelos HLA por estado a alta resolución, se encontraron igualmente algunos alelos con tendencias a estar expresados en cierto estados con mayor o menor frecuencia que en los demás.

En la Figura 6.13 se observan los alelos HLA que tuvieron  $p < 0.01$ , como lo es el alelo A\*2301 el cual se encontró en todos los individuos de Sonora y en bajas frecuencias en individuos provenientes de otros estados (Figura 6.13.a).



**Figura 6.12 Alelos HLA-C con frecuencias diferentes en las tres regiones del país.** Las comparaciones se realizaron con el programa estadístico “R” v2.8.1, usando el paquete q value v1.1. Las gráficas se obtuvieron con el programa StatView. Las gráficas muestran en color amarillo

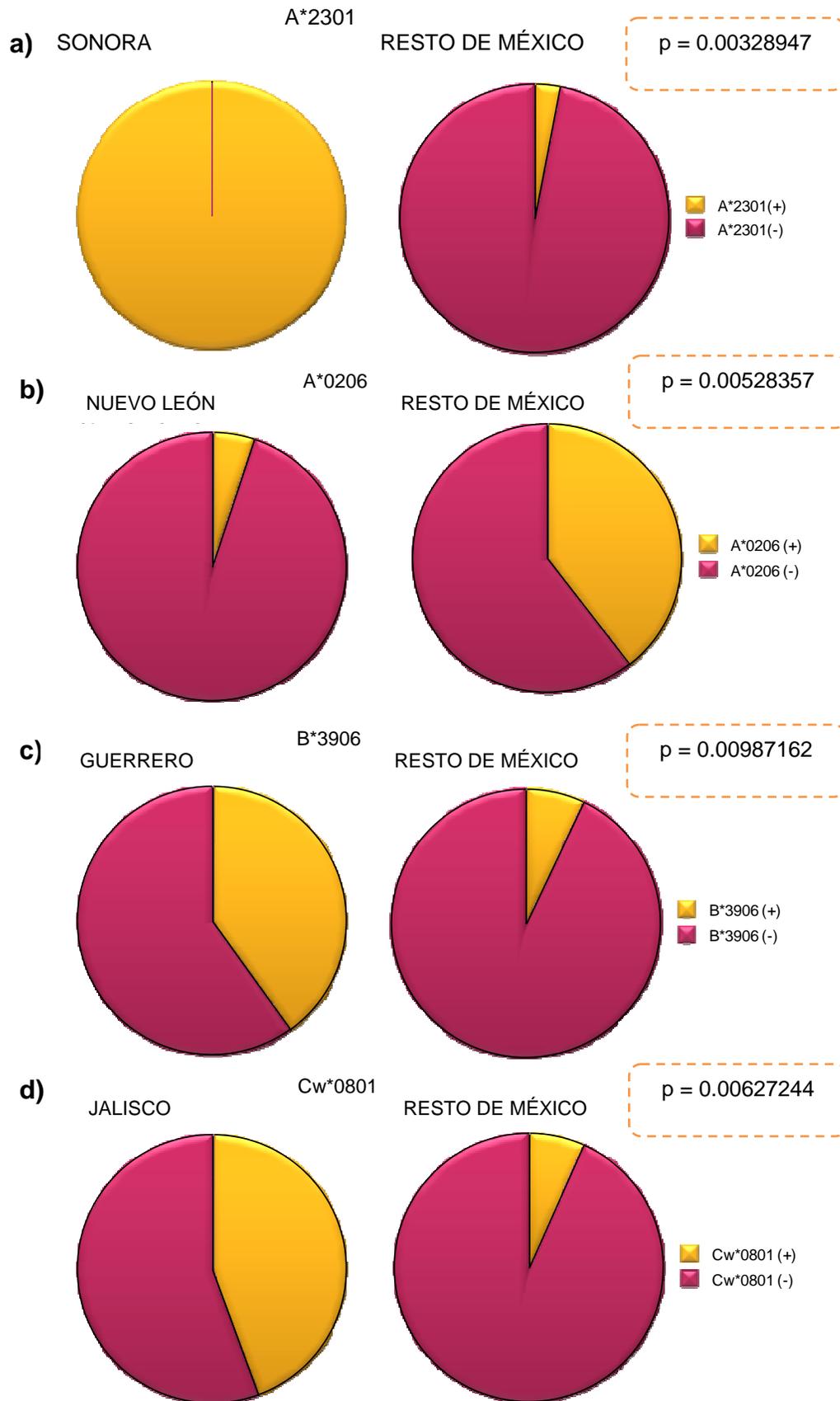
la porción de individuos que presentaron el alelo HLA-C y en fucsia la porción de individuos que no presentaron el alelo HLA-C. Las gráficas de la izquierda muestran la proporción de individuos que presentaron el alelo exclusivamente en la región con diferencia significativa; las gráficas de la derecha muestran la proporción de individuos que presentaron el alelo en el resto del país. Las gráficas se encuentran ordenadas del valor de p más bajo al más alto, en los recuadros se muestra el valor de p.

El alelo A\*0206, observado previamente con frecuencias menores en la región Norte del país, con respecto a la región Sur (Figura 6.7.a) y el resto del país (Figura 6.9), se observó con frecuencias menores en el estado de Nuevo León (Figura 6.13.b) Esta tendencia es importante, ya que casi la mitad de los individuos que proceden del Norte en la cohorte del estudio fueron originarios del estado de Nuevo León (47%, ver Metodología).

El alelo B\*3906 (Figura 6.13.c), se observó en mayor frecuencia en el estado de Guerrero en relación al resto del país. Un resultado similar se observó para el alelo Cw\*0801 el cual se encontró en mayor frecuencia en el estado de Jalisco (Figura 6.13.d), con respecto al resto de los estados que se incluyeron en el estudio.

En el análisis con los alelos HLA a baja resolución, tampoco se encontraron diferencias significativas, aunque los valores de p obtenidos para algunos alelos ( $p < 0.01$ ), sugieren tendencias de expresión diferencial en zonas específicas del país. Estudios posteriores con cohortes más grandes de individuos permitirán definir la existencia de estas diferencias en frecuencias HLA entre las poblaciones de distintas regiones de México.

Para el estado de Sonora, se encontró nuevamente una alta frecuencia del alelo A\*23 (Figura 6.14.a), que se había observado ya en el análisis a alta resolución (A\*2301) como un alelo prevalente en el estado (Figura 6.13.a).

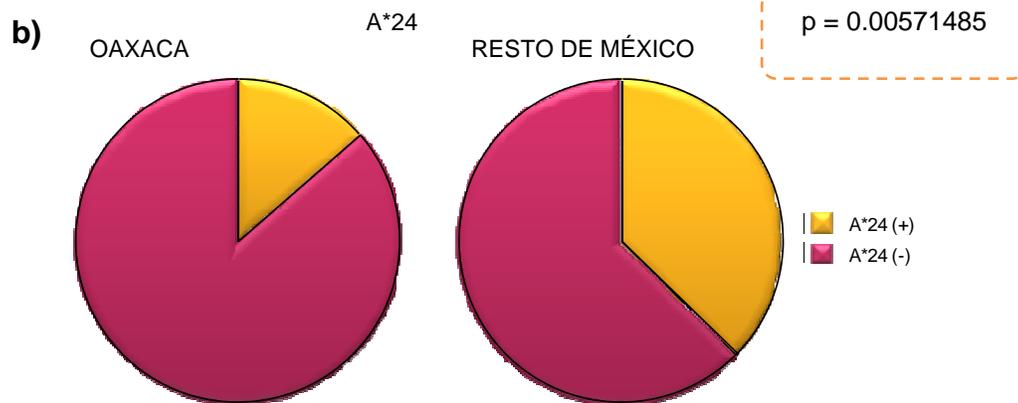
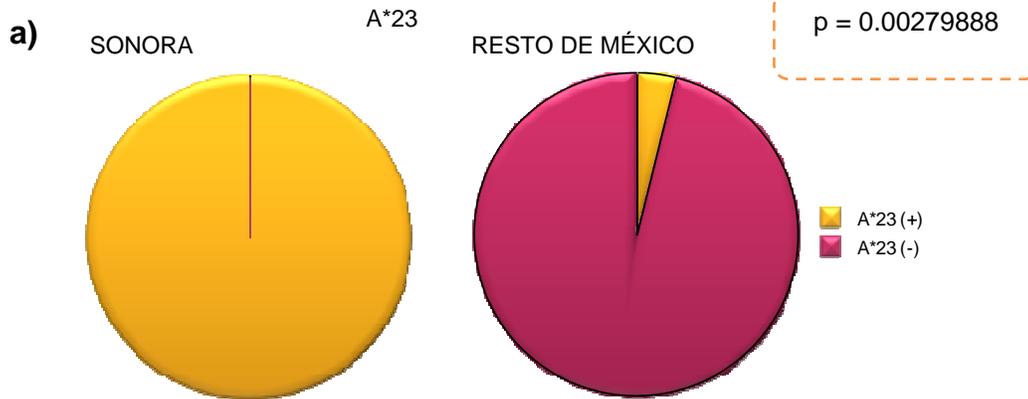


**Figura 6.13 Alelos HLA con tipificación a alta resolución con frecuencias diferentes en las tres regiones del país.** Las gráficas muestran en color amarillo la porción de individuos que presentaron el alelo HLA-C y en fucsia la porción de individuos que no presentaron el alelo HLA-C. Las gráfica de la izquierda muestran la proporción de individuos que presentaron el alelo exclusivamente en la región con diferencia significativa; las gráficas de la derecha muestran la proporción de individuos que presentaron el alelo en el resto del país. Las graficas se encuentran ordenadas del valor de p más bajo al más alto y por locus HLA. En los recuadros se muestran los valores de p para cada comparación. Las comparaciones se realizaron con el programa estadístico "R" v2.8.1, usando el paquete q value v1.1. Las gráficas se obtuvieron con el programa StatView.

El alelo A\*24, se observó en menor frecuencia en los individuos del estado de Oaxaca (Figura 6.14.b), en comparación con el resto de los estados incluidos en el estudio.

Para el D. F. se encontraron dos alelos expresados con mayor frecuencia que en el resto de los estados: el alelo B\*07 y el alelo Cw\*05 (Figura 6.14.c y d respectivamente). El alelo B\*07 se observó previamente sobre expresado en la región Centro, con respecto al resto del país (Figura 6.11.a). Este resultado se explica por tanto por la alta prevalencia del alelo en el D. F.

El alelo Cw\*05 se observó más expresado en el D. F. que en resto de los estados. Previamente se mostró que ningún individuo con el alelo Cw\*05 se encontró en la región Sur, pero sí en las dos regiones restantes (Figura 6.12.a). Con este análisis se sabe entonces que este alelo se encuentra principalmente en el D. F. con respecto a los demás estados incluidos en la estudio.



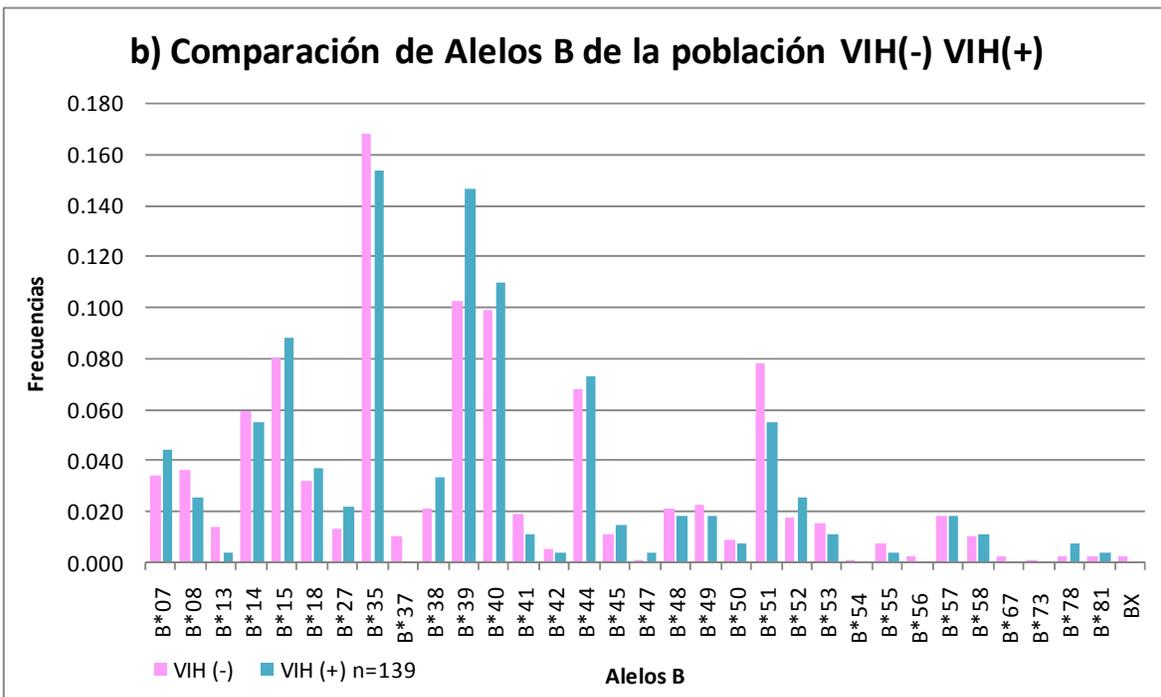
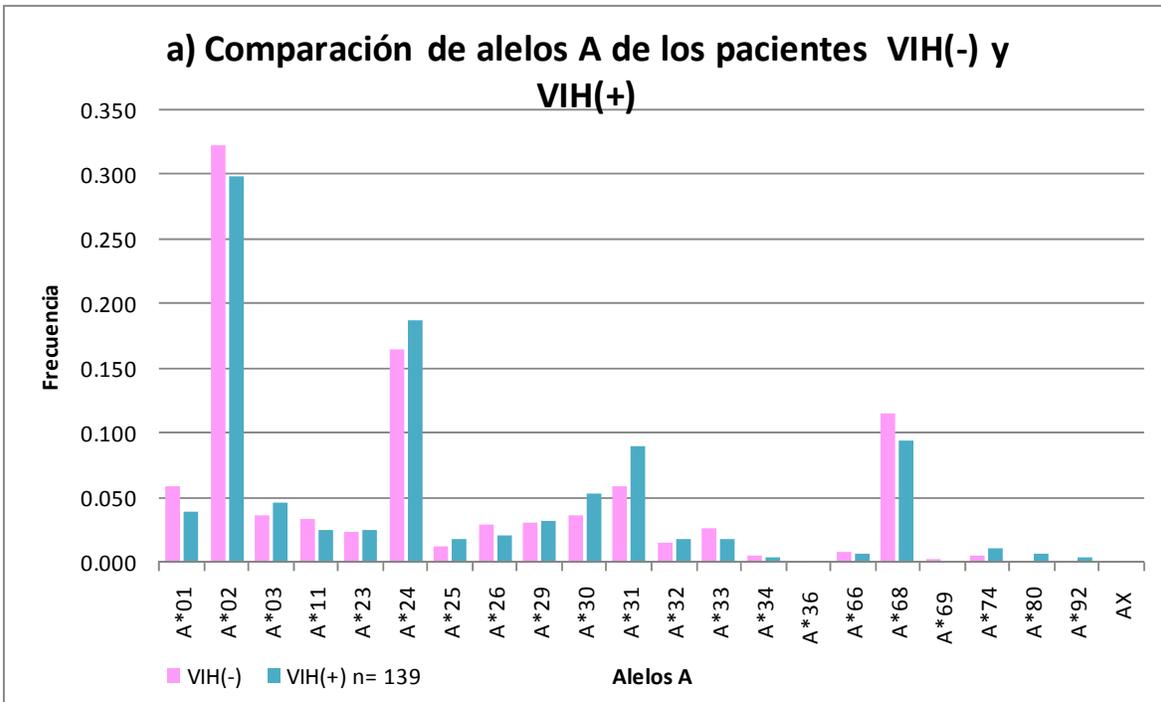
**Figura 6.14 Alelos HLA con tipificación a baja resolución con frecuencias diferentes en las tres regiones del país.** Las gráficas muestran en color amarillo la porción de individuos que presentaron el alelo HLA-C y en fucsia la porción de individuos que no presentaron el alelo HLA-C. Las gráfica de la izquierda muestran la proporción de individuos que presentaron el alelo exclusivamente en la región con diferencia significativa; las gráficas de la derecha muestran la proporción de individuos que presentaron el alelo en el resto del país. Las graficas se encuentran ordenadas del valor de p más bajo al más alto y por locus HLA. En los recuadros se muestran los valores de p para cada comparación. Las comparaciones se realizaron con el programa estadístico "R" v2.8.1, usando el paquete q value v1.1. Las gráficas se obtuvieron con el programa StatView.

### ***COMPARACIÓN DE FRECUENCIAS ALÉLICAS DE LA COHORTE DE INDIVIDUOS VIH POSITIVOS CON UNA COHORTE DE INDIVIDUOS VIH NEGATIVOS.***

Se realizó una comparación de las frecuencias HLA de clase I de la cohorte de individuos VIH positivos del presente estudio y una cohorte de 381 individuos no infectados con VIH descrita previamente por Barquera y colaboradores [7] (Figura 6.15). Esta cohorte incluye 191 Familias mexicanas de la región Norte y Centro de la República Mexicana, en su mayoría los individuos provienen de los estados: Sinaloa, Distrito Federal y Puebla, pero incluye también individuos de los estados de : Sonora, Durango, San Luis Potosí, Aguascalientes, Querétaro y Estado de México.

Los alelos A\*02, A\*24 y A\*68 que ya se habían observado con una alta frecuencia en la población estudiada, se encontraron también con frecuencias altas en la población VIH (-), sin observarse diferencias entre las frecuencias alélicas HLA en los dos grupos (Figura 6.15.a). Lo mismo sucedió con los alelos HLA-B (Figura 6.15.b) más frecuentes en la población en estudio, como B\*39, B\*35, B\*15. En el caso del alelo B\*39 la frecuencia observada en individuos VIH (+) fue mayor que en la población VIH negativa.

Algunos alelos como A\*36, A\*69, B\*37, B\*54, B\*56, B\*67, B\*73, que sólo se observaron en la población VIH (-), posiblemente no se encontraron en la población VIH (+) del presente estudio por el número de individuos incluidos en la cohorte. Los alelos A\*80 y A\*92 se observaron exclusivamente de población VIH (+), aunque a baja frecuencia.



**Figura 6.15 Comparación de las frecuencias de HLA de la población VIH(+)** en estudio y una población VIH(-). Las frecuencias alélicas de los individuos VIH(+) (barras azules) corresponden a las observadas en el total de individuos incluidos en la cohorte del estudio (n=139). Las frecuencias alélicas para la cohorte de individuos VIH(-) (barras en color rosa) corresponden a las observadas en una cohorte de 381 individuos de 191 familias Mexicanas reportadas previamente por Barquera y colaboradores [7]. La comparación se hizo para los alelos **a)** HLA-A y **b)** HLA-B. Todas las tipificaciones fueron hechas a baja resolución como se muestra en la sección de Metodología.

### **Comparación de las frecuencias HLA de los individuos VIH (+) de la región Norte con las observadas en una población VIH (-)**

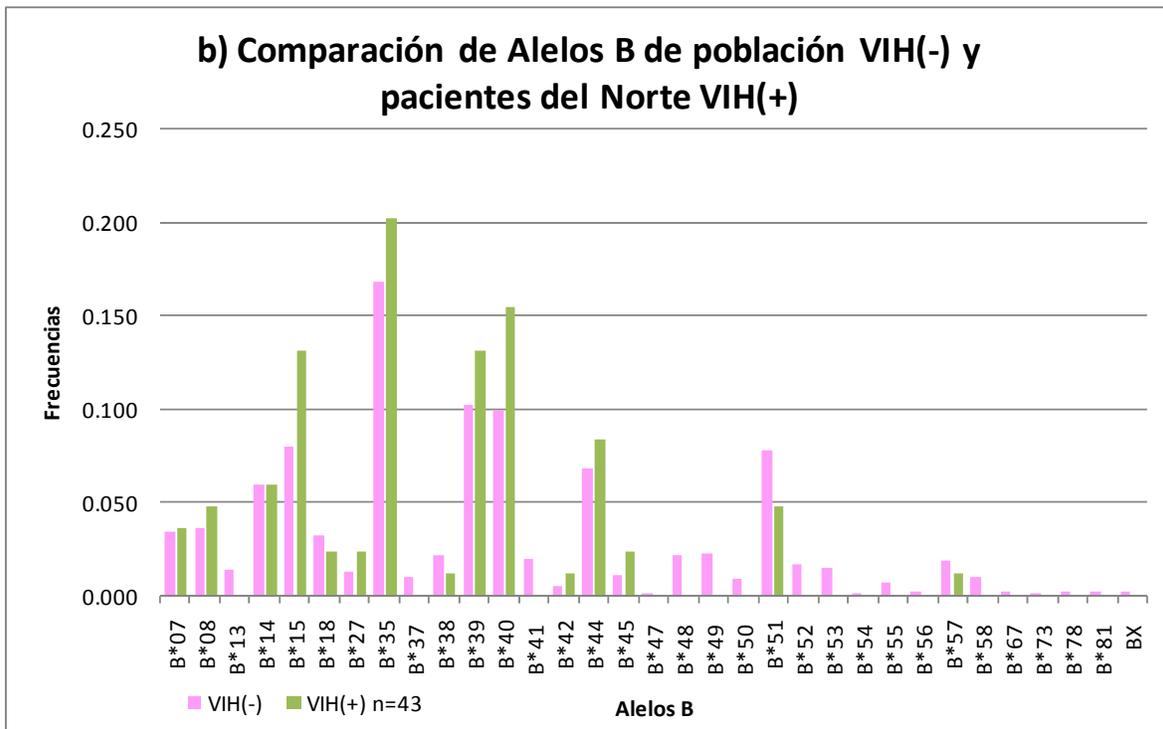
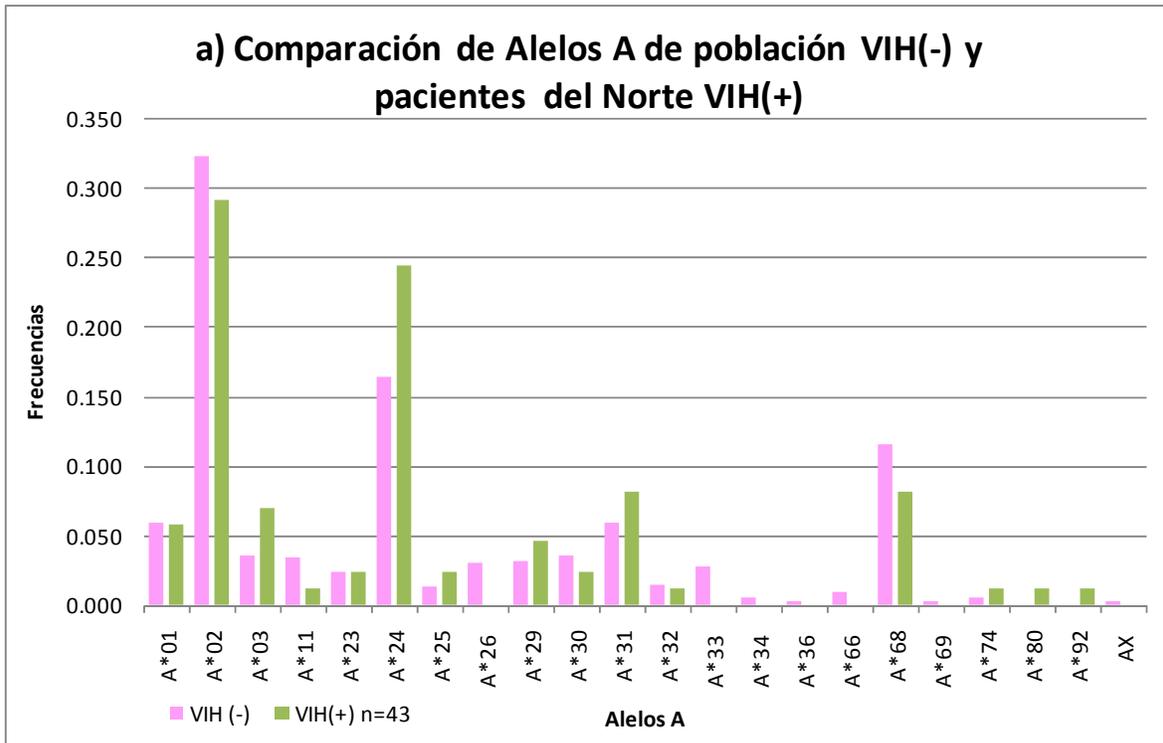
Al comparar la población VIH (-) (n=381) reportada por Barquera y colaboradores [7] con la población de este estudio proveniente de la región Norte del país (n=43) (Figura 6.16), se encontraron alelos como A\*02 y A\*68 que se observaron con mayor frecuencia en la población VIH (-) que en la población VIH (+), pero hay alelos como A\*24, A\*03 y A\*31 que se encontraron con mayor frecuencia en la población VIH (+) del Norte del país (Figura 6.16.a). Para los alelos HLA-B (Figura 6.16.b), solo el alelo B\*51 se encontró con mayor frecuencia en la población VIH (-) que en los pacientes del Norte, mientras que los alelos B\*15, B\*35, B\*39, B\*40 y B\*44 se observaron con mayor frecuencia en los pacientes VIH (+) de dicha región.

Se observaron algunos alelos que fueron exclusivos sólo de la población VIH(-) como son: A\*26, A\*33, A\*34, A\*36, A\*66, A\*69, B\*13, B\*37, B\*41, B\*47, B\*48, B\*49, B\*50, B\*52, B\*53, B\*54, B\*55, B\*56, B\*58, B\*67, B\*73, B\*78 y B\*81. Los alelos exclusivos para los individuos VIH (+) de la región norte fueron A\*80 y A\*92.

### **Comparación de las frecuencias HLA de los individuos VIH (+) de la región Centro con las observadas en una población VIH (-).**

Comparando la población VIH(-) reportada por Barquera y colaboradores [7] con los pacientes VIH(+) de este estudio provenientes del Centro (Figura 6.17), tanto los alelos HLA-A como los alelos HLA-B se encontraron distribuidos de manera homogénea. No se encontraron alelos HLA con frecuencias distintas en las dos poblaciones. Posiblemente, este resultado refleje la inclusión mayoritaria de individuos del centro del país en el análisis de Barquera y colaboradores.

Alelos como A\*01, A\*02 y A\*68 se encontraron con frecuencias ligeramente mayores en la población VIH (-), mientras que alelos como A\*24, A\*30 y A\*31 se observaron con mayor frecuencia en la población VIH (+) del centro del país (Figura 6.17.a).



**Figura 6.16 Comparación de las frecuencias de los alelos HLA-A y HLA-B de los individuos VIH (+) del Norte con una cohorte de individuos VIH (-).** Se muestra una comparación de frecuencias HLA entre la cohorte de individuos VIH (-) descrita por Barquera y colaboradores [7] (n=381) (barras rosas) y los individuos VIH(+) de este estudio, provenientes de la región Norte del país (n=43) (barras verdes) para los alelos HLA-A (a) y HLA-B (b).

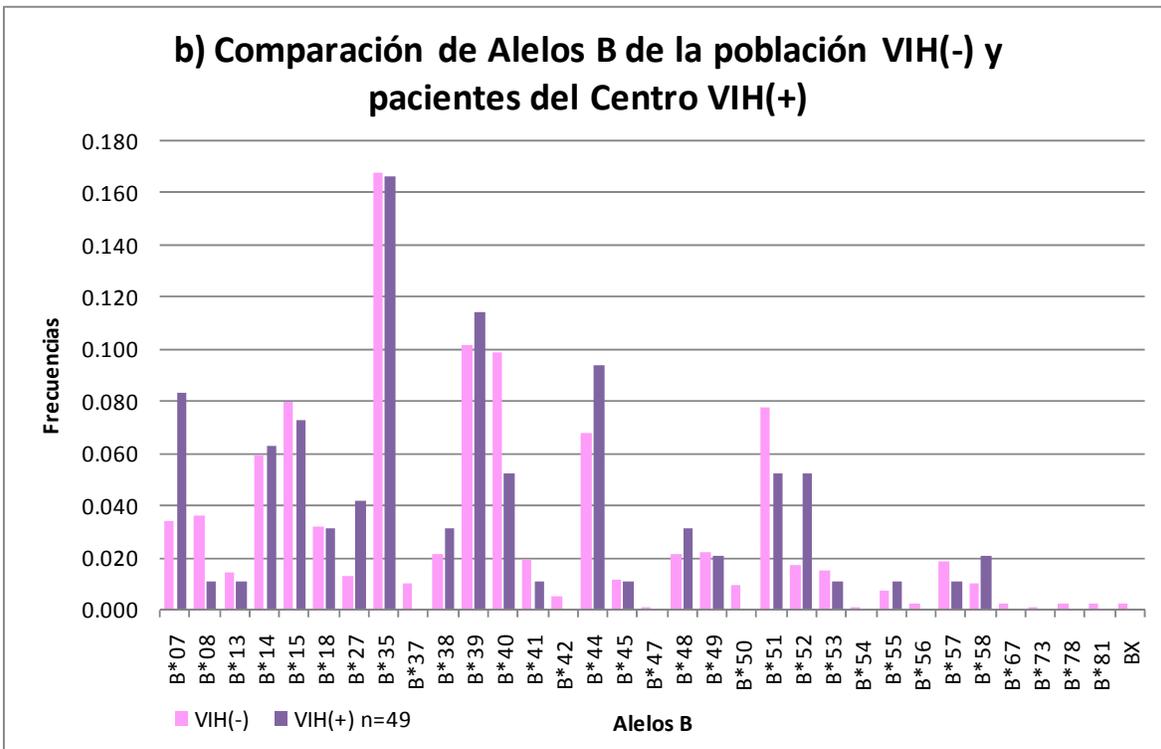
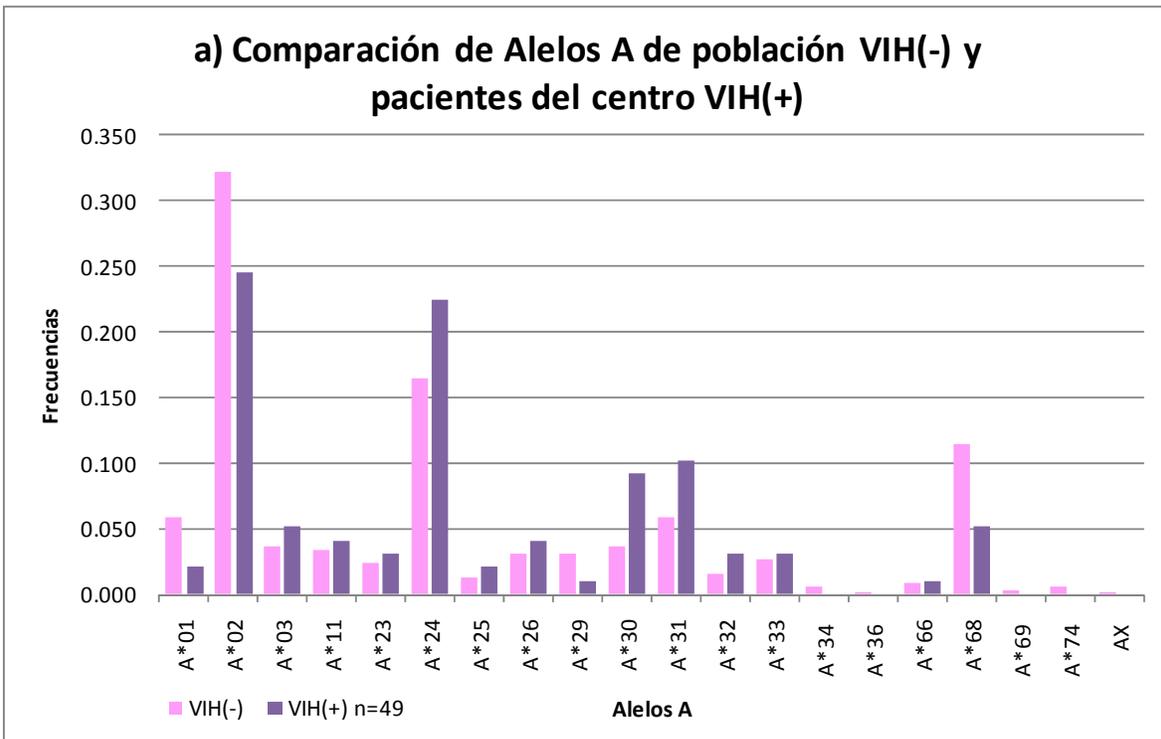
Para los alelos HLA-B (Figura 6.17.b), alelos como B\*08, B\*15 y B\*51 se observaron más frecuentemente en los individuos VIH (-), mientras que B\*07, B\*27, B\*44 y B\*52 se observaron con mayor frecuencia en los individuos VIH (+) del Centro.

Algunos alelos como A\*34, A\*36, A\*69, A\*74, B\*37, B\*42, B\*47, B\*50, B\*54, B\*56, B\*67, B\*73, B\*78 y B\*81 se observaron sólo en la población VIH (-), esto posiblemente es debido al número de individuos del Centro (n=49) incluidos en el presente estudio.

### **Comparación de las frecuencias HLA de los individuos VIH (+) de la región Sur con las observadas en una población VIH (-).**

En la comparación de las frecuencias alélicas de los loci HLA-A y HLA-B (Figura 6.18) de la población VIH(-) reportada por Barquera y colaboradores [7] con las observadas en los individuos VIH (+) del presente estudio provenientes de la región Sur del país (n=47), se observó que los alelos A\*02, A\*31 y A\*68 (Figura 6.18.a) fueron altamente frecuentes en ambas poblaciones. Sin embargo, contrastando con las comparaciones realizadas para las otras regiones en estudio, estos alelos fueron más frecuentes en la población VIH (+) de la región Sur que en la población VIH (-).

Dentro de los alelos HLA-B (Figura 6.18.b), B\*35, B\*39 y B\*40 fueron altamente frecuentes en ambas poblaciones, pero B\*39 y B\*40 fueron más frecuente para los individuos VIH (+) provenientes del Sur del país. Lo contrario ocurrió con los alelos B\*07, B\*35, B\*44 y B\*51 que fueron más frecuentes en la población VIH (-).



**Figura 6.17 Comparación de las frecuencias de los alelos HLA-A y HLA-B de los individuos VIH (+) del Centro con una cohorte de individuos VIH (-).** Se muestra una comparación de frecuencias HLA entre la cohorte de individuos VIH (-) descrita por Barquera y colaboradores [7]

(n=381) (barras rosas) y los individuos VIH(+), provenientes de la región Centro del país (n=49) (barras moradas) para los alelos HLA-A (a) y HLA-B (b).

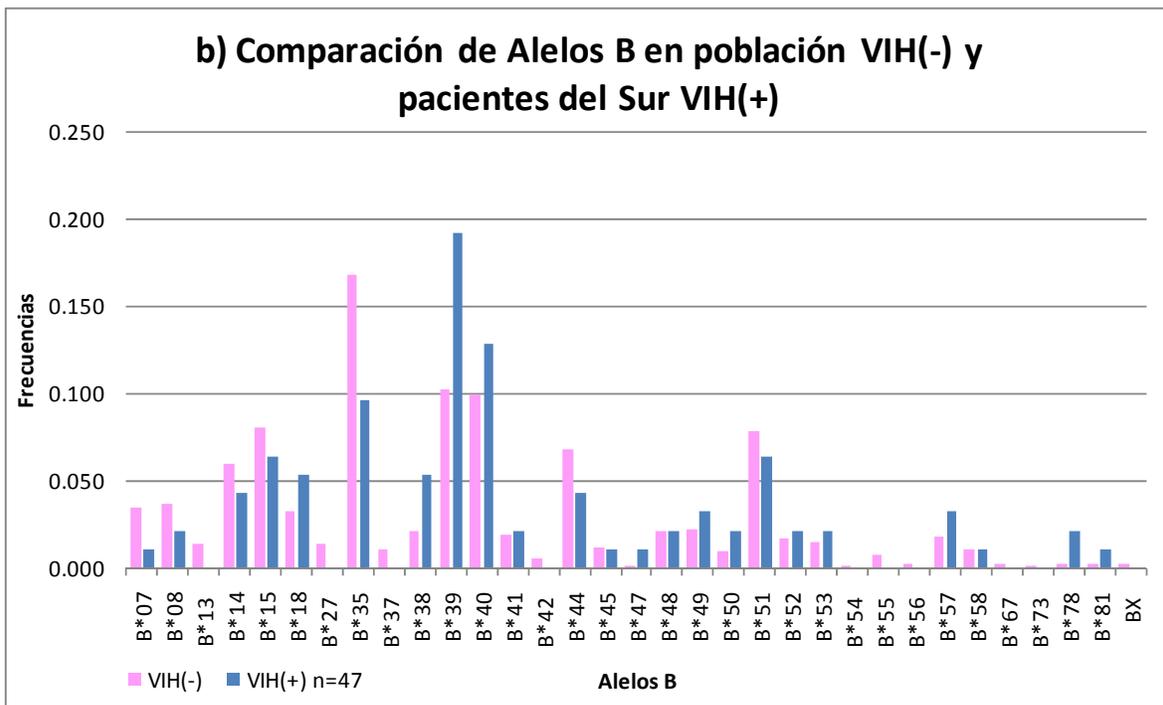
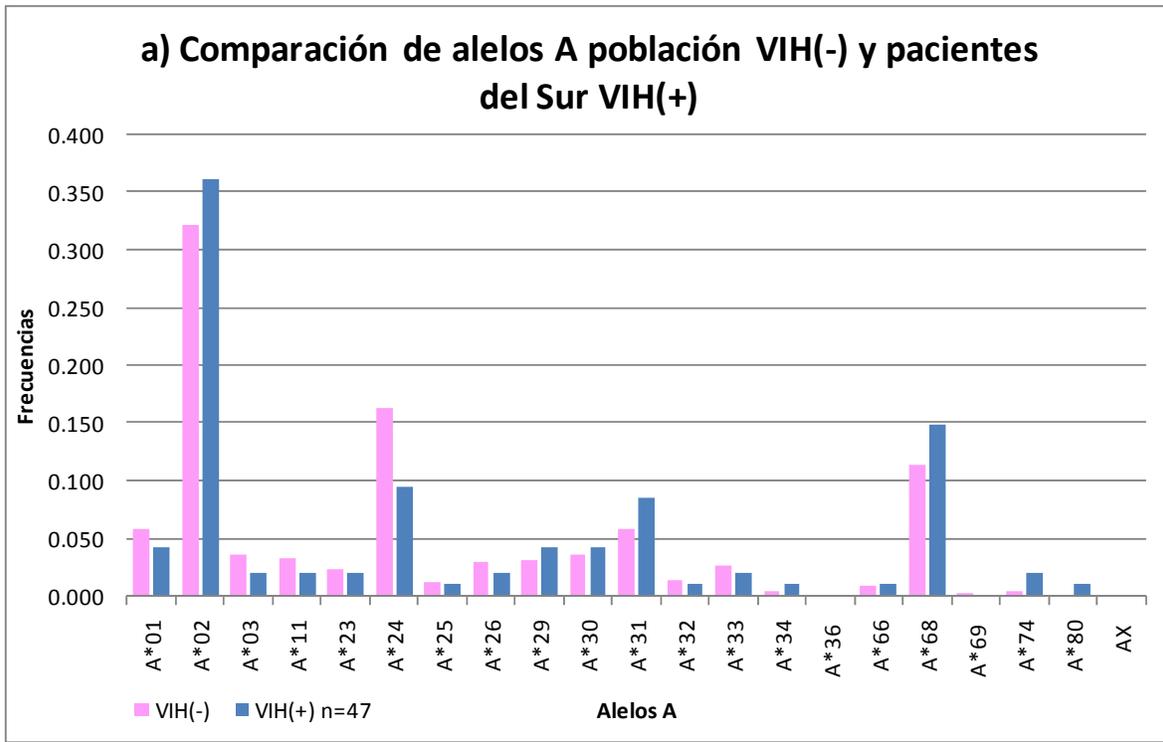


Figura 6.18 Comparación de las frecuencias de los alelos HLA-A y HLA-B de los individuos VIH (+) del Sur con una cohorte de individuos VIH (-). Se muestra una comparación de

frecuencias HLA entre la cohorte de individuos VIH (-) descrita por Barquera et al [7] (n=381) (barras rosas) y los individuos VIH (+) de este estudio, provenientes de la región Sur del país (n=47) (barras azules) para los alelos HLA-A **(a)** y HLA-B **(b)**.

En esta comparación los alelos exclusivos de la población VIH (-) fueron: A\*36, A\*69, B\*13, B\*27, B\*42, B\*54, B\*55, B\*56, B\*67, B\*73; mientras que sólo el alelo A\*80 fue exclusivo para los individuos VIH (+) de la región Sur del país.

### ***COMPARACIÓN DE LAS FRECUENCIAS ALÉLICAS HLA DE LA POBLACIÓN MEXICANA VIH (+) EN ESTUDIO CON UNA POBLACIÓN VIH (+) CAUCÁSICA.***

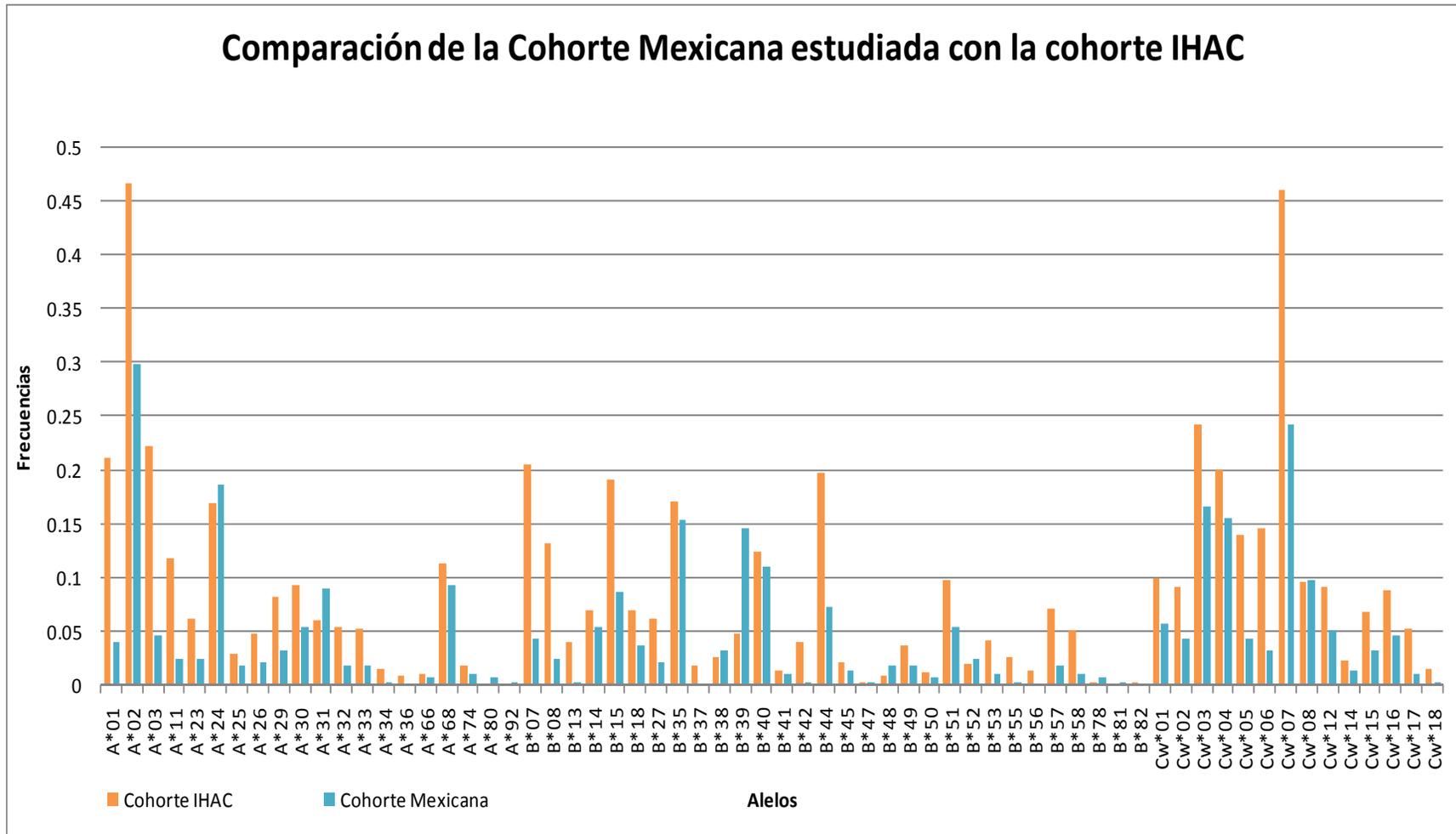
Se hizo una comparación de frecuencias alélicas HLA entre la cohorte multicéntrica IHAC (International HIV Adaptation Combined) que incluye 1845 individuos donde predominan los caucásicos de: Canadá, poniente de Australia y los Estados Unidos [30, 39] y la cohorte de individuos mexicanos en estudio (Figura 6.19).

Los alelos A\*01, A\*02, A\*03, A\*11, B\*07, B\*08, B\*15, B\*44, Cw\*03, Cw\*04, Cw\*05, Cw\*06 y Cw\*07, fueron más prevalentes en los individuos de la cohorte IHAC, los que son consistentes con los alelos con mayor frecuencia encontrados en poblaciones caucásicas [5, 7].

Mientras que para la cohorte mexicana incluida en este estudio, los alelos que ya se han mencionado antes, con mayor frecuencia en esta población fueron: A\*02, A\*24, A\*68, B\*35, B\*39, B\*40, Cw\*07, Cw\*03 y Cw\*08, consistentes también con con los típicos alelos amerindios [5, 7].

Los alelos A\*24, A\*31 y B39 presentaron frecuencias mayores en la cohorte mexicana que en la cohorte IHAC, lo que es de esperarse ya que son considerados alelos amerindios [7].

Además, se observaron alelos que sólo se encontraron en la cohorte IHAC como A\*36, B\*37, B\*56 y B\*82, mientras que los alelos A\*80, A\*92 y B\*81 sólo se encontraron en la cohorte VIH (+) mexicana.



**Figura 6.19 Comparación de frecuencias HLA entre la cohorte mexicana en estudio y una cohorte de individuos VIH (+) predominantemente caucásicos.** Se compararon frecuencias HLA poblacionales de la cohorte de individuos VIH (+) mexicanos del presente estudio (n=139) (barras azules), con una cohorte multicéntrica de individuos VIH (+) predominantemente caucásicos provenientes de Canadá, Estados Unidos y Australia (cohorte IHAC) (n=1845) (barras amarillas). Se muestran las frecuencias poblacionales para los alelos HLA-A, B y C observados en ambas cohortes ordenadas por alelo HLA. [30, 39].

## ***ANÁLISIS DE DESEQUILIBRIO DE LIGAMIENTO***

Con ayuda de las herramientas de inmunología de la página de Los Álamos (ver Metodología), se realizó un análisis de desequilibrio de ligamiento entre los tres alelos HLA, para la cohorte en estudio, en alta y baja resolución. El fenómeno de desequilibrio de ligamiento es común en la herencia de los genes HLA, haciendo que ciertos alelos tiendan a heredarse juntos. El conocer este fenómeno en la cohorte de individuos mexicanos del presente estudio era valioso para los análisis evolutivos, ya que el desequilibrio de ligamiento puede confundir asociaciones entre la expresión de alelos HLA determinados y la presencia de polimorfismos característicos en la secuencia viral al obtenerse asociaciones espurias entre el alelo HLA en desequilibrio de ligamiento con el alelo verdaderamente selector y polimorfismos en la secuencia viral. El análisis se llevó a cabo calculando la frecuencia observada en la población de tener un alelo HLA P(HLA1) u otro P(HLA2); la frecuencia de tener ambos P(HLA1 & HLA2) y la frecuencia esperada si estos alelos se heredan de manera independiente uno de otro E(P(HLA1 & HLA2)), evaluando posteriormente si la frecuencia observada de tener ambos alelos HLA era significativamente diferente a la frecuencia esperada para dichos alelos en caso de que se heredaran de manera independiente, evaluando el nivel de significancia mediante pruebas exactas de Fisher.

En el análisis a alta resolución se encontró que los alelos B\*3417/Cw\*0401 y A\*6803/B\*3905 (Tabla 6.7), tienden a heredarse juntos en la cohorte mexicana en estudio. Para el análisis en con los alelos en baja resolución se encontró desequilibrio de ligamiento entre los alelos B\*35/Cw\*04, B\*40/Cw\*03, B\*14/Cw\*08, B\*15/Cw\*01 (Tabla 6.7). Ninguno de los alelos que se observaron en desequilibrio de ligamiento a alta resolución coincidieron con los observados a baja resolución, posiblemente porque algunos individuos de la región Centro del país se tipificaron exclusivamente a baja resolución, excluyéndose por lo tanto del análisis a alta resolución. Además, al aumentar la resolución de la tipificación HLA se multiplica el número de grupos y de comparaciones posibles, disminuyendo el poder de las pruebas estadísticas para detectar asociaciones entre las frecuencias de dos alelos HLA particulares en cohortes pequeñas.

Tabla 6.7 Alelos HLA con desequilibrio de ligamiento en la cohorte mexicana estudiada†

Alelos HLA a alta resolución										
HLA1	HLA2	Valor de p	HLA1(+) HLA(+)	HLA1(+) HLA2(-)	HLA1(-) HLA2(+)	HLA1(-) HLA2(-)	P (HLA1).	P(HLA2)	P(HLA1 & HLA2)	E (P(HLA1 & HLA2))
B*3517	Cw*0401	2.38E-07	9	0	11	70	0.1	0.2222	0.1	0.0222
A*6803	B*3905	7.42E-06	5	1	3	81	0.0667	0.0889	0.0556	0.0059
Alelos HLA a baja resolución										
HLA1	HLA2	Valor de p	HLA1(+) HLA(+)	HLA1(+) HLA2(-)	HLA1(-) HLA2(+)	HLA1(-) HLA2(-)	P (HLA1).	P(HLA2)	P(HLA1 & HLA2)	E (P(HLA1 & HLA2))
B*35	Cw*04	3.09E-11	17	4	4	65	0.2333	0.2333	0.1889	0.0544
B*40	Cw*03	6.69E-10	20	3	10	57	0.2556	0.3333	0.2222	0.0852
B*14	Cw*08	2.83E-09	9	0	5	76	0.1	0.1556	0.1	0.0156
B*15	Cw*01	4.51E-05	7	9	2	72	0.1778	0.1	0.0778	0.0178

† El análisis se realizó con las herramientas de Inmunología de la Base de Datos de Los Álamos con tipificación HLA a alta y baja resolución, en la cohorte de individuos VIH (+) mexicanos en estudio (n=139). Número de comparaciones de dos vías =  $A*B+A*C+B*C = 1191$ , valor de p para dos vías =  $0.57(A*B+A*C+B*C)=4.58e-05$  para un nivel de confianza del 95%. Signo (+): se tiene el alelo HLA, (-) no se tiene el alelo HLA.

P (HLA) Frecuencia observada en la población de tener el alelo HLA1.

P (HLA2) Frecuencia observada en la población de tener HLA2.

P (HLA1 &HLA2) Frecuencia observada en la población de que tienen HLA1 y HLA2.

E (P (HLA1 & HLA2)) Frecuencia esperada de tener HLA1 y HLA2 si son independientes uno de otro.

## ***ANÁLISIS DE LA EXPRESIÓN DE ALELOS ESPECÍFICOS HLA CON LOS VALORES DE CARGA VIRAL (CV).***

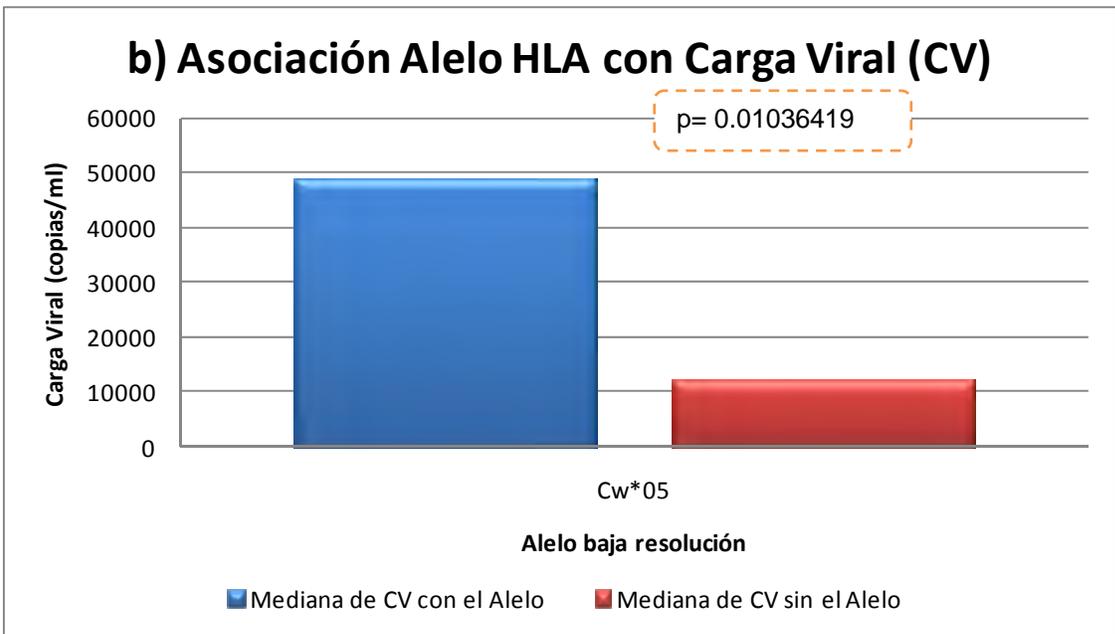
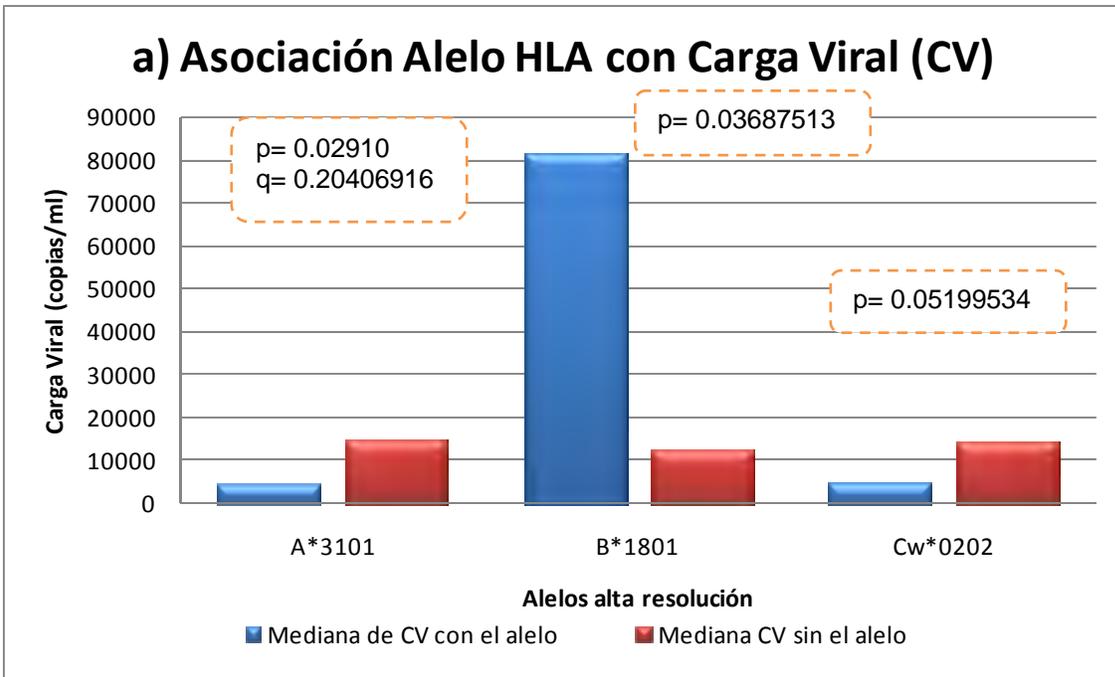
Debido a que la cohorte en estudio se conformó con individuos con cuentas de células T CD4+/ $\mu\text{l}$  > 500, se puede estimar que la mayor parte de los individuos participantes se encontraba en etapas tempranas de la infección, con valores de CV cercanos al punto de equilibrio (set-point) establecido al inicio de la etapa crónica. Debido a la correlación existente entre valores altos y bajos de CV en el punto de equilibrio y la expresión de alelos HLA específicos, observada previamente para distintas poblaciones [8], se buscaron correlaciones HLA – CV en la cohorte mexicana. Este análisis permitió buscar alelos HLA protectores o de riesgo para el avance de la enfermedad por VIH en el contexto de la población mexicana.

El análisis de correlación HLA – CV se realizó con el programa “R” v2.8.1, aplicando pruebas de Man – Whitney a la mediana de CV en presencia y ausencia de cada alelo HLA a alta y baja resolución (Figura 6.20).

En el análisis de correlación HLA – CV a alta resolución (Figura 6.20.a) se encontraron tres alelos con  $p < 0.05$ , pero uno sólo mostró estar cerca del nivel de significancia establecido por tasa de descubrimientos falsos ( $q < 0.2$ ). El alelo A\*3101 se encontró asociado con una mediana de CV más baja que la del resto de la cohorte ( $p=0.0291$ ,  $q=0.2041$ ), correspondiendo así a un posible alelo protector en el contexto de la población mexicana.

El alelo B\*1801 mostró una tendencia de asociación con valores de CV más altos que la mediana de la cohorte ( $p=0.0369$ ), constituyendo un posible alelo de riesgo para el avance de la enfermedad por VIH en la población mexicana. (Figura 6.20.a).

Tanto el alelo A\*3101 como B\*1801, se encontraron más expresados en la región Sur del país (Figura 6.7.a y b respectivamente).



**Figura 6.20 Correlación HLA - CV a alta y baja resolución para la cohorte en estudio.** Se muestran las correlaciones cercanas al nivel de significancia establecido ( $p < 0.05$ ,  $q < 0.2$ ) que se encontraron en el análisis de correlación HLA - CV para los alelos HLA en **a)** alta resolución y **b)** baja resolución. El análisis se realizó con pruebas de Man-Whitney comparando la mediana de Carga Viral (CV) con la presencia (barras azules) o ausencia (barras rojas) del alelo HLA, utilizando el programa "R" v2.8.1, con el paquete q value v1.1.

El alelo Cw\*0202 también presentó una tendencia de correlación con valores de CV bajos ( $p=0.0520$ ) (Figura 6.20.a), por lo que podría constituir un alelo protector en la cohorte en estudio. En la Figura 6.7.c se muestra que este alelo se encontró con mayor frecuencia en la región Norte del país.

No se encontró ningún alelo HLA a baja resolución correlacionado significativamente con valores altos o bajos de CV. Sin embargo, el alelo Cw\*05, presentó una tendencia de asociación con valores de CV altos, constituyendo un posible grupo alélico de riesgo en la población mexicana (Figura 6.20.b). Las Figuras 6.8 y 6.12 muestran que este alelo no es prevalente en la región Sur del país. En la Figura 6.14, se muestra una alta prevalencia del alelo Cw\*05

## ***ANÁLISIS EVOLUTIVOS***

### **Análisis de selección positiva por modelos de codones**

Utilizando secuencias del gen *pol* de 123 individuos incluidos en el estudio, (ya que de los 16 individuos restantes no se contaba con su secuencia del gen *pol* en la base de datos del CIENI); se realizó una búsqueda de posiciones con evidencia de estar sujetas a selección positiva (Tabla 6.8), utilizando el modelo molecular de evolución de codones, basado en tasas de sustituciones sinónimas (dS) y no sinónimas (dN), disponible en el servidor Datamonkey ([www.datamonkey.org](http://www.datamonkey.org)), utilizando el análisis SLAC (ver Metodología).

Tabla 6.8 Posiciones del gen *pol*/viral bajo selección positiva en la cohorte mexicana.\*

POSICIÓN	NORTE		CENTRO		SUR		TODOS	
	dN - dS Normalizado	Valor de p	dN - dS Normalizado	Valor de p	dN - dS Normalizado	Valor de p	dN - dS Normalizado	Valor de p
12 PR	2.58253	0.0555528	2.26551	0.0167857	1.81305	0.0386478	1.9876	0.0001862
19 PR	NS	-	NS	-	NS	-	1.25565	0.0252072
35 PR	NS	-	NS	-	NS	-	1.07412	0.0843293
37 PR	3.53533	0.0709183	NS	-	2.42243	0.0720365	2.43782	0.00125653
39 PR	NS	-	NS	-	NS	-	0.65311	0.0379266
57 PR	NS	-	NS	-	1.63038	0.070359	0.868864	0.0645422
63 PR	NS	-	NS	-	2.35695	0.0958666	1.73174	0.0373661
77 PR	2.50781	0.0834536	2.40105	0.0216181	NS	-	2.40734	9.97E-05
93 PR	NS	-	2.44775	0.0128487	2.90818	0.00571011	2.07504	3.38E-05
35 RT	3.36838	0.0497455	2.03579	0.0565163	2.77256	0.012811	2.06128	0.00054879
102 RT	NS	-	NS	-	NS	-	0.643473	0.0532981
135 RT	NS	-	3.35988	0.0682082	NS	-	1.5931	0.0880826
165 RT	NS	-	NS	-	NS	-	0.485579	0.0884551
200 RT	4.04402	0.0122319	2.47348	0.0296661	2.68224	0.0227467	2.49982	7.64E-05
211 RT	3.54855	0.043934	NS	-	NS	-	2.24927	0.00436106
272 RT	NS	-	3.15891	0.00335145	NS	-	2.43853	7.55E-05
286 RT	2.89002	0.0390184	NS	-	NS	-	NS	-
288 RT	NS	-	2.26889	0.0375699	NS	-	NS	-
292 RT	NS	-	NS	-	1.55079	0.0712001	NS	-
293 RT	NS	-	NS	-	2.31055	0.0486011	1.71358	0.0168941
297 RT	4.4236	0.0150888	NS	-	4.67748	0.000835209	2.92472	7.33E-05
322 RT	NS	-	NS	-	NS	-	0.943855	0.0553404
334 RT	NS	-	3.65289	0.0363833	4.01929	0.0092875	2.41851	0.00865297

\* Estableciendo un nivel de significancia de  $p < 0.1$ . NS: posición que no tienen la significancia en dicha región.

El análisis de detección de posiciones en el genoma viral bajo presión selectiva positiva se realizó para cada una de las regiones del país en estudio (Norte, Centro y Sur) y en la cohorte completa, con el objeto de determinar posibles huellas de evolución característica del VIH en las diferentes regiones geográficas de México. Es sabido que una proporción importante de las presiones selectivas que actúan sobre el VIH están asociadas a respuestas restringidas por HLA [25-27, 40]. Por esta razón, es factible pensar que diferentes frecuencias de alelos HLA en las diferentes regiones del país podrían determinar características únicas de evolución viral dentro del VIH circulante en México.

En el análisis en el que se incluyen las tres regiones en estudio, se encontraron 20 diferentes posiciones del gen *pol* bajo selección positiva, de las cuales 9 estaban dentro del gen de la Proteasa (PR) y 11 dentro del gen de la retrotranscriptasa (RT). Seis de estas posiciones aparecieron sólo cuando se incluían a todas las regiones, es decir, no se comparten con ninguna región, lo cual puede explicarse posiblemente por un aumento en el poder estadístico para detectar selección positiva al incluir un número mayor de secuencias en el análisis. Para la región Norte, se encontraron 8 posiciones bajo presión selectiva positiva con significancia, 3 de las cuales se encontraron dentro del gen de la PR, mientras que las 5 restantes se encontraron dentro del gen de la RT. Dos de estas ocho posiciones fueron únicas en esta región.

Para la región Centro, se encontraron 9 posiciones bajo selección positiva, tres de ellas dentro del gen de la PR y las 6 restantes dentro del gen de la RT. Dentro de estas 9 posiciones, solo 3 fueron específicas de la región, las cuales se encuentran dentro de la RT.

En la región Sur se encontraron 11 posiciones bajo selección positiva: 5 dentro de la PR y 6 dentro de la RT. Cuatro de estas posiciones fueron exclusivas de la región Sur del país, de las cuales: 2 se encuentran en la PR y 2 en la RT.

### **Análisis de asociaciones alelo HLA – polimorfismos del VIH.**

La diversidad de distribuciones de frecuencias de alelos HLA en distintas poblaciones es bien conocida [8]. Estudios previos han reportado diferencias en la distribución de frecuencias de los alelos HLA incluso entre diferentes regiones dentro del territorio nacional [7]. En el presente estudio se mostró también la existencia de posibles diferencias en las distribuciones de frecuencias de los genes HLA en las regiones norte, centro y sur del país (Figuras 6.4, 6.5 y 6.6). Recientemente, el grupo de investigación del CIENI demostró que diferencias en las distribuciones de frecuencias HLA pueden determinar características únicas de evolución del VIH en diferentes poblaciones infectadas al comparar la evolución mediada por HLA del VIH en una cohorte mexicana y una cohorte predominantemente caucásica [5]. Considerando estas observaciones, la existencia de diferencias en las distribuciones de alelos HLA dentro de las poblaciones mestizas mexicanas de diferentes regiones del país podría determinar diferencias en la evolución mediada por HLA del virus dentro del territorio nacional. Por esta razón, en el presente estudio se realizaron análisis de asociación entre la expresión de alelos HLA específicos y la presencia de polimorfismo en las posiciones del gen *pol* del VIH, en una cohorte de individuos mexicanos del norte, centro y sur de México. Las comparaciones regionales se hicieron con pruebas exactas de Fisher para los alelos HLA en alta y baja resolución.

En la Tabla 6.9 se muestran los alelos HLA asociados a una posición en el gen *pol* del VIH. En total se identificaron 24 asociaciones HLA – codón de VIH a alta resolución y 33 a baja resolución, para el total de los individuos de la cohorte en estudio con  $p < 0.01$ . Ningún alelo en alta resolución alcanzó el valor de razón de descubrimientos falsos (FDR) establecido para determinar significancia ( $q < 0.2$ ), posiblemente debido al número limitado de individuos incluidos en la cohorte. Sin embargo, varias posiciones del gen *pol* viral mostraron tendencias importantes de asociación con alelos HLA que deberán confirmarse en estudios posteriores.

De las posiciones identificadas bajo selección positiva (Tabla 6.8), sólo se observaron dos que se pudieron explicar por presión por HLA a alta resolución: RT 135 y 292, asociadas a los alelos HLA A\*2301 y B\*4002 y 3 posiciones por presión

por HLA a baja resolución: PR 19, para el alelo A\*11 y Cw\*12 y en la RT 297 por el alelo Cw\*06

En el análisis de asociación con alelos HLA a baja resolución, sólo se encontró una posición en el gen *pol* viral asociada significativamente a HLA. En la posición 283 de la RT se encontró asociada al alelo A\*03 con la presencia de una arginina ( $p=2.00e-05$  y  $q=0.11568042$ ) (Tabla 6.9). Esta posición no se identificó bajo presión selectiva positiva en el análisis por modelo de codones.

**Tabla 6.9. Asociaciones HLA - polimorfismo del VIH, con alelos HLA en alta y baja resolución**

Alelo HLA	AA	Posición	HLA(+) AA(+)	HLA(+) AA(-)	HLA(-) AA(+)	HLA(-) AA(-)	Valor de p
<b>Cw*0401</b>	E	PR 23	16	3	77	0	0.00678191
<b>Cw*0801</b>	T	PR 29	4	6	72	14	0.00492465
<b>Cw*0304</b>	K	PR 49	15	3	78	0	0.00571109
<b>B*4403</b>	T	PR 80	3	3	6	84	0.0101246
<b>Cw*0303</b>	L	PR 80	5	1	14	76	0.00099508
<b>B*3906</b>	G	PR 85	8	2	86	0	0.00986842
<b>Cw*0802</b>	V	PR 88	5	6	7	78	0.00399185
<b>B*1801</b>	V	RT 88	4	4	8	80	0.00779674
<b>Cw*0304</b>	K	RT 120	13	5	74	4	0.01059194
<b>A*2301</b>	I	RT 135	3	3	4	86	0.00443739
<b>Cw*0801</b>	D	RT 140	2	8	58	26	0.00391701
<b>A*2301</b>	I	RT 152	0	6	57	32	0.00317645
<b>Cw*0202</b>	I	RT 152	0	6	57	32	0.00317645
<b>B*4002</b>	I	RT 212	11	3	81	1	0.00928639
<b>A*2301</b>	T	RT 217	0	6	64	25	0.00084717
<b>B*4403</b>	A	RT 217	4	2	13	76	0.00879313
<b>B*3501</b>	E	RT 224	4	4	6	80	0.00386245
<b>B*1801</b>	Q	RT 224	3	5	74	12	0.00440539
<b>A*0206</b>	V	RT 262	10	12	57	17	0.00764311
<b>B*4002</b>	K	RT 292	11	3	81	1	0.00928639
<b>Cw*0602</b>	E	RT 308	4	2	90	0	0.00328947
<b>A*0301</b>	K	RT 328	3	3	87	3	0.00259988
<b>A*0301</b>	R	RT 328	3	3	4	86	0.00443739
<b>ALELOS HLA BAJA RESOLUCIÓN</b>							
<b>A*11</b>	V	PR 19	5	2	14	101	0.00095437
<b>Cw*12</b>	V	PR 19	6	6	13	96	0.00350911
<b>B*44</b>	E	PR 41	8	10	81	23	0.00743478
<b>Cw*04</b>	M	PR 42	21	9	84	8	0.00638204

<b>A*01</b>	R	PR 47	1	5	87	30	0.00710471
<b>B*07</b>	M	PR 52	9	2	111	0	0.00745156
<b>A*24</b>	T	PR 69	8	30	4	81	0.00812819
<b>A*11</b>	T	PR 97	5	2	115	1	0.00816533
<b>Cw*03</b>	K	RT 109	26	7	86	3	0.00394021
<b>A*23</b>	I	RT 124	3	4	5	111	0.00569935
<b>A*24</b>	D	RT 129	18	20	63	20	0.00326132
<b>A*23</b>	I	RT 141	0	7	71	44	0.00172841
<b>A*23</b>	T	RT 141	5	2	25	90	0.00975153
<b>A*03</b>	V	RT 147	8	2	112	0	0.00609673
<b>B*18</b>	T	RT 148	4	5	7	106	0.00370314
<b>Cw*07</b>	A	RT 164	43	5	74	0	0.00826051
<b>B*15</b>	S	RT 168	9	11	79	23	0.00559425
<b>B*44</b>	S	RT 168	18	0	70	34	0.0030527
<b>A*68</b>	M	RT 184	8	15	10	90	0.00585868
<b>Cw*08</b>	I	RT 185	5	19	3	95	0.00741422
<b>A*23</b>	T	RT 206	1	6	84	30	0.00275434
<b>Cw*06</b>	I	RT 206	3	2	10	105	0.00891865
<b>Cw*07</b>	A	RT 206	14	33	7	66	0.00644234
<b>Cw*03</b>	K	RT 281	29	4	89	0	0.00465879
<b>A*03</b>	R	RT 283	10	1	27	85	2.00E-05 ■
<b>A*03</b>	K	RT 283	2	9	87	25	0.00014104
<b>A*30</b>	R	RT 287	5	9	6	103	0.00287112
<b>Cw*05</b>	R	RT 287	4	6	7	105	0.00587175
<b>Cw*06</b>	E	RT 297	4	2	116	0	0.00203224
<b>A*03</b>	Q	RT 300	3	8	3	108	0.00956575
<b>Cw*04</b>	T	RT 302	26	4	92	0	0.00312009
<b>A*11</b>	V	RT 323	3	4	110	6	0.00071505
<b>A*11</b>	A	RT 323	4	3	9	107	0.00228392

♪ Los análisis de asociación se realizaron con el programa estadístico R<sup>®</sup> v2.8.1, usando el paquete q value v1.1. Los análisis de asociación se realizaron con alelos HLA tipificados a alta y baja resolución como se indica en la sección de Metodología, en un fragmento del gen *pol*, que incluye 99 posiciones de la proteasa (PR) y 335 posiciones la retrotranscriptasa (RT), en 123 individuos mexicanos incluidos en este estudio. HLA (+): presenta el alelo HLA, HLA (-): no presenta el alelo HLA, AA (+): que presenta el aminoácido y AA (-): no presenta el aminoácido en la posición en cuestión.

■ es el único alelo HLA que tiene una  $q < 0.2$  ( $q = 0.11568042$ ).

Interesantemente, varios de los alelos HLA asociados a polimorfismos de VIH específicos estuvieron asociados a la ausencia de un AA particular en la posición en cuestión. Por ejemplo: A\*2301 se asoció con la ausencia de treonina en la posición RT 217 y Cw\*0202 se asoció con la ausencia de isoleucina en la posición 152. Es posible que existan asociaciones complementarias que permitan

establecer asociaciones positivas entre la presencia de estos mismos HLA con la presencia de otros AA en la posición de VIH en cuestión. Sin embargo, estas asociaciones pueden no haberse observado por el poder estadístico del análisis, dado el número limitado de individuos incluidos en la cohorte.

Para los alelos a baja resolución se encontraron alelos asociados a la presencia de un AA en una posición particular. Por ejemplo: el alelo Cw\*04 se encuentra asociado a la presencia de metionina en la posición 42 de la PR y el alelo Cw\*07 se encontró asociado a la presencia de alanina en la posición 164 de la RT.

Como ya se ha mencionado antes, estos resultados muestran sólo una tendencia de asociaciones de alelo HLA con un codón en particular del VIH, ya que el número de individuos incluidos en el estudio es limitado para el análisis de evolución.

### **Análisis de asociación entre la presencia de un polimorfismo en el genoma viral y la procedencia de los individuos.**

Se buscaron alguna asociaciones entre la presencia de un aminoácido (AA) y el Estado o la Región de procedencia de los individuos de la cohorte, utilizando el programa estadístico R" v2.8.1. Ninguna de las asociaciones AA – estado ó AA - región fue significativa, ya que los valores de q no alcanzaron el umbral de significancia establecido ( $q < 0.2$ ), sin embargo, se observaron varias posibles tendencias de asociación con valores bajos de de p ( $p < 0.01$ ).

Se encontraron 10 polimorfismos del VIH con una tendencia a expresarse preferentemente en alguno de los estados incluidos en el estudio. Similarmente, se encontraron 5 polimorfismos expresados con mayor frecuencia en alguna de las tres regiones del país analizadas (Tabla 6.10). Por ejemplo: el ácido glutámico en la posición 314 de la RT se encontró mayoritariamente en la región Centro en comparación con las otras dos regiones. Y este mismo AA se encontró con mayor frecuencia en el D. F.

Tabla 6.10 Polimorfismos en el gen *pol* viral con mayor frecuencia en estados o regiones específicas de México  $\alpha$ .

AA	POSICIÓN	ESTADO	AA(+) EDO(+)	AA(+) EDO(-)	AA(-) EDO(+)	AA(-) EDO(-)	Valor de p
D	PR 54	Baja California	4	9	4	106	0.004287969
V	PR 89	DF	4	4	12	103	0.010007
E	RT 140	Jalisco	7	33	2	81	0.005326995
D	RT 140	Jalisco	2	79	7	35	0.007402725
K	RT 228	Oaxaca	6	46	23	48	0.009336655
F	RT 231	Guerrero	4	95	6	18	0.003763911
L	RT 231	Guerrero	6	20	4	93	0.006014459
I	RT 262	Oaxaca	6	3	23	91	0.005388446
E	RT 314	DF	14	49	2	58	0.002350385
K	RT 314	Baja California	5	17	3	98	0.004644164

AA	POSICIÓN	REGIÓN	AA(+) REG(+)	AA(+) REG(-)	AA(-) REG(+)	AA(-) REG(-)	valor de p
K	PR 58	SUR	20	17	24	62	0.007621082
K	RT 228	SUR	11	41	33	38	0.004375511
T	RT 303	NORTE	14	71	17	21	0.00146905
A	RT 303	NORTE	17	22	14	70	0.003149467
E	RT 314	CENTRO	32	31	16	44	0.009333062

$\alpha$  Análisis realizado con el programa "R" v2.8.1, usando el paquete q value v1.1 en la cohorte de n=123 individuos VIH positivos mexicanos del presente estudio. Presencia de aminoácidos (AA) con respecto al estado (EDO) o Región (REG) de los individuos incluidos en el estudio. AA(+) EDO/REG(+): tiene el AA y se encuentra en el Estado/Región; AA(+) EDO/REG(-): tiene el AA, pero no se encuentra en el Estado/Región; AA(-) EDO/REG(+): no tienen el AA, pero se encuentra en el Estado/Región y AA(-) EDO/REG(-): No tiene el AA, ni tampoco se encuentra en el Estado o Región.

### Confirmación de las asociaciones HLA – polimorfismo de VIH mediante mapeo de epítomos de CTLs.

Con ayuda de una secuencia consenso del gen *pol* tomada de [27], con 99 posiciones de la PR y 335 posiciones de la RT (Figura 6.21), se construyó un mapa de epítomos de CTLs donde se colocaron las asociaciones HLA – codón del VIH obtenidas en el presente trabajo (Tabla 6.9). Se muestran epítomos óptimos de CTLs [http://www.hiv.lanl.gov/content/immunology/tables/ctl\\_summary.html](http://www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html) y epítomos predichos por posiciones de anclaje para alelos HLA (Base de datos de los Álamos, Búsqueda de motivos de anclaje, [http://www.hiv.lanl.gov/content/immunology/motif\\_scan/motif\\_scan](http://www.hiv.lanl.gov/content/immunology/motif_scan/motif_scan)).

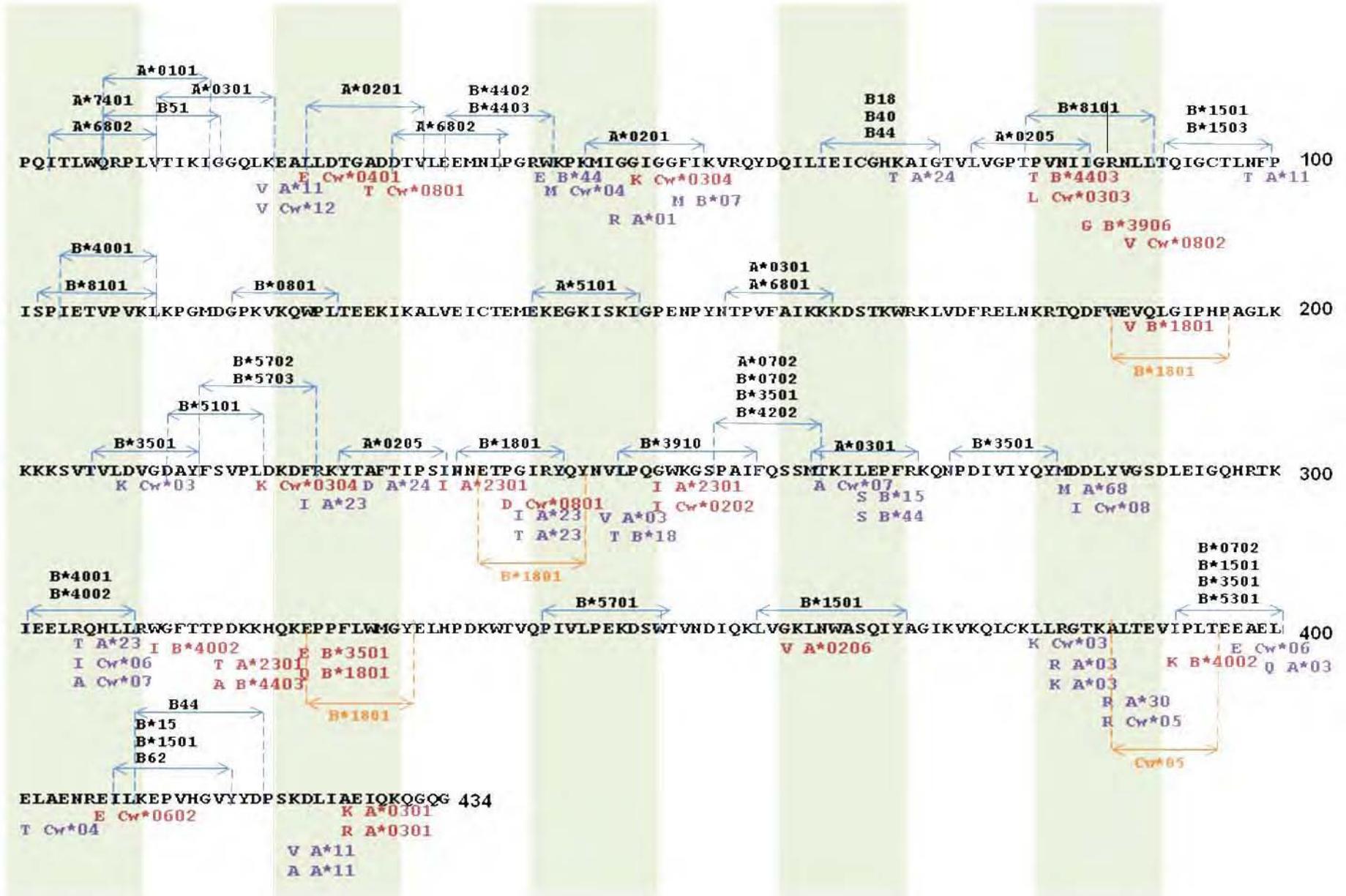
La posición 41 de la PR que se encontró asociada al alelo B\*44 con ausencia del ácido glutámico, esta posición fue la única que coincidió de manera precisa con el epítotope óptimo para el alelo B\*4402 y B\*4403.

Existen tres posiciones que se encontraron muy cercanas a los epítotope óptimos. Una de ellas es la posición 80 de la PR asociada al alelo B\*4403 con la presencia de treonina, muy cercana al epítotope óptimo de B\*44; otra posición es la 148 de la RT asociada al alelo B\*18 con la ausencia del treonina en dicha posición y muy cercana al epítotope óptimo de B\*1801; la última posición muy cercana a un epítotope óptimo es la posición 292 de la RT, asociada al alelo B\*4002 con la presencia de isoleucina y muy cercana al epítotope óptimo B\*4002.

Ninguna de las posiciones encontradas con desequilibrio de ligamiento en esta cohorte, coincidieron con los epítotope óptimos.

Se predijeron epítotope para el alelo B\*1801 que fue un alelo que se observó asociado con CV alta en la cohorte de este estudio (Figura 6.20.a) y que ya se encuentra reportado como un alelo de riesgo en otras poblaciones [8]. Dos de los epítotope predichos para el alelo B\*1801 que se encuentran en las posiciones 87 – 96 y 224 – 223 de la RT, coinciden con las posiciones que se encontraron asociadas a este alelo (Tabla 6.9): la posición 88 de la RT asociada a la ausencia de valina y la posición 224 asociada a la ausencia de glutamina (Figura 6.21). Sin embargo el epítotope predicho de las posiciones 138 – 146 de la RT, no coincidió con la posición 148 de la RT asociada al alelo B\*18 con la ausencia de treonina. Esta última posición puede ser una mutación por procesamiento que ocurrió muy cercana al epítotope óptimo.

Se predijeron epítotope para el alelo Cw\*05, el cual se encontró también como un alelo de riesgo para la cohorte del presente trabajo, al estar asociado con CV alta (Figura 6.20.b). La ausencia de arginina en la posición 287 de la RT asociada con el alelo Cw\*05, coincide con el epítotope predicho de las posiciones 287 – 296 de la RT para este mismo alelo (Figura 6.21).



**Figura 6.21 Mapa de epítopes de CTLs donde se muestran las asociaciones HLA – polimorfismo de VIH observadas en el presente estudio.** La secuencia consenso de las 99 posiciones de la PR y 335 posiciones de la RT del gen *pol* del VIH fue tomada de [27]. Se muestran epítopes óptimos de CTLs con flechas azules por encima de la secuencia, obtenidos de la Base de Datos de Inmunología de Los Álamos ([http://www.hiv.lanl.gov/content/immunology/tables/ctl\\_summary.html](http://www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html)). Por debajo de la secuencia se muestran las posiciones que se encontraron asociadas con alelos HLA en alta resolución (rojo) y las posiciones asociadas a los alelos HLA en baja resolución (morado) en el presente estudio. Se predijeron epítopes con la herramienta de Búsqueda de motivos de anclaje de la base de datos de Inmunología de Los Álamos ([http://www.hiv.lanl.gov/content/immunology/motif\\_scan/motif\\_scan](http://www.hiv.lanl.gov/content/immunology/motif_scan/motif_scan)) los cuales se muestran con flechas de color naranja por debajo de la secuencia.

## 7. DISCUSIÓN

### ***CARACTERÍSTICAS DE LA COHORTE ESTUDIADA***

Para el presente trabajo, se seleccionaron 139 individuos infectados con VIH-1, de la cohorte de Epidemiología y Evolución del Centro de Investigación en Enfermedades Infecciosas, con el objetivo de profundizar en el conocimiento de la evolución dependiente de HLA del VIH en la población mexicana y de buscar posibles asociaciones entre la presencia de alelos HLA específicos y efectos positivos o negativos en la velocidad de avance de la enfermedad por VIH en el contexto inmunogenético de la población mexicana. Los criterios de selección de los individuos participantes en el presente estudio fueron presentar cuentas de células T CD4+ > 500 células/ $\mu$ l de sangre, no estar bajo tratamiento antirretroviral y pertenecer equitativamente a las regiones: Norte, Centro y Sur del país. El valor umbral de cuentas de células T CD4+, garantizó la formación de una cohorte enriquecida en individuos en etapas tempranas de la infección y por tanto, con valores de CV cercanos al *set-point* establecido en el inicio de la fase crónica de la infección. El incluir individuos en etapas tardías de la infección podría haber influido en las asociaciones carga viral – HLA, ya que es sabido que en etapas tardías de la infección por VIH, los valores de carga viral tienden a aumentar al disminuir el control inmune de la replicación viral (Figura 1.7). Bajo estas condiciones el promedio de CD4+ obtenido fue de 590.842 células/ $\mu$ l y de CV de 49352.782 copias de RNA/ml de plasma. Esta cohorte es muy diferente a las cohortes típicas de América Latina [41], donde los individuos infectados por VIH se encuentran en muy malas condiciones de salud, pues acuden a los centros de salud tardíamente con la infección ya muy avanzada, con cuentas de células T CD4+ muy bajas, (cuentas de células T CD4+ < 200 células/  $\mu$ l de sangre) y muchas veces con enfermedades oportunistas, debido a la falta de acceso a los servicios de salud, la mala información con que se cuenta y la discriminación que sufren los individuos infectados.

En total se incluyeron 139 individuos provenientes de las regiones: Norte (n=43); Centro (n=49) y Sur (n=47) de la República Mexicana. Se escogieron estas tres regiones ya que previamente se ha demostrado que dentro del territorio nacional existen características inmunogenéticas diferentes, con un predominio de genes caucásicos en la región Norte del país y un predominio de genes amerindios en las regiones Centro y Sur de la República Mexicana [6, 7]. Dividir la corte en estas regiones permite una visión más amplia de la diversidad inmunogenética en el territorio nacional. El número de individuos incluido en el estudio es pequeño en comparación a otros estudios de diversidad genética que se han hecho hasta el momento [6, 7], por lo que los resultados pueden no tener el impacto esperado, aunque sí mostrar una tendencia de estas diferencias en diversidad genética en distintas regiones del país. Otra limitación con el número de individuos es la que existe en el análisis evolutivo. El tamaño de la muestra es de gran importancia para poder observar las asociaciones HLA – polimorfismos del VIH, ya que al disminuir el tamaño de muestra, disminuye también el poder de detección de asociaciones HLA [42]. Ampliaciones posteriores del número de individuos en la cohorte, permitirán confirmar las observaciones obtenidas en el presente estudio. Con base en un análisis filogenético de un fragmento de la secuencia del gen *pol* de cada uno de los individuos participantes (obtenida de la base de datos del CIENI), se observó que todos los individuos incluidos en el estudio estaban infectados con el VIH-1 subtipo B, lo cual apoya observaciones anteriores que identifican a este subtipo viral como el predominante en los países de América Latina, incluyendo a México [14]. Sin embargo, la presencia de otros subtipos y formas recombinantes no debe descartarse. El número limitado de individuos incluidos en el presente estudio puede reducir la posibilidad de encontrar subtipos y formas recombinantes presentes en bajas frecuencias. Estudios previos con cohortes mexicanas más grandes, como la cohorte del CIENI descrita en Avila – Ríos (n=280), han mostrado la existencia de formas recombinantes en bajas frecuencias entre los virus circulantes en el país [5]. La misma observación se ha realizado en otro estudio del CIENI con una cohorte de 1500 individuos VIH (+), con alta representatividad a nivel nacional (datos no publicados).

El conocer el/los subtipo(s) prevalentes en una población es importante para los análisis de evolución del VIH mediada por HLA a nivel poblacional, ya que la presencia de polimorfismos naturales asociados a subtipos virales específicos pueden llegar a producir asociaciones espurias con alelos HLA sobre representados en individuos infectados con ese subtipo en particular dentro de la cohorte [42]. En el caso de la cohorte mexicana utilizada en el presente estudio, la homogeneidad en el subtipo viral presente facilita el análisis evolutivo.

### ***COMPARACIÓN DE FRECUENCIAS HLA DE CLASE I***

La gran variabilidad de las distribuciones de frecuencias HLA en diferentes poblaciones del mundo hace necesario el estudio de la evolución del VIH en distintas cohortes con diferentes características inmunogenéticas [5]. La riqueza genética de la población mexicana, con su mestizaje característico, abre la posibilidad de estudiar la evolución del virus en poblaciones inmunogenéticamente diferentes dentro del territorio nacional. En la primera fase del presente estudio se estudiaron las frecuencias alélicas HLA a alta y/o baja resolución en una cohorte de individuos VIH positivos mexicanos y se compararon estas frecuencias HLA en subcohortes provenientes de diferentes regiones geográficas del país. Como se ha observado previamente [6], la distribución de frecuencias HLA de la población mexicana refleja un claro mestizaje genético, observándose tanto alelos típicos de poblaciones caucásicas como de poblaciones amerindias y negras [7].

Algunos ejemplos de alelos que se encontraron con mayor frecuencia en la cohorte mexicana en este estudio (Figura 6.3) y que muestran la presencia de esta diversidad alélica son: A\*02 y A\*68 que ya se han reportado como alelos amerindios pero que son predominantes también en la población africana y caucásica [43] y el alelo A\*24 que se observa con mayor frecuencia en poblaciones nativas americanas y asiáticas [43]. Para los alelos HLA-B que se encontraron muy frecuentes en la cohorte en estudio, también se observa este mestizaje. Los alelos B\*35 y B\*39 son alelos considerados amerindios y son muy frecuentes en los hispanos y nativos americanos, mientras que el alelo B\*40 se observa más frecuentemente en población asiática. Para los alelos HLA-C, el

alelo Cw\*07 se encuentra con mayor frecuencia en la población Africana; el alelo Cw\*03 es un alelo que predomina en la población hispana y es considerado un alelo amerindio y el alelo Cw\*04 es un alelo asiático [43].

Estos alelos que se comparten con otras poblaciones son resultado del mestizaje que ha sufrido la población mexicana a lo largo de su historia, ya que México ha sufrido intervenciones por parte de países europeos como Francia y España, los cuales contaban con gente africana como esclavos. Así, que al permanecer en el país aportaron sus genes a los nativos mexicanos.

Recientemente, se reportaron diferencias entre las distribuciones de frecuencias HLA en la población mestiza general de diferentes regiones dentro del territorio nacional, con un enriquecimiento en alelos caucásicos en la población norte del país y un predominio de alelos amerindios en la población del sur [7]. Esta tendencia no fue evidente en la cohorte del presente estudio, probablemente debido al número limitado de individuos incluidos en la cohorte o a sesgos de muestreo ya sea en el estudio de Barquera y colaboradores o en el nuestro. Sin embargo, se observaron tendencias que sugieren fuertemente la existencia de diferencias en las características inmunogenéticas de las poblaciones de las distintas regiones geográficas del país.

En el análisis de frecuencias de alelos HLA en baja resolución en la cohorte donde se incluyeron a todos los individuos en estudio (n=139), se observaron como alelos HLA más frecuentes a A\*02, A\*24, A\*68, A\*31, B\*35, B\*39, B\*40, B\*15 y Cw\*07, Cw\*03, Cw\*04, Cw\*08, los cuales son alelos típicamente amerindios, siendo consistentes con las frecuencias HLA observadas en otras cohortes mexicanas [5, 7].

Interesantemente, en las distribuciones de frecuencias alélicas HLA de la región Norte, con tipificación a alta resolución de la cohorte mexicana del presente estudio se observaron alelos HLA típicamente amerindios como los más predominantes en esta región como son los alelos: A\*0201, A\*2402, A\*3101, B\*4002, B\*3517, Cw\*0401, Cw\*0304, Cw\*0702, Cw\*0701, mientras que no se observó una frecuencia mayor de alelos HLA predominantes en poblaciones caucásicas. En la región Centro del país, los alelos HLA a baja resolución

observados con mayor frecuencia fueron consistentemente alelos amerindios. Lo mismo ocurrió para los alelos HLA a alta resolución de los individuos de la región Sur del país.

Estos datos podrían reflejar un fenómeno socioeconómico en cuanto a los individuos más frecuentemente afectados por la infección por el VIH en la población mexicana. Las características de la cohorte del presente estudio reflejan las características típicas de las cohortes de América Latina [41], donde la epidemia se concentra en grupos vulnerables incluyendo individuos con bajos recursos económicos y migrantes a los Estados Unidos de América. Es posible que exista un sesgo de muestreo debido a la concentración de este tipo de individuos en los centros de salud implicados en este estudio. Esta población se caracteriza por estar compuesta mayoritariamente por individuos en su mayoría mestizos que pertenecen a sectores indígenas o rurales de los estados incluidos en el estudio, los cuales, lamentablemente son los que cuentan con menos información a cerca de la infección por VIH.

En la comparación de las regiones Norte y Sur del país, se encontraron diferencias significativas en las frecuencias de algunos alelos HLA como A\*0206, predominante en la región Sur del país y con muy poca frecuencia en la región Norte del país, especialmente en el Estado de Nuevo León y el alelo A\*2402, sobre expresado en la región Norte del país, ambos alelos considerados típicamente amerindios [7]. Dos alelos HLA-C se encontraron sobre expresados en la región Norte del país, el alelo Cw\*0401 es una alelo amerindio, mientras que el alelo Cw\*0501 es predominante de la población caucásica [5]. En esta comparación algunos alelos como: A\*0307, A\*0312, B\*0705, B\*2705, B\*4402, B\*4446, y Cw\*0501, que mostraron tener una tendencia a estar sólo expresados en la región Norte del país, son alelos presentes en cohortes caucásicas, mientras que algunos de los alelos que sólo se presentaron en la región Sur A\*0236, A\*2403, A\*2428, A\*3102, A\*6805, B\*1510, B\*1530, B\*1548, B\*3514, B\*3903, B\*3913, B\*4006, B\*4008, Cw\*0305, Cw\*0704 son alelos amerindios [7].

La marcada diferencia de expresión de alelos HLA observada entre las dos regiones comparadas, muestra que dentro del territorio nacional existen regiones

con características inmunogenéticas diferentes que pueden estar compuestas por alelos típicos de poblaciones amerindias y/o caucásicas. Esta diversidad genética puede determinar características únicas en la evolución del VIH mediada por HLA en diferentes regiones del país. Además, es muy probable que las diferencias encontradas en este estudio para las dos regiones no sean las únicas existentes, ya que el número de individuos por cada región es limitado y no permitiría alcanzar el poder estadístico necesario para detectarlas.

Al comparar las tres regiones incluidas en este estudio, nuevamente se observaron alelos amerindios como predominantes en las regiones Sur y Centro del país. El alelo Cw\*05 se observó con mayor frecuencia en el D. F. a diferencia de lo que se esperaba, no se encontró sobre expresado en los estados de la región Norte del país. Lo mismo ocurrió con el alelo B\*07, considerado como alelo típicamente caucásico, que también estuvo presente con mayor frecuencia en la región Centro del país, con respecto a las otras dos regiones incluidas en el estudio. Sin embargo, el hecho de haber encontrado diferencias entre las distribuciones alélicas de los individuos de diferentes regiones del país presenta la oportunidad de estudiar la posible existencia de características únicas de adaptación del VIH en el contexto de poblaciones inmunogenéticamente distintas dentro del territorio nacional.

En el análisis de desequilibrio de ligamiento se encontraron dos haplotipos con desequilibrio a alta resolución y cuatro haplotipos a baja resolución. Sólo B\*3517/Cw\*0401 en alta resolución coincide con los haplotipos en baja resolución (B\*35/Cw\*04). Los alelos A\*68/B\*39, B\*35/Cw\*04, B\*40/Cw\*03, B\*15/Cw\*01 y B\*14/Cw\*08, encontrados en baja resolución con desequilibrio de ligamiento son algunos de los haplotipos reportados previamente para la población VIH(+) mexicana dentro del estudio de Avila-Ríos [5]. Estas observaciones sugieren que la diversidad de alelos en la población también está influenciada por el fenómeno de desequilibrio de ligamiento, es decir, algunos alelos HLA tienden a heredarse juntos, lo que aumenta la probabilidad de que estos alelos estén presentes en la población mexicana. Asimismo, el comprobar la existencia de alelos HLA en la cohorte mexicana del presente estudio en desequilibrio de ligamiento, hace

importante el considerar este fenómeno en el análisis de evolución mediada por HLA en el VIH a nivel poblacional, ya que algunas asociaciones indirectas HLA - codón de VIH podrían observarse con alelos HLA en desequilibrio de ligamiento con el alelo verdaderamente selector.

Al hacer el análisis de comparación de frecuencias HLA de la cohorte mexicana en su totalidad y de las subcohortes regionales incluidas en el presente estudio, con las observadas en una población de 381 individuos VIH(-), descrita por Barquera y colaboradores [7], se observó la misma tendencia de predominio de los alelos amerindios en la población mexicana descrita arriba. Cabe señalar que el alelo B\*39 se encontró con mayor frecuencia en la población infectada con VIH-1, principalmente en los individuos de la región Sur del país, lo cual apoya las observaciones de Avila – Ríos y colaboradores, en una cohorte de cerca de 300 individuos infectados por VIH del Centro-Sur del país [5]. La posibilidad de que B\*39 represente un factor de riesgo para adquirir la infección por VIH en la población mexicana aún no puede comprobarse, pero la influencia de este altamente frecuente alelo HLA en la evolución mediada por HLA en la población mexicana deberá ser estudiada más a fondo, ya que podría determinar características únicas de evolución del virus en nuestra población. De cualquier manera, la posibilidad de un sesgo de muestreo en las cohortes mexicanas descritas hasta la fecha aún no puede descartarse y deberá comprobarse en estudios futuros con cohortes con mayor tamaño y representatividad.

La ausencia de algunos alelos HLA en la población VIH (+) del presente estudio con respecto a la población VIH (-) del estudio de Barquera y colaboradores podría deberse a que la cohorte VIH (-), es considerablemente más grande que la cohorte VIH (+) en estudio y el poder para observar la presencia de dichos alelos HLA poco frecuentes no fue suficiente. Asimismo, las diferencias entre las dos cohortes podrían deberse a la proveniencia de los individuos que las conforman ya que los estados que se incluyeron en el estudio de la cohorte VIH (-) son predominantemente: Sinaloa, D. F. y Puebla, mientras que en este estudio, se incluyen más estados de cada una de las regiones del país.

Con el fin de poner en evidencia el contexto inmunogenético único de la población mexicana con respecto a otras poblaciones del mundo, se compararon las frecuencias HLA de la cohorte de individuos VIH (+) mexicanos del presente estudio y de una cohorte donde predominaban los individuos VIH (+) caucásicos [30], la cohorte IHAC (n=1845). Como era esperado, se observó una mayor frecuencia de alelos típicamente caucásicos como A\*01, A\*03, A\*11, B\*07, B\*08, B\*13, B\*27, B\*44, B\*51, B\*57, Cw\*05, Cw\*06 y Cw\*07 en la cohorte IHAC, mientras que alelos típicamente amerindios como A\*24 y B\*39 fueron más frecuentes en la cohorte mexicana. Estas observaciones son consistentes con la comparación ya reportada por Avila-Ríos y colaboradores [5].

Las diferencias en frecuencias de alelos HLA entre la cohorte mexicana y la cohorte IHAC comprueba la existencia de características inmunogénicas diferentes y únicas de cada cohorte. La cohorte mexicana caracterizada por un mestizaje genético notable, mostró una marcada diferencia de expresión alelos HLA típicos amerindios dentro del territorio nacional. Esta diferencia con otras cohortes internacionales bien estudiadas, demuestra la importancia de estudios de evolución de VIH mediada por HLA en esta cohorte inmunogenéticamente única.

### ***FACTORES DE RIESGO O PROTECCIÓN DENTRO DE LA COHORTE EN ESTUDIO***

Con la diversidad de alelos HLA dentro de la cohorte en estudio mostrada arriba y la ventaja de que la cohorte estaba compuesta por individuos en etapas tempranas de la infección, se buscó la existencia de asociaciones HLA con valores altos o bajos de CV, para así poder encontrar alelos de riesgo o alelos de protección en la progresión de la infección por VIH [8] en el contexto inmunogenético de la población mexicana. Dentro de la cohorte estudiada se encontraron los alelos A\*3101 ( $p=0.022910$ ,  $q=0.20406916$ ) y Cw\*0202 asociados con CV baja, por lo que podrían estar funcionando como alelos de protección dentro de esta cohorte mexicana. Ninguno de estos alelos se encuentra reportado como alelo de protección.

Se encontró al alelo B\*1801 asociado con medianas de CV altas, por lo que se considera un alelo de riesgo, con una progresión rápida de la infección por VIH. La región donde se encontró mayoritariamente expresado fue la región Sur del país. Este alelo es consistente con las observaciones en otras poblaciones, donde se ha observado asociado con CV alta y una progresión rápida de la enfermedad [8].

Para los alelos a baja resolución se encontró el alelo Cw\*05, asociado a medianas de CV altas, pudiendo representar un alelo de riesgo. Este alelo se observó expresado principalmente en la región Norte de nuestro país, lo que significa que los individuos que provienen de esta región tienen el riesgo de presentar una progresión rápida de la infección. Este alelo no está reportado como un alelo de riesgo o un alelo de protección en otras poblaciones.

Los alelos que generalmente se encuentran asociados con efecto importantes en el control de la replicación y por tanto a velocidades de progresión rápidas o lentas de la enfermedad por VIH son típicamente los alelos HLA-B. Esta capacidad que tienen los alelos HLA-B, aun no es totalmente clara, pero una razón importante es que existe una mayor cantidad de epítopes diferentes que pueden ser expresados por estas moléculas en comparación con las moléculas HLA-A y HLA-C. Otra razón podría ser que los epítopes expresados por las moléculas HLA-B, pueden tener un mejor reconocimiento por los CTLs y provocar respuestas de CTLs más polifuncionales, es decir, que tengan una mayor producción de citocinas y una capacidad mayor de proliferación que las respuestas restringidas por HLA-A y HLA-C [8].

## ***EVOLUCIÓN DEL VIH***

Debido al alto polimorfismo de los genes HLA y a las diferencias en distribuciones de frecuencias HLA en las tres regiones de este estudio, se analizó la posibilidad de encontrar diferencias en la evolución poblacional del VIH dependiente de alelos HLA en el interior del territorio nacional ya que previamente se ha comprobado la existencia de características únicas en la evolución del VIH dependiente del HLA

en una población del Centro – Sur de México, en comparación con una cohorte predominantemente caucásica [5].

El gen que se utilizó para los estudios evolutivos fue un fragmento del gen *pol*. Se utilizó este gen debido a que es con el que se contaba en la base de datos del CIENI, ya que en este gen se buscan las mutaciones de resistencia a los antirretrovirales. Además de esto, existen estudios donde observan que las asociaciones HLA – polimorfismo se dan a lo largo del genoma del VIH [27, 30, 44], y en México Avila – Ríos y cols. [5] observaron posiciones bajo presión positiva dependiente del HLA dentro del gen *pol* en una cohorte mexicana.

Para evidenciar la existencia de posibles diferencias en la evolución mediada por HLA del VIH en las diferentes regiones del país, se realizó una búsqueda de posiciones en el genoma viral bajo presión selectiva positiva, utilizando un modelo de codones basado en tasas de sustituciones sinónimas y no sinónimas.

Se observaron algunas posiciones dentro del fragmento de la secuencia del gen *pol* bajo selección positiva, algunas de las cuales fueron exclusivas para regiones específicas del país. Debido al número limitado de individuos incluidos en este estudio, no se puede determinar si estos resultados son reales o sólo son artefactos de los métodos estadísticos usados para este análisis. En la búsqueda de las posiciones bajo presión positiva de cada región se trabajó con menos de 50 individuos, por lo que las conclusiones del análisis son limitadas.

En la búsqueda de las posiciones bajo selección positiva en el que se incluyeron 123 secuencias del gen *pol* de los individuos de la cohorte mexicana. Se observaron algunas posiciones que son consistentes con las reportadas previamente en el estudio de Avila – Ríos y colaboradores [5], lo que aumenta la confianza en los resultados del presente estudio. Es claro que debe aumentarse el número de individuos en la cohorte para cada región en estudio, con el objetivo de tener análisis más robustos para la búsqueda de diferencias en la evolución del VIH en diferentes regiones del país. Sin embargo, las diferencias observadas en posiciones bajo presión selectiva en los virus de las diferentes regiones sugieren la posible existencia de características únicas en la evolución del virus en poblaciones inmunogenéticamente distintas dentro del territorio nacional. Esta

hipótesis es apoyada por las observaciones de estudios previos [5, 23] que han mostrado que cerca del 80% de las posiciones bajo presión selectiva en el genoma viral, pueden explicarse por presión selectiva por HLA.

Con estas limitaciones, se buscó explicar la presencia de dichas posiciones bajo selección positiva por la presión selectiva que ejercen los alelos HLA sobre el VIH ya que se sabe que gran parte de las presiones positivas son restringidas por los alelos HLA [25-27, 40]. Sin embargo, en el análisis de asociación entre la expresión de alelos HLA específicos y la presencia de polimorfismo en codones particulares del VIH, no se observaron coincidencias entre las posiciones bajo selección positiva y la expresión de alelos HLA de manera significativa. Sólo dos posiciones de la secuencia del gen *pol* mostraron una tendencia a estar asociadas con alelos HLA: la posición 135 de la RT, asociada al alelo HLA A\*2301 exclusivamente en la región Centro del país; y la posición 292 de la RT, asociada al alelo B\*4002 exclusivamente en la región Centro del país.

En los alelos en baja resolución se encontró solo una posición asociada significativamente a HLA: la posición 283 de la RT ( $p=2.00e-05$  y  $q=0.11568042$ ), la cual tampoco fue identificada como una posición bajo presión selectiva positiva en el análisis de selección por modelos de codones.

El hecho de que no exista una correlación entre las posiciones encontradas bajo presión positiva por el modelo de codones y las asociaciones HLA no implica que éstas no existan. Previamente se ha observado que la mayoría de las posiciones bajo presión selectiva positiva pueden explicarse en el contexto de alelos HLA [5]. Sin embargo, la falta de correlación obtenida en el presente estudio no es sorprendente ya que el número de individuos incluidos en la cohorte es pequeño, limitando la potencia estadística para detectar asociaciones HLA – codón de VIH. Se ha reportado que estudios evolutivos de asociación HLA – codón del VIH, requieren cohortes que incluyan un número grande de individuos, pues al aumentar el número de individuos aumenta también la potencia para detectar asociaciones [42]. Carlson y colaboradores han comprobado que cohortes tan grandes como 1000 individuos permiten observar un 80% de las asociaciones

existentes. Además, es importante mencionar que existen factores que pueden confundir la detección de asociaciones HLA – codón de VIH, ocasionando la detección de asociaciones espurias. Estos factores incluyen 1) efectos de la filogenia de las secuencias, en los que eventos fundadores asociados a subpoblaciones de individuos enriquecidas en algún alelo HLA provocarían la observación de asociaciones entre determinados polimorfismos virales y la expresión de dicho alelo HLA; 2) la existencia de desequilibrio de ligamiento entre los alelos HLA, pudiéndose observar asociaciones entre polimorfismos específicos del virus y alelos HLA en desequilibrio de ligamiento con el alelo selector; 3) co-evolución de codones del VIH, pudiéndose observar asociaciones indirectas entre cambios en un codón del virus y la expresión de alelos HLA, cuando en realidad los cambios en dicho codón del virus serían producidos por cambios en otro codón.

Por tanto, el análisis de asociación HLA – codón de VIH tiene limitaciones importantes que deben ser consideradas en su interpretación. Así, varias de las asociaciones observadas podrían representar asociaciones espurias, considerando además, que la mayoría de ellas no alcanzaron el umbral de significancia establecido de  $q < 0.02$ , el cual nos daría la confianza de que al menos un 80% de las asociaciones observadas como significativas serían asociaciones reales. Sin embargo, este estudio puede formar parte de otros estudios más amplios, en los que se tengan cohortes más grandes que proporcionen un mayor poder estadístico para detectar asociaciones HLA – codones del VIH reales que describan mejor la evolución del virus mediada por HLA dentro del territorio nacional.

Se encontraron asociaciones HLA – codón del VIH para alelos como A\*2301 (217 Treonina RT) y Cw\*0202 (135 Isoleucina RT), los cuales están asociados de manera negativa, es decir, si el alelo HLA está presente es muy probable que no se encuentre el AA blanco, o dicho de otra manera, la presencia de dicho AA en la posición indicada en el genoma del VIH, implica muy probablemente que no se tiene dicho alelo HLA. Esta asociación es importante ya que aunque se observa que la presencia del alelo HLA está asociada con la falta de un AA, muy

probablemente también está asociada con la presencia de otro (u otros). Esta asociación posiblemente no fue observada porque el poder estadístico fue menor del necesario requerido para reconocerla, ya que es común que la presencia del alelo HLA se asocie a cambios múltiples de AA's. Por ejemplo: la presencia de A\*2301 podría provocar un cambio del AA1 al AA2 ó al AA3, en cuyo caso el método estadístico y el limitado número de individuos incluidos en el estudio podría no ser suficiente para reconocer las asociaciones con el AA2 y el AA3 y sólo pudo observarse la asociación a la ausencia el AA1.

Sólo una de las asociaciones encontradas de alelos HLA – codón del VIH pudo comprobarse en el mapa de epítopes. La posición 41 de la PR asociada con el alelo B\*44 con la ausencia de ácido glutámico, coincide con el epítope óptimo que comprende las posiciones 34 – 42 de la PR. Es importante considerar las limitaciones ya mencionadas para la cohorte en estudio y para el análisis estadístico utilizado, pero también es de gran importancia comprobar esta posición de manera experimental, para observar si los individuos que presentan este alelo, tienen CTLs capaces de respondan a este epítope.

Sin embargo, algunas de las asociaciones ya se observaron en el estudio publicado con una cohorte del CIENI por Avila – Ríos y colaboradores [5], con 280 individuos VIH (+) de la región Centro – Sur del país. Con un estudio PDN (red de dependencia filogenética), se encontraron un total de 43 asociaciones HLA-codones del VIH y 251 asociaciones entre los codones del VIH (mutaciones compensatorias), con las cuales explican muchas de las posiciones bajo selección positiva por la presión por HLA. Algunas de las posiciones son: la posición 12 de la PR que está asociada al alelo B\*51 y con la presencia de una serina y el alelo Cw\*03 asociado a la presencia de una prolina en esa misma posición.

En este estudio, ninguno de los alelos que se encontró con desequilibrio de ligamiento, pudo comprobarse con el mapa de epítopes. Sin embargo, haciendo una búsqueda de los alelos con desequilibrio de ligamiento publicados en el estudio previo con la cohorte del CIENI [5], se encontraron asociaciones de alelos HLA – codón del VIH en posiciones que coinciden con la posición de los epítopes óptimos de los alelos con que se encuentran en desequilibrio de ligamiento. Por

ejemplo: el alelo Cw\*07 que en este trabajo se encontró asociado con la posición 164 de la RT con la presencia de una alanina. Esta posición se encuentra dentro del epítotope óptimo para B\*0702 que se encuentra en las posiciones 156 – 164 de la RT. Estos dos alelos (B\*07/Cw\*07) presentan desequilibrio de ligamiento en la cohorte de los 288 individuos mexicanos incluidos en el estudio de Avila – Ríos y colaboradores [5].

Otra asociación que puede ser comprobada por desequilibrio de ligamiento fue la expresión del alelo Cw\*07 y la ausencia de alanina en la posición 206 de la RT. Esta posición se encuentra dentro del epítotope óptimo del alelo B\*4001, alelo que se ha reportado en desequilibrio de ligamiento con Cw\*07 en la cohorte del CIENI [5].

Es claro que es de gran importancia conocer el desequilibrio de ligamiento en una población, ya que esto permite encontrar asociaciones indirectas de alelos HLA – codón de VIH. Estas asociaciones que a primera vista parecen reflejar la presión selectiva de un alelo en una posición determinada, pueden no ser reales y deberse a la presión de otro alelo en desequilibrio de ligamiento con el alelo HLA asociado. A nivel poblacional esto es muy importante, ya que los alelos que se encuentran en desequilibrio de ligamiento, pueden ser alelos con expresiones diferenciales en una región específica que pueden determinar características únicas de la evolución del VIH en una región en particular.

Dos de las tres posiciones que se encontraron asociadas al alelo B\*1801 (88 de la RT y 224 de la RT), pudieron comprobarse con la predicción de epítotos para este alelo con motivos de anclaje. Ambas posiciones se encontraron dentro de los posibles epítotos, lo que le da más validez a este resultado. Lo mismo ocurrió con la posición 287 de la RT que también se encuentra dentro del epítotope predicho para Cw\*05 identificado como el alelo selector.

Esta coincidencia de las posiciones de estos dos alelos con los epítotos predichos, apoya la validez de estas asociaciones, descartando la posibilidad de que representen asociaciones espurias. Es de gran importancia poder confirmar las asociaciones de estos alelos en particular mediante experimentos *in vitro*, ya que son alelos que encontraron como alelos de riesgo en la cohorte en estudio.

La posición 148 de la RT que se encontró asociada también al alelo B\*18, no coincidió de manera precisa con el epítotope óptimo ni con el epítotope predicho. Se sabe que algunas posiciones que no se pueden explicar con epítotope óptimos se pueden considerar como posiciones con mutaciones de procesamiento que ocurren a distancia del sitio del epítotope óptimo, mutaciones compensatorias que se asocian a cambios en otro u otros codones del VIH, un epítotope inusual, o simplemente es una posición falsa positiva [27].

Es necesario realizar experimentos para demostrar que estos individuos tienen CTLs capaces de responder a estos epítotope predichos para así tener la seguridad de que se está ejerciendo una presión selectiva en dicha posición por parte de estos alelos de igual manera, estudios de co-evolución de codones de VIH podrían determinar posibles patrones mutacionales que explicarían la presencia de asociaciones indirectas.

Se hizo un análisis de asociación entre la presencia de un AA y la región y/o el estado de procedencia de los individuos incluidos en el estudio, con lo que se pudo observar que el residuo del ácido glutámico en la posición 314 de la RT, se encuentra principalmente presente en el Centro, y que el estado que mostró mayoritariamente la presencia de este mismo AA en dicha posición fue el D. F. Esto probablemente se deba a la presencia diferencial mayoritaria de un posible alelo HLA selector en esta región del país. Con ayuda del mapa de epítotope se observó que la posición RT 314 se encuentra muy cercana a un epítotope óptimo de los alelos B\*44. El alelo B\*44 se encontró con una expresión un poco mayor en la región Centro en comparación con resto de las regiones incluidas en este estudio (Figura 6.8). Esto sugiere que la presencia de B\*44 podría explicar la expresión mayoritaria de ácido glutámico en la posición RT 314, en la región centro y más específico en el D.F. Estas posiciones que se encuentran con diferentes frecuencias en los distintos estados del país, muestra que la evolución del VIH se está dando de manera diferente en cada región dentro del territorio nacional y que además esta evolución podría ser dependiente de la expresión de alelos HLA selectores en distintas frecuencias.

## 8. CONCLUSIONES

Para el desarrollo de este proyecto se trabajó con una cohorte de 139 individuos tomada de la cohorte de Epidemiología y Evolución del Centro de Investigación en Enfermedades Infecciosas. Los criterios de inclusión fueron: individuos VIH (+) con una cuenta de linfocitos T CD4+ > 500 células/μl de sangre y sin tratamiento antirretroviral. Se seleccionaron 43 individuos de la región Norte del país, 49 de la región Centro y 47 individuos de la región Sur de país.

El análisis de las frecuencias alélicas HLA de la cohorte mexicana, mostró una diferencia en la presencia de alelos HLA de clase I, en las distintas regiones en estudio, con la presencia marcada de alelos como A\*02, A\*24, A\*68, B\*35, B\*39, B\*40, B\*15, Cw\*07, Cw\*03, Cw\*04, y Cw\*08, considerados principalmente alelos Amerindios, pero también se comparten con poblaciones africanas y caucásicas [5, 7, 43]. Los alelos considerados amerindios tienen una mayor frecuencia en las regiones Centro y Sur del país.

Se encontró alelos con una diferencia significativa en la expresión de alelos HLA en las regiones Norte y Sur del país. El alelo A\*0206 predominante en la región Sur del país, y los alelos A\*2402, Cw\*0401 y Cw\*0501 mostraron estar presentes con mayor frecuencia en la región Norte del país. Para los alelos HLA-B no se encontró ninguna diferencia significativa en esta comparación, sólo la tendencia de encontrar alelos como: B\*1401, B\*3517 y B\*4002 a expresarse más en la región Norte del país y los alelos B\*1801, B\*3906 y B\*5102 se encuentran principalmente en la región Sur del territorio nacional.

Los alelos que se encontraron en desequilibrio de ligamiento (B\*3517/ Cw0401, A\*6803/B\*3905, B\*35/Cw\*04, B\*40/Cw\*03, B\*14/Cw\*08 y B\*15/Cw\*01) coincidieron con los encontrados con una cohorte del CIENI en estudios previos [5]. Además estos alelos coincidieron con los haplotipos previamente reportados como los más prevalentes en la población mexicana [5, 7]. Este fenómeno es importante, ya que afecta directamente la distribución de frecuencias de estos alelos HLA dentro del territorio nacional. El fenómeno de desequilibrio de

ligamiento también afecta a las asociaciones observadas de alelos HLA – polimorfismos del VIH.

En las asociaciones HLA-CV se encontró evidencia de la existencia de alelos HLA protectores y de riesgo en la cohorte mexicana, ya que dos alelos en alta resolución se encontraron relacionados con CV bajas: A\*3101 y Cw\*0202, pudiendo representar alelos protectores dentro de la población mexicana. Por otro lado, el alelo B\*1801 se encontró asociado con la progresión de la enfermedad. Esto es consistente con lo ya reportado para otras poblaciones [8], ya que este alelo se ha encontrado relacionado con alta CV en otras poblaciones. En el análisis a baja resolución solo se encontró el alelo Cw\*05 asociado con alta carga viral, por lo que se le considera un alelo de riesgo.

Con la limitación del número de individuos incluidos en este estudio, las posiciones encontradas bajo presión positiva no se pudieron explicar bajo el contexto de los alelos HLA. Sin embargo, estas posiciones fueron consistentes con las encontradas en el estudio de Avila – Ríos [5], lo que muestra que al aumentar el número de pacientes, se aumentará el poder de detección de asociaciones y por lo tanto dichas posiciones bajo presión positiva se podrán explicar por alelos HLA.

En el análisis de HLA – polimorfismo del VIH, se encontraron 24 posiciones del gen *pol* asociadas a un alelo HLA a alta resolución y 33 asociaciones HLA – codón del VIH a baja resolución. Un ejemplo es la posición 164 de la RT asociada al alelo Cw\*07 con presencia de alanina. De manera muy interesante se encontraron asociaciones negativas, en las que la presencia del alelo HLA se encontró asociada a la ausencia de un aminoácido en una posición del VIH particular. Tal fue el caso de la posición 217, asociada con el alelo A\*2301 donde la presencia del alelo selector se asoció a la ausencia de treonina en dicha posición.

Se espera que existan otras asociaciones positivas complementarias que no se observaron por falta de potencia estadística debido al número limitado de individuos incluidos en el estudio.

Solo se encontró una posición que puede ser explicada directamente con un epítotope óptimo: la posición 41 de la PR que coincidió con el epítotope asociado al alelo B\*44.

Coincidieron también en el mapa de epítotos de este estudio, dos de los alelos con desequilibrio de enlace observados en el estudio de Avila – Ríos y colaboradores [5], los cuales fueron: Cw\*07 de la posición 164 de la RT que coincidió con el epítotope óptimo del alelo B\*0702 y el alelo Cw\*07 de la posición 206 de la RT con el epítotope óptimo del alelo B\*4001.

Dos de los epítotos predichos para los alelos B\*1801 y Cw\*05 coincidieron con las posiciones asociadas observadas en este estudio (88 y 224 de la RT para el alelo B\*1801 y 287 de la RT para Cw\*05) Pudiendo validar las asociaciones HLA – polimorfismo del VIH encontradas.

Las posiciones que no coincidieron con ningún epítotope óptimo o predicho pueden corresponder a mutaciones de procesamiento, mutaciones compensatorias, epítotos inusuales o simplemente falsos positivos. Aumentar el número de individuos incluidos en el estudio, aumentará el poder de detección de asociaciones HLA- Codón del VIH, y se podrán comprobar o rechazar las asociaciones observadas en este estudio.

Con las limitaciones ya mencionadas, el presente estudio muestra que la evolución del VIH dependiente de los alelos HLA, en la población mexicana tiene características únicas. La presencia de alelos con mayor frecuencia en una región determinada del país marcará el rumbo de la evolución del VIH, por lo que es necesario conocer estas diferencias para entender la evolución dependiente de HLA del virus en nuestro país. Este conocimiento será de utilidad para la investigación en vacunas.

## 9. PERSPECTIVAS

La realización de análisis evolutivos poblacionales dentro del territorio nacional, cobra gran importancia debido a la evidencia de que existen diferencias entre las características inmunogenéticas de la población mexicana, así como de la presión selectiva que ejerce el HLA sobre la evolución del VIH lo que lleva a una adaptación del virus a nivel poblacional.

Este estudio es sólo una parte de estudios posteriores que mostrarán la marcada adaptación del VIH que existe a lo largo del territorio nacional, lo que lleva a seguir con análisis similares, pero con cohortes más grandes que puedan mostrar los datos reales de lo que está sucediendo con dicha evolución. Por lo que es necesario: 1) aumentar el número de pacientes para así definir las tendencias que existen dentro de este estudio y contar una cohorte que tenga el poder suficiente para observar todas las asociaciones de alelos HLA – polimorfismos del VIH, disminuyendo así la presencia de falsos positivos o espurias provocados por los análisis estadísticos; 2) comprobar de manera experimental las posiciones encontradas en este estudio o en estudios posteriores son epítopes reales que producen respuestas de CTLs. Por lo que es necesario hacer experimentos con los CTLs de los individuos incluidos y validar así la existencia de dichas asociaciones HLA con estas posiciones del VIH; y por último 3) Incluir otros genes del VIH para la búsqueda de asociaciones HLA – polimorfismos del VIH a lo largo del genoma del virus en contexto de la población mexicana.

Cumpliendo con estas metas, se tendrá un panorama real y más amplio de la evolución del VIH a nivel poblacional dentro de nuestro país.



## 10. REFERENCIAS

1. ONUSIDA, *Informe sobre la epidemia mundial del SIDA 2008*. 2008.
2. Kindt J. Thomas, R.A.G., Barbara A. Osborne, *Inmunología de Kuby*. sexta ed. 2007, México: McGraw Hill Interamericana. 574.
3. José Ángel Córdova Villalobos, S.P.d.L.R., José Luis Valdespino, ed. *25 años de SIDA en México. Logros, desaciertos y retos*. ed. P. Edición. 2008, Instituto Nacional de Salud Publica: México. 182.
4. Peterlin, B.M. and D. Trono, *Hide, shield and strike back: how HIV-infected cells avoid immune eradication*. *Nat Rev Immunol*, 2003. **3**(2): p. 97-107.
5. Avila-Rios, S., et al., *Unique features of HLA-mediated HIV evolution in a Mexican cohort: a comparative study*. *Retrovirology*, 2009. **6**: p. 72.
6. Gorodezky, C., et al., *The genetic structure of Mexican Mestizos of different locations: tracking back their origins through MHC genes, blood group systems, and microsatellites*. *Hum Immunol*, 2001. **62**(9): p. 979-91.
7. Barquera, R., et al., *HLA class I and class II haplotypes in admixed families from several regions of Mexico*. *Mol Immunol*, 2008. **45**(4): p. 1171-8.
8. Goulder, P.J. and D.I. Watkins, *Impact of MHC class I diversity on immune control of immunodeficiency virus replication*. *Nat Rev Immunol*, 2008. **8**(8): p. 619-30.
9. (OMS), P.c.d.I.N.U.s.e.V.S.O.y.O.M.d.I.S., *Situación de la epidemia de sida: informe especial sobre la prevención del VIH: Diciembre de 2007*. 2007, ONUSIDA/OMS: Suiza.
10. David M. Knipe, P.P.M.H., MD; Diane E. Griffin, MD, PhD; Robert A. Lamb, PhD, ScD; Malcolm A. Martin, MD; Bernard Roizman, ScD; Stephen E. Straus, MD, *Field' Virology*. 5 ed. 2007.
11. Simon, V., D.D. Ho, and Q. Abdool Karim, *HIV/AIDS epidemiology, pathogenesis, prevention, and treatment*. *Lancet*, 2006. **368**(9534): p. 489-504.

12. Operativa, C.D.d.I., *El VIH/SIDA en México. Datos al 14 de Noviembre de 2008*, S.d.S. Pública, Editor. 2008, Centro Nacional para la Prevención y el Control del VIH/SIDA: México.
13. Carlos Magis Rodríguez, M.H.Á. (2007) *Epidemiología del SIDA en México. Volume*,
14. Hemelaar, J., et al., *Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004*. AIDS, 2006. **20**(16): p. W13-23.
15. Murray, P.R., *Microbiología Médica*. Cuarta ed. 2004, España: Elsevier Science. 658 - 665.
16. Hoffmann Christian, J.K.R.y.B.S.K., *HIV Medicine 2007*. 15 ed. 2007, Paris: Flying Publisser.
17. Grossman, Z., et al., *Pathogenesis of HIV infection: what the virus spares is as important as what it destroys*. Nat Med, 2006. **12**(3): p. 289-95.
18. Dyer, W.B., et al., *Mechanisms of HIV non-progression; robust and sustained CD4+ T-cell proliferative responses to p24 antigen correlate with control of viraemia and lack of disease progression after long-term transfusion-acquired HIV-1 infection*. Retrovirology, 2008. **5**: p. 112.
19. Abul K. Abbas, A.H.L.y.J.S.P., *inmunología Celular y Molecular*. Tercera ed. 1999, España: McGraw-Hill Interamericana. 552.
20. Middleton, D., et al., *New allele frequency database: <http://www.allelefrequencys.net>*. Tissue Antigens, 2003. **61**(5): p. 403-7.
21. Leslie, A., et al., *Differential selection pressure exerted on HIV by CTL targeting identical epitopes but restricted by distinct HLA alleles from the same HLA supertype*. J Immunol, 2006. **177**(7): p. 4699-708.
22. Hendel, H., et al., *New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS*. J Immunol, 1999. **162**(11): p. 6942-6.
23. Carlson, J.M. and Z.L. Brumme, *HIV evolution in response to HLA-restricted CTL selection pressures: a population-based perspective*. Microbes Infect, 2008. **10**(5): p. 455-61.

24. Allen, T.M., et al., *Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution.* J Virol, 2005. **79**(21): p. 13239-49.
25. Moore, C.B., et al., *Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level.* Science, 2002. **296**(5572): p. 1439-43.
26. Ahlenstiel, G., et al., *Selective pressures of HLA genotypes and antiviral therapy on human immunodeficiency virus type 1 sequence mutation at a population level.* Clin Vaccine Immunol, 2007. **14**(10): p. 1266-73.
27. Brumme, Z.L., et al., *Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1.* PLoS Pathog, 2007. **3**(7): p. e94.
28. Altfeld, M., et al., *The majority of currently circulating human immunodeficiency virus type 1 clade B viruses fail to prime cytotoxic T-lymphocyte responses against an otherwise immunodominant HLA-A2-restricted epitope: implications for vaccine design.* J Virol, 2005. **79**(8): p. 5000-5.
29. Leslie, A., et al., *Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA.* J Exp Med, 2005. **201**(6): p. 891-902.
30. Brumme, Z.L., et al., *HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins.* PLoS One, 2009. **4**(8): p. e6687.
31. Storey, J.D. and R. Tibshirani, *Statistical significance for genomewide studies.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9440-5.
32. Storey, J.D. and R. Tibshirani, *Statistical methods for identifying differentially expressed genes in DNA microarrays.* Methods Mol Biol, 2003. **224**: p. 149-57.
33. Tamura, K., et al., *MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.* Mol Biol Evol, 2007. **24**(8): p. 1596-9.
34. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence*

- weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994. 22(22): p. 4673-80.*
35. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol, 1987. 4(4): p. 406-25.*
  36. Kimura, M., *A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol, 1980. 16(2): p. 111-20.*
  37. Feisenstein, J., *Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 1985. 39: p. 783 - 791.*
  38. Kosakovsky Pond, S.L. and S.D. Frost, *Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol, 2005. 22(5): p. 1208-22.*
  39. Brumme ZL, J.M., Brumme CJ. Carlson J. Haubrich R, Riddler S, Swenson L, Tao I, Szeto S, Chan D, et al, *Relationship between HLA class I-driven evolution in Gag, Pol and Nef and clinical markers of HIV disease:a multi-center collaborative study. AIDS Vaccine, Abstract P09-01, 2008.*
  40. Mueller, S.M., et al., *Dual selection pressure by drugs and HLA class I-restricted immune responses on human immunodeficiency virus type 1 protease. J Virol, 2007. 81(6): p. 2887-98.*
  41. Bastos, F.I., et al., *AIDS in Latin America: assessing the current status of the epidemic and the ongoing response. Int J Epidemiol, 2008. 37(4): p. 729-37.*
  42. Carlson, J.M., et al., *Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. PLoS Comput Biol, 2008. 4(11): p. e1000225.*
  43. Cao, K., et al., *Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. Hum Immunol, 2001. 62(9): p. 1009-30.*

44. Frahm, N., et al., *Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities*. J Virol, 2004. **78**(5): p. 2187-200.