



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**FACULTAD DE QUÍMICA**

**“Impacto de mutaciones en el gen *gag* en la capacidad replicativa del VIH en el contexto inmunogenético de la población mexicana.”**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE  
QUÍMICA FARMACÉUTICA BIÓLOGA**

**PRESENTA**

**SELMA NAYELY ALVA HERNÁNDEZ**



**MÉXICO, D.F.**

**2012**



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:** FRANCISCO JAVIER PLASENCIA DE LA PARRA \_\_\_\_\_

**VOCAL:** ROCÍO GABRIELA TIRADO MENDOZA \_\_\_\_\_

**SECRETARIO:** SANTIAGO ÁVILA RÍOS \_\_\_\_\_

**1er. SUPLENTE:** JULIO CÉSAR MARTÍNEZ ÁLVAREZ \_\_\_\_\_

**2° SUPLENTE:** LORENZO ISLAS VÁZQUEZ \_\_\_\_\_

**SITIO DONDE SE DESARROLLÓ EL TEMA:**

**CENTRO DE INVESTIGACIÓN EN ENFERMEDADES INFECCIOSAS (CIENI), INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS (INER).**

\_\_\_\_\_  
DR. SANTIAGO ÁVILA RÍOS.  
ASESOR DEL TEMA.

\_\_\_\_\_  
SELMA NAYELY ALVA HERNÁNDEZ.  
SUSTENTANTE.

¡Esta es tu vida! Haz lo que te gusta y hazlo seguido. Si algo no te gusta, cámbialo. Si no te gusta tu trabajo, déjalo. Si no tienes tiempo, deja de perderlo. Lee más. Si estas buscando el amor de tu vida, detente que él te está esperando. Comienza a hacer lo que amas. Deja de analizar, la vida es simple. Todas las emociones son hermosas, cuando comas aprecia cada bocado. Abre tu mente, tus brazos y tu corazón a nuevas cosas y personas. Estamos unidos en nuestras diferencias, pregunta a la próxima persona que veas cuál es su pasión y comparte tu inspiración con ellos. Viaja más seguido, perderte te ayudará a encontrarte contigo mismo. Algunas oportunidades pasan una sola vez, aprovéchalas. La vida se trata de conocer personas y de las cosas que creas con ellas, entonces ve afuera y comienza a crear. ¡La vida es corta, vive tus sueños y comparte tu pasión!

## AGRADECIMIENTOS.

A mis padres: Alfonso Alva y Fermina Hernández, ya que sin su apoyo y ejemplo no habría tenido motivación para seguir adelante. Por aguantar tantas noches que se convirtieron en mañanas. Por hacer innumerable cantidad de sacrificios para conseguir libros, materiales, cursos, BATAS y tantas otras cosas. Han sido, son y serán siempre un gran ejemplo y motor en mi vida. Este logro también es suyo.

A mi hermana Perla por aguantar tantas cosas científicas, aún odiándolas. Por soportar días de mal genio y estrés. Por hacerme reír en los días difíciles. Por decirme que soy una “ñoña”. Por ayudarme a estudiar y que junto con Nala me acompañaron durante las noches de hacer reportes de laboratorio.

A mi familia: mis abuelos, tíos y primos por nunca perder la confianza en mí. Por disculpar mis ausencias físicas y mentales debido a este “capricho”. Por apoyarme en todo momento e impulsarme para seguir adelante. Por aligerar los momentos difíciles y las cargas pesadas.

A mis amigos: Fernanda, Paola, Berenice, Miguel, Roberto, Hugo, Laurita, Aracely, Emiliano, Viry, Gerardo, Hermann, tantos momentos vividos y tan poco espacio para escribirlos. Cuantas experiencias compartidas: risas, viajes, fiestas, juegos, pero también enojos, tristeza, estrés y exámenes. ¡Me han enseñado tanto! Gracias por su apoyo y ánimos constantes. Ustedes son mi familia elegida.

A cada uno de los profesores que he tenido durante mi vida académica, los buenos y los malos, pues de todos he aprendido. Ustedes me han dado las herramientas para salir adelante. En especial gracias a la Dra. Perla Castañeda, un gran apoyo en estos casi 5 años.

Al Dr. Santiago Ávila Ríos que además de ser un excelente profesor y tutor, es una excelente persona. Gracias por la paciencia y el conocimiento compartido. Gracias al CIENI por abrir sus puertas a los estudiantes y permitirnos tener los recursos para seguir aprendiendo; a todas las personas que me encontré en sus instalaciones durante el tiempo que duró esta tesis, por compartir sus conocimientos conmigo y por tener la serenidad al enseñarme tantas cosas.

A la Facultad de Química y a la UNAM por abrirme sus puertas y permitirme ser de los privilegiados que pueden terminar una carrera en una de las mejores universidades a nivel mundial.

Gracias a la vida por cada una de las personas y experiencias que me ha puesto en el camino, pues cada una ha influido para que yo llegara a estar donde estoy ahora.

¡México, Pumas, Universidad!

## ÍNDICE.

RESUMEN .....	vii
LISTADO DE ABREVIATURAS. ....	viii
ÍNDICE DE FIGURAS .....	ix
ÍNDICE DE TABLAS.....	xi
CAPÍTULO I. INTRODUCCIÓN. ....	1
<b>1. Historia del VIH.....</b>	<b>1</b>
1.1. Origen de la enfermedad.....	1
1.2. Actualidad y futuro. ....	2
1.2.1. En el Mundo.....	2
1.2.2. En México.....	4
<b>2. Generalidades del VIH.....</b>	<b>6</b>
2.1. Origen y epidemiología molecular del VIH .....	6
2.2. Estructura del VIH: genoma y proteínas.....	8
2.2.1. Proteínas estructurales.....	10
2.2.2. Proteínas reguladoras.....	13
2.2.3. Proteínas accesorias. ....	13
2.3 Ciclo de replicación. ....	14
2.4. Historia natural de la enfermedad.....	17
<b>3. Respuesta inmune contra el VIH.....</b>	<b>19</b>
3.1. Respuesta innata .....	19
3.2. Linfocitos T CD4 <sup>+</sup> .....	20
3.3. Anticuerpos.....	21
3.4. Linfocitos T CD8 <sup>+</sup> .....	22
3.5. Presión selectiva por HLA. ....	24
CAPÍTULO II. PLANTEAMIENTO DEL PROBLEMA.....	28
<b>1. Planteamiento del problema.....</b>	<b>28</b>
CAPÍTULO III. HIPÓTESIS Y OBJETIVOS.....	31
<b>1. Hipótesis.....</b>	<b>31</b>
<b>2. Objetivos. ....</b>	<b>32</b>
2.1. Generales. ....	32

2.2. Particulares.....	32
<b>CAPÍTULO IV. METODOLOGÍA.....</b>	<b>33</b>
1. Selección de pacientes.....	33
2. Extracción de RNA.....	33
3. Amplificación del gen <i>gag</i> de virus libre en plasma. ....	33
4. Visualización de los productos de amplificación por electroforesis en gel de agarosa. ....	35
5. Purificación de los productos de amplificación. ....	35
6. Cuantificación de los productos de amplificación. ....	36
7. Secuenciación .....	36
8. Análisis filogenético. ....	37
9. Análisis de variabilidad. ....	37
10. Análisis Evolutivo. ....	38
11. Análisis de presión selectiva por HLA e impacto de mutaciones seleccionadas por HLA en la capacidad replicativa viral. ....	38
<b>CAPÍTULO V. RESULTADOS.....</b>	<b>41</b>
1. Descripción de la cohorte. ....	41
2. Análisis filogenético. ....	43
3. Análisis de variabilidad de las secuencias de VIH de pacientes mexicanos. ....	45
4. Determinación de sitios bajo presión selectiva por razón de sustituciones sinónimas y no sinónimas. ....	50
5. Análisis de Presión selectiva por HLA. ....	54
6. Impacto de la presión selectiva mediada por HLA en la capacidad replicativa del VIH. ....	64
7. Obtención de mapa de epítomos específico de la población mexicana. ....	80
<b>CAPÍTULO VI. DISCUSIÓN.....</b>	<b>82</b>
<b>CAPÍTULO VII. CONCLUSIÓN Y PERSPECTIVAS.....</b>	<b>88</b>
<b>1. Conclusión. ....</b>	<b>88</b>
<b>2. Perspectivas.....</b>	<b>89</b>
<b>CAPÍTULO VIII. BIBLIOGRAFÍA.....</b>	<b>90</b>

## RESUMEN

**ANTECEDENTES.** Las respuestas de CTLs (restringida por los alelos HLA-I), dada su efectividad para el control parcial de la viremia, ejercen una fuerte presión selectiva que dirige la evolución del VIH. Como consecuencia de esta presión evolutiva se seleccionan virus con mutaciones de escape. La presencia de estas mutaciones de escape inmune está muchas veces asociada a un costo en la capacidad replicativa del virus (fitness), lo que conduce a una reversión al transmitirse a un individuo carente del HLA selector. Aunado a esto, la existencia de la evolución del VIH mediada por HLA a nivel poblacional se ha demostrado por asociaciones entre la presencia de polimorfismos y la expresión de alelos HLA específicos. Debido al alto polimorfismo de los genes HLA, el presente estudio pretende por primera vez describir el impacto de la presión selectiva ejercida en un contexto inmunogenético mexicano sobre la variabilidad de la proteína Gag en virus circulantes en población mexicana. Dada la importancia previamente descrita de las respuestas de CTLs dirigidas a Gag para el control inmune del VIH, los resultados de este estudio contribuirán para apoyar la investigación sobre vacunas en el contexto global, y ayudarán en la identificación de posibles sitios blanco para la generación de inmunógenos que pudieran ser útiles para controlar el virus en nuestros pacientes.

**METODOLOGÍA.** Se eligieron al azar 200 muestras de plasma de pacientes mexicanos infectados con VIH-I B. Se realizó la extracción de RNA y la amplificación de *gag* utilizando una PCR anidada con transcripción inversa. Se realizó electroforesis en gel de Agarosa al 1%, se purificó y se cuantificó para ajustar la concentración. Se realizó la placa de secuenciación con el reactivo Big Dye Terminator v.3.1 y 8 primers, y se purificó. La secuenciación se realizó con el método de Sanger. El llamado de bases y ensamblaje fueron con el software Sequencing Analysis v5.4 y SeqScape v2.7. Con el software MEGA 5.05 se realizó el alineamiento y un árbol filogenético (referencias la secuencia HXB2 y un subtipo B), se utilizó el método Neighbor-Joining, el modelo de dos parámetros de Kimura y 1000 réplicas de bootstrap. La variabilidad se calculó estimando la entropía de Shannon. Con el servidor DataMonkey se analizó por el método SLAC obteniéndose codones bajo presión selectiva por la razón dN/dS. La presencia de epítomos de CTLs reportados previamente se identificó con la Base de Datos de Los Álamos. Utilizando los datos de HLA en la Base de Datos de Epidemiología Molecular y Evolución del VIH en México del CIENI se evaluaron alelos frecuentes que pudieran explicar la presencia de selección positiva. Con la literatura se identificaron mutaciones de escape seleccionadas por HLA con costos en capacidad replicativa, para estimar el impacto en la cohorte (58 secuencias con valores de capacidad replicativa). La significancia se evaluó por pruebas de Fisher con el programa GraphPad Prism 5.04. Un análisis más específico para el contexto inmunogenético de la población mexicana del impacto de la presión selectiva por HLA se realizó con el modelo de PDN.

**RESULTADOS.** La cohorte analizada fue altamente representativa de la población infectada en México y estuvo enriquecida en pacientes en etapas tempranas de la infección. Al comparar árboles filogenéticos de los virus basados en el gen *gag* con árboles del gen *pol* se observó una mayor divergencia en las secuencias del gen *gag* probablemente debido a su alta inmunogenicidad. Los estudios de variabilidad del gen *gag* de VIH circulante en México concordaron con lo reportado por Brumme, et al, refiriendo a las regiones p17 y p27p1p6 como las más variables del gen. La alta conservación de p24 podría explicarse debido a su importancia funcional y estructural para el virus. Se manifestó la existencia de una gran cantidad de posiciones bajo presión selectiva positiva, muy probablemente asociadas a selección por HLA, dentro de las cuales se encontraron algunas previamente ya reportadas (poblaciones caucásicas) y que podrían representar rasgos comunes en la evolución del virus en diferentes poblaciones. Sin embargo, también se observaron posiciones previamente no reportadas, que podrían representar nuevas asociaciones en el contexto inmunogenético de los mexicanos. Aunque las diferencias no fueron significativas en la mayoría de los casos, se observaron polimorfismos con tendencias interesantes en la influencia sobre la capacidad replicativa viral. Conjuntando las observaciones anteriores, hemos mostrado evidencia de un impacto importante de las respuestas inmunes restringidas por HLA en la determinación de la variabilidad del gen *gag* del VIH circulante en México, expresado mediante un mapa del gen *gag* que resume las asociaciones con presión selectiva positiva, presión selectiva negativa y selección por HLA en el contexto inmunogenético específico de esta población. Siendo el primer trabajo que trata de describir y explicar la variabilidad de este gen en virus circulantes en el país será una importante base para continuar con el análisis de posiciones con asociaciones de interés, tomando en cuenta para posteriores estudios el uso de metodologías distintas y aumentando la cantidad de sujetos en la cohorte.



## LISTADO DE ABREVIATURAS.

a.a.	Aminoácido.
APC	Célula Presentadora de Antígeno, en inglés.
CRF	Formas Recombinantes Circulantes, en inglés.
CTLs	Linfocitos T Citotóxicos, en inglés.
dN	Sustituciones No Sinónimas, en inglés.
DNA	Ácido Desoxirribonucleico, en inglés.
dS	Sustituciones Sinónimas, en inglés.
HLA	Antígeno Leucocitario Humano, en inglés.
HSH	Hombres que tienen sexo con hombres.
MHC	Complejo Mayor de Histocompatibilidad, en inglés.
OMS	Organización Mundial de la Salud.
ORF	Fragmentos de Lectura Abierta, en inglés.
PDN	Redes de Dependencia Filogenética, en inglés.
RNA	Ácido Ribonucleico, en inglés.
RT	Transcriptasa Inversa, en inglés.
SIDA	Síndrome de Inmunodeficiencia Adquirida.
TCR	Receptor de células T.
UDI	Usuarios de drogas inyectables.
VIH	Virus de Inmunodeficiencia Humana.

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Estimado de adultos y niños que viven con VIH en el mundo, reportado por la OMS en 2010. ....	3
<b>Figura 2.</b> Distribución de casos de SIDA en hombres de México. Datos al 30 de septiembre del 2011 según CENSIDA. ....	5
<b>Figura 3.</b> Distribución de casos de SIDA en mujeres de México. Datos al 30 de septiembre del 2011 según CENSIDA. ....	5
<b>Figura 4.</b> Distribución global de subtipos de VIH-I y sus recombinantes (CRF).....	8
<b>Figura 5.</b> Organización del genoma del VIH-I. ....	9
<b>Figura 6.</b> Representación de una partícula madura del VIH.....	10
<b>Figura 7.</b> Descripción esquemática de los eventos que ocurren durante el ciclo de replicación del VIH. ....	16
<b>Figura 8.</b> Historia natural de la enfermedad causada por VIH. ....	18
<b>Figura 9.</b> Representación esquemática de las funciones de los dos tipos de Linfocitos T CD4+ (Th1 y Th2). ....	21
<b>Figura 10.</b> Esquema que representa los efectos de la activación de un Linfocito T CD8+ (CTL) sobre una célula infectada. ....	23
<b>Figura 11.</b> Gráfico mostrando los 15 alelos HLA-B más comunes en la población mexicana y sus frecuencias alélicas. ....	25
<b>Figura 12.</b> Ejemplo de un gel de agarosa al 1%. ....	35
<b>Figura 13.</b> A) Distribución geográfica de la cohorte estudiada. B) La proporción de pacientes por cada estado se muestra en el gráfico circular. ....	42
<b>Figura 14.</b> Árboles filogenéticos para los genes <i>gag</i> (a) y <i>pol</i> (b) de los virus mexicanos. ....	44
<b>Figura 15.</b> Árbol filogenético basado en el gen <i>gag</i> de los 200 virus mexicanos.....	45
<b>Figura 16.</b> Frecuencia de polimorfismos en los distintos codones de <i>gag</i> , en VIH circulante en México.....	47
<b>Figura 17.</b> Análisis de entropía de Shannon en el gen <i>gag</i> de VIH circulante en México. ....	48
<b>Figura 18.</b> Comparación de entropía de Shannon para virus de pacientes en etapas tempranas y tardías de la infección en México. ....	49

<b>Figura 19.</b> Distribución de posiciones bajo presión selectiva positiva y negativa en el alineamiento del gen <i>gag</i> para las 200 secuencias mexicanas. ....	54
<b>Figura 20.</b> Mapas de epítomos para p17 y posiciones bajo presión selectiva positiva.....	55
<b>Figura 21.</b> Mapas de epítomos para p24 y posiciones bajo presión selectiva positiva.....	56
<b>Figura 22.</b> Mapas de epítomos para p24 y posiciones bajo presión selectiva positiva.....	57
<b>Figura 23.</b> Mapas de epítomos para p15 y posiciones bajo presión selectiva positiva.....	58
<b>Figura 24.</b> Frecuencia de mutaciones con costo en capacidad replicativa en una cohorte de 200 pacientes mexicanos.....	69
<b>Figura 25.</b> Árbol filogenético de las 58 muestras con valores asignados de capacidad replicativa. ....	71
<b>Figura 26.</b> Gráfico mostrando la razón de mutaciones de fase aguda, de fase crónica y mutaciones compensatorias, reportadas previamente en la literatura, en los 29 virus de alta capacidad replicativa y los 29 de baja capacidad replicativa de la cohorte estudiada.....	73
<b>Figura 27.</b> Gráfico de caja y bigotes de las posibles mutaciones compensatorias y de escape encontradas por el análisis de razones de sustituciones no sinónimas / Sinónimas (DataMonkey).....	76
<b>Figura 28.</b> Gráfico de caja y bigotes de las posibles mutaciones compensatorias y de escape encontradas por el análisis de Redes de Dependencia Filogenética (PDN). ....	79
<b>Figura 29.</b> Mapa de escape inmunológico de <i>gag</i> en VIH circulante en México. ....	81

## ÍNDICE DE TABLAS.

<b>Tabla 1.</b> Grupos, subtipos y distribución geográfica del VIH.....	7
<b>Tabla 2.</b> Características demográficas y clínicas de la cohorte. ....	41
<b>Tabla 3.</b> Sitios bajo presión selectiva negativa por razones de sustitución dN/dS en el gen <i>gag</i> de VIH circulante en México. ....	50
<b>Tabla 4.</b> Sitios bajo presión selectiva positiva por razones de sustitución dN/dS en el gen <i>gag</i> de virus circulante en México. ....	52
<b>Tabla 5.</b> Asociaciones HLA-polimorfismo de VIH en el gen <i>gag</i> en una cohorte de 500 pacientes mexicanos, según el modelo de PDN*. ....	60
<b>Tabla 6.</b> Resumen de resultados de asociaciones HLA-polimorfismo de VIH por diferentes análisis. ....	63
<b>Tabla 7.</b> Mutaciones en Gag reportadas como importantes en la capacidad replicativa del VIH. ....	65
<b>Tabla 8.</b> Relación de capacidad replicativa viral de 58 muestras y mutaciones con costo en capacidad replicativa reportadas en literatura.....	72
<b>Tabla 9.</b> Relación de posiciones bajo presión selectiva (DataMonkey) y capacidad replicativa viral.....	75
<b>Tabla 10.</b> Codones bajo presión selectiva con posibilidad de impacto en la capacidad replicativa viral.....	77
<b>Tabla 11.</b> Relación de codones bajo presión selectiva (PDN) y capacidad replicativa viral.....	78
<b>Tabla 12.</b> Codones con asociación a HLA en la cohorte mexicana por el modelo PDN, no reportadas en la literatura, con posibilidad de impacto en la capacidad replicativa viral.....	80

## CAPÍTULO I. INTRODUCCIÓN.

### 1. Historia del VIH

#### 1.1. Origen de la enfermedad.

Hace ya 30 años algunos grupos clínicos estadounidenses enfrentaron una enfermedad de causa desconocida que provocaba, inevitablemente, la muerte por enfermedades oportunistas pues causaba la destrucción del sistema inmunológico [11]. Al inicio se observó que la enfermedad afectaba principalmente a hombres homosexuales y personas que usaban drogas intravenosas pero después de algunos meses se descubrió que también eran afectados otros grupos, incluyendo: hemofílicos, personas que habían recibido transfusiones y más significativamente, parejas sexuales e hijos de miembros de los distintos grupos de riesgo [21].

El patrón epidemiológico observado permitió proponer que la transmisión de este nuevo patógeno era por vía sanguínea (contacto con sangre contaminada) o por vía sexual (relaciones con un individuo infectado). La enfermedad, hasta ese entonces sin nombre, fue llamada: Síndrome de Inmunodeficiencia Adquirida (SIDA) [21].

Posteriormente, en 1983 el Dr. Luc Montagnier y su grupo de colaboradores del Instituto Pasteur en Francia, aislaron de los nódulos linfáticos de un hombre que presentaba linfadenopatía un agente que llamaron virus asociado a linfadenopatía (LAV). Al hacer cultivos celulares observaron que se producían altos títulos de viriones que contenían actividad de Transcriptasa Inversa (RT, Reverse Transcriptase) dependiente de magnesio; lo cual, les permitió clasificarlo como un retrovirus. Aunado a lo anterior, se observó que el tropismo del virus era hacia células mononucleares de sangre periférica (PBMCs) y más específicamente hacia linfocitos T CD4<sup>+</sup> [21].

Simultáneamente, otro grupo de investigación dirigido por el Dr. Robert Gallo en el National Cancer Institute en Estados Unidos, aislaron al mismo retrovirus y demostraron que era el agente causal del SIDA. Este virus fue posteriormente llamado Virus de Inmunodeficiencia Humana (VIH) [21]. Para ese momento, casos de la enfermedad ya

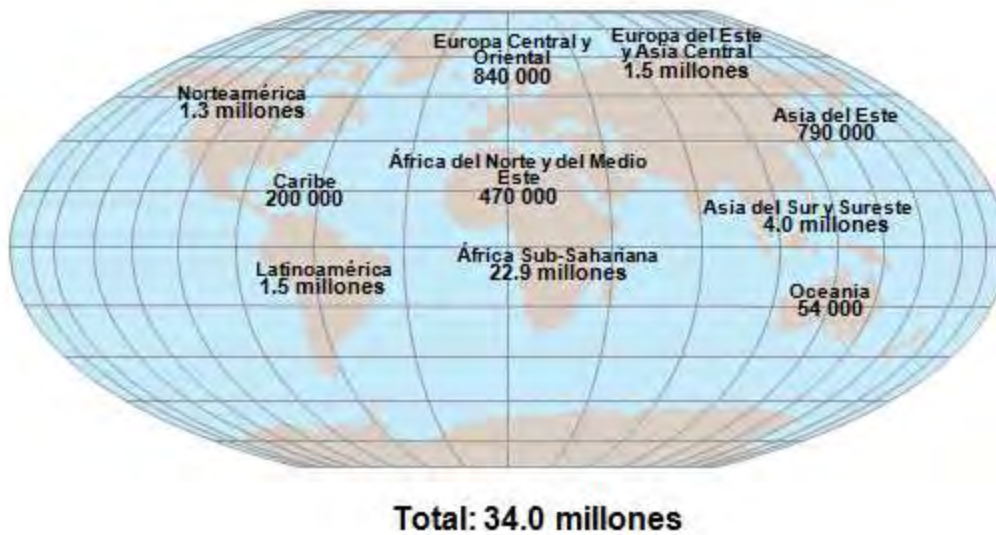
habían sido detectados en varias partes del mundo fuera de Estados Unidos, causando el interés de diversos equipos de investigadores; los cuales, empezaron el desarrollo de un tratamiento [4].

## 1.2. Actualidad y futuro.

### 1.2.1. En el Mundo.

El aumento de personas infectadas por el VIH creció de forma alarmante, lo cual promovió que gobiernos de distintos países tomaran medidas para poder frenar este avance y mejoraran la calidad de vida de los pacientes existentes. Entre las medidas tomadas están: la realización de campañas de educación y concientización sobre esta enfermedad, mejoramiento de la infraestructura y atención hospitalaria, ampliación en la cobertura y optimización del Tratamiento Antirretroviral (TAR). Estas medidas permitieron una disminución en el número de decesos de 2.1 millones reportados en 2004 a 1.8 millones en 2010 [11].

A pesar de esta cifra alentadora, el número de persona infectadas sigue en aumento con 34.0 millones de personas viviendo con el virus a nivel mundial y 2.7 millones de nuevas infecciones reportadas en 2011 [12], teniendo como principales causas en la actualidad, las prácticas sexuales de alto riesgo y el consumo de drogas intravenosas. Entre las zonas geográficas más afectadas por la enfermedad se encuentra el continente africano, más precisamente la zona Sub-Sahariana y en segundo lugar el sureste asiático [21] (**Fig. 1**).



**Figura 1.** Estimado de adultos y niños que viven con VIH en el mundo, reportado por la OMS en 2011 [Modificado de 24].

El uso del TAR ha permitido tener un panorama un poco más alentador sobre el manejo de la infección por VIH y el SIDA. Por ejemplo, tan sólo en Estados Unidos, a partir de la implementación de este tratamiento se obtuvo una reducción de más del 60% de la tasa de mortalidad en apenas 2 años (1995 a 1997). Cabe mencionar, que el tratamiento tiene sus limitaciones pues dependerá del acceso y del apego al mismo [11]. Tomando en cuenta esto se conoce que de los 34.0 millones de personas infectadas, 30.0 millones viven en países de mediano a bajo ingreso y sólo aproximadamente 6.6 millones están en la posibilidad de obtener el TAR [42].

Todo lo anterior nos lleva a pensar que la investigación en torno al VIH/SIDA deberá obtener los recursos materiales y humanos necesarios para permitir hacer frente al gran reto que representa. Para tal fin sigue siendo necesario ampliar el conocimiento sobre la biología del virus y su relación con el sistema inmune del hospedero; así como la identificación de nuevos blancos para la terapia antirretroviral, nuevas estrategias de prevención y control, y la búsqueda de posibles inmunógenos que sean efectivos en los contextos biológicos, inmunogenéticos y sociales de las diversas poblaciones del mundo.

### 1.2.2. En México.

Desde 1983, con la detección de los primeros casos de VIH/SIDA en México, el gobierno y diversas asociaciones no gubernamentales del país han tomado medidas para hacerle frente a esta enfermedad mediante el desarrollo de distintos planes nacionales. Fue en 1986 cuando se creó el primer Comité Nacional para Coordinar los Esfuerzos en VIH/SIDA, al cual le siguió la formación del Consejo Nacional para la Prevención y Control del VIH/SIDA (CONASIDA) y del Centro Nacional para la Prevención y Control del VIH/SIDA (CENSIDA). Aunado a lo anterior se logró la creación de una red de clínicas para el cuidado y la prevención del VIH/SIDA y de Infecciones de Transmisión Sexual (CAPASITS) y la implementación de campañas de prevención especialmente orientadas a grupos vulnerables, como la campaña “Brigada Callejera de Apoyo a la Mujer Elisa Martínez”, orientada a sexoservidores [11].

En cuanto al tratamiento antirretroviral, fue hasta 1991 que la Zidovudina (AZT) estuvo disponible en México y en 1993 se distribuyó por instituciones de seguridad social. En cuanto a la Terapia Antirretroviral (TAR) el panorama fue más difícil, pues no fue sino entre los años 1997 y 1998 que se logró la aprobación en la compra de los fármacos y hasta 2003 se logró su distribución a personas fuera de los grupos con seguridad social [11].

Según las estimaciones y proyecciones de VIH del Onusida para México existe un intervalo de 151,438 - 223,704 personas viviendo con VIH. En cuanto al Registro Nacional de casos SIDA que existe en México, al 30 de septiembre de 2011 se habían diagnosticado y registrado 151,614 casos acumulados de SIDA, de los cuales 82% eran hombres y 18% mujeres; y se habían registrado 36,714 casos acumulados de VIH, de los cuales 72.8% eran hombres y 27.2% mujeres. Se estima una prevalencia de VIH en adultos (15 a 49 años) de 0.3%, valor por debajo del porcentaje en América latina (0.6%) y en el mundo (0.8%). Las muertes por SIDA ascienden a 4,192 (2011) y la cobertura de antirretrovirales es del 78% total [13,43]. Sin embargo, estas cifras pueden estar significativamente



subestimadas, debido al subregistro característico de las estadísticas nacionales y a que una proporción importante de la población infectada no conoce su estado serológico.

Del total de los casos registrados de SIDA en el país, el 44% se encuentran en el Distrito Federal (24,198), Estado de México (16,686), Veracruz (13,958) y Jalisco (11,323), siendo estas las entidades Federativas con más casos registrados [12]. En cuanto a las vías de transmisión, la vía sexual es la principal en el país (**Fig. 2 y 3**), y la mayoría de los casos se encuentra concentrada en hombres que tienen sexo con hombres, personas que realizan trabajo sexual y personas que usan drogas inyectables [12,13].

Distribución de casos de sida por categoría de transmisión en Hombres (1983-2011)



Figura 2. Distribución de casos de SIDA en hombres de México. Datos al 30 de septiembre del 2011 según CENSIDA. UDI (usuarios de drogas inyectables), HSH (hombres que tienen sexo con hombres) [Obtenida de 12].

Distribución de casos de SIDA por categoría de transmisión en mujeres (1983-2011)

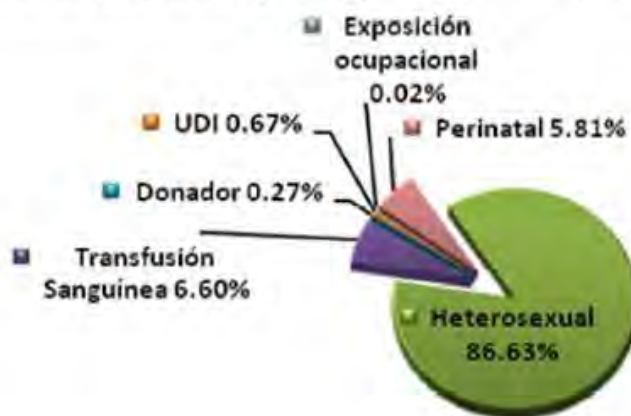


Figura 3. Distribución de casos de SIDA en mujeres de México. Datos al 30 de septiembre del 2011 según CENSIDA. UDI (usuarios de drogas inyectables) [Obtenida de 12].

A pesar de los grandes avances que ha logrado México en materia de prevención y tratamiento de la enfermedad es necesario continuar con campañas de gran impacto a nivel nacional para concientizar a la población (sobre todo a grupos vulnerables y/o marginados: usuarios de drogas inyectables, mujeres embarazadas, etc.) sobre la gravedad de la infección y la progresión a SIDA; enfatizando la importancia de la prevención y de la detección temprana. Así mismo, es de suma importancia luchar contra la estigmatización de la enfermedad que sigue caracterizando la mayoría de los escenarios de las epidemias en países en vías de desarrollo, así como continuar con la investigación sobre VIH/SIDA, en el contexto de la población mexicana, buscando estrategias de control adecuadas a las características sociales e inmunogenéticas de nuestra población.

## **2. Generalidades del VIH.**

### **2.1. Origen y epidemiología molecular del VIH.**

La epidemia global causada por el VIH-1 es el resultado de saltos zoonóticos de retrovirus entre monos y humanos. Ahora es bien sabido que el VIH-1 se originó a partir de un lentivirus de chimpancés, el virus de inmunodeficiencia simiana (SIV<sub>cpz</sub>). Una vez que el VIH empezó a ser analizado, la principal característica que llamó la atención de los investigadores fue el alto grado de diversidad y variabilidad genética que lo caracteriza. Para 1986 se recuperó un retrovirus de personas que vivían en países de África Oriental, el cual estaba relacionado con el agente antes descubierto pero era distinto inmunológicamente. Este virus fue llamado VIH-2. El origen evolutivo de este virus es distinto al del VIH-1, siendo su antecesor el SIV del mono mangabey ahumado. El VIH-2 presenta una progresión a la enfermedad más lenta, periodos asintomáticos más largos y una tasa de mortalidad menor [21].

A partir de estudios de las secuencias nucleotídicas y de proteínas, además de análisis filogenéticos y de reloj molecular, se ha determinado que existen cuatro grupos de HIV-1: el grupo M (“main”), el O (“outlier”), el N (“new”) [22] y el recientemente agregado grupo P, cada uno de los cuales asociado a saltos zoonóticos distintos de SIVcpz al humano. De estos grupos, el M es el más extendido en el mundo y presenta nueve subtipos (A, B, C, D,

F, G, H, J, K), los cuales presentan entre sí algunas diferencias en su comportamiento biológico y en su distribución geográfica [21] (**Tabla 1**).

Tabla 1. Grupos, subtipos y distribución geográfica del VIH [22].

TIPO	GRUPO	SUBTIPOS	DISTRIBUCIÓN
	M	A - K	UNIVERSAL y endémicos en Africa
HIV-1	O	-	CAMERUN GABON CONGO GUINEA-ECUATORIAL
	N	-	CAMERUN
HIV-2		A - F	Todos endémicos en Africa

Actualmente se han caracterizado además algunas variantes que incluyen secuencias provenientes de la recombinación de genomas de distintos subtipos por una co-infección gracias al fenómeno de elección de copia de la RT (la enzima puede cambiar el templado usado durante su actividad, debido a las 2 copias presentes en el genoma del virus); de las cuales, algunas se han diseminado dando lugar a las formas recombinantes circulantes (CRF). En los últimos años se han asociado 51 CRFs a distintas regiones del mundo [22] (**Fig. 4**).

Aunado a lo anterior, se ha descrito también que gracias a la gran capacidad de variación del VIH se generan múltiples variantes ligeramente distintas entre sí, llamadas cuasiespecies virales, que coexisten en cualquier momento en un individuo infectado.

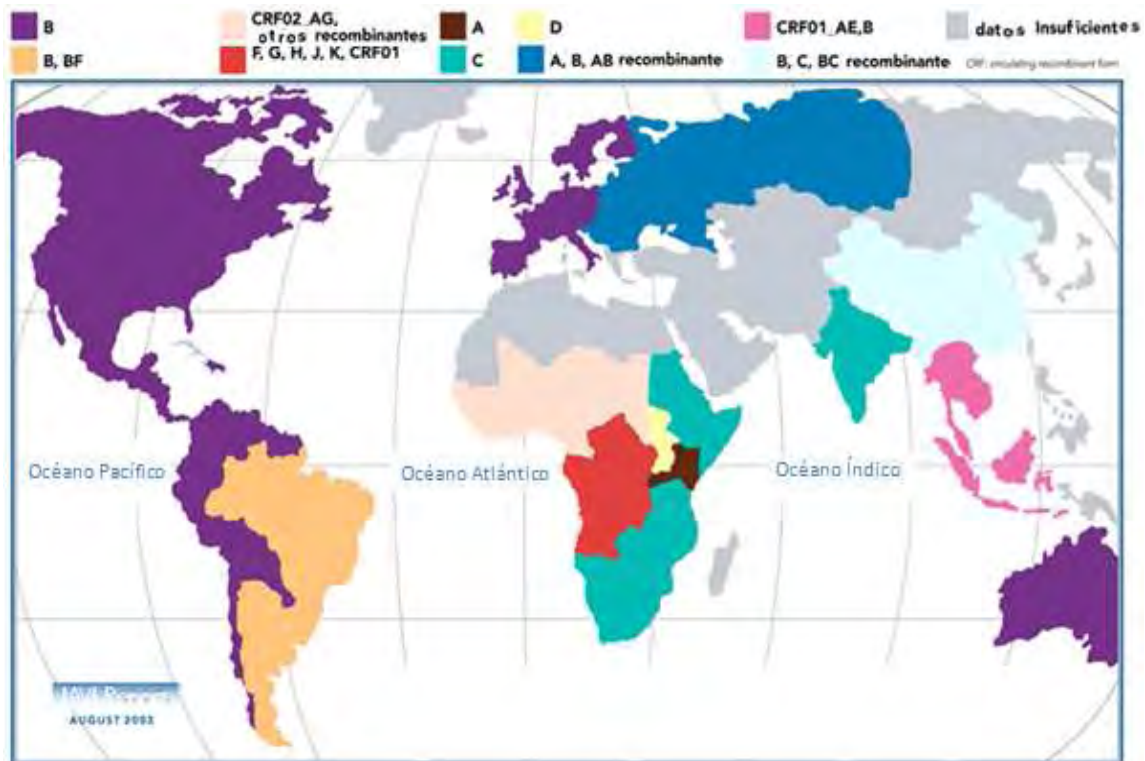


Figura 4. Distribución global de subtipos de VIH-1 y sus recombinantes (CRF). Se observa que el subtipo C es uno de los más predominantes a nivel mundial por su presencia en poblaciones con gran prevalencia de infecciones (África sub-Sahariana e India). Por otro lado, en América el subtipo predominante es el B [Obtenido de 33].

## 2.2. Estructura del VIH: genoma y proteínas.

El VIH es un virus envuelto que pertenece a la familia Retroviridae, dentro del género de los lentivirus. Esta clasificación se debe a que se replican a través de una forma intermedia de DNA bicatenario gracias a una enzima dependiente de magnesio llamada Transcriptasa Inversa, que es capaz de transcribir su genoma de RNA de cadena sencilla positiva a DNA en la célula del hospedero, y por otro lado a que la enfermedad que causa tiene una progresión lenta. Las células blanco de este virus, como las de otros lentivirus, son algunas líneas de células hematopoyéticas, particularmente linfocitos y macrófagos diferenciados [14].

El genoma del Virus de Inmunodeficiencia Humana incluye a los tres genes característicos de los retrovirus que codifican para las proteínas estructurales y enzimas necesarias para infectar: *gag* (gen antígeno grupo-específico), *pol* (polimerasa) y *env* (envoltura). Este

genoma también contiene marcos abiertos de lectura (ORF) que están traslapados y que codifican para proteínas reguladoras o accesorias: Vif, Vpu, Vpr, Tat, Rev y Nef, las cuales ahora se sabe que cumplen funciones fundamentales para la replicación del virus *in vivo*. Las secuencias largas repetitivas (LTRs) se localizan en cada extremo y sirven, entre otras cosas, de enlace con el genoma celular y para permitir la unión de factores para la transcripción de los genes virales [14, 21] (Fig. 5).

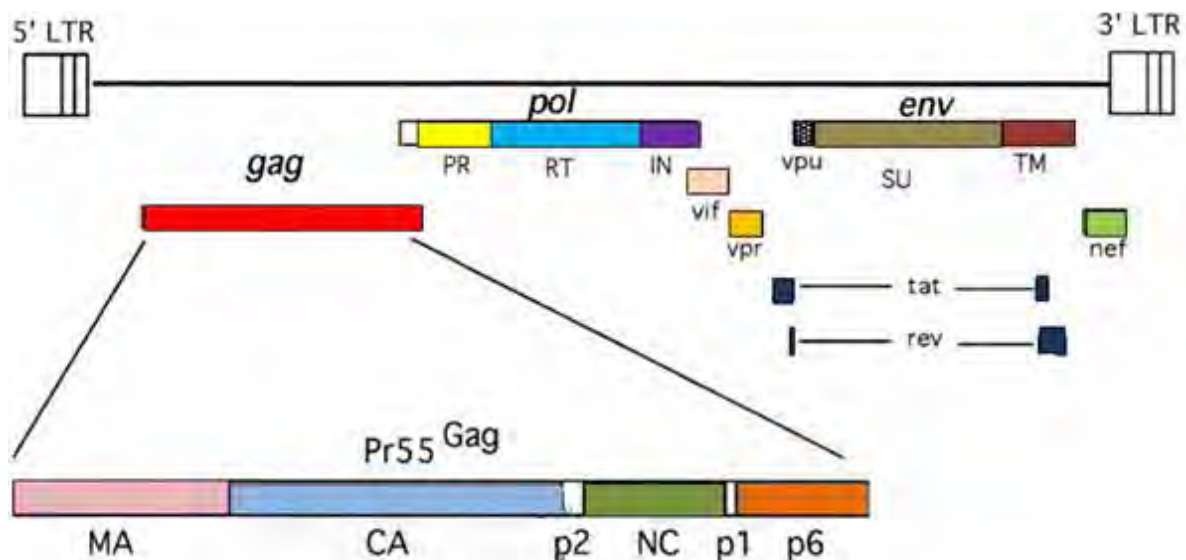


Figura 5 Organización del genoma del VIH-I. Se observan tres marcos de lectura abiertos (ORF) principales que comprenden a los genes *gag*, *pol* y *env*. Además, se observan seis ORF más que codifican para las proteínas accesorias/reguladoras. El fragmento de *gag* está ampliado. Además existen dos secuencias de repetición terminal que flanquean el genoma, LTR5' y LTR3' que contienen secuencias promotoras y reguladoras de transcripción [Obtenido de 20]

Haciendo un análisis en microscopio electrónico, se observa que las partículas envueltas de este virus miden de 100 a 120 nm de diámetro y tienen una bicapa lipídica proveniente de la membrana celular [14]. En esta membrana se localizan dos glicoproteínas en forma de espículas: gp120, localizada en la superficie externa de la membrana y gp41, localizada de forma transmembranal. En la parte interna de esta membrana se encuentra la proteína de matriz (p17, MA), encargada de proveer estructura al virus. Dentro de la matriz está una cápside cónica (p24, CA) y un núcleo que contiene dos cadenas, no complementarias, de RNA (+) de 9.3 Kb que forman el genoma del virus y que se encuentran unidas a la

nucleoproteína p7 (NC), y las enzimas del virus: integrasa (p32, IN), proteasa (p10, PR) y transcriptasa inversa (p66/p51, RT) [21] (**Fig. 6**).

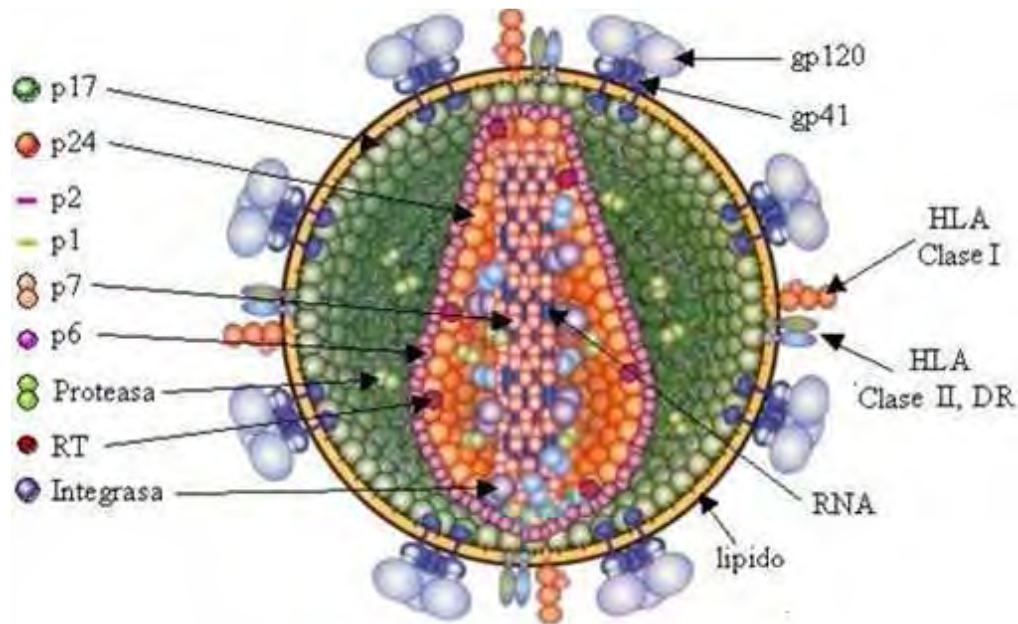


Figura 6. Representación de una partícula madura del VIH. Se muestran la localización aproximada de las proteínas estructurales, las enzimas y el genoma de este virus [Obtenido de 17].

Como con otros retrovirus, los productos de la traducción se sintetizan inicialmente como precursores poli-proteínicos que posteriormente serán procesados por las proteasas virales o celulares hasta obtener proteínas maduras. En total, los genes del VIH codifican para nueve proteínas clasificadas en tres clases: proteínas estructurales: Gag, Pol y Env, que formarán las partículas virales; proteínas reguladoras: Tat y Rev, que participan en los procesos tempranos de regulación de la transcripción de los genes virales; y proteínas accesorias: Vpu, Vpr, Nef y Vif, que favorecen los procesos para la internalización del genoma y permiten la replicación del virus *in vivo* al antagonizar la acción de factores antirretrovirales naturalmente presentes en las células del hospedero [21].

### 2.2.1. Proteínas estructurales.

El gen *gag* codifica para la poliproteína precursora Gag (Pr55<sup>Gag</sup>) cuyo orden de proteínas maduras es el siguiente: p17/p24/p2/p7/p1/p6 [14]. Al momento de la salida de los nuevos viriones, la proteasa viral es la encargada de realizar la proteólisis necesaria para

obtener proteínas individuales. Además, Gag está organizada en 3 dominios necesarios para la formación de las partículas virales: el dominio de unión a membrana (M), el dominio de interacción (I) y el dominio tardío (L) [34].

MA se localiza en la superficie interna de la bicapa lipídica y es responsable de la estabilidad del virus; además, la presencia de ácido mirísitico en su porción amino terminal permite la unión a la membrana plasmática. MA también se ha relacionado con la incorporación de la glicoproteína Env a los viriones y a otros eventos tempranos post-entrada, como el favorecimiento del complejo de preintegración (PIC) al núcleo [20, 21].

CA es la cápside del virus, y está compuesta por 2 dominios: el dominio N-terminal “nuclear”, que es importante para la maduración del virus y tiene un asa flexible que permite la incorporación de la ciclofilina A, la cual es esencial en la pérdida de la cápside, ya que la ausencia de esta interacción inhibe la replicación vírica; y el dominio C-terminal de “dimerización”, que favorece las interacciones Gag-Gag [20, 21].

NC facilita la transcripción inversa (protege de nucleasas e incorpora un RNAt para el inicio del proceso) y es necesaria para el empaquetamiento del genoma del virión (aunado a las 4 asas de la “señal de encapsidación” o  $\Psi$ ), constituyendo la nucleocápside del virus. Estas proteínas presentan unos motivos de “dedos de zinc”, los cuales permiten la interacción con el RNA [20, 21].

La proteína p6 favorece la liberación de los viriones por interacción de dominios terminales de la proteína con componentes del sistema de proteínas ESCRT del hospedero. ESCRT se encarga de dirigir proteínas al sistema de transporte vesicular en el interior de la célula. El VIH tiene la capacidad de usurpar al sistema ESCRT para promover la salida de viriones nuevos de la célula. Las proteínas p1 y p2 son espaciadoras del genoma; sin embargo, estudios recientes han demostrado que son esenciales para el ensamblaje correcto e infectividad [20, 21, 36].

El gen *pol* produce una poliproteína conocida como Gag-Pol (Pr160<sup>Gag-Pol</sup>) y que por autocatálisis deriva en tres enzimas: proteasa, transcriptasa inversa e integrasa [14].

PR es una proteasa aspártica, un homodímero encargado de cortar los polipéptidos durante la gemación del nuevo virión y produciendo cambios estructurales y morfológicos conocidos como maduración viral [21].

RT es una DNA-polimerasa dependiente de RNA y DNA con una actividad RNAsa H que degrada la cadena original de RNA viral pero sin actividad exonucleasa 3' – 5'. Por lo tanto presenta una baja fidelidad que da lugar a una alta tasa de mutación ( $3.4 \times 10^{-5}$  errores/base/ciclo de replicación) y además, genera eventos de recombinación durante el proceso de retro-transcripción al cambiar entre las dos cadenas de RNA genómico viral (elección de copia) [16, 21].

IN permite la inserción del genoma viral dentro del material genético del hospedero mediante tres actividades enzimáticas: actividad exonucleasa (permitiendo la eliminación de dos nucleótidos terminales del extremo 3' del DNA viral), actividad endonucleasa de doble cadena (permite el corte del DNA de la célula en el sitio donde es incorporado el DNA viral) y ligasa (permite la unión covalente de cada extremo del DNA proviral con el DNA de la célula) [21].

El producto del gen *env* es un polipéptido de 160 kDa llamado gp160 o Env, que es cortado por una proteasa celular en dos fragmentos: gp120 (SU) que se localiza en la superficie del virus, y gp41 (TM) localizada en la región transmembranal. Ambas glicoproteínas interactúan mediante enlaces no covalentes formando trímeros que producen las llamadas espículas de la envoltura del virus. SU (la proteína con más variabilidad del VIH) tiene los motivos de unión al receptor CD4, y los correceptores de quimiocinas CCR5 y CXCR4. La unión de gp120 al receptor y correceptor produce cambios conformacionales que exponen el dominio de fusión en TM, permitiendo la unión de la membrana viral con la del hospedero y por tanto la entrada del virus a la célula del hospedero [19, 21].



### 2.2.2. Proteínas reguladoras.

La proteína reguladora Tat es requerida para el inicio de la transcripción inversa por su actividad de transactivador transcripcional. Tat se une al RNA del VIH en un asa de la región 5' terminal (elemento TAR) del genoma viral. La unión de Tat con su elemento TAR activa e incrementa la transcripción de genes del VIH. Cuando es secretada actúa sobre células no infectadas como transductor de señales facilitando una nueva infección [19, 21].

La segunda proteína reguladora es Rev que es un homotetrámero de 13 kDa e induce la transición de la fase temprana a la tardía en la expresión de los genes del VIH. Facilita el transporte de RNA sin *splicing* del núcleo hacia el citoplasma. Se une a una secuencia específica del genoma en una región llamada elemento de reconocimiento de Rev (RRE) [19, 21].

### 2.2.3. Proteínas accesorias.

En sistemas *in vivo* las proteínas accesorias son factores determinantes en la replicación y liberación de nuevas partículas virales, así como en la virulencia del VIH [14].

Nef, una proteína de aproximadamente 27kDa, se encuentra asociada a la cara citoplasmática de la membrana celular. Disminuye la expresión de CD4 en la superficie de células infectadas, favoreciendo su endocitosis. Disminuye la expresión de moléculas HLA de clase I y por ese medio permite la evasión del sistema inmune [19, 21]. Además, Nef interactúa con múltiples proteínas celulares, interfiriendo con cascadas de señalización intracelular con el efecto final de mantener a la célula infectada en un estado ideal para la producción de nuevos viriones.

Vpr es una proteína de 14 kDa que está presente en un complejo de pre-integración (PIC) en el núcleo celular y actúa como un factor de transporte del DNA viral al núcleo del hospedero. Además Vpr tiene la habilidad de arrestar a las células en la fase G2 del ciclo celular, inhibiendo la apoptosis. Recientemente se ha sugerido que Vpr actúa como un

factor soluble que es secretado por células infectadas induciendo la muerte de linfocitos CD4 no infectados, aumentando la depleción de estos [19, 21].

Vpu es una fosfoproteína de 16 kDa exclusiva de VIH-I que reduce la expresión de la molécula CD4 y favorece la liberación de los viriones de las células infectadas. Vpu interacciona con la proteína celular llamada teterina (proteína transmembranal que une los viriones recién sintetizados a la membrana celular de la célula infectada, evitando su salida) dirigiéndola hacia degradación endosomal. La ausencia de Vpu provoca que las partículas virales sean retenidas en vesículas intracelulares y en la membrana plasmática por la actividad de teterina [19, 21].

Vif tiene un peso de 23 kDa y es fundamental para la replicación del VIH en linfocitos T y en macrófagos, además de ser importante para el adecuado empaquetamiento de las proteínas virales. Estudios recientes sugieren que Vif tiene la capacidad de inhibir la actividad de ciertos factores celulares de la familia de APOBEC3G (A3G), que de forma natural bloquean la replicación viral induciendo hipermutación en el genoma viral [19, 21]. Vif sirve de enlace entre A3G y proteínas de ubiquitinación, provocando consecuentemente la degradación proteosomal de A3G.

### 2.3 Ciclo de replicación.

Para su replicación, el VIH debe integrar su genoma al del hospedero y así poder hacer uso de su maquinaria celular para producir nuevos viriones. Por tanto, la replicación del VIH inicia con la unión de gp120 y la molécula CD4 de la célula blanco, la cual funciona como receptor principal y está presente en linfocitos T ayudadores, macrófagos CD4<sup>+</sup> y algunas células dendríticas. Posteriormente la entrada del VIH requiere de la unión de la región V3 de gp120 a un correceptor. Existen varias moléculas que se han identificado como posibles correceptores del virus, siendo los dos principales los receptores de quimiocinas CXCR4 y CCR5. Estas moléculas determinan el tropismo del virus clasificando al virus en M-trópicos o R5 si el correceptor usado es CCR5 y T-trópicos o X4 si el correceptor usado es CXCR4. Se han aislado algunos virus cuya afinidad es intermedia, definiéndose como virus de tropismo dual. En la mayoría de los casos, al principio de la infección predominan los

virus R5 (etapa temprana) y en etapas posteriores predominan virus X4 [11, 14, 21]. Esta unión producirá cambios conformacionales exponiendo a gp41 que permitirá el acercamiento y fusión de la membrana viral con la membrana de la célula blanco. De esta manera la cápside viral entra al citoplasma del hospedero. A continuación el RNA viral se transcribe a DNA de doble cadena por medio de la RT. El dominio con actividad de RNAsa H de la RT degrada la porción de RNA del híbrido RNA-DNA formado, dejando una cadena sencilla de DNA que se complementará para formar las dos hebras listas para su inserción en el genoma de la célula hospedera. Esta inserción se logra mediante la migración al núcleo del complejo de preintegración (PIC) que incluye a IN, NC, Vpr y el DNA recién sintetizado y que presenta señales de localización nuclear en varias de las proteínas del complejo. Una vez que el PIC se encuentra dentro del núcleo, la IN forma extremos cohesivos y la ligasa une el DNA viral con el DNA de la célula hospedera. Al DNA viral insertado en el genoma de la célula se le conoce como provirus y funciona como un gen celular durante el tiempo de vida de la célula infectada [14, 21].

Después de la integración y para comenzar la replicación del VIH es necesaria la inducción de sus genes por medio de factores de transcripción celular (que son producidos en forma constitutiva por la célula o inducidos por señales de activación como citocinas) y proteínas virales reguladoras como Tat. Los primeros RNAm virales producidos sufren varios eventos de corte y empalme (splicing), formando, en la etapa temprana, las proteínas Tat, Rev, Nef, Vif, Vpr y Vpu. En la fase tardía de la infección se favorece la concentración de Rev, lo que favorece la salida de los RNAm virales del núcleo, evitando así los eventos de corte y empalme. En consecuencia, los RNAm virales de la etapa tardía se caracterizan por tener uno o ningún evento de corte y empalme, lo que conlleva a la producción de las poliproteínas Pol, Env y Gag y los RNAs que se empaquetarán en las nuevas partículas virales [14, 21].

Los precursores Gag, Pol, Env y el nuevo genoma de RNA se dirigen a la superficie celular y son ensamblados para estructurar viriones. El transporte y salida de las nuevas partículas virales son dependientes de Vpu y las proteínas celulares del sistema ESCRT en unión con p6. Al gemar los nuevos viriones, éstos adquieren parte de la membrana lipídica del

hospedero y los productos del gen *env* (gp120 y gp41). Durante la maduración del nuevo virión, la enzima proteasa (autocatalizada) corta específicamente cada poliproteína para formar las proteínas virales funcionales. El VIH maduro será capaz de infectar a otras células y comenzar nuevamente su ciclo de replicación [14, 21] (Fig. 7).

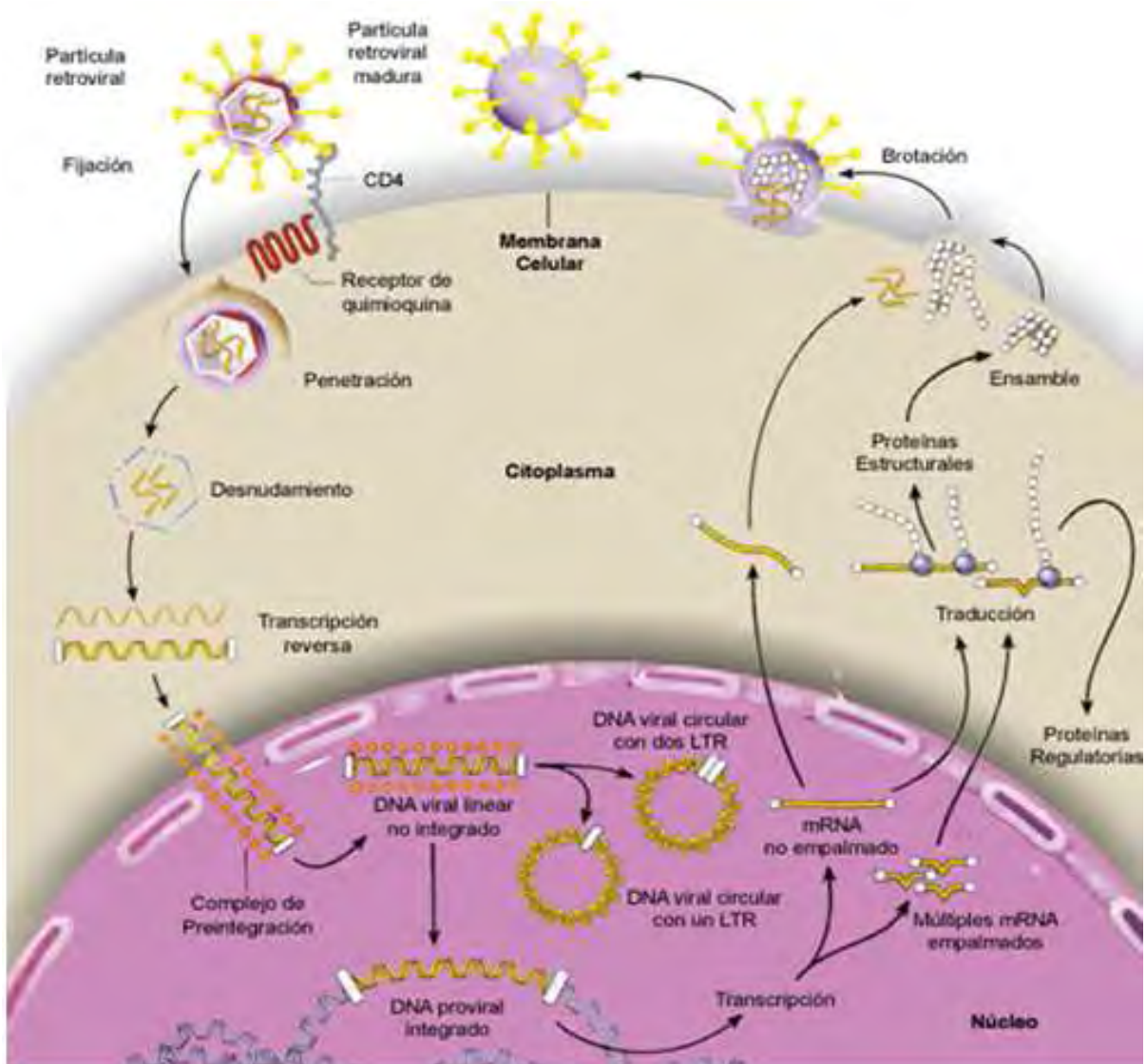


Figura 7. Descripción esquemática de los eventos que ocurren durante el ciclo de replicación del VIH. Interacción entre gp120, CD4 y el receptor de quimiocinas (CCR5 o CXCR4). Fusión de membranas gracias a gp41, entrada de la CA, desnudamiento del virión, transcripción inversa del genoma viral, formación del complejo de preintegración. Transporte del complejo de preintegración al interior del núcleo e integración del DNA viral formando el provirus. Durante la transcripción temprana del DNA viral se generan las proteínas reguladoras, mientras que en las etapas tardías, se producen las proteínas estructurales y los genomas de RNA viral que se empaquetarán en las nuevas partículas virales. Transporte de RNA viral y poliproteínas a la membrana plasmática, ensamblaje y liberación de nuevos viriones por gemación. Maduración del virión por el corte de precursores gracias a la PR. Inicio de nuevo ciclo de replicación [Obtenido de 17].

#### 2.4. Historia natural de la enfermedad.

Durante la infección por el VIH se distinguen diversas etapas que son inducidas y diferenciadas por la interacción entre el virus y la respuesta inmune del hospedero. La historia natural de la enfermedad se divide en las siguientes etapas sucesivas: etapa de infección primaria o aguda caracterizada por altas cargas virales, asociadas a una replicación descontrolada del virus, así como una caída temporal de la cuenta de Linfocitos T CD4<sup>+</sup> en sangre; etapa intermedia o de latencia clínica, caracterizada por una estabilización de la carga viral y la cuenta de linfocitos T CD4<sup>+</sup> en sangre, con concentraciones entre 200 a 500 células/ $\mu$ l; etapa tardía o crónica donde los pacientes presentan concentraciones por debajo de 200 células T CD4<sup>+</sup>/ $\mu$ l, perdiéndose el control parcial de la replicación viral con aumentos en la carga viral en sangre e incrementándose el riesgo de adquisición de infecciones oportunistas. Es en esta última etapa cuando se presenta el SIDA y resulta en la muerte del hospedero [14, 28].

Durante las primeras seis semanas después de la exposición, los individuos infectados desarrollan un síndrome agudo, caracterizado por valores elevados de viremia (carga viral) en la sangre periférica y síntomas de tipo gripal. La fuente principal de viremia parece ser los LT CD4<sup>+</sup> CCR5<sup>+</sup> de memoria, distribuidas principalmente en las mucosas. Al mismo tiempo ocurre una notable reducción de las LT CD4<sup>+</sup> en circulación; pero no tan acentuada como sucede en las mucosas [28].

Posteriormente (6 a 8 semanas) se monta una respuesta inmune específica, principalmente celular, contra el VIH que provoca un descenso importante en la viremia plasmática, aumentando también los niveles de LT CD4<sup>+</sup> en sangre, pero muy rara vez hasta niveles de referencia. En la fase subsiguiente, los individuos entran en una etapa asintomática clínicamente estable que, dependiendo de factores inmunológicos, virológicos y genéticos, puede durar un número variable de años [28]. En esta etapa las concentraciones del virus en la sangre periférica son relativamente bajas, aunque la producción de nuevos viriones, la muerte y renovación de células blanco es constante y

ocurre a alto nivel. La replicación del virus e infección de células hospederas son procesos dinámicos que se equilibran por los procesos de depuración viral, muerte de células infectadas y abastecimiento de nuevos LT CD4<sup>+</sup>, lo cual mantiene niveles plasmáticos estables de viremia y de LT CD4<sup>+</sup> [28]. Sin embargo, esta viremia plasmática persistente aunada a varias estrategias del virus como la formación de reservorios de VIH latentes (LT CD4<sup>+</sup> y macrófagos), su gran capacidad de variación genética, el cambio de fenotipo biológico durante el curso de la enfermedad y la extraordinaria tasa de replicación, provocan la reducción gradual de los LT CD4<sup>+</sup>, que invariablemente conduce a SIDA [14] (Fig. 8).

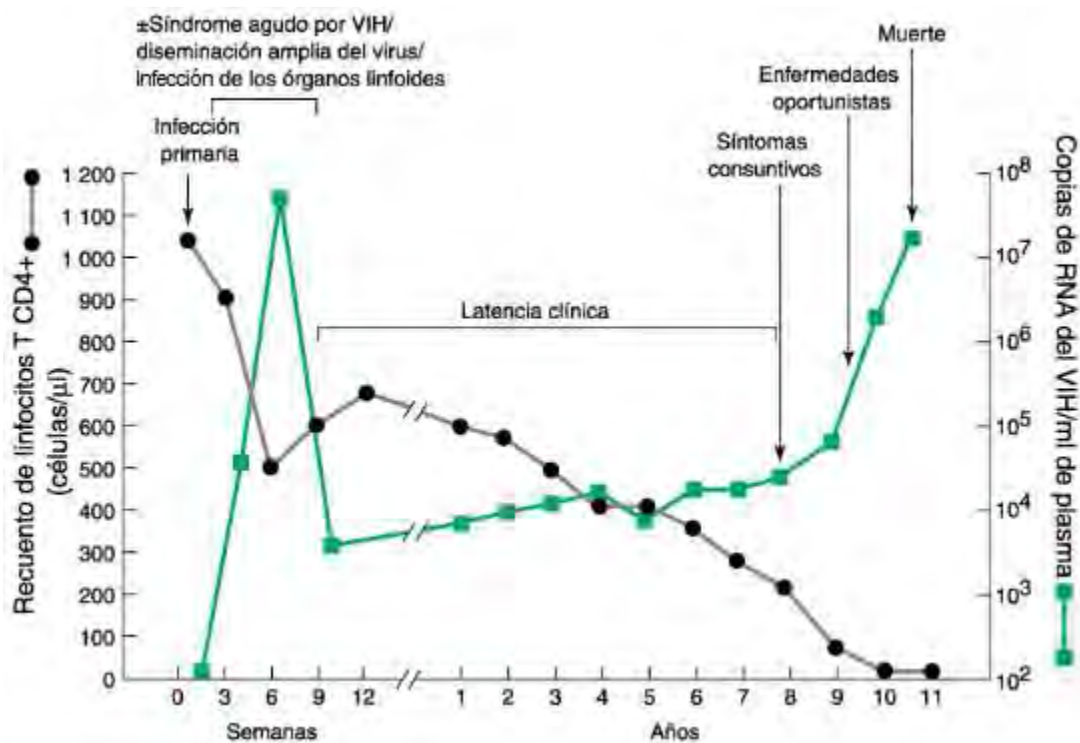


Figura 8. Historia natural de la enfermedad causada por VIH. Se observa la variación de los linfocitos T CD4<sup>+</sup> y del número de copias de RNA del VIH a través de las diversas etapas en las que se divide la enfermedad [Obtenida de17].

La fase de SIDA se caracteriza por una inmunodeficiencia grave, lo que conlleva a una mayor susceptibilidad de contraer enfermedades oportunistas como tuberculosis, neumonía o histoplasmosis, que eventualmente ocasionan la muerte del individuo

infectado. Con el objetivo de controlar la replicación del VIH, en la actualidad se utilizan combinaciones de fármacos que interfieren con diversos eventos en el ciclo de replicación del virus [17,28]. La administración oportuna y ordenada de estos fármacos permite un aumento en el número de LT CD4<sup>+</sup> en sangre y una disminución hasta niveles indetectables de la carga viral en plasma, lo que detiene la progresión de la enfermedad y disminuye significativamente la incidencia de enfermedades oportunistas con una mejoría notable en la calidad de vida del individuo [14, 28].

### **3. Respuesta inmune contra el VIH.**

#### **3.1. Respuesta innata**

La inmunidad innata es la primera línea de defensa durante la fase temprana de la infección por VIH. Las células del sistema inmune implicadas son: células asesinas naturales (NKs), monocitos, linfocitos T gamma/delta (LT  $\gamma\delta$ ) y células dendríticas (DCs) [28]. Estas células tienen inicialmente funciones efectoras y posteriormente adquieren funciones reguladoras, ayudando a las células de la inmunidad adaptativa. Aunado a lo anterior, las células de la inmunidad innata tienen la capacidad de secretar citocinas como IL-2, IL-12, IL-4, IL-10, IL-15 e IFN- $\gamma$ , TNF- $\alpha$ , y quimiocinas como MCP-2, MIP-1 $\alpha$ , MIP-1  $\beta$  e IFN $\alpha/\beta$  que ayudan al control de la infección [1].

Las NKs tienen un papel crítico en el control de las infecciones virales, siendo células citotóxicas efectoras de la respuesta inmune innata capaces de eliminar de forma directa o por mecanismos mediados por anticuerpos, a las células infectadas. Las células NKs además son productoras de IFN- $\gamma$  [1].

Los linfocitos T  $\gamma\delta$  se localizan en las mucosas y generalmente no reconocen antígenos peptídicos presentados por moléculas HLA, sino que interactúan directamente con antígenos no peptídicos o con proteínas de estrés celular. Los LT  $\gamma\delta$  producen citocinas tipo T<sub>H</sub>1 y T<sub>H</sub>2, además lisan a células blanco infectadas por VIH. Estas células también generan factores antivirales supresores como RANTES (Regulador durante la Activación

Normal de linfocitos T Expresados y Secretados), MIP-1 $\alpha$  y MIP-1 $\beta$  que se unen al co-receptor del virus, CCR5, disminuyendo su expresión [1].

Las DCs expresan co-receptores de quimiocinas que son usados por el VIH para entrar a la célula y que las atraen a órganos linfoides. Las DCs son importantes como mediadores del reconocimiento de antígenos y activan respuestas innatas y adaptativas mediante la secreción de citocinas (IL-7, IL-1, IFN, TNF- $\alpha$  e IL-12) [14, 28].

### 3.2. Linfocitos T CD4<sup>+</sup>.

Los linfocitos T CD4<sup>+</sup> o linfocitos T cooperadores (Th) son una población celular cuya función está determinada de acuerdo a las citocinas que producen y secretan. Por tanto, se pueden dividir en Th1, que producen citocinas con efectos proinflamatorios y citotóxicos; y las Th2, que producen citocinas promotoras de la síntesis de inmunoglobulinas, cambio de isotipo y reacciones anafilácticas. En el desarrollo de una respuesta inmune se favorece uno u otro tipo; en el caso del VIH, se ven favorecidas las Th1 [1, 28] (**Fig. 9**).

Cuando una célula presentadora de antígeno (APC) procesa un antígeno y lo presenta a un Th en el contexto de moléculas del Complejo Mayor de Histocompatibilidad (MHC) tipo II también llamadas Antígeno Leucocitario Humano (HLA) tipo II, en el caso de los humanos, se llevan a cabo tres procesos en los que las citocinas secretadas juegan un papel fundamental [1]:

- a) Citotoxicidad mediada por linfocitos T CD8<sup>+</sup>.
- b) Activación de macrófagos por citocinas tipo Th1.
- c) Producción de anticuerpos por linfocitos B activados por citocinas tipo Th2.



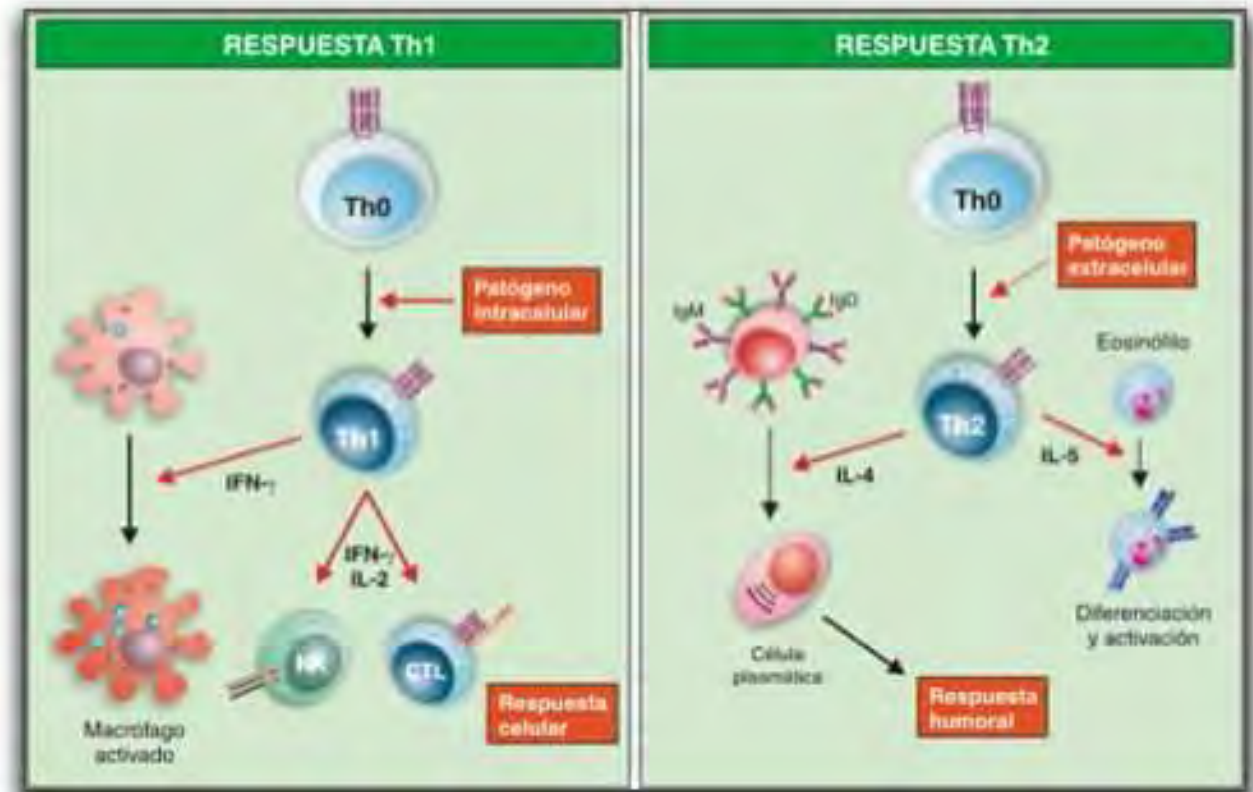


Figura 9. Representación esquemática de las funciones de los dos tipos de Linfocitos T CD4+ (Th1 y Th2). Se muestran las células y citocinas asociadas a cada actividad [Obtenida de 1].

### 3.3. Anticuerpos.

La respuesta inmune por parte de los linfocitos B se desarrolla dentro de la primera semana después de encontrar valores detectables de viremia y son determinados de forma inicial como complejos antígeno-anticuerpo. A esta fase le sigue la aparición de anticuerpos anti-gp41 y anticuerpos anti-pg120. Sin embargo, estos anticuerpos no tienen efectos importantes sobre la viremia.

Por otro lado, no es sino hasta varios meses después que los anticuerpos neutralizantes (NABs) aparecen. Los NABs actuarán contra la cepa infectante y no son capaces de neutralizar virus divergentes. A pesar de que los NABs juegan un papel en el control de infecciones virales, la habilidad del VIH para evadir la respuesta inmune y su amplia diversidad antigénica provoca que el control ejercido sea limitado.

Adicionalmente, los anticuerpos pueden actuar en forma conjunta con células efectoras para provocar citotoxicidad mediada por anticuerpos, eliminando de esta manera a las células infectadas. En cuanto al rol protector que juegan los anticuerpos en las mucosas se conoce poco y es necesario realizar más estudios. En general, más investigación básica es necesaria para entender el impacto de los anticuerpos en la transmisión y patogénesis del VIH, ayudando a lograr un diseño racional de vacunas contra este virus [38].

#### 3.4. Linfocitos T CD8<sup>+</sup>.

Los linfocitos T CD8<sup>+</sup> o citotóxicos (CTLs) son esenciales para el control de infecciones virales. Los CTLs se estimulan por antígenos presentados a su receptor de células T (TCR), en el contexto de moléculas HLA clase I de una APC. La APC reconoce a un antígeno endógeno extraño, como puede ser un péptido viral, lo procesa y lo presenta en forma de péptidos de 8 a 10 aminoácidos en moléculas HLA clase I. El TCR de los CTLs reconoce este complejo antígeno-HLA clase I por medio de un mecanismo tipo “llave-cerradura” y dispara las funciones efectoras de los CTLs. Estas funciones pueden dividirse en dos tipos: producción de factores solubles (citocinas y quimiocinas) y función citotóxica o lisis de la célula infectada [1].

En el caso de la liberación de factores solubles, los linfocitos T CD8<sup>+</sup> pueden secretar IFN- $\gamma$  que tiene actividad antiviral pues, por una compleja cascada de activación, provoca que las células cercanas se vuelvan relativamente resistentes a una infección viral productiva. Además los CTLs también producen y secretan  $\beta$ -quimiocinas (MIP-1 $\alpha$ , MIP-1 $\beta$  y RANTES) que se unen a la molécula CCR5, bloqueando así la unión del VIH con su co-receptor, promoviendo la internalización de este co-receptor y disminuyendo la habilidad del VIH para infectar las células blanco [1, 28].

Por otro lado, se sabe que los CTLs median la muerte de la célula blanco por dos vías principales: la primera es la vía gránulo-dependiente, donde son liberadas moléculas citotóxicas como perforinas (ayudan a la permeabilización de la membrana de la célula infectada) y granzimas (activan cascadas de señalización de apoptosis), causando la muerte de la célula infectada. La segunda vía es independiente de gránulos citotóxicos y

en ésta la molécula FasL (ligando de Fas) presente en la membrana de los CTLs encuentra al receptor Fas (CD95) en la célula infectada. Al unirse a su ligando el receptor es capaz de inducir la muerte por apoptosis de la célula [1] (**Fig. 10**).

Los CTLs por tanto, requieren establecer un contacto íntimo con la célula blanco para establecer su acción citotóxica. Además, esta interacción requiere de dos condiciones para ser eficiente [1]:

- 1.- El CTL debe ser viable y estar metabólicamente activo pues su adherencia con la célula blanco requiere de energía.
- 2.- El medio extracelular debe contener cationes  $Ca^{2+}$  y  $Mg^{2+}$  pues sus enzimas son dependientes de estos cationes.

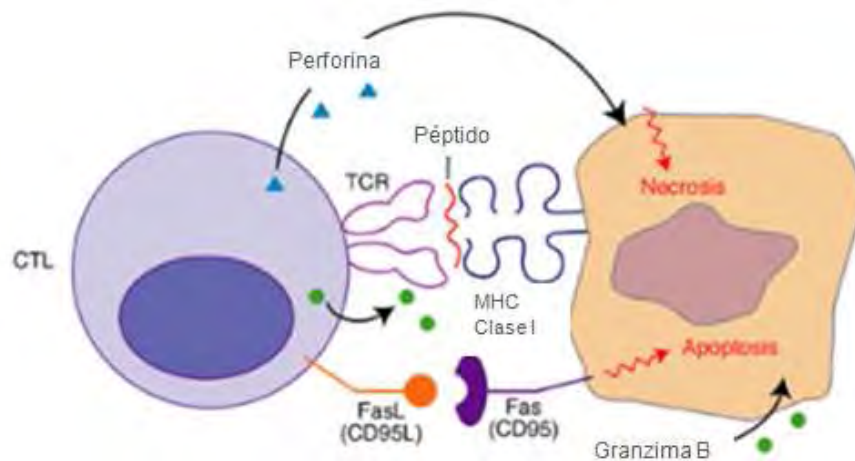


Figura 10. Esquema que representa los efectos de la activación de un Linfocito T CD8+ (CTL) sobre una célula infectada. Se observa la vía gránulo-dependiente y la vía independiente de gránulos citotóxicos (vía Fas-FasL). Ambas provocando la muerte de la célula infectada [Modificado de 40].

Quizá una de las evidencias más sólidas sobre la importancia de la respuesta de CTLs en la infección por VIH proviene de estudios realizados en simios infectados con SIV. Al depletar, *in vivo*, los CTLs de macacos *rhesus* se observó una pérdida del control en la replicación del SIV, demostrándose que cuando la duración de esta depleción es mayor a 28 días, la viremia primaria nunca es controlada y los macacos mueren con un síndrome parecido al SIDA. Además, la depleción transitoria de las CTLs en macacos *rhesus* infectados crónicamente con SIV, se asocia con un aumento sustancial de la replicación

viral, que regresa a sus niveles basales cuando la población de linfocitos T CD8<sup>+</sup> reaparece [5, 35].

### 3.5. Presión selectiva por HLA.

Como se ha mencionado anteriormente, los CTLs son sumamente importantes para el control de la viremia, especialmente en las etapas tempranas de la enfermedad. El proceso por el cual los CTLs reconocen a células infectadas se ha descrito anteriormente, mencionando la importancia de las moléculas HLA clase I. Aunado a lo anterior, se ha descrito también que el VIH presenta una gran capacidad de variación, generando múltiples variantes ligeramente distintas entre sí, llamadas cuasiespecies virales, que coexisten en cualquier momento en un individuo infectado. Las respuestas de CTLs, al controlar parcialmente la viremia, ejercen una fuerte presión selectiva sobre el virus [46]. Dada la gran capacidad de variación del VIH, es posible que se seleccionen variantes con mutaciones dentro de o adyacentes a los fragmentos de las proteínas virales que son presentados por las moléculas HLA clase I a los CTLs. Estas mutaciones, llamadas de escape inmune, provocarán que el péptido viral se vuelva incapaz de unirse a la molécula HLA o que el epítipo viral en cuestión ya no sea reconocido por el TCR del CTL; aunque también puede ser que la mutación se encuentre en los extremos del epítipo y entonces es el procesamiento de antígeno el que se ve afectado [16]. De esta manera, el virus puede evadir de manera muy eficiente las respuestas inmunes de CTLs, definidas y restringidas por los genes HLA clase I [41].

Ahora es bien sabido que el VIH es capaz de adaptarse a la respuesta de CTLs, mediante la selección de mutaciones de escape que evitan el procesamiento, la presentación y/o el reconocimiento de epítipes virales presentados por moléculas HLA clase I [16]. De los tres genes que codifican para moléculas HLA de clase I clásicas: A, B y C, capaces de presentar antígeno, se ha observado que el gen B, el más polimórfico de los tres, es el más importante para el control viral [26]. En efecto, respuestas restringidas por moléculas HLA-B específicas, frecuentemente se asocian a protección o riesgo para el avance de la enfermedad por el virus. Más aún, las respuestas protectoras generalmente se asocian a

epítomos presentes en la proteína Gag viral [23, 27]. En este sentido, varios estudios han mostrado la asociación de HLA-B\*5701 y HLA-B\*2705 con un control eficiente de la infección y un avance lento de la enfermedad, mientras que HLA-B\*3502 y HLA-B\*3503 consistentemente se asocian a un control deficiente de la replicación y un avance rápido de la enfermedad [16, 27].

La adaptación del VIH a las respuestas de CTLs puede también observarse a nivel poblacional. Estudios previos han demostrado que los genes HLA están dejando una huella poblacional característica en la evolución del virus al seleccionarse mutaciones de escape inmune que se ven reflejadas en la mayoría de los virus circulantes, de acuerdo a la frecuencia alélica HLA de la población [5, 6, 7, 8]. Dado el alto polimorfismo de los genes HLA de clase I y la distribución altamente variable de los diferentes alelos de estos genes en distintas poblaciones del mundo, existe ya evidencia de que la evolución dirigida por HLA del VIH es distinta en poblaciones inmunogenéticamente diferentes. Nuestro grupo de investigación ha demostrado que las características inmunogenéticas únicas de la población mexicana han determinado características únicas en la adaptación del virus en nuestra población, al seleccionarse mutaciones de escape en los virus circulantes por alelos HLA que son altamente frecuentes en los mexicanos y que difieren de los observados en otras poblaciones como las caucásicas [6] (**Fig. 11**).

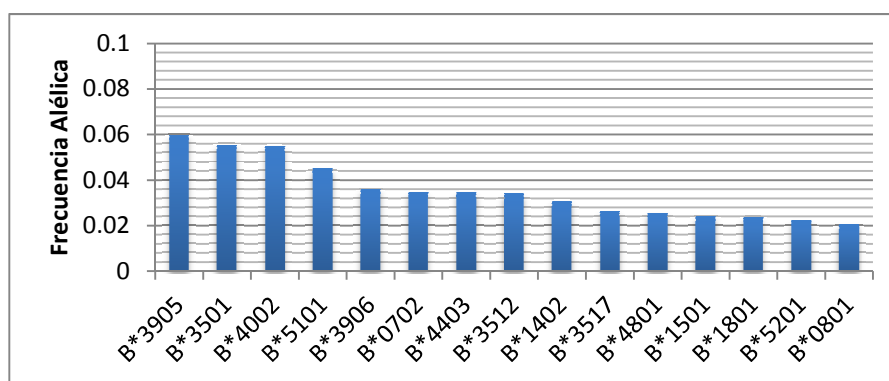


Figura 11. Gráfico mostrando los 15 alelos HLA-B más comunes en la población mexicana y sus frecuencias alélicas. Cohorte formada por 785 pacientes provenientes de diferentes estados de la república; incluyendo, DF, Estado de México, Hidalgo, Puebla, Chiapas, Oaxaca, Guerrero, Baja California, Sonora, Nuevo León, Tamaulipas, Sinaloa, Jalisco, Morelos y Veracruz. Las edades van de los 16 hasta los 60 años, siendo la media 20 años. Con valores de 78.6% hombres y 21.4% mujeres. Con valores medios de CV de 39,500 copias/ $\mu$ L y T CD4+ de 409.33 células/ $\mu$ L [Obtenido del CIENI].

Aunque las mutaciones de escape a CTLs confieren una ventaja al virus en contra las defensas naturales del cuerpo, su presencia puede cambiar la estructura y función de las proteínas virales que las presentan, provocando un costo en la capacidad replicativa del virus. En este sentido, se han demostrado asociaciones discordantes entre la presencia de respuestas de CTLs dirigidas a diferentes proteínas virales y los niveles de carga viral [27]. Entre los muchos epítopes a lo largo de todo el proteoma viral que pueden ser presentados a CTLs, son los epítopes presentes en Gag, especialmente p24 por ser una secuencia muy conservada e inmunogénica, los que tienen un mayor impacto en la capacidad replicativa viral y por tanto en la disminución de la viremia [2, 16, 23, 25, 26, 32, 42, 46].

El costo en la capacidad replicativa del virus debido a mutaciones de escape se ha relacionado a las distintas evoluciones que puede presentar la enfermedad en los pacientes [27, 30]. Por ejemplo, los alelos B\*5701 y B\*2705 están relacionados con una progresión lenta a SIDA, seleccionando mutaciones de escape en Gag, con altos costos en la capacidad replicativa viral [23]. En la literatura se han descrito distintas mutaciones con costo en la capacidad replicativa en la proteína Gag, como: 12 (E→Q), 34 (I →L) y 371 (Delección de T), todas en fase aguda; y 58 (R→G), 218 (V→Q) y 482 (E→D) en fase crónica [7].

La permanencia de estas mutantes dependerá de la presencia de presión selectiva y del costo de las mismas en la capacidad replicativa del virus. La presencia del alelo HLA selector hará que permanezcan las mutaciones, sin importar si son de alto o bajo costo viral; por otro lado, la ausencia de este alelo provocará que las mutantes con alto costo reviertan y las de bajo costo permanezcan, volviéndose el consenso poblacional con el avance de la epidemia [3, 29].

Por último, se debe considerar que conforme progresa la enfermedad hacia la fase crónica, el virus se irá adaptando y equilibrando esta disminución en la capacidad replicativa por medio de la selección de las llamadas mutaciones compensatorias [9, 15, 37, 39]. Estas mutaciones restablecen la capacidad replicativa del virus y se pueden

localizar dentro o fuera del epítotope donde está la mutación con costo en capacidad replicativa; como ejemplos se tiene: 219 (H→Q) y 248 (G→A) [7]. La mayoría de los mecanismos por los cuales actúan estas compensaciones aún no se encuentran bien definidos [16].

## CAPÍTULO II. PLANTEAMIENTO DEL PROBLEMA.

### 1. Planteamiento del problema.

Los CTLs reconocen células infectadas por VIH mediante el reconocimiento de péptidos virales presentados por moléculas HLA de clase I, a través de su TCR. Las células infectadas reconocidas son destruidas por la acción de perforinas (provocan poros en la membrana) y granzimas (serín proteasas que inducen la apoptosis), liberadas por los CTLs [1]. Por tanto, la respuesta de los linfocitos T citotóxicos se encuentra definida y restringida por los HLA de clase I. Dada su efectividad para el control parcial de la viremia, especialmente en etapas tempranas de la infección, las respuestas de CTLs ejercen una fuerte presión selectiva que dirige la evolución del VIH [6]. Como consecuencia de esta presión evolutiva se seleccionan virus con mutaciones de escape que evitan la unión de péptidos virales a moléculas HLA, que afectan el procesamiento de los epítomos antes de su presentación o que disminuyen la afinidad de la unión TCR-HLA/péptido [2]. La presencia de estas mutaciones de escape inmune está muchas veces asociada a un costo en la capacidad replicativa del virus, lo que conduce a una reversión al transmitirse a un individuo carente del HLA selector [23]. Sin embargo, existe evidencia de que algunas mutaciones de escape pueden ser transmitidas y mantenidas en el hospedero, favoreciéndose así su acumulación y fijación en la mayoría de los virus circulantes en una población, al tener bajos costos en la capacidad replicativa viral o gracias a la aparición de mutaciones compensatorias (mutaciones que restablecen la capacidad replicativa) [3, 28].

Diversos factores a nivel poblacional pueden impactar en la permanencia de las mutaciones de escape en los virus circulantes, lo cual resalta la complejidad de la adaptación viral a la respuesta inmune dentro de una población [18]. A pesar de esta complejidad hay evidencia de que existe una huella de adaptación del VIH a las respuestas restringidas por HLA a nivel poblacional que pueden variar según los contextos inmunogenéticos poblacionales distintos [7, 23]. La existencia de la evolución del VIH mediada por HLA a nivel poblacional se ha demostrado por asociaciones entre la presencia



de polimorfismos en posiciones determinadas de alguna proteína del virus y la expresión de alelos HLA específicos [7, 18, 23].

Debido al alto polimorfismo de los genes HLA, la frecuencia de alelos en distintas poblaciones es diferente, pudiendo impactar de manera distinta en la evolución del VIH según el contexto inmunogenético de la población de que se trate [16]. Por lo anterior, es probable que ciertas asociaciones HLA-polimorfismo se observen preferentemente en cierta población, determinando características únicas de la evolución del VIH en diferentes grupos humanos. Varios estudios sobre el impacto que tienen las mutaciones seleccionadas por HLA en el VIH han sido conducidos en diferentes poblaciones alrededor del mundo, observándose que el patrón de aparición de estas mutaciones varía según las combinaciones específicas de HLA presentes en cada una de las cohortes estudiadas [16, 18]. Por tanto, estudios de la evolución dependiente de HLA del VIH circulante en población mexicana son importantes debido a las características inmunogenéticas únicas que presenta nuestra población [5]. Estos estudios permitirán conocer de qué manera está evolucionando el VIH en este contexto inmunogenético. Más aún, estudios en poblaciones específicas sobre la adaptación del virus a la respuesta inmune podrían ayudar a identificar y entender los correlatos de protección o de riesgo de avance de la enfermedad por VIH. El conocimiento de la evolución mediada por HLA del VIH en México tiene aplicación en la investigación en vacunas candidatas que pudieran tener efectividad en el contexto de nuestra población, y podría ayudar en la optimización de estrategias de control del virus.

El presente estudio pretende por primera vez describir el impacto de la presión selectiva por HLA sobre la variabilidad de la proteína Gag en virus circulantes en población mexicana. A partir de secuencias del gen gag de VIH de una cohorte de pacientes mexicanos, se evaluó la diversidad poblacional del gen, así como la posibilidad de que esta diversidad esté asociada a la presión selectiva inmune a nivel poblacional sobre el virus, en el contexto inmunogenético de la población mexicana. Dada la importancia previamente descrita de las respuestas de CTLs dirigidas a Gag para el control inmune del VIH, los resultados de este estudio contribuirán para apoyar la investigación sobre

vacunas en el contexto global, y ayudarán en la identificación de posibles sitios blanco para la generación de inmunógenos que pudieran ser útiles para controlar el virus en nuestros pacientes.

### CAPÍTULO III. HIPÓTESIS Y OBJETIVOS.

#### 1. Hipótesis.

Dada la distribución de frecuencias característica de los alelos HLA en la población mexicana, se espera que el VIH circulante en México presente características adaptativas únicas a la presión selectiva inmune impuesta por CTLs.

Se espera que parte de la variabilidad inherente de la proteína Gag del VIH circulante en México pueda ser explicada por la presencia de presión selectiva inmune sobre el virus.

Algunas de las mutaciones de escape seleccionadas por alelos HLA frecuentes en México en el gen *gag* del virus circulante en el país podrían tener impacto en la capacidad replicativa del virus, explicando posibles correlatos de control del virus en el contexto inmunogenético de los mexicanos.

## 2. Objetivos.

### 2.1. Generales.

Describir la variabilidad del gen *gag* del VIH circulante en México, evaluando el impacto de la presión selectiva inmune sobre esta variabilidad y sobre su capacidad para replicarse, en el contexto inmunogenético de nuestra población.

### 2.2. Particulares.

- Obtener las secuencias del gen *gag* de pacientes mexicanos infectados con VIH-I.
- Determinar las relaciones filogenéticas de los virus presentes en los pacientes con respecto a virus de referencia, a través del gen *gag*.
- Evaluar la variabilidad de las diversas proteínas codificadas por el gen *gag* en el virus circulante en México.
- Asociar la variabilidad en la secuencia del gen *gag* de virus circulante en México con la presencia de presión selectiva positiva, mediante el uso de razones de sustituciones sinónimas y no sinónimas.
- Identificar la presencia de mutaciones adaptativas seleccionadas por alelos HLA en el gen *gag* de virus circulantes en México.
- Determinar la presencia de mutaciones de escape previamente reportadas como mutaciones con impacto en la capacidad replicativa viral, en el gen *gag* de virus circulante en México.
- Determinar la presencia de posibles nuevas mutaciones en el gen *gag* con posible impacto en la capacidad replicativa viral, en el contexto inmunogenético característico de la población mexicana.

## CAPÍTULO IV. METODOLOGÍA.

### 1. SELECCIÓN DE PACIENTES.

La cohorte estudiada consistió en 200 pacientes mexicanos infectados con VIH-I subtipo B, los cuales eran vírgenes a tratamiento antirretroviral y presentaban una progresión típica de la enfermedad. Las muestras de plasma fueron elegidas al azar del total de muestras existentes en el banco de Plasmas del Proyecto de Epidemiología Molecular y Evolución del VIH en México del Centro de Investigación en Enfermedades Infecciosas (CIENI), constatando que se cumplieran los criterios de inclusión antes mencionados.

### 2. EXTRACCIÓN DE RNA.

A partir de las muestras de plasma seleccionadas se realizó la extracción de RNA viral con el estuche de reactivos QIAamp® Viral RNA Purification Kit (QIAGEN, Valencia, CA). Brevemente, en esta técnica se utilizó una solución de lisis de tiocianato de guanidina, para liberar el RNA viral, permitiendo eliminar las RNasas presentes. Después, la muestra fue colocada en una columna de sílice lo cual permitió que el RNA viral presente se uniera a la columna. La columna fue lavada dos veces con distintas soluciones amortiguadoras para eliminar posibles contaminantes. Finalmente, el RNA viral fue eluido utilizando una solución de baja fuerza iónica.

### 3. AMPLIFICACIÓN DEL GEN *gag* DE VIRUS LIBRE EN PLASMA.

Para amplificar la muestra de RNA extraído se utilizó una Reacción en Cadena de la Polimerasa (PCR) anidada con previa transcripción inversa. Esta técnica permitió amplificar un fragmento de *gag* de aproximadamente 1500 pb. Para ambas rondas de amplificación se utilizó un termociclador Veriti 96 well Thermal Cycler (Life Technologies, Carlsbad, CA).

Para la primera ronda de amplificación se utilizaron los primers “forward” 623Fi universal (5'-AAATCTCTAGCAGTGCGCCCGAACAG-3') y “reverse” 2cRx (5'-

TAACCCTGCGGGATGTGGTATTCC-3'); así como, la enzima SuperScript® III One-Step RT-PCR Platinum® Taq High Fidelity (Invitrogen, Carlsbad, CA).

Componente	Volumen por muestra
dH2O	8.5 µL
Buffer 2x	12.5 µL
Primer "forward" (20 pmol/µL)	0.5 µL
Primer "reverse" (20 pmol/µL)	0.5 µL
RT-Platinum Taq polimerasa	0.5 µL
RNA viral	2.5 µL
Volumen Final	25 µL

Se aplicó el siguiente programa para el termociclador: 1 ciclo de 30 min a 55°C y 2 min a 94°C, seguido de 35 ciclos de 15 s a 94°C, 30 s a 55°C y 2 min a 68°C, finalmente se aplicó 1 ciclo de 5 min a 68°C.

Por otro lado, para la segunda ronda de amplificación se utilizaron los primers "forward" G1 (5'-GCAGGACTCGGCTTGCTGAA-3') y "reverse" G10 (5'-TATCATCTGCTCCTGTATC-3'), así como, la enzima Platinum® Taq DNA polymerase (Invitrogen, Carlsbad, CA).

Componente	Volumen por muestra
dH2O	16.25 µL
Buffer 10x	2.5 µL
MgCl (50 mM)	1.0 µL
dNTP's mix (10 mM)	0.5 µL
Primer "forward" (20 pmol/µL)	1.0 µL
Primer "reverse" (20 pmol/µL)	1.0 µL
Platinum Taq DNA polimerasa	0.25 µL
RNA viral	2.5 µL
Volumen Final	25 µL

Se aplicó el siguiente programa para el termociclador: 1 ciclo de 3 min a 94°C, seguido de 35 ciclos a de 30 s a 94°C, 30 s a 56°C y 2 min a 72°C, finalmente se aplicó un ciclo de 5 min a 72°C.

#### 4. VISUALIZACIÓN DE LOS PRODUCTOS DE AMPLIFICACIÓN POR ELECTROFORESIS EN GEL DE AGAROSA.

Para comprobar que se llevó a cabo una exitosa amplificación del fragmento deseado se realizó una electroforesis en un gel de Agarosa al 1%, utilizando como marcador de peso molecular a  $\phi$ x 174 RF DNA/*Hae* III Fragments (Invitrogen, Carlsbad, CA) y seleccionando como positivas aquellas bandas que se encontraran aproximadamente en las 1500 pb (Fig. 12).

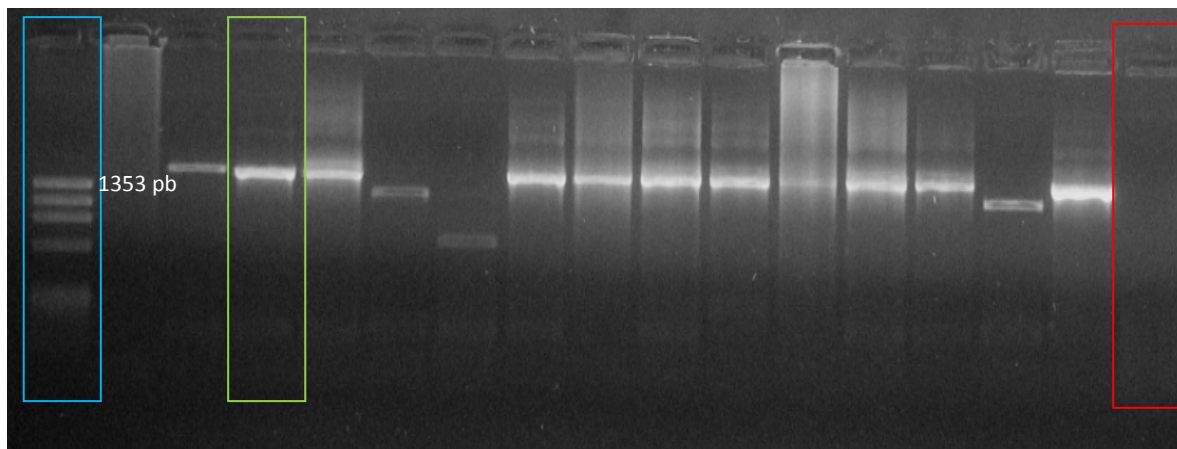


Figura 12. Ejemplo de un gel de agarosa al 1%. En un rectángulo azul se muestran los marcadores de peso molecular. En un rectángulo verde se observa una de las muestras que corresponden al producto de amplificación del gen *gag* (aproximadamente 1500 pb). En un rectángulo rojo se muestra el control negativo.

#### 5. PURIFICACIÓN DE LOS PRODUCTOS DE AMPLIFICACIÓN.

Los productos de amplificación con un tamaño cercano a 1500 pb fueron purificados por medio del estuche de reactivos QIAquick® PCR Purification Kit (QIAGEN, Valencia, CA), el cual permite que el DNA se una a una columna de sílice por un fenómeno de adsorción en presencia de una solución de saturada de sales. Posteriormente se lavó la columna para eliminar contaminantes y se eluyó el DNA en presencia de una solución de baja fuerza iónica y pH alcalino (solución de Tris).

## 6. CUANTIFICACIÓN DE LOS PRODUCTOS DE AMPLIFICACIÓN.

Se determinó la concentración del producto de PCR purificado por medio de un espectrofotómetro NanoDrop® ND-1000 (Thermo Scientific, Wilmington, DE) para poder ajustar la concentración de DNA en la reacción de secuenciación dentro del rango óptimo (10 – 40 ng/μL). La pureza del DNA purificado fue evaluada por la relación entre ácidos nucleicos/proteínas (absorbancia 260/280), buscando valores mayores a 1.8.

## 7. SECUENCIACIÓN

Posteriormente se realizó la placa de secuenciación utilizando el reactivo Big Dye Terminator v.3.1 cycle-sequencing (Life Technologies, Carlsbad, CA); 4 primers “reverse”: GS1 (5'-TTATCTAAAGCTTCCTTGGTGTCT-3'), GAS4 (5'-GGTTCTCTCATCTGGCCTGG-3'), GAS6 (5'AAAATAGTCTTACAATCTGG-3'), GA2274R (5'-TCTTTATTGTGACGAGGGGTCG-3'); y 4 primers “forward”: G2 (5'-GCGGCGACTGGTGAGTA-3'), GAS3 (5'-CATCAATGAGGAAGCTGCAG-3'), GAS5 (5'-CTCTAAGAGCCGAGCAAGCT-3'), HPR1977 (5'-GTAAAGTGTTCATTGTGG-3').

Componente	Volumen por pozo y por primer
dH2O	3.0 μL
Sequencing Buffer 5x	2.0 μL
Big Dye v3.1	1.0 μL
Primers (2 pmol/ μL)	2.0 μL
DNA viral (10-40 ng totales)	2.0 μL
Volumen Final	10 μL

Se utilizó el siguiente programa en el termociclador: 40 ciclos de 10 s a 96°C, 5 s a 50°C y 1 min a 60°C.

A continuación la placa fue purificada agregando 1μL de acetato de sodio y 50 μL de etanol al 100% por pozo, se mezcló con un vórtex y se procedió a realizar una centrifugación a 2000g por 20 minutos. Posteriormente, se eliminó la mayor cantidad de



etanol mediante la inversión en papel y se centrifugó a 150 g por 1 minuto. Se continuó con otra serie de limpieza, agregando 150  $\mu$ L de etanol al 70%, mezclando con vórtex y centrifugando a 2000 g por 5 minutos. A continuación se llevo a cabo la eliminación del etanol por una inversión en papel y se centrifugó a 150 g por un minuto; este último paso fue repetido una vez más para asegurar la eliminación total del etanol. Para finalizar, se resuspendió el precipitado en 15  $\mu$ L de formamida, se mezcló con el uso del vórtex y se centrifugo a 150 g por 1 minuto para bajar toda la suspensión que hubiera quedado adherida a las paredes superiores del pozo.

La secuenciación se realizó en un secuenciador capilar ABI Prism 3730xl DNA Genetic Analyzer (Life Technologies, Carlsbad, CA), mediante el método de Sanger. El llamado de bases y la preparación de las secuencias para análisis fueron realizados con el software Sequencing Analysis v5.4 (Life Technologies, Carlsbad, CA). Posteriormente, las 8 secuencias obtenidas para cada paciente fueron ensambladas con el software SeqScape v2.7 (Life Technologies, Carlsbad, CA), revisándose manualmente los alineamientos y obteniéndose la secuencia consenso del gen *gag* del virus para cada paciente.

## 8. ANÁLISIS FILOGENÉTICO.

A través del software MEGA 5.05 se realizó el alineamiento de las secuencias consenso obtenidas para el gen *gag* de los virus de cada uno de los pacientes participantes en el estudio. El alineamiento fue editado manualmente, y posteriormente se construyó un árbol filogenético usando como referencias la secuencia consenso HXB2 (usada como punto de comparación para la numeración de los genes del VIH) y un virus de referencia de subtipo B (Base de datos de VIH de los Alamos). Se utilizó el método Neighbor-Joining, calculando las distancias evolutivas con el modelo de dos parámetros de Kimura. La significancia del árbol fue calculada mediante 1000 réplicas de bootstrap.

## 9. ANÁLISIS DE VARIABILIDAD

La variabilidad de las secuencias del gen *gag* de virus mexicanos se calculó estimando la entropía de Shannon para cada una de las posiciones del alinamiento, utilizando la

herramienta de la base de Datos de VIH de los Alamos llamada Entropy (<http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy/entropy.html>) en su versión Entropy One como estrategia para cuantificar la variabilidad de las secuencias en el alineamiento completo y Entropy Two para realizar comparaciones de entropía nucleótido a nucleótido entre dos grupos de secuencias con una estimación de apoyo estadístico de 1000 aleatorizaciones con reemplazo, tomando en cuenta la frecuencia (e.g. una posición con 50% A y 50% T tiene una mayor entropía de Shannon que una posición con 90% A y 10% T) y el número de posibilidades (e.g. una posición con 90% A, 5% T y 5% G tiene una mayor entropía de Shannon que una posición con 90% A y 10% T). La entropía de Shannon es una medida simple de incertidumbre en un conjunto de datos. Aplicado a secuencias, la entropía de Shannon podría considerarse como la habilidad de predecir qué aminoácido se encontraría en cada posición de la secuencia, basándose en un muestreo previo.

## 10. ANÁLISIS EVOLUTIVO.

Utilizando el servidor DataMonkey ([www.datamonkey.org](http://www.datamonkey.org)), el alineamiento obtenido fue analizado por el método SLAC (Single-Likelihood Ancestor Counting), basado en la construcción de secuencias ancestrales sitio por sitio hasta obtener la de mayor congruencia, estimando valores de variación a cada posición. De esta manera se obtuvieron posibles codones bajo presión selectiva por razones de sustituciones no sinónimas y sustituciones sinónimas.

## 11. ANALISIS DE PRESIÓN SELECTIVA POR HLA E IMPACTO DE MUTACIONES SELECCIONADAS POR HLA EN LA CAPACIDAD REPLICATIVA VIRAL.

Se analizó la posibilidad de que las posiciones encontradas como bajo presión selectiva positiva utilizando los resultados del servidor DataMonkey estuvieran relacionadas con selección por HLA. Para esto se evaluó si los sitios bajo presión selectiva en los virus mexicanos se encontraban dentro de epítomos de CTLs previamente reportados y comprobados, mediante el uso de mapas de epítomos de la Base de Datos de Inmunología de Los Álamos (<http://www.hiv.lanl.gov/content/index>). Además, utilizando los datos de

tipificación HLA de clase I para los genes A, B y C disponibles en la Base de Datos de Epidemiología Molecular y Evolución del VIH en México del CIENI, INER (Valenzuela-Ponce, et al, datos no publicados), se evaluó la posibilidad de que alelos HLA frecuentes en la cohorte mexicana, pudieran explicar la presencia de selección positiva en las secuencias de *gag* analizadas.

Haciendo uso de la literatura se identificaron mutaciones reportadas previamente como mutaciones de escape seleccionadas por HLA, con altos costos en la capacidad replicativa [7,23]. Las posiciones correspondientes a estas mutaciones de escape fueron identificadas en el alineamiento de las 200 secuencias de *gag* de la cohorte mexicana y la variabilidad en cada una fue evaluada. Para estimar el posible impacto de diferentes variantes en estas posiciones en la capacidad replicativa, se identificaron los virus en la cohorte para los cuales existían valores de capacidad replicativa relativa estimados previamente en ensayos fenotípicos de replicación viral en los laboratorios del CIENI, INER (Soto-Nava, et al, datos no publicados) y que forman parte de la Base de Datos de Epidemiología y Evolución del VIH en México del CIENI, INER. En total se utilizaron las secuencias de 58 virus para los cuales existían valores de capacidad replicativa, correlacionando la presencia de las mutaciones previamente descritas en la literatura con posibles valores altos o bajos de capacidad replicativa de los virus mexicanos.

Además, se buscó la posibilidad de que variantes observadas en las posiciones identificadas como bajo presión selectiva positiva por el servidor Data Monkey pudieran asociarse a valores altos o bajos de capacidad replicativa viral. Para ello se utilizaron las secuencias de los 58 virus para los cuales había valores de capacidad replicativa disponibles dentro de la Base de Datos de Epidemiología Molecular y Evolución del VIH en México del CIENI, INER. La significancia de dichas asociaciones se evaluó mediante pruebas de Fisher con el programa GraphPad Prism 5.04.

Con el fin de hacer un análisis más específico del impacto de la presión selectiva por HLA sobre la variabilidad de la secuencia del gen *gag* del VIH en el contexto inmunogenético único de la población mexicana, se utilizaron datos de asociaciones entre la expresión de

alelos HLA específicos y la presencia de polimorfismos en el virus, previamente estimados para la cohorte mexicana usando un modelo evolutivo complejo desarrollado por Carlson y colaboradores [10]. Este modelo llamado Redes de Dependencia Filogenética (PDN) estima la existencia de asociaciones HLA-polimorfismo de VIH, corrigiendo para los tres principales confusores para evaluar la evolución viral mediada por HLA: la covariación de codones de VIH, el desequilibrio de ligamiento de los genes HLA y los efectos confusores de la filogenia del VIH. Para determinar estas asociaciones, se hace uso de un modelo gráfico que representa las dependencias probabilísticas de atributos blanco (presencia o ausencia de aminoácidos específicos en cada uno de los codones de una proteína del VIH), y una serie de atributos predictores (presencia o ausencia de aminoácidos específicos en todos los codones de la proteína en estudio adicionales al codón blanco y/o la presencia de todos los posibles alelos HLA), tomando en cuenta a su vez la influencia por la filogenia de las secuencias. Dentro de este modelo se compara la verosimilitud de un modelo nulo que supone que la secuencia blanco no está bajo la presión del atributo predictor, contra la de un modelo alternativo que supone que la variable blanco si se encuentra bajo esta presión. Al final se definen cuatro distribuciones posibles: atracción, tener predictor hace más probable tener el blanco; repulsión, no tener el predictor hace menos probable tener el blanco; escape, tener el predictor hace menos probable tener el blanco; reversión, no tener el predictor hace más probable tener el blanco. Las correlaciones positivas están determinadas por el par atracción/repulsión y las correlaciones negativas por el par escape/reversión [10].

Las asociaciones HLA-polimorfismo de VIH identificadas por el modelo PDN aplicado previamente a una cohorte de más de 700 pacientes (Juárez et al, datos no publicados) se analizaron en el contexto de las 200 secuencias del gen *gag* del presente estudio, con el fin de evaluar de manera más específica el papel de la selección por HLA en la evolución del virus a nivel poblacional en México. Más aún, el impacto de la presión selectiva por HLA en los virus mexicanos fue evaluado buscando correlaciones entre la presencia de diversas variantes en las posiciones en *gag* asociadas a HLA y la observación de valores altos o bajos de capacidad replicativa.

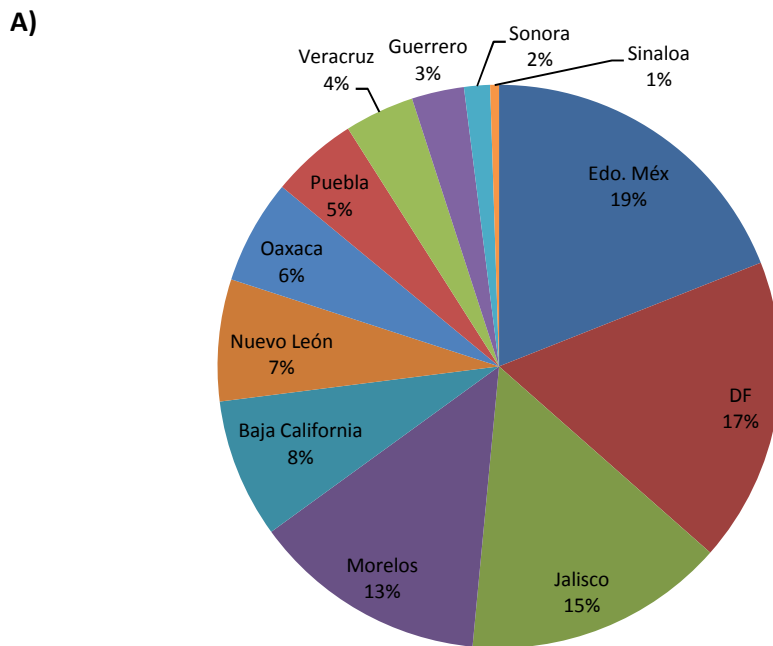
## CAPÍTULO V. RESULTADOS.

### 1. Descripción de la cohorte.

Se obtuvieron 200 muestras de pacientes infectados con VIH-I subtipo B, vírgenes al tratamiento con antirretrovirales y en la etapa crónica de la infección. Las muestras provinieron de 12 estados diferentes de la República Mexicana: Baja California, Distrito Federal, Estado de México, Guerrero, Jalisco, Morelos, Nuevo León, Oaxaca, Puebla, Sinaloa, Sonora y Veracruz (**Fig. 13 A**). Se calculó también la proporción de muestras obtenidas de cada estado, determinándose que el Estado de México, Distrito Federal y Jalisco son los que proporcionaron un mayor número de pacientes con cifras de 19%, 17% y 15%, respectivamente (**Fig. 13 B**). En cuanto a las proporciones por género, el 75.5% de pacientes fueron varones y el 24.5% fueron mujeres. Las edades variaron entre los 17 y los 61 años de edad con una media de 29 años. Por otro lado, en cuanto a los valores de carga viral el intervalo se encontró entre 104 copias/ $\mu$ L y  $>1,000,000$  copias/ $\mu$ L, siendo la media 103,988 copias/ $\mu$ L. Finalmente, los valores de linfocitos T CD4+ variaron de entre 12 células/ $\mu$ L hasta 1,310 células/ $\mu$ L con una media de 403 células/ $\mu$ L (**Tabla 2**).

Tabla 2. Características demográficas y clínicas de la cohorte.

<b>N</b>	<b>200</b>
<b>Edad media (años,[Intervalo])</b>	29 [17:61]
<b>Proporción de hombres (n, [%])</b>	151 [75.5%]
<b>Carga Viral Media (copias/<math>\mu</math>L, [Intervalo])</b>	103,988 [104: $>1,000,000$ ]
<b>Conteo de T CD4+ medio (células/<math>\mu</math>L, [Intervalo])</b>	403 [12:1,310]



B)

Figura 13. A) Distribución geográfica de la cohorte estudiada. 200 pacientes mexicanos infectados con VIH-I subtipo B, vírgenes a tratamiento antirretroviral, procedentes de 12 diferentes estados (en rojo), con una línea negra se marca la división territorial utilizada para posteriores menciones, señalando a los estados del norte y del sur. B) La proporción de pacientes por cada estado se muestra en el gráfico circular.

## 2. Análisis filogenético.

Con los datos obtenidos del alineamiento de las secuencias del gen *gag* de las 200 muestras presentes en el estudio se realizó un árbol filogenético para evaluar las relaciones de distancia genética y los patrones de agrupamiento entre las secuencias. A su vez, con fines comparativos se realizó un árbol con secuencias del gen *pol* obtenidas para los virus de los mismos pacientes, a partir de la Base de Datos de Epidemiología Molecular y Evolución del VIH en México del CIENI, INER (García-Morales, et al, datos no publicados) (**Fig. 14**). Al comparar el árbol obtenido para *gag* y el obtenido para *pol*, se puede observar que las ramas del primero están más alejadas del centro del árbol (teórico primer ancestro común) y las del segundo se encuentran más cercanas al centro del árbol; lo cual sugiere una mayor variabilidad y diversificación del gen *gag* y una mayor conservación del gen *pol* en los virus mexicanos. Para el caso de las secuencias de *gag* fueron evidentes algunos agrupamientos característicos, los cuales no se observaron para el gen *pol*. Estos agrupamientos tienen en común que pertenecen preferentemente a zonas del norte o del sur, respectivamente. Específicamente en el ejemplo mostrado en la figura 14, el agrupamiento marcado con el círculo rojo incluye secuencias de pacientes provenientes de estados del sur del país y que además presentan de los conteos de linfocitos T CD4+ más bajos.

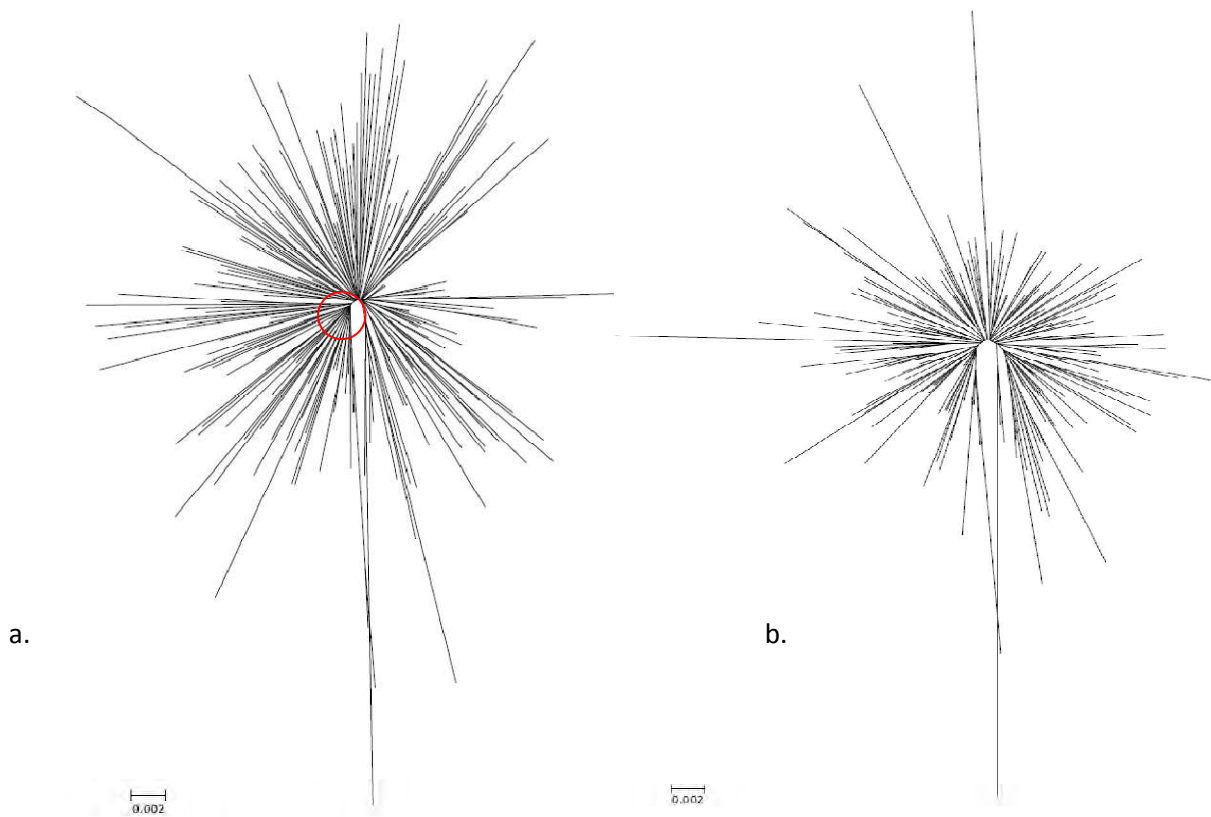


Figura 14. Árboles filogenéticos para los genes *gag* (a) y *pol* (b) de los virus mexicanos. Se construyeron árboles filogenéticos por el método de Neighbor Joining con el modelo de dos parámetros de Kimura a partir de alineamientos de 200 secuencias del gen *gag* y *pol* del VIH de los pacientes participantes en el estudio. La confiabilidad de los árboles se evaluó mediante 1000 repeticiones de bootstrap. Se observa que las ramas del árbol del gen *gag* son en general más largas, demostrando una mayor diversificación entre las muestras analizadas. Aunado a lo anterior, se encuentra marcado en círculo rojo un ejemplo de cómo las muestras se organizan en clusters distintos, los cuales no fueron evidentes en el árbol del gen *pol*.

Posteriormente, realizando un análisis específico en el árbol obtenido para *gag*, se diferenciaron las muestras provenientes del norte del país de aquellas provenientes del sur (**Fig. 15**). Interesantemente, se observó que las muestras tienden a agruparse en clusters que incluyen específicamente muestras del norte o del sur, respectivamente. Esto concuerda con lo observado en las comparaciones de los árboles correspondientes a *gag* y *pol*, indicando que únicamente en las secuencias correspondientes a *gag*, los virus que presentan similitud tenderán a agruparse.



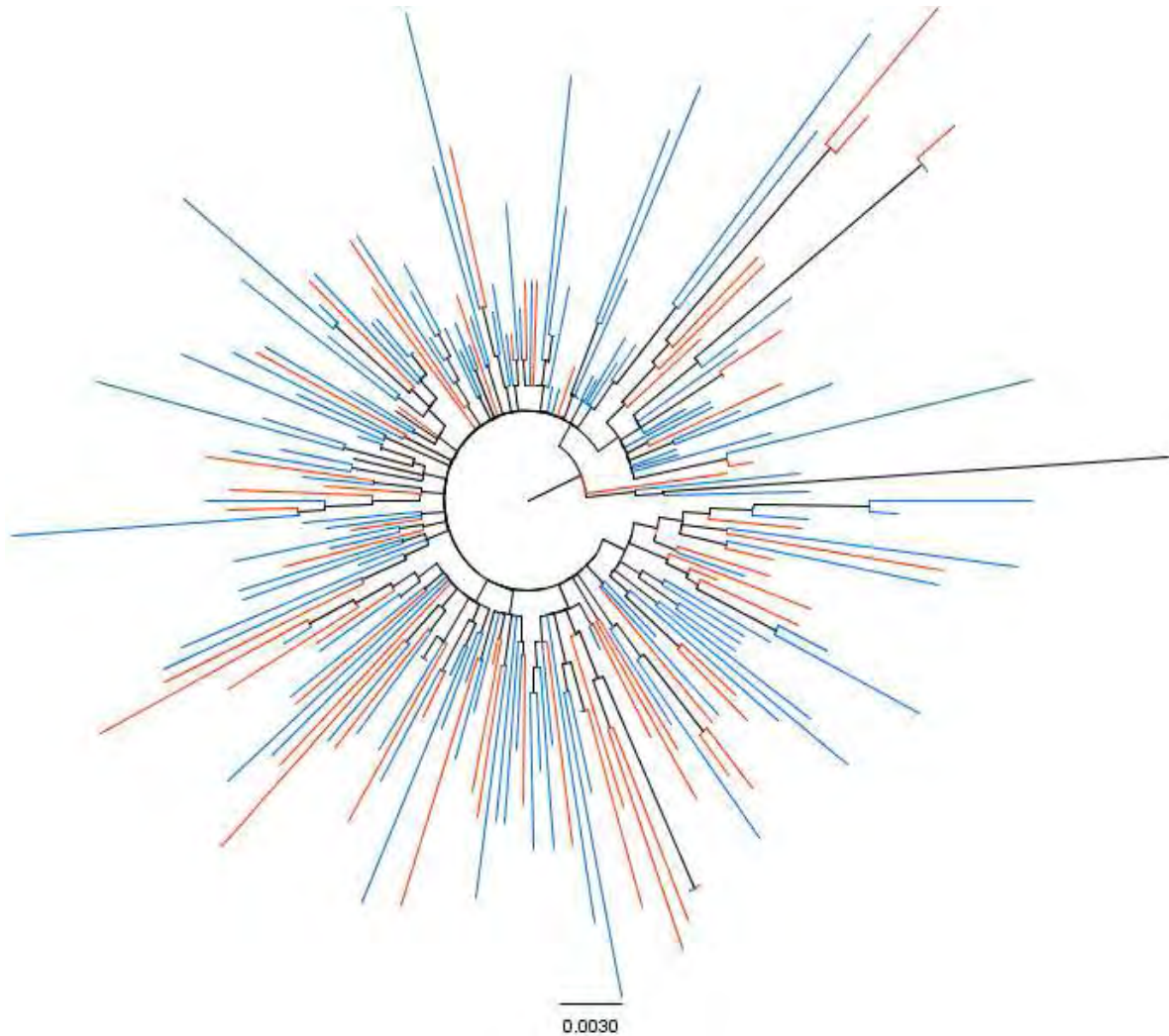


Figura 15. Árbol filogenético basado en el gen *gag* de los 200 virus mexicanos. El árbol se construyó por el método de Neighbor Joining con el modelo de dos parámetros de Kimura a partir de alineamientos de 200 secuencias del gen *gag* del VIH de los pacientes participantes en el estudio. La confiabilidad de los árboles se evaluó mediante 1000 repeticiones de bootstrap. Marcados en azul se observan virus de pacientes que habitan en el sur de país y los marcados en anaranjado corresponden a virus de pacientes del norte del país.

### 3. Análisis de variabilidad de las secuencias de VIH de pacientes mexicanos.

Para explorar la variabilidad del gen *gag* de los virus mexicanos se revisaron los polimorfismos presentes en las 200 secuencias obtenidas y se realizó un gráfico que relaciona cada codón con la frecuencia de los polimorfismos que presenta con respecto a la secuencia de referencia HXB2 (**Fig. 16**). Según el alineamiento, la zona correspondiente a p17 abarca del codón 1 al codón 132 de *gag*, la zona de p24 del codón 133 al codón 364 y la zona de p2p7p1p6 del codón 365 al codón 500.

Se observó que las zonas correspondientes a p17 y a p2p7p1p6 son las que presentaron mayor variabilidad, teniendo una mayor cantidad de codones con polimorfismos (75% y 87%, respectivamente). Por el contrario la zona central, correspondiente a p24, se mostró como la más conservada con una frecuencia de posiciones polimórficas de 44%. Además, el 80.5% de los codones que presentaron valores >50% de presencia de polimorfismos se encontraron en las zonas de p17 y p2p7p1p6, situándose así las posiciones más polimórficas en estas dos regiones del gen.

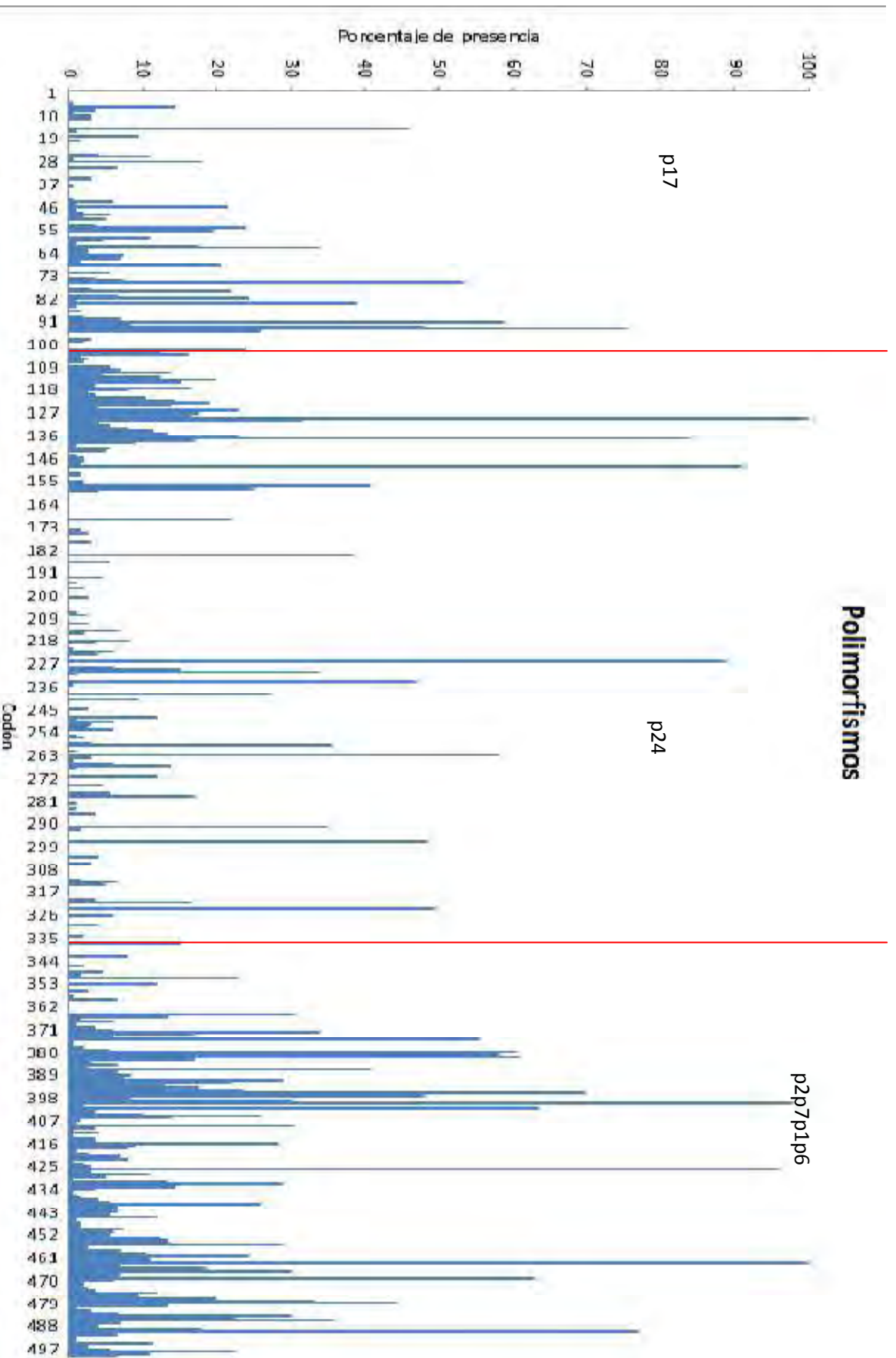


Figura 16. Frecuencia de polimorfismos en los distintos codones de *gag*, en VIH circulante en México. Se muestra la frecuencia de polimorfismo por codón para 200 virus mexicanos, con respecto a la secuencia de referencia HXB2. El gráfico señala el inicio y final de los codones pertenecientes a las distintas proteínas que se encuentran codificadas en este gen (p17, p24 y p2p7/p1p6). Se observa que la parte final del gen es la que presenta mayor variabilidad, mientras que la zona correspondiente a p24 es la más conservada.

Para estudiar con más detalle la variabilidad del gen *gag* en los virus mexicanos, se calculó el parámetro de entropía de Shannon para cada posición de aminoácidos en el alineamiento de 200 secuencias. La entropía de Shannon es una medida de aleatoriedad asociada a cada posición de la secuencia [31]. Entre más polimórfica sea una posición, la entropía asociada será mayor. En el gráfico obtenido para el alineamiento de las 200 secuencias trabajadas, se observó que la zona correspondiente a p24 fue la que presentó la menor cantidad de zonas con alta entropía y por el contrario las zonas de p17 y p2p7p1p6 presentaron más zonas con alta entropía. De esta manera se corroboran las observaciones obtenidas por el análisis de polimorfismos, teniéndose a p24 como la zona más conservada.

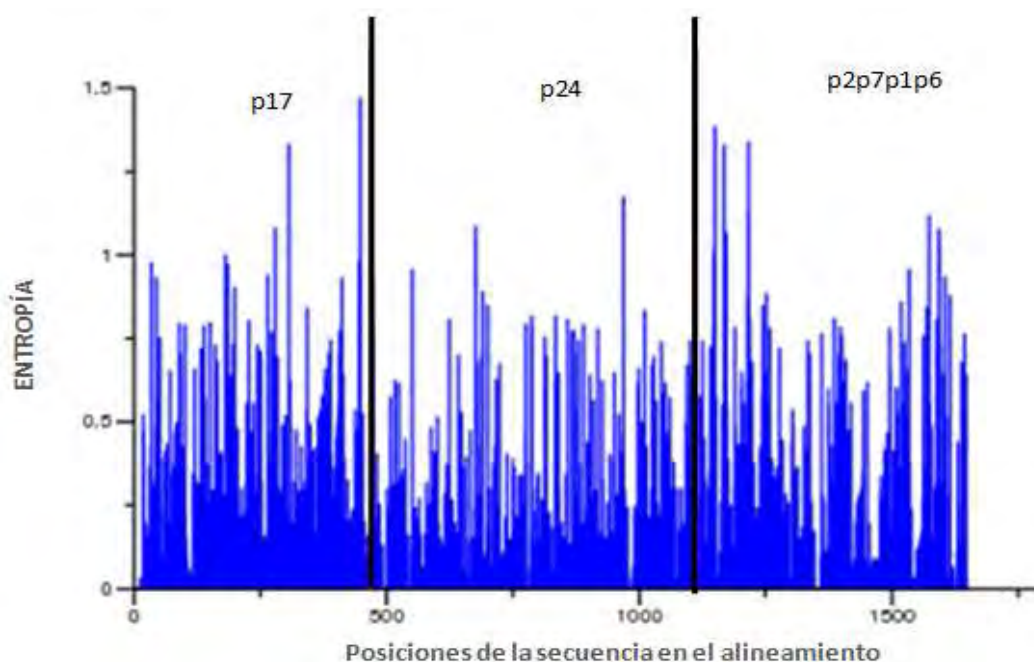


Figura 17. Análisis de entropía de Shannon en el gen *gag* de VIH circulante en México. Se calculó la entropía de Shannon para un alineamiento de 200 secuencias del gen *gag* de virus circulantes en México mediante las herramientas de la Base de Datos de VIH de Los Alamos. Las proteínas Gag se encuentran divididas con líneas rojas las zonas correspondientes a p17, p24 y p2p7p1p6, de izquierda a derecha.

Con el objeto de explorar la variabilidad temporal de las secuencias, se realizó una comparación de la entropía de Shannon de cada posición del gen *gag* de virus de pacientes mexicanos con infección temprana y tardía, clasificando a las secuencias en dos grupos. El primero corresponde a todas las muestras que tuvieron un conteo de T CD4+

mayor a 500 células/ $\mu\text{L}$  (muestras en fase temprana de la enfermedad) y el segundo corresponde a las muestras que tuvieron un conteo de T CD4+ menos a 200 células/ $\mu\text{L}$  (muestras en la fase crónica de la enfermedad). Se esperaba que las secuencias de virus de pacientes con infección tardía presentaran más entropía debido a la diversificación de los virus a lo largo de la infección.

Como era esperado, se observó una mayor cantidad de sitios con mayor entropía en las secuencias de pacientes con infección tardía que en secuencias de pacientes con infección temprana (34 vs 19). Sin embargo, también se observaron posiciones con mayor entropía en el grupo de secuencias de virus de pacientes en etapas tempranas. Estos sitios podrían reflejar la reversión temprana de mutaciones de escape seleccionadas en hospederos previos o, alternativamente, la aparición de mutaciones adaptativas tempranas.

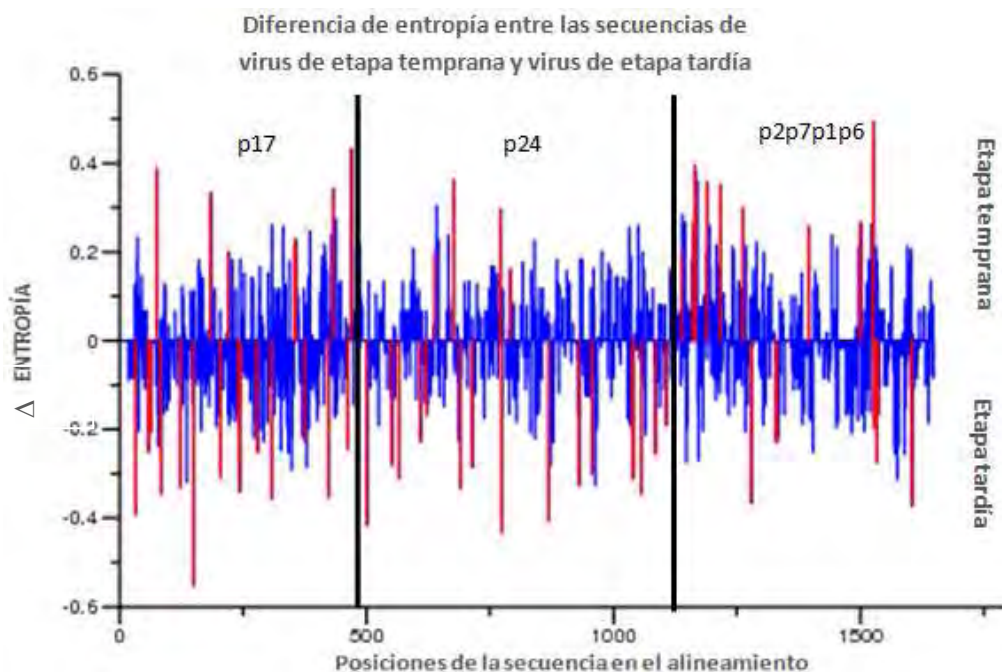


Figura 18. Comparación de entropía de Shannon para virus de pacientes en etapas tempranas y tardías de la infección en México. Utilizando las herramientas de la Base de Datos de VIH de Los Álamos, se comparó la entropía de las secuencias de virus de pacientes con un conteo de T CD4+ >500 células/ $\mu\text{L}$  (infección temprana) y de las secuencias de virus de pacientes con un conteo <200 células/ $\mu\text{L}$ . (infección tardía) Las barras hacia arriba muestran más entropía en virus de etapas tempranas y las barras hacia abajo muestran más entropía en virus de etapas tardías. Las barras en rojo muestran diferencias estadísticamente significativas y las negras las zonas correspondientes a cada una de las proteínas codificadas por *gag*.

#### 4. Determinación de sitios bajo presión selectiva por razón de sustituciones sinónimas y no sinónimas.

El paso siguiente fue determinar si la variabilidad observada en *gag* podría estar relacionada con presión selectiva positiva. Para tal fin se utilizó el servidor DataMonkey [www.datamonkey.org], con el fin de estimar las razones de sustituciones sinónimas y no sinónimas (dN/dS) en las 200 secuencias de *gag* de los virus mexicanos. De esta manera, se pudieron determinar zonas bajo posible presión selectiva positiva (tendencia a cambio con respecto al consenso), donde la razón de sustituciones no sinónimas fue significativamente mayor que la de sustituciones sinónimas (dN>dS) y zonas bajo posible selección negativa (tendencia a conservación con respecto al consenso), donde la razón de sustituciones no sinónimas fue significativamente mayor a la de sustituciones sinónimas (dS>dN). En total se obtuvieron 301 sitios seleccionados negativamente y 40 sitios seleccionados positivamente, con un nivel de significancia de  $p < 0.05$  (Tabla 3 y 4).

Tabla 3. Sitios bajo presión selectiva negativa por razones de sustitución dN/dS en el gen *gag* de VIH circulante en México.

Codón Alineado	dN-dS Normalizada	Valor p	Codón Alineado	dN-dS Normalizada	Valor p	Codón Alineado	dN-dS Normalizada	Valor p
6	-1.95	1.42E-12	180	-0.51	0.00724507	333	-1.74	4.12E-11
9	-0.40	0.0071623	181	-0.33	0.00413001	334	-0.56	0.00166158
10	-0.82	0.000167432	183	-0.43	0.00169605	335	-2.54	5.11E-15
11	-0.76	0.000480344	185	-0.20	0.037037	337	-2.96	5.25E-12
13	-0.77	3.04E-05	186	-0.77	2.96E-05	338	-2.37	1.33E-15
14	-1.08	1.38E-07	188	-0.88	1.35E-06	339	-0.20	0.0371118
17	-3.63	1.11E-16	189	-0.33	0.0048114	340	-0.86	2.99E-06
20	-0.52	0.00310394	192	-0.39	0.0013754	341	-1.08	3.04E-07
21	-1.26	6.40E-08	194	-0.54	0.000349321	342	-2.59	3.21E-13
22	-2.08	4.22E-15	195	-0.67	0.00108917	343	-2.08	3.55E-11
24	-2.25	4.44E-16	196	-2.57	-4.44E-16	344	-1.72	8.87E-13
27	-0.41	0.00556187	197	-0.62	0.0018622	345	-1.59	1.15E-09
29	-1.18	1.21E-07	199	-0.90	1.24E-06	347	-0.33	0.00413009
31	-1.84	1.54E-09	200	-2.72	-6.66E-16	348	-2.38	3.96E-14
32	-1.14	3.39E-08	201	-0.33	0.0386201	349	-1.51	3.14E-11
33	-0.43	0.00169605	202	-0.46	0.000457247	350	-0.56	0.00131405
39	-0.98	1.08E-07	203	-0.26	0.0123508	352	-1.05	2.84E-08
40	-3.14	6.66E-16	205	-1.40	1.07E-09	354	-1.49	3.43E-10
41	-0.76	2.94E-05	206	-1.02	9.13E-06	355	-1.14	2.16E-08



42	-1.11	2.38E-07	208	-2.28	2.29E-14	356	-1.05	5.24E-06
43	-0.80	4.00E-05	210	-1.22	8.56E-08	360	-0.33	0.00413001
44	-4.38	2.00E-15	212	-0.63	0.000204936	361	-1.61	1.37E-10
45	-0.88	2.22E-06	213	-0.53	0.00031349	362	-0.56	0.000540245
46	-1.72	1.44E-05	214	-1.37	9.04E-06	363	-0.26	0.0141535
47	-3.70	2.22E-16	215	-0.29	0.0424224	364	-1.71	6.16E-13
48	-0.95	7.42E-06	216	-1.57	1.16E-11	365	-2.12	2.66E-15
50	-1.08	2.53E-07	217	-0.21	0.0411832	366	-1.05	4.19E-08
52	-0.96	3.32E-07	218	-0.62	0.00394483	367	-2.24	-1.55E-15
56	-0.79	2.11E-06	219	-1.92	1.29E-13	368	-0.81	0.0284077
57	-1.40	3.13E-09	220	-0.33	0.00411523	369	-1.50	2.01E-10
59	-0.97	0.0011126	224	-0.75	1.42E-05	370	-1.91	1.48E-13
60	-0.71	0.0056041	225	-0.90	6.70E-07	372	-1.02	5.30E-06
63	-1.82	2.60E-07	226	-0.86	0.0443582	373	-1.57	1.89E-07
64	-1.28	1.70E-07	227	-1.40	1.07E-09	374	-1.06	1.52E-06
65	-1.98	7.51E-08	228	-0.26	0.0124514	375	-2.03	1.44E-15
66	-1.36	6.76E-07	229	-1.40	0.0005047	376	-2.10	9.55E-15
68	-0.55	0.0016894	232	-0.26	0.0123543	377	-0.52	0.000152416
70	-0.46	0.000463562	233	-0.39	0.00137544	379	-1.82	1.10E-09
71	-0.66	2.57E-05	235	-0.56	0.000355637	393	-5.89	0.0309903
73	-0.79	5.12E-06	237	-0.79	2.13E-06	394	-12.43	0.000428272
74	-0.76	0.000444531	238	-1.46	4.75E-09	396	-0.97	2.69E-06
77	-1.25	1.34E-08	240	-0.82	2.18E-06	398	-0.90	0.000655861
78	-1.19	5.87E-06	241	-0.77	0.00454324	399	-0.75	0.00274356
79	-0.82	0.0439496	242	-0.26	0.0123457	401	-0.97	0.0113478
80	-0.86	3.11E-06	244	-0.26	0.0139945	405	-2.49	1.33E-07
83	-0.33	0.0196832	245	-0.46	0.0110162	408	-2.51	3.33E-16
86	-0.75	3.77E-05	246	-1.43	1.58E-09	409	-1.95	1.97E-12
87	-0.40	0.00967326	248	-0.33	0.0041685	410	-0.76	1.50E-05
88	-3.71	-1.78E-15	250	-0.61	0.00890938	411	-1.08	1.25E-07
90	-0.84	0.0098932	251	-0.69	4.72E-05	412	-0.32	0.0106472
96	-1.29	4.90E-09	252	-0.59	0.00762423	413	-1.44	3.19E-11
97	-0.20	0.037037	254	-0.52	0.000175492	414	-0.54	0.0277683
98	-1.77	2.47E-10	255	-0.64	0.0016044	416	-1.80	3.68E-13
99	-0.45	0.00621397	256	-1.29	1.96E-08	417	-1.94	4.43E-13
100	-1.39	1.48E-10	257	-0.88	9.65E-06	419	-1.48	2.07E-10
101	-0.53	0.000761726	262	-0.33	0.00412018	421	-1.64	5.56E-09
102	-3.60	2.02E-12	268	-0.26	0.0452376	422	-0.86	5.49E-06
103	-1.37	7.50E-05	269	-0.85	6.27E-07	424	-1.04	2.47E-08
105	-0.66	0.000322273	270	-0.39	0.00165505	425	-0.66	1.79E-05
106	-1.05	5.51E-06	272	-2.68	-1.11E-15	426	-3.20	1.33E-15
107	-2.00	3.31E-13	273	-0.86	8.80E-06	427	-1.01	0.000121562
108	-0.35	0.00950432	274	-0.44	0.00127807	429	-0.52	0.000154958
109	-0.72	0.00140476	278	-1.76	5.97E-09	430	-1.39	7.69E-09
114	-3.52	3.39E-13	280	-1.38	2.14E-10	432	-0.81	5.57E-05
116	-1.10	2.52E-05	281	-0.82	1.54E-05	436	-0.72	0.00101636
118	-0.52	0.039407	282	-0.83	2.15E-05	437	-0.33	0.00485108
123	-2.38	0.0388631	285	-0.78	1.88E-06	438	-1.94	3.89E-13

129	-0.88	0.0389369	286	-2.55	-1.78E-15	439	-0.40	0.0132508
137	-1.04	0.0357648	288	-0.64	0.000161523	440	-6.63	0.00810105
140	-0.73	9.35E-05	289	-5.24	-4.44E-16	441	-0.49	0.00168862
143	-0.53	0.0261241	290	-1.13	1.22E-08	442	-1.76	1.37E-08
144	-0.52	0.000155276	292	-0.70	0.000347177	443	-2.48	-8.88E-16
145	-1.61	1.66E-06	293	-0.93	2.34E-07	446	-2.62	1.60E-11
146	-1.85	8.72E-14	294	-1.55	2.58E-10	453	-0.26	0.012352
147	-1.70	4.02E-08	295	-1.07	1.18E-07	454	-3.73	-2.22E-16
148	-1.60	6.30E-10	296	-1.75	3.82E-08	456	-16.26	0.00290392
149	-2.60	1.94E-06	298	-0.87	8.91E-06	461	-0.75	0.0110015
150	-0.89	7.36E-06	299	-1.31	5.62E-10	462	-2.93	4.88E-15
151	-1.09	4.86E-08	301	-2.81	6.66E-16	464	-1.72	8.43E-13
152	-0.36	0.0340434	302	-0.88	1.35E-06	465	-2.53	2.22E-16
153	-4.57	0.00312288	303	-1.51	6.49E-09	466	-4.37	3.44E-15
154	-0.33	0.0041254	304	-0.65	7.37E-05	467	-4.35	8.88E-16
155	-0.32	0.00840554	305	-0.74	6.99E-06	468	-1.07	0.0010578
156	-1.83	6.03E-10	306	-3.35	-2.22E-16	469	-0.32	0.0134952
160	-0.85	6.69E-07	308	-0.33	0.00411523	471	-1.72	2.04E-12
161	-0.67	2.24E-05	309	-2.92	-4.44E-16	473	-1.36	1.73E-06
162	-0.33	0.00428246	310	-0.96	1.28E-07	481	-0.72	5.80E-06
163	-0.39	0.0054366	311	-0.33	0.00856732	484	-0.75	8.53E-05
164	-0.53	0.000425555	312	-0.37	0.0293641	499	-1.57	0.000169492
167	-0.72	5.82E-06	314	-0.53	0.0242349	501	-0.50	0.0292819
168	-0.35	0.00459328	315	-0.90	1.24E-06	504	-0.76	0.00977147
169	-0.52	0.000156817	316	-0.60	7.37E-05	506	-4.40	-2.00E-15
171	-0.80	6.08E-06	317	-2.42	-1.11E-15	511	-0.79	0.00152952
172	-3.19	1.11E-15	318	-0.73	1.16E-05	521	-1.13	0.0015063
173	-0.90	9.43E-06	319	-0.53	0.000952573	531	-2.59	1.06E-07
174	-0.85	1.21E-05	320	-1.05	1.53E-05	534	-2.16	1.93E-13
175	-3.71	-6.66E-16	322	-1.37	1.21E-08	546	-1.55	7.12E-07
176	-1.06	1.15E-06	324	-0.91	2.09E-07	548	-1.20	0.000113577
177	-0.46	0.000457969	325	-0.44	0.00121856	549	-2.65	1.10E-11
178	-1.23	7.01E-09	331	-0.59	5.08E-05			
179	-1.25	2.57E-07	332	-0.41	0.00257113			

Tabla 4. Sitios bajo presión selectiva positiva por razones de sustitución dN/dS en el gen *gag* de virus circulante en México.

Codón Alineado	dN-dS Normalizada	Valor p	Codón Alineado	dN-dS Normalizada	Valor p	Codón Alineado	dN-dS Normalizada	Valor p
7	0.76	0.000406802	234	1.15	0.00184324	470	1.12	9.04E-05
12	1.73	1.18E-06	259	1.47	0.000118167	472	0.47	0.0414225
54	1.49	5.37E-07	263	1.13	0.0248523	482	0.84	0.000141911
62	1.31	0.00100537	291	1.37	0.00194824	505	0.33	0.0375984
69	1.19	0.000393183	297	0.78	0.0270342	508	0.79	0.0049563
76	0.81	0.0228049	321	0.60	0.0184007	509	1.05	0.00132055



82	0.85	0.00172049	351	0.78	0.00500474	512	1.32	0.00115465
84	1.91	4.38E-06	384	2.37	1.74E-06	532	1.81	2.44E-05
94	0.84	0.00372483	390	1.78	0.000231882	537	0.26	0.0379875
95	0.78	0.0235837	391	1.11	0.0132477	539	1.61	5.21E-07
115	0.49	0.0380871	406	3.60	9.79E-15	544	0.63	0.00538148
127	0.89	0.000112033	420	1.64	6.87E-08	547	1.31	5.87E-07
184	2.18	4.69E-11	445	1.25	0.000179859			
230	0.95	0.00809466	459	1.65	7.79E-07			

Se observó que hay más posiciones que tienden a conservar el aminoácido presente en la secuencia consenso que aquéllas que tienden a cambiar. Esto es esperado, dada la importancia funcional de Gag, pues cambios discretos en estas proteínas pueden tener consecuencias graves en la viabilidad del virus al impedir el correcto ensamblaje de la matriz, cápside y nucleocápside virales.

A continuación, se procedió a determinar qué proteína tenía mayor impacto de presión selectiva en Gag. Por tanto, en el gráfico obtenido del servidor DataMonkey se marcaron los codones pertenecientes a las diferentes proteínas codificadas por *gag* (**Fig. 19**). Al analizar el gráfico se observó que el 52.8% del total de las posiciones bajo presión selectiva negativa se encontraron en p24, mientras que un 77.5% del total de las posiciones bajo presión selectiva positiva se encontraron en p17 y p2p7p1p6. Al tomar en cuenta el total de codones correspondientes a cada proteína se tiene que p24 presentó un 3.9% de las posiciones bajo presión selectiva positiva y un 68.2% de las posiciones bajo presión selectiva negativa, mientras que p17 y p2p7p1p6 presentaron un 9% y 13.9% de las posiciones bajo presión selectiva positiva y un 51.5% y 54% de las posiciones bajo presión selectiva negativa, respectivamente. Así, las proteínas se ordenaron de la siguiente manera por frecuencia de posiciones bajo presión selectiva positiva: p2p7p1p6>p17>p24 y por frecuencia de posiciones bajo presión selectiva negativa: p24>p2p7p1p6>p17. Debido a esta mayor presencia de selección positiva en la zona de p17 y p2p7p1p6, esta gráfica refuerza la idea proporcionada por el gráfico de polimorfismos de que existe una mayor tendencia a variaciones en la zona inicial y final de gag (p17 y p2p7p1p6, respectivamente); mientras que existe una mayor tendencia a la conservación en la zona correspondiente a p24.

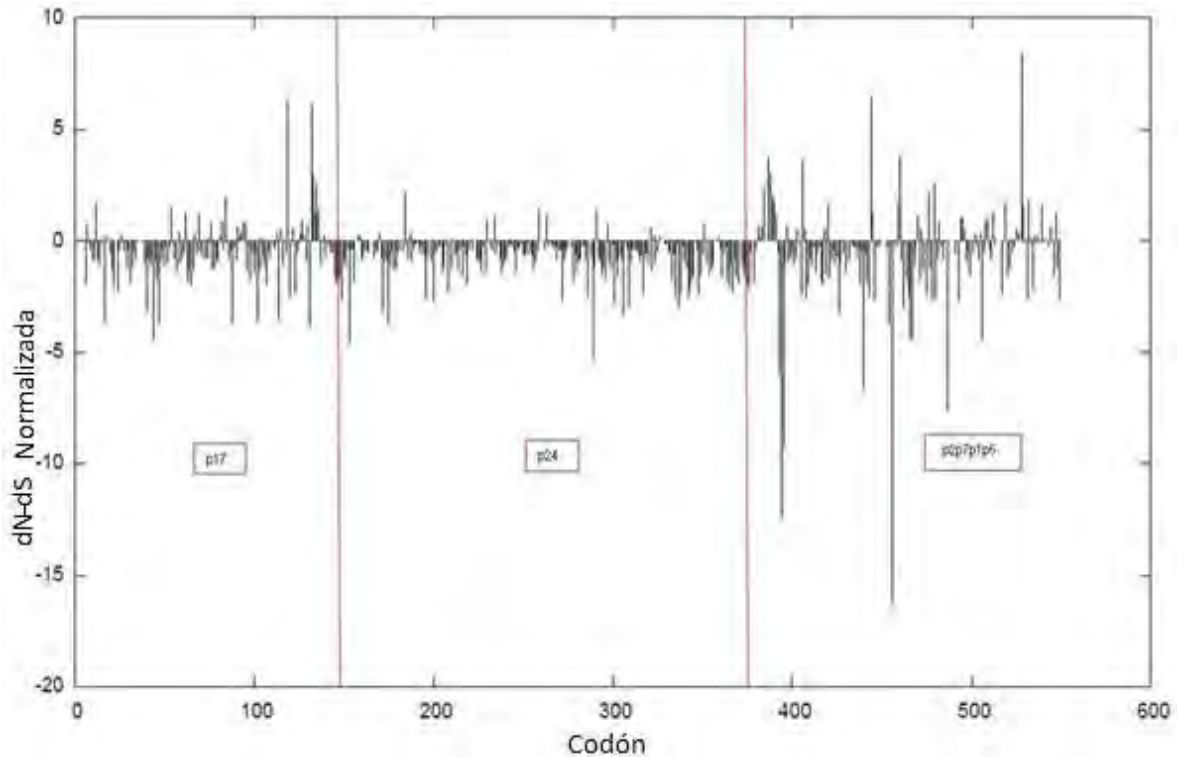


Figura 19. Distribución de posiciones bajo presión selectiva positiva y negativa en el alineamiento del gen *gag* para las 200 secuencias mexicanas. Los valores positivos indican que  $dN > dS$  por lo que las posiciones tienden a variar, definiéndose como bajo presión selectiva positiva. Los valores negativos indican que  $dS > dN$ , por lo que las posiciones tienden a conservarse, definiéndose como bajo presión selectiva negativa. Los valores  $dN-dS$  se obtuvieron mediante la herramienta de análisis evolutivo en línea DataMonkey.

## 5. Análisis de presión selectiva por HLA.

Las 40 posiciones con selección positiva encontradas por medio del servidor DataMonkey son posibles zonas que podrían asociarse con presión selectiva por HLA, pues su variabilidad podría coincidir con la presencia de mutaciones de escape seleccionadas por la respuesta inmune de CTLs. Para determinar si alguna de estas posiciones podía relacionarse con alelos HLA selectores y por tanto ser zonas bajo la influencia de presión inmune por CTLs, se hizo uso de los mapas de epítomos presentes en la base de datos de los Álamos [<http://www.hiv.lanl.gov/content/index>] (Fig. 20, 21, 22 y 23). Se muestran los distintos codones correspondientes a cada una de las proteínas codificadas en *gag*, los codones bajo selección positiva determinada por DataMonkey y los alelos que podrían estar seleccionando las mutaciones.



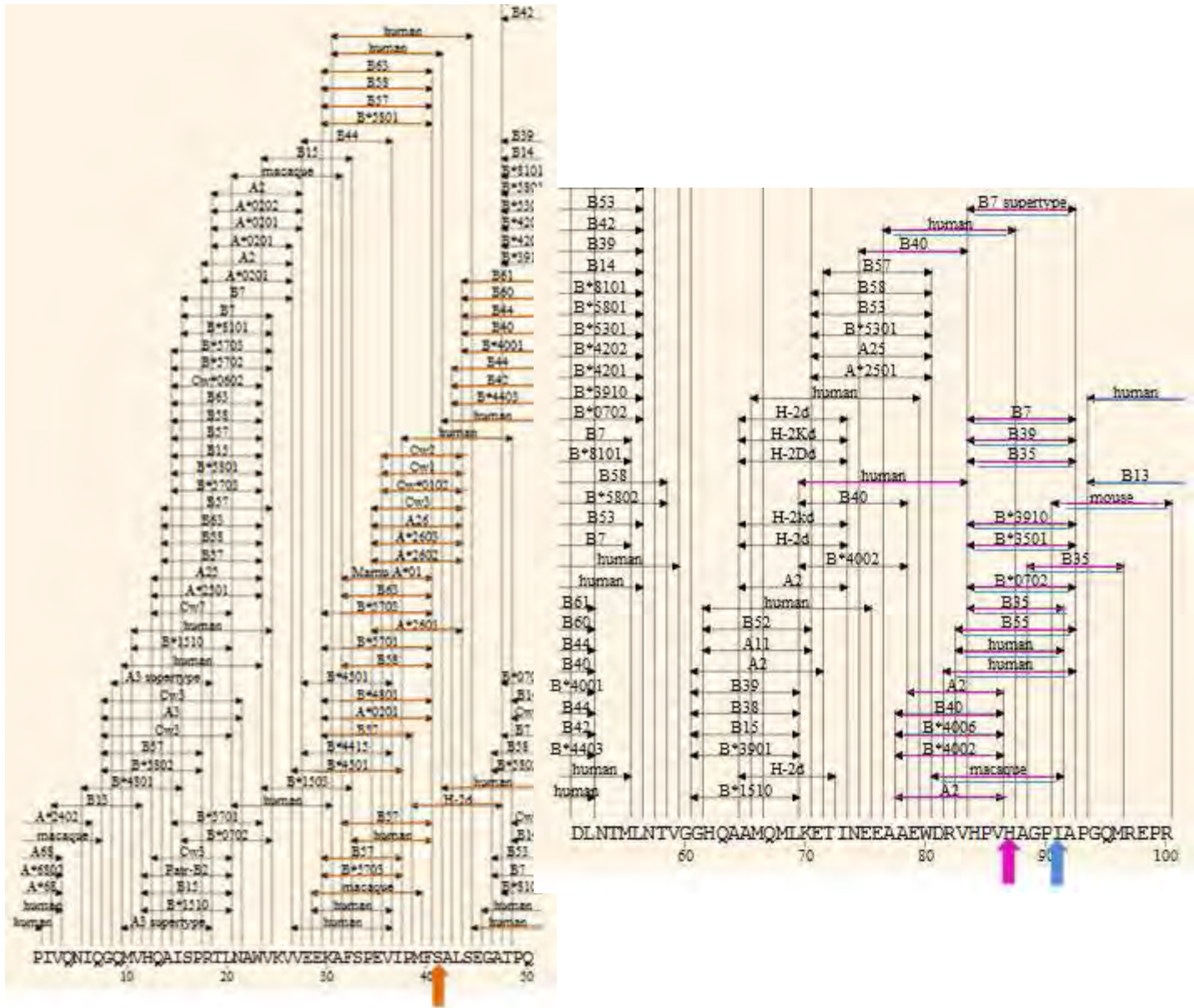


Figura 21. Mapas de epítomos para p24 y posiciones bajo presión selectiva positiva. Se muestran los distintos codones de *gag* correspondientes a la zona que codifica para los primeros 100 aminoácidos de p24. Con las flechas de colores se muestran las diferentes posiciones identificadas como bajo presión selectiva positiva por la herramienta en línea DataMonkey y en líneas del color correspondiente los alelos que podrían estar relacionados con la selección de esa mutación de acuerdo a la posición de epítomos restringidos por los mismos. Los mapas de epítomos muestran exclusivamente epítomos confirmados experimentalmente, incluidos en la Base de Datos de Inmunología de VIH de Los Álamos [31].



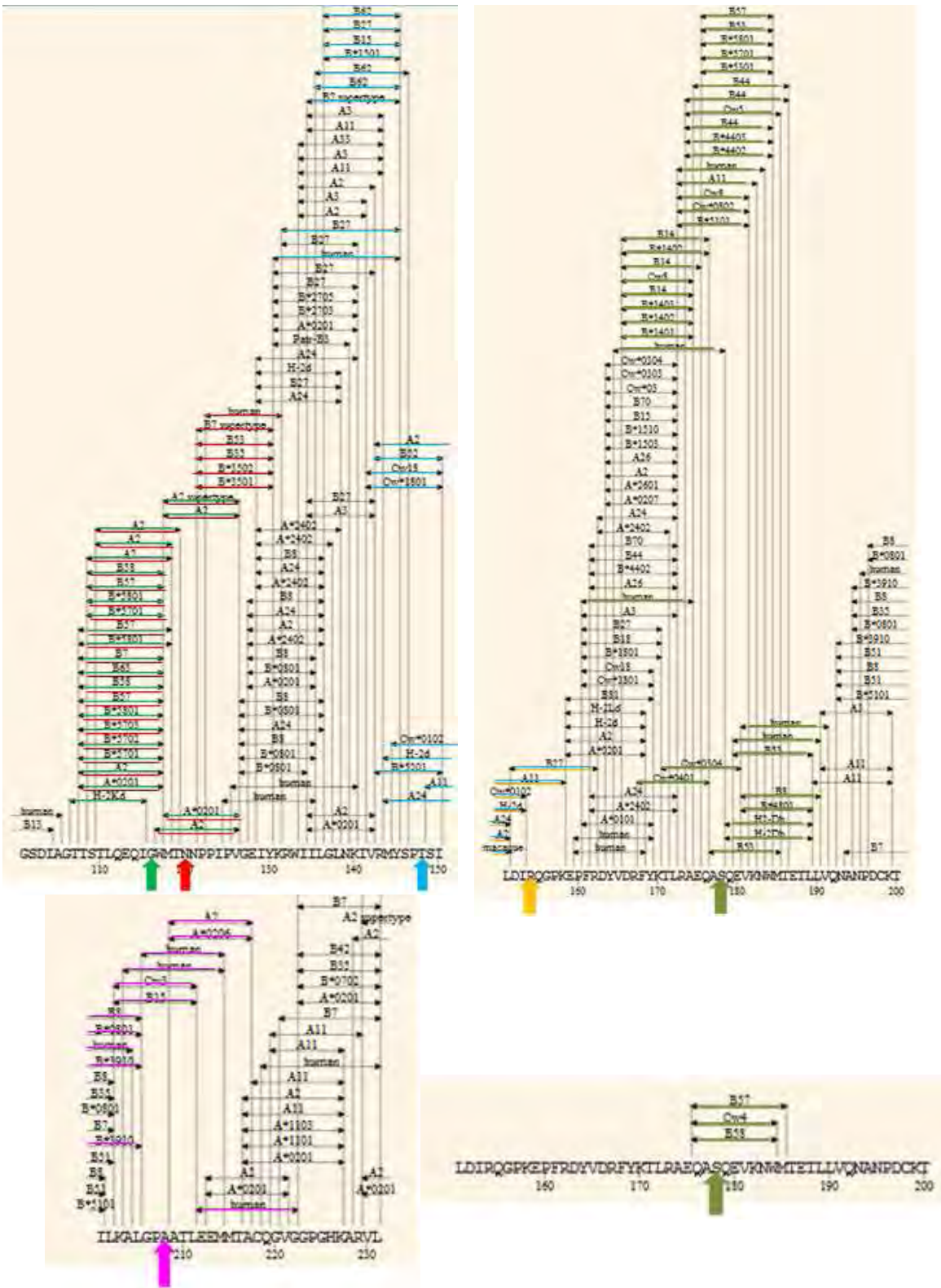
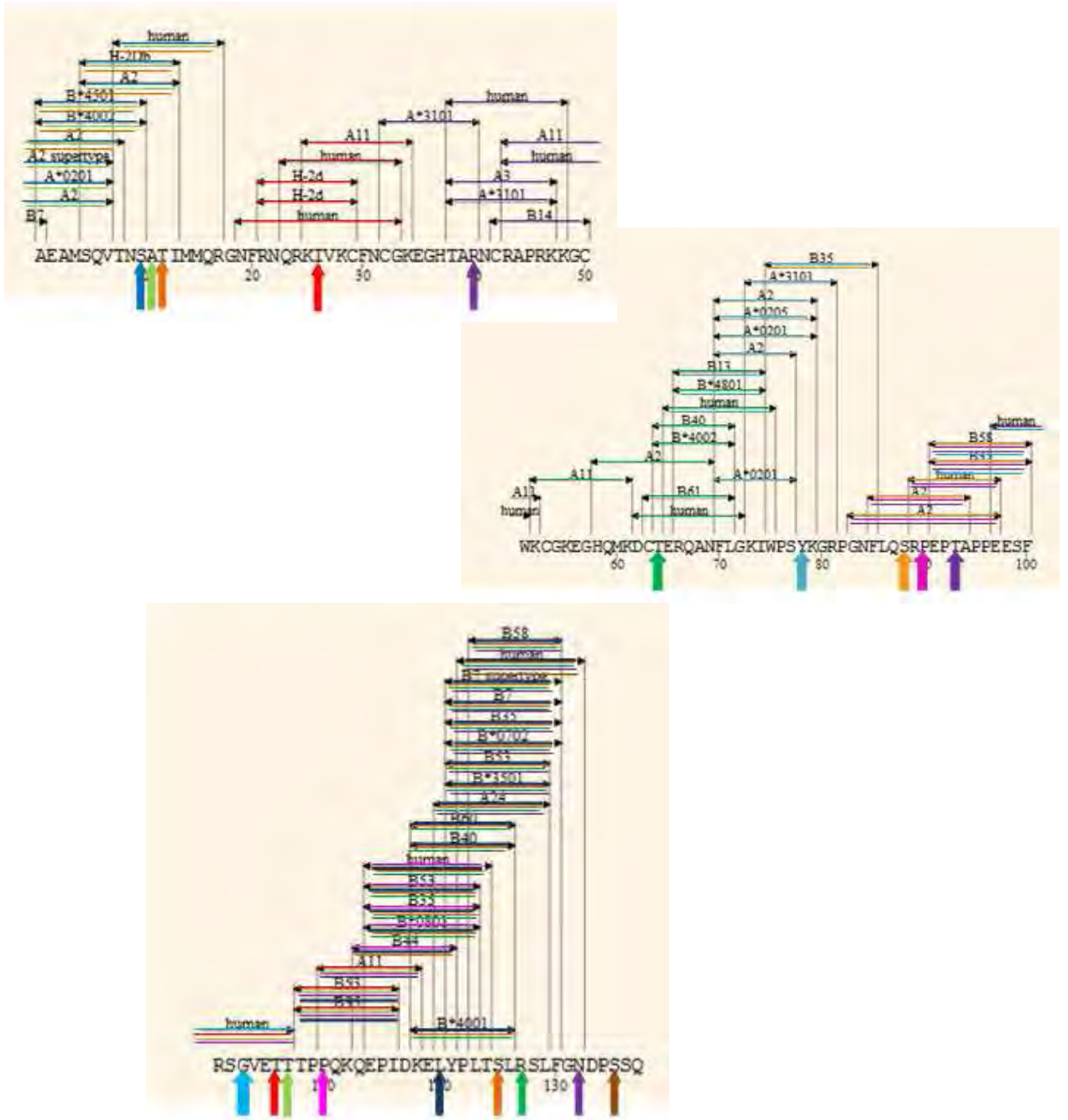


Figura 22. Mapas de epítomos para p24 y posiciones bajo presión selectiva positiva. Se muestran los distintos codones de *gag* correspondientes a la zona que codifica para los aminoácidos 101 al 231 de p24. Con las flechas de colores se muestran las diferentes posiciones identificadas como bajo presión selectiva positiva por la herramienta en línea DataMonkey y en líneas del color correspondiente los alelos que podrían estar relacionados con la selección de esa mutación de acuerdo a la posición de epítomos restringidos por los mismos. Los mapas de epítomos muestran exclusivamente epítomos confirmados experimentalmente, incluidos en la Base de Datos de Inmunología de VIH de Los Álamos [31].



En conjunto, estos análisis muestran que prácticamente todas las posiciones encontradas como bajo presión selectiva positiva por razón de dN/dS podrían estar directamente relacionadas con alelos HLA selectores. En particular se pueden mencionar algunos alelos HLA altamente frecuentes en la población mexicana los cuales podrían estar relacionados con selección de mutaciones de escape en algunas de las posiciones en específico. Tal es el caso del alelo B\*4002 que está relacionada con las posiciones 7 y 12 de p17, y las posiciones 10, 11, 12 y 64 de p2p7p1p6. También está el caso de A\*0201 relacionado con las posiciones 82 y 84 de p17, las posiciones 41 y 116 de p24, y la posición 78 de p2p7p1p6. Otro alelo frecuente es B\*3501 que se encuentra relacionado con las posiciones 122 de p17, y las posiciones 87, 91 y 120 en p24. Finalmente, se encontraron también a los alelos B\*1501 y A\*3101 relacionados con las posiciones 148 de p24 y 40 de p2p7p1p7, respectivamente.

Para continuar, se compararon las posiciones encontradas como bajo presión selectiva positiva por el servidor DataMonkey con posiciones reportadas previamente en la literatura como posiciones bajo presión selectiva inmune asociada a HLA tipo I en una cohorte de más de 1500 pacientes principalmente caucásicos de Estados Unidos, Canadá y Australia (**Tabla 6**) [7,47]. Se encontraron 12 posiciones (12, 122, 219, 223, 248, 280, 286, 374, 389, 403, 427 y 498) reportadas en la literatura que también fueron identificadas en el análisis por sustituciones dN/dS realizado por el servidor DataMonkey. Estas posiciones podrían representar sitios compartidos de variación debida a presión selectiva por HLA en el VIH circulante en Norteamérica tanto en población caucásica como en población mestiza mexicana. El resto de las posiciones identificadas como bajo presión selectiva en la cohorte mexicana, podría representar sitios de evolución característica del virus en México, no identificados en la cohorte caucásica.

Para identificar de una manera más específica y representativa de la población mexicana en estudio codones bajo presión selectiva por alelos HLA, se utilizaron datos previamente obtenidos de la aplicación del modelo de Redes de dependencia filogenética (PDN) [10] a una cohorte de pacientes mexicanos infectados por VIH. El modelo PDN identifica asociaciones HLA-polimorfismo de VIH, corrigiendo para los principales confusores de este

tipo de análisis evolutivo: efectos filogenéticos del virus, desequilibrio de ligamiento de los alelos HLA y coevolución de codones de VIH. Aplicando previamente el modelo PDN a una base de datos de 500 pacientes mexicanos con secuencias de *gag* de VIH y tipificación HLA de la base de datos de Epidemiología y Evolución del VIH en México del CIENI, INER, se obtuvieron como resultado 61 asociaciones HLA-polimorfismo de VIH presentes en 36 codones distintos de *gag* (Tabla 5).

Tabla 5. Asociaciones HLA-polimorfismo de VIH en el gen *gag* en una cohorte de 500 pacientes mexicanos, según el modelo de PDN\*.

Codón	Alelo HLA	a.a. no escapado	a.a. escapado	p	q
18	C*03:04	K	R	6.60E-05	0.12086234
18	C*03:04	K	K	8.93E-05	0.13550503
26	B*15:01	K	R	8.22E-15	2.56E-10
26	B*15:01	K	K	2.35E-09	2.93E-05
28	A*03	K	R	4.44E-05	0.1022498
28	A*03:01	K	K	4.08E-15	2.54E-10
28	A*03:01	K	Q	2.90E-09	3.01E-05
28	A*30:01	K	K	0.00016127	0.19717921
28	B*13:02	K	R	3.15E-05	0.08486136
30	A*33:03	K	Q	0.00012438	0.16129404
35	B*48:01	V	V	6.90E-05	0.12122472
54	B*15	S	S	0.00011154	0.15119477
69	C*07	Q	K	8.08E-05	0.12567122
81	B*15:03	T	T	0.00010327	0.14608691
81	C*14	T	T	2.20E-08	0.00019538
81	C*14	T	A	1.16E-06	0.00515521
104	B*49:01	I	V	4.86E-05	0.09277452
116	B*08:01	Q	Q	1.20E-05	0.0300012
120	C*08	A	A	4.92E-05	0.09277452
127	A*11:01	Q	P	1.32E-05	0.03153906
128	C*16:01	V	V	3.69E-05	0.07926894
130	A*11	A	Q	5.84E-06	0.01816107
130	A*11	A	H	7.09E-06	0.02053571
146	C*07	I	A	0.00010101	0.14608691
147	B*13	V	I	4.23E-06	0.01464078
147	B*13	V	L	8.72E-06	0.0235552



168	A*26:01	E	V	5.46E-06	0.01789656
168	A*26:01	E	I	4.25E-05	0.08811503
203	C*17:01	V	D	5.34E-05	0.11088017
203	C*17:01	V	E	5.34E-05	0.11088017
215	B*51:01	V	M	7.81E-05	0.12567122
218	C*01	T	A	4.90E-05	0.09277452
242	B*57	G	N	7.41E-11	1.54E-06
242	B*57	G	T	2.18E-10	3.40E-06
242	B*58	G	N	2.80E-06	0.01024889
242	B*58	G	T	7.26E-06	0.02053571
248	C*04:04	T	G	2.01E-05	0.04470171
264	B*27	R	R	6.69E-05	0.13418993
264	C*01:02	R	K	8.95E-05	0.15925689
280	B*15	K	T	9.79E-05	0.14511817
280	B*52	K	S	5.67E-08	0.00044101
280	B*52	K	T	3.08E-07	0.0021284
286	C*01	F	K	7.01E-05	0.12122472
286	C*01	F	R	0.00016156	0.19717921
302	B*14:01	G	K	0.00011416	0.15119477
302	B*14:01	G	R	0.00011416	0.15119477
335	A*26	N	K	3.41E-05	0.08486136
335	A*26	N	R	3.41E-05	0.08486136
357	B*07:02	A	S	1.35E-06	0.00560883
357	B*07:02	A	G	9.08E-06	0.0235552
372	C*15	M	N	7.47E-05	0.12564881
374	B*27	N	N	1.38E-05	0.03170569
378	C*03:04	V	I	0.00011204	0.18220908
385	C*05:01	K	P	8.01E-05	0.12567122
390	B*27:05	E	I	0.00014199	0.18036866
397	A*31	K	R	4.48E-07	0.00278724
397	A*31	K	K	6.52E-07	0.0033801
398	B*40:02	E	E	5.97E-07	0.0033774
398	C*17:01	E	I	7.15E-05	0.13478697
403	A*31	R	R	1.08E-06	0.005149
403	A*31	R	K	2.37E-06	0.00921593

\*Asociaciones obtenidas por el modelo de redes de dependencia filogenéticas de Carlson y colaboradores [10].

Al hacer la comparación entre los 36 codones de VIH asociados a presión selectiva por HLA por el modelo de PDN en la cohorte mexicana y asociaciones reportadas previamente en

la literatura en pacientes caucásicos, se encontraron 21 codones que coincidían en las dos cohortes: 26, 28, 30, 81, 104, 127, 146, 147, 218, 242, 248, 264, 280, 286, 302, 357, 372, 374, 397, 398 y 403 [47]. Por tanto, 15 posiciones en *gag* mostraron asociación con HLA en la cohorte mexicana que no han sido reportadas con anterioridad en otras cohortes, mostrando características particulares de la evolución asociada a HLA del VIH en México.

A continuación, al comparar los codones obtenidos por el modelo PDN con los datos obtenidos del servidor DataMonkey, se observó que 6 de ellos coincidían con posiciones identificadas como bajo presión selectiva negativa; es decir, tienden a conservar el aminoácido presente en la secuencia consenso, y fueron los codones: 116, 248, 280, 286, 374 y 403. Interesantemente, 50% (3/6) de estos codones se encontraron dentro de la proteína p24. De los codones restantes, 2 coincidieron con posiciones bajo selección positiva determinada individualmente por ambos métodos, y fueron los codones: 54 y 69. Interesantemente, ninguno de estas posiciones se encontró dentro de la proteína p24.

Finalmente, utilizando el modelo de PDN se encontraron 12 codones asociados a HLA que no han sido reportados previamente en la literatura y que no fueron detectados por el servidor DataMonkey, y son los codones: 18, 35, 120, 128, 130, 168, 203, 215, 335, 378, 385 y 390. En conjunto, a partir de todos los análisis de selección positiva y asociación con HLA realizados, se observaron 14 codones que no han sido reportados previamente en la literatura que tienden a cambiar el aminoácido presente en la secuencia consenso, que se encuentran bajo presión selectiva mediada por HLA y que podrían pertenecer específicamente a la población mexicana (**Tabla 6**).

Tabla 6. Resumen de resultados de asociaciones HLA-polimorfismo de VIH por diferentes análisis.

Codón	ANÁLISIS		
	Posiciones previamente reportadas asociadas con HLA <sup>1</sup>	Posiciones bajo presión selectiva por razones dN/dS <sup>2</sup>	Posiciones asociadas a HLA por modelo PDN <sup>3</sup>
7		*	
12	*	*	
18			*
26	*		*
28	*		*
30	*		*
34	*		
35			*
53	*		
54		*	*
55	*		
58	*		
62		*	
67	*		
69		*	*
71	*		
75	*		
76		*	
81	*		*
82		*	
84		*	
94		*	
95		*	
104	*		*
115		*	
116		*	*
120			*
122	*	*	
127	*		*
128			*
130			*
146	*		*
147	*		*
148	*		
168			*
176		*	
203			*
215			*
218	*		*
219	*	*	
223	*	*	
228	*		
242	*		*

Codón	ANÁLISIS		
	Posiciones previamente reportadas asociadas con HLA <sup>1</sup>	Posiciones bajo presión selectiva por razones dN/dS <sup>2</sup>	Posiciones asociadas a HLA por modelo PDN <sup>3</sup>
248	*	*	*
252		*	
260	*		
264	*		*
280	*	*	*
286	*	*	*
302	*		*
310		*	
335			*
340		*	
342	*		
357	*		*
371	*		
372	*		*
373		*	
374	*	*	*
375		*	
378			*
383	*		
385			*
389	*	*	
390			*
397	*		*
398	*		*
403	*	*	*
427	*	*	
441		*	
451		*	
453		*	
456		*	
466		*	
469		*	
470		*	
473		*	
479	*		
482	*		
483		*	
484	*		
488		*	
490		*	
495		*	
498	*	*	

\* Se encontró relación con alelos HLA-I.

1. Posiciones reportadas previamente en la literatura [7, 23, 25, 26, 46].
2. Análisis realizado por medio del servidor DataMonkey [www.datamonkey.org].
3. Análisis desarrollado por Carlson y colaboradores [10].

En amarillo codones con polimorfismos asociados a HLA-I que coinciden en la literatura y en el análisis de DataMonkey y que tienden a cambiar el aminoácido de la secuencia consenso. En rosa codones con polimorfismos asociados a HLA-I que coinciden en la literatura y el análisis por PDN y que tienden a cambiar el aminoácido de la secuencia consenso. En azul codones con polimorfismos asociados a HLA-I que coinciden en el análisis de DataMonkey y el análisis por PDN y que tienden a cambiar el aminoácido de la secuencia consenso. En verde codones con polimorfismos asociados a HLA-I que tienden a conservar el aminoácido presente en el consenso y que coinciden en el análisis de DataMonkey y el análisis por PDN. En anaranjado codones con polimorfismos asociados a HLA-I que tienden a conservar el aminoácido presente en el consenso y que coinciden en el análisis de DataMonkey, el análisis por PDN y la literatura.

Aunque los análisis de colocación de posiciones bajo presión selectiva positiva por razón  $dN/dS$  con epítomos de CTLs previamente comprobados mostraron que prácticamente todas las posiciones podían estar asociadas a selección por HLA, muchos de los alelos HLA reportados en la Base de Datos de Inmunología de VIH de Los Álamos no son comunes en la población mexicana y por tanto las mutaciones no serían específicas del contexto inmunogenético de esta población. Para explorar la influencia característica del contexto inmunogenético mexicano en la evolución del VIH se estudió la frecuencia de los alelos HLA asociados a los 14 codones, no reportados previamente y bajo presión selectiva positiva por HLA en la cohorte mexicana, de acuerdo a los resultados del modelo de PDN. Los alelos HLA asociados a estos codones fueron comparados con los 10 alelos HLA-A, los 10 HLA-B y los 10 HLA-C más comunes en la población mexicana. Los datos de frecuencias HLA, previamente obtenidos en el laboratorio, fueron obtenidos de la Base de Datos de Epidemiología Molecular y Evolución del VIH del CIENI. Interesantemente, 7 de los codones bajo presión selectiva positiva (18, 69, 120, 130, 215, 378 y 385) estuvieron relacionados con alelos HLA que se encuentran entre los HLA- A, B o C más comunes de la población mexicana ( A\*11, B\*51:01, C\*07, C\*08 y C\*03:04) y podrían representar asociaciones HLA-polimorfismo de VIH no detectadas en la cohorte mexicana por el modelo de PDN.

#### **6. Impacto de la presión selectiva mediada por HLA en la capacidad replicativa del VIH.**

Para explorar el impacto de mutaciones seleccionadas por HLA en la funcionalidad de las proteínas del gen *gag*, se buscaron en la literatura posiciones previamente asociadas con costos en la capacidad replicativa viral. Se encontraron en total 31 mutaciones en *gag* relacionadas con alelos HLA y que tienen relevancia en la capacidad replicativa [7, 16, 26] las cuales están distribuidas de la siguiente manera: 12 mutaciones con disminución en la capacidad replicativa viral en fase aguda de infección, 15 mutaciones con disminución en la capacidad replicativa en fase crónica de infección y 4 mutaciones reportadas como compensatorias y que por tanto, restablecen la capacidad replicativa. Estas posiciones relacionadas con la capacidad replicativa fueron identificadas en cada una de las muestras de la cohorte analizada. Los resultados se compilaron en la **Tabla 7 y la Figura 24**.

Tabla 7. Mutaciones en Gag reportadas como importantes en la capacidad replicativa del VIH.

Codón	12	28	30	34	53	55	58	67	71	75	122	146	147	148	218	218	219	223	228	242	248	260	342	371	383	389	427	479	482	484	498			
LRC-AGUDA	E->Q	K->T	R->K	L	S->T	Q	R->G	S->A	V	L	T->I	A->P	I->L	S->T	V	V->Q	H->Q	I->V	M->I	T->N	G->A	E->D	S->X	DEL.T	F->T	T->I	T->P	T	E->D	Y->X	L			
LRC-CHRONICA	E	H	M	L	T	E	R	S	G	L	-	A	I	S	V	V	H	I	M	S	A	D	S	N	F	I	I	T	R	E	-	-		
C-HRC	K	K	K	K	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	S	A	D	S	N	F	I	I	T	R	E	-	-		
B023804	K	H	M	L	T	E	R <td>S</td> <td>G</td> <td>L</td> <td>-</td> <td>A</td> <td>I</td> <td>S</td> <td>V</td> <td>V</td> <td>H</td> <td>I</td> <td>M</td> <td>S <td>A</td> <td>D</td> <td>S</td> <td>N</td> <td>F</td> <td>I</td> <td>I</td> <td>T <td>R</td> <td>E</td> <td>-</td> <td>-</td> </td></td>	S	G	L	-	A	I	S	V	V	H	I	M	S <td>A</td> <td>D</td> <td>S</td> <td>N</td> <td>F</td> <td>I</td> <td>I</td> <td>T <td>R</td> <td>E</td> <td>-</td> <td>-</td> </td>	A	D	S	N	F	I	I	T <td>R</td> <td>E</td> <td>-</td> <td>-</td>	R	E	-	-		
#0196	Q	K	R	R	T	G	R	S	G	L	T	P	I	S	V	V	H	V	M	T	G	E	T	T	F	V	I	T	I	D	Y	S		
#0205	K	K	R	R	T	E	R	S	G	L	T	P	I	S	V	V	P	I	M	T	G	E	T	T	F	I	T	I	K	Y	S			
#0209	K	K	R	?	T	E	R	S	G	L	T	P	I	S	V	V	P	I	M	T	G	E	T	T	F	I	T	I	I	E	Y	S		
#0255	E	K	Q	I	T	D	R	S	G	L	A	P	I	S	V	V	H	I	I	M	T	A	E	T	F	P	T	A	E	Y	S			
#0258	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	P	V	M	T	G	E	T	T	F	F	P	T	A	E	Y	L		
#0259	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	E	T	T	F	F	?	N	K	E	Y	S		
#0261	K	K	R	L	T	E	R	S	G	L	A	A	I	S	V	V	Q	V	M	T	A	E	T	T	F	F	I	T	I	E	Y	S		
#0281	K	K	K	L	T	E	R	S	G	L	A	P	?	S	V	V	H	I	M	T	A	D	E	T	T	F	?	T	R	E	Y	S		
#0291	K	K	R	I	T	E	R	S	A	L	T	A	L	S	V	V	Q	V	M	T	G	E	T	T	F	F	I	T	I	E	Y	S		
#0299	E	K	R	I	T	T	R	S	G	L	T	A	L	S	V	V	Q	V	M	T	G	E	T	T	F	F	A	T	R	E	Y	S		
#0311	E	K	R	I	T	T	R	S	G	L	T	A	L	S	V	V	Q	V	M	T	G	E	T	T	F	M	F	I	T	N	?	?	S	
#0322	Q	K	?	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	I	T	I	E	Y	S		
#0331	E	K	R	I	T	E	R	S	?	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	I	T	I	E	Y	S		
#0340	E	?	R	I	T	G	K	S	?	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	R	E	Y	S			
#0341	E	K	K	I	T	S	R	S	G	L	T	P	I	S	V	V	H	V	M	T	G	E	T	T	F	F	A	T	I	E	Y	L		
#0355	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	I	T	R	E	Y	S		
#0360	K	K	K	I	T	G	R	S	G	L	A	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	S	T	R	E	Y	S		
#0361	K	K	R	I	T	E	R	S	G	L	V	P	?	S	V	V	H	I	L	M	T	G	E	T	T	F	P	N	K	E	Y	S		
#0362	E	K	R	I	T	E	R	S	G	L	?	T	P	I	S	V	Q	V	I	M	T	A	E	T	T	F	P	N	M	G	Y	S		
#0365	K	K	R	I	T	E	R	?	G	L	T	A	I	S	V	A	?	I	M	T	G	E	T	T	F	F	I	T	I	E	Y	S		
#0377	E	Q	K	I	T	E	R	S	G	L	T	A	I	S	V	A	H	P	V	M	T	G	D	T	N	F	I	N	I	G	Y	S		
#0390	E	K	H	I	T	G	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	E	T	T	F	F	A	T	T	E	Y	S		
#0393	E	Q	K	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	I	?	M	T	A	D	T	T	F	?	T	T	E	Y	S		
#0403	E	K	R	I	T	?	R	S	G	L	T	A	I	S	V	?	H	A	I	M	T	G	E	T	T	F	?	T	I	E	Y	L		
#0412	Q	T	Q	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	A	T	R	E	Y	L		
#0413	E	K	R	L	T	Q	R	S	G	L	?	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	V	T	V	G	Y	S		
#0420	E	K	R	I	T	Q	K	?	G	L	?	P	I	S	V	V	H	I	M	T	G	E	T	T	F	F	?	T	I	E	Y	S		
#0422	E	K	R	I	T	Q	K	?	G	L	?	P	I	S	V	V	H	I	M	T	G	E	T	T	F	F	?	T	I	E	Y	S		
#0425	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	T	T	T	E	Y	S		
#0427	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	T	T	T	E	Y	S		
#0428	E	K	R	I	?	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	T	T	T	E	Y	S		
#0429	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	A	?	T	T	F	F	I	S	I	E	Y	S		
#0441	K	L	K	L	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	A	A	E	T	T	F	P	T	T	E	Y	L		
#0443	E	K	K	I	T	G	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	T	I	E	Y	S		
#0449	A	K	R	V	T	E	R	S	G	L	T	A	I	S	V	A	Q	I	I	M	T	G	E	T	T	F	I	T	I	E	Y	S		
#0472	E	K	K	I	T	Q	R	S	A	L	T	?	P	I	S	V	A	Q	I	M	T	G	E	T	T	F	T	T	N	F	E	Y	S	
#0478	E	K	K	I	T	E	R	S	G	L	V	P	I	S	V	V	H	V	V	M	T	G	D	S	N	F	T	S	I	E	Y	S		
#0494	G	K	R	I	T	E	R	S	G	L	I	A	I	S	V	V	H	V	V	M	T	G	E	T	T	F	L	P	N	V	E	Y	L	
#0514	E	K	R	L	T	E	R	S	G	L	T	A	L	S	V	V	Q	V	M	T	G	E	T	T	F	F	I	T	I	E	Y	S		
#0526	E	Q	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	T	T	I	E	Y	S		
#0538	?	K	R	L	T	?	R	S	A	G	?	A	I	S	V	V	H	I	M	T	G	E	T	T	F	F	?	N	I	E	Y	S		
#0545	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	?	Q	V	L	M	T	G	E	T	T	F	S	T	T	E	H	?		
#0577	E	R	R	I	T	E	R	S	G	L	T	A	I	S	V	P	?	Q	V	L	M	T	G	E	T	T	F	S	T	T	E	H	?	
#0581	K	K	K	I	T	D	R	A	G	L	P	P	I	?	V	V	H	V	L	M	T	A	D	T	T	F	P	N	M	E	Y	L		

Tabla 7. Continuación.

#0586	K	Q	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	V	I	N	Q	E	T	N	F	?	T	I	E	Y	S	
#0591	?	K	R	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	-	M	N	I	D	Y	S
#0592	K	K	K	K	T	E	R	S	G	L	T	A	L	S	?	?	H	I	M	T	G	E	T	T	F	P	T	?	E	Y	S	
#0594	E	K	R	R	T	G	R	S	G	L	A	A	L	S	A	A	Q	V	I	T	G	E	T	T	F	I	T	I	E	Y	S	
#0595	E	K	R	R	T	E	R	?	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	?	T	I	E	Y	S	
#0596	E	K	R	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#0605	E	K	Q	Q	T	E	R	A	G	L	T	A	L	S	?	?	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0606	E	K	R	Q	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0616	E	K	K	K	T	E	R	V	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0623	K	K	Q	Q	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0651	K	R	R	K	T	E	R	S	G	L	T	P	P	?	V	Q	V	I	M	T	G	E	T	T	F	I	T	I	E	Y	W	
#0654	E	K	R	R	T	E	R	S	G	L	T	P	P	?	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S		
#0680	E	K	R	R	T	E	R	S	G	L	T	P	P	?	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S		
#0687	E	K	R	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0688	D	Q	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0692	K	Q	R	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0695	E	K	Q	R	T	E	R	K	S	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	L	
#0708	E	K	R	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	L	
#0717	E	K	K	K	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	L	
#0724	E	K	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0743	E	K	?	?	V	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0760	E	Q	K	K	T	E	R	?	S	L	T	A	L	S	Q	Q	V	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0769	E	?	?	?	T	E	R	?	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0781	E	K	K	R	T	E	R	?	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0797	E	K	R	R	T	E	R	?	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0807	?	?	?	?	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0829	E	K	K	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0832	E	K	K	R	T	E	R	S	G	L	T	A	L	S	?	?	Q	?	?	?	?	E	T	T	F	I	T	I	E	Y	S	
#0842	E	K	H	H	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0884	?	?	R	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0889	E	K	K	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0901	E	K	R	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0917	E	K	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0918	E	Q	H	H	T	E	R	A	G	L	T	A	L	S	A	A	Q	V	I	M	T	G	E	T	T	F	I	T	I	E	Y	S
#0920	K	K	K	R	T	E	R	K	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0926	E	K	K	K	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0937	E	K	K	K	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0941	E	K	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0949	E	K	K	K	T	E	R	K	S	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0959	E	K	K	K	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0967	E	K	R	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0985	E	?	R	R	T	E	R	K	S	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0989	E	?	R	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#0991	E	K	K	R	T	E	R	?	S	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1001	E	K	K	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1010	E	Q	K	K	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1012	E	K	K	R	T	E	R	A	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1016	E	K	K	R	T	E	R	Q	S	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1024	E	K	K	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1031	E	K	?	?	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	
#1033	E	K	K	R	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	T	F	I	T	I	E	Y	S	

Tabla 7. Continuación.

#1054	E	K	K	I	T	E	R	S	G	L	T	P	I	S	V	V	H	V	M	T	G	D	T	F	?	T	I	E	Y	S		
#1056	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	A	M	T	G	D	T	F	P	D	-	-	Y	S		
#1062	E	K	R	I	T	E	R	S	G	L	A	A	L	S	V	V	Q	I	M	T	T	E	T	F	P	I	T	I	E	Y	S	
#1063	E	Q	R	I	T	E	K	A	G	L	P	P	L	S	V	V	Q	?	V	T	A	E	T	F	P	T	D	I	E	Y	S	
#1084	E	Q	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	A	E	T	F	T	I	I	E	Y	S		
#1087	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	A	E	T	F	T	I	I	E	Y	S		
#1000	K	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	S	T	T	I	I	E	Y	S		
#1077	G	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	Q	I	M	T	T	E	T	F	I	T	I	E	Y	S		
#1111	D	Q	K	I	T	E	R	S	G	L	A	A	L	S	V	V	H	I	M	T	G	E	T	F	T	T	I	E	Y	S		
#1126	K	R	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	S	T	T	I	I	E	Y	S		
#1129	K	K	K	I	T	E	R	S	G	L	T	S	I	S	V	V	Q	I	M	T	G	E	S	T	T	I	I	E	Y	S		
#1130	K	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	S	T	T	I	I	E	Y	S		
#1138	K	R	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	Q	E	T	F	A	T	I	E	Y	S		
#1140	E	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	H	I	M	T	A	E	T	F	P	T	I	E	Y	S		
#1143	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	A	E	T	F	P	T	I	E	Y	S		
#1149	Q	Q	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	F	?	N	I	E	Y	S		
#1166	Q	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	F	A	T	I	E	Y	S		
#1181	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	F	I	T	I	E	Y	S		
#1185	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	?	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1187	K	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	?	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1190	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1191	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1197	Q	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1203	E	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	H	I	M	T	A	E	T	F	?	T	I	E	Y	S		
#1209	K	?	K	I	T	E	R	S	G	L	T	P	I	S	V	V	Q	I	?	?	?	E	T	N	G	S	R	I	E	Y	S	
#1210	E	K	R	I	T	E	R	S	G	L	F	A	L	S	V	V	H	?	M	T	N	G	E	T	F	I	N	I	E	Y	S	
#1218	K	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	A	E	T	F	I	N	I	E	Y	S		
#1229	E	K	K	I	T	E	R	S	G	L	T	P	I	S	V	V	H	I	M	T	G	D	T	F	I	N	I	E	Y	S		
#1232	E	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	H	I	M	T	A	E	T	F	I	N	I	E	Y	S		
#1239	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	F	I	N	I	E	Y	S		
#1257	K	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	T	F	?	T	I	E	Y	S		
#1301	K	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	H	I	M	T	?	G	E	T	F	N	T	R	E	Y	S	
#1319	E	Q	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	M	T	G	E	T	F	P	T	I	E	Y	S		
#1328	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	V	M	T	A	E	S	T	F	S	T	M	I	E	Y	S
#1335	Q	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	A	M	T	G	E	T	F	I	T	I	E	Y	S		
#1346	K	K	R	I	T	E	R	S	G	L	F	A	L	S	V	V	H	V	M	T	G	E	T	F	I	T	I	E	Y	S		
#1349	?	?	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	N	T	G	E	T	F	I	T	I	E	Y	S		
#1367	D	K	K	I	T	E	R	S	G	L	T	?	L	S	V	V	H	I	M	T	G	E	T	F	P	T	I	E	Y	S		
#1369	E	K	R	I	T	E	R	S	G	L	T	P	I	S	V	V	?	H	I	M	T	T	E	S	T	?	T	I	E	Y	S	
#1374	E	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	V	M	T	G	E	T	F	I	T	I	E	Y	S		
#1376	?	A	K	I	T	E	R	S	G	L	T	P	I	S	V	V	?	H	I	M	T	G	E	T	F	I	T	I	E	Y	S	
#1377	E	K	R	I	T	E	R	S	G	L	T	?	L	S	V	V	Q	I	N	T	A	E	T	F	S	N	N	M	E	Y	S	
#1378	K	K	R	I	T	E	R	S	G	L	T	A	L	S	V	V	H	V	M	T	A	E	T	F	I	T	I	E	Y	S		
#1380	E	R	R	I	T	E	R	S	G	L	F	A	L	S	V	V	Q	I	M	T	A	E	T	F	I	T	I	E	Y	S		
#1387	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	I	M	T	G	E	D	T	I	I	E	Y	S			
#1392	K	K	H	I	T	E	R	S	G	L	T	P	L	S	V	V	H	I	M	T	A	E	T	F	S	T	I	E	Y	S		
#1394	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	V	M	T	A	E	T	F	I	T	I	E	Y	S		
#1395	E	K	K	I	T	E	R	S	G	L	T	A	L	S	V	V	H	V	M	T	G	E	T	F	I	T	I	E	Y	S		
#1396	E	Q	K	I	T	E	R	S	G	L	T	A	L	S	V	V	Q	I	?	?	T	E	T	F	?	?	T	I	E	Y	S	
#1398	E	K	K	I	T	E	R	S	G	L	T	P	I	S	V	V	Q	I	M	T	?	E	T	N	Y	T	T	I	E	Y	S	

Tabla 7. Continuación.

#1400	E	K	R	L	T	E	R	S	G	L	T	A	I	S	A	V	Q	V	M	T	G	E	?	T	F	P	T	I	E	Y	S	
#1401	E	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	?	F	-	T	I	E	Y	S	
#1402	K	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	A	E	S	T	F	I	A	G	E	Y	S	
#1406	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	A	E	T	T	F	P	T	I	E	Y	S	
#1407	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	A	E	T	T	F	P	T	I	E	Y	S	
#1408	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	D	T	T	F	I	I	E	Y	S		
#1409	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	R	D	T	T	F	S	A	D	R	E	Y	S
#1410	K	C	K	L	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	E	T	T	F	A	D	M	G	E	Y	S
#1412	K	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	?	E	T	T	F	N	T	I	E	Y	S	
#1415	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	D	T	T	F	V	T	I	E	Y	S	
#1419	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	D	T	T	F	V	T	I	E	Y	S	
#1421	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	D	T	T	F	V	T	I	E	Y	S	
#1424	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	E	T	T	F	P	I	D	I	E	Y	S
#1426	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	G	E	T	T	F	P	I	D	I	E	Y	S	
#1427	Q	K	Q	I	T	A	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	?	E	T	T	F	P	K	T	I	E	Y	S
#1428	Q	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	?	E	T	T	F	P	K	T	I	E	Y	S
#1436	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1440	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1441	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1444	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1447	Q	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	N	G	E	T	F	A	T	I	E	Y	S	
#1463	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1469	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1484	K	Q	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1488	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1514	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1518	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1519	K	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1520	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1526	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1528	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1530	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1531	E	Q	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1534	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1535	Q	R	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1536	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1538	E	K	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1540	K	K	?	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1550	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1552	Q	T	K	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	?	E	T	T	F	A	T	I	E	Y	S	
#1553	K	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1561	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1599	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1620	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	A	T	I	E	Y	S	
#1621	D	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	N	A	E	T	T	F	A	T	I	E	Y	S
#1640	Q	K	K	L	T	E	R	S	G	L	T	A	I	S	V	V	Q	V	M	T	A	E	T	T	F	P	T	I	E	Y	S	
#1649	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1654	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1656	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1682	K	K	Q	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1694	K	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1698	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1703	K	Q	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#1714	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#CT21	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	
#CT2a	E	K	R	I	T	E	R	S	G	L	T	A	I	S	V	V	H	I	M	T	G	E	T	T	F	P	T	I	E	Y	S	

Las posiciones marcadas en azul corresponden a mutaciones de aparición en fase aguda, las posiciones que aparecen en amarillo corresponden a mutaciones de aparición en fase crónica y las posiciones marcadas en anaranjado corresponden a mutaciones compensatorias. Adicionalmente, se marcan en rojo otros polimorfismos presentes pero no reportados como importantes en la capacidad replicativa del VIH. Los signos de interrogación corresponden a codones que presentan ambigüedad en sus nucleótidos. Los guiones representan deleciones.



Se observó que los cambios 53T, 75L y 218V, asociados a disminución de capacidad replicativa viral en fase aguda, estuvieron presentes en un 96.5%, 89.5% y 84.9% de los virus, respectivamente (**Fig. 24**). Así, estas mutaciones se han acumulado en la población viral hasta volverse el consenso poblacional. Los costos en capacidad replicativa asociados a estas mutaciones podrían estar compensados con la aparición de mutaciones compensatorias en los virus que las poseen. Adicionalmente, se observó que la totalidad de mutaciones compensatorias encontradas se localizaron en la zona correspondiente a p24.

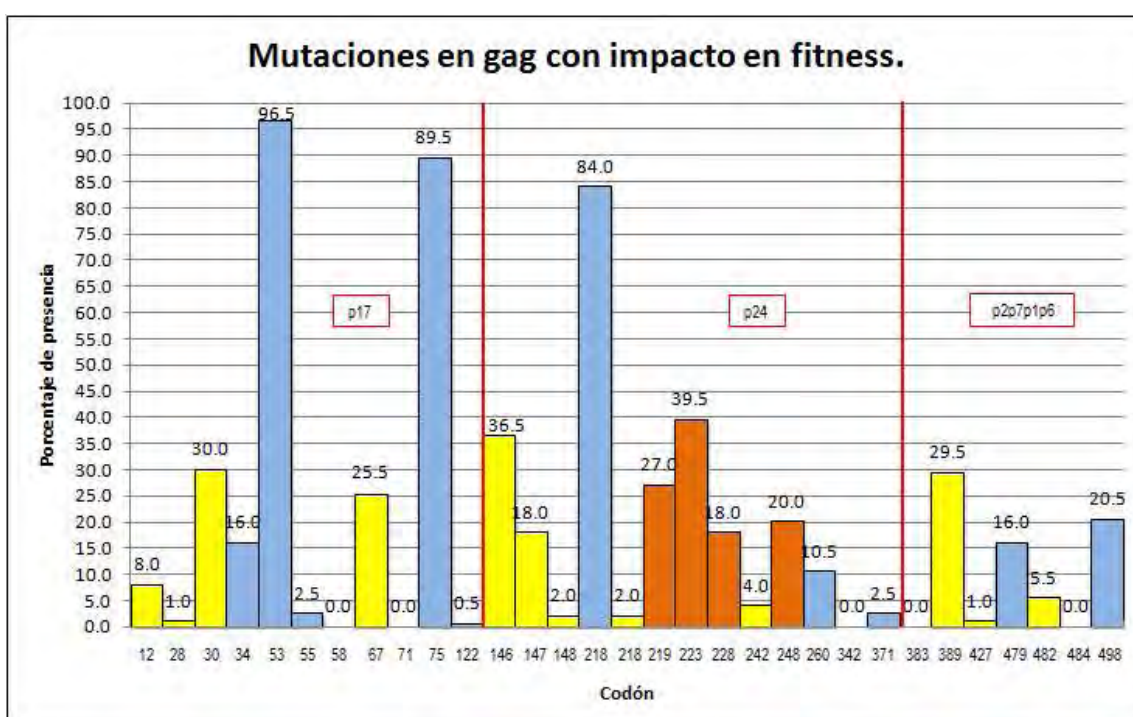


Figura 24. Frecuencia de mutaciones con costo en capacidad replicativa en una cohorte de 200 pacientes mexicanos. En azul se muestran las mutaciones en fase aguda, en amarillo las mutaciones de fase crónica y en anaranjado las mutaciones compensatorias. Además se señalizan las ubicaciones de las diferentes proteínas codificadas en este gen.

Por otro lado, 5 mutaciones asociadas a costos en capacidad replicativa reportadas en literatura: R58G, 71V, S342X, F383T y Y484X, no se encontraron en ninguna de las secuencias analizadas. Esto es interesante y podría representar características únicas de adaptación en el VIH circulante en México, muy probablemente debido a que el alelo HLA selector para estas mutaciones no se encuentra presente en los pacientes analizados. El resto de las mutaciones se presentaron en diferentes frecuencias en los virus mexicanos.

La mayor cantidad de codones con mutaciones que afectan la capacidad replicativa, así como los cambios con impacto en la capacidad replicativa con mayor frecuencia en la población mexicana se encontraron en p17; sin embargo, se observó que la totalidad de las mutaciones compensatorias se encontraban localizadas en p24. Este resultado es lógico dado que p24 es una de las zonas más conservadas del genoma del virus y frecuentemente cambios en esta zona provocarían una disminución considerable de la capacidad replicativa, por lo cual es necesario contrarrestar este efecto mediante las mutaciones compensatorias en esa zona en específico.

Para evaluar más específicamente el impacto de variantes observadas en la población mexicana en la capacidad replicativa, se seleccionaron virus para los cuales se tenían mediciones experimentales de capacidad replicativa. Los datos de medición de capacidad replicativa se obtuvieron previamente en los laboratorios del CIENI. Se obtuvieron los valores de capacidad replicativa para 58 de los 200 virus incluidos en el estudio. De estas secuencias se obtuvo un árbol filogenético y se identificó en distintos colores los virus correspondientes a valores de alta capacidad replicativa y de baja capacidad replicativa (**Fig. 25**). Se observó que las secuencias tienden a formar grupos correspondientes a secuencias que tienen valores de alta o baja capacidad replicativa, respectivamente. Éste es un primer acercamiento que muestra la similitud entre las secuencias analizadas y por consiguiente, podría estar mostrando la similitud entre posibles mutaciones debidas a presión selectiva por HLA que estén provocando un aumento o una disminución en la capacidad replicativa de los virus.

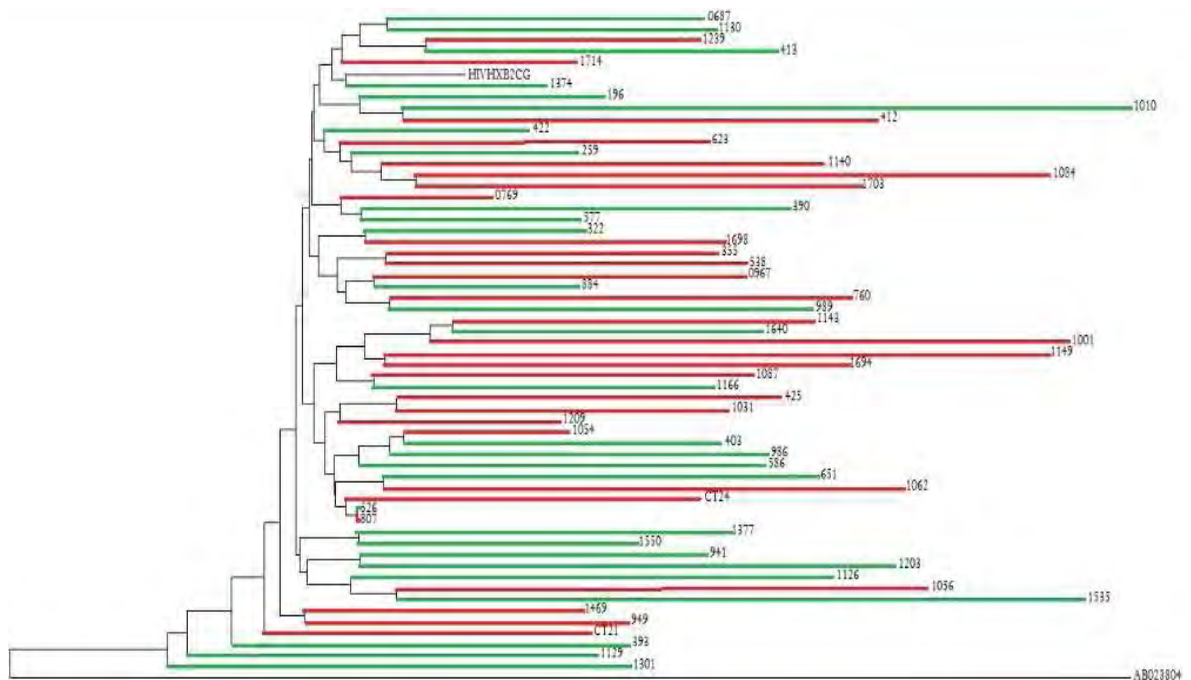


Figura 25. Árbol filogenético de las 58 muestras con valores asignados de capacidad replicativa. En verde se muestran las secuencias correspondientes a altos valores de capacidad replicativa y en rojo las correspondientes a bajos valores de capacidad replicativa. Se observa la formación de grupos específicos según la clasificación por alta o baja capacidad replicativa.

Posteriormente, los virus fueron ordenados en orden creciente de capacidad replicativa, marcando a su vez las mutaciones con costo en capacidad replicativa antes mencionadas (**Tabla 8**). Se observó que la cantidad de mutaciones con costo en capacidad replicativa presentes en las 29 muestras con capacidad replicativa baja fue similar a la cantidad presente en las 29 muestras con capacidad replicativa más alta. En cuanto al tiempo de aparición de las mutaciones con costo en capacidad replicativa se observaron las razones: 3.45 mutaciones de etapa aguda/virus y 1.34 mutaciones de etapa crónica/virus, en virus de menor capacidad replicativa; y 3.52 mutaciones de etapa aguda/virus y 1.72 mutaciones de etapa crónica/virus, en virus de mayor capacidad replicativa. Por el contrario, las mutaciones compensatorias se observaron en frecuencia mayor en los virus con mayor capacidad replicativa, con las razones: 0.69 mutaciones compensatorias/virus, en virus de menor capacidad replicativa; y 1.24 mutaciones compensatorias/virus, en los de mayor capacidad replicativa (**Fig. 26**). Esta distribución de mutaciones compensatorias permite a los virus recuperar pérdidas en capacidad replicativa provocadas por mutaciones de escape inmune.



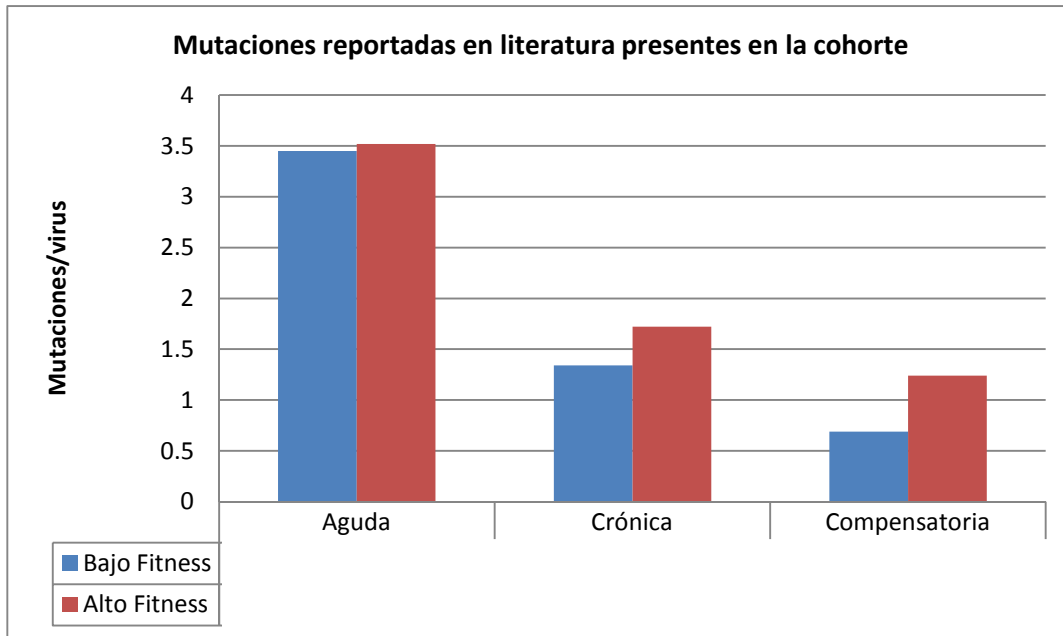


Figura 26. Gráfico mostrando la razón de mutaciones de fase aguda, de fase crónica y mutaciones compensatorias, reportadas previamente en la literatura, por virus, en los 29 virus de alta capacidad replicativa y los 29 de baja capacidad replicativa de la cohorte estudiada.

Adicionalmente, se observaron varios casos que respaldan la idea de que las mutaciones en la zona de p24 afectan en un mayor grado la capacidad replicativa que aquellas en p17 o en p2p7p1p6. Tal es el caso de las muestras 1374-HCF, 1301-ATCD y 1377-VCRC (las 3 secuencias incluidas en el grupo de mayor capacidad replicativa). La primera presentó la menor capacidad replicativa de las tres con 3 mutaciones en la zona de p24 y 1 mutación compensatoria; la segunda presentó mayor valor de capacidad replicativa que la anterior a pesar de tener 4 mutaciones que disminuyen la capacidad replicativa en p17 y ninguna compensación; y la última es la que presentó el mayor valor de capacidad replicativa de las 3 con 3 mutaciones en la zona de p24 y 4 mutaciones compensatorias.

Posteriormente, se hizo la relación de capacidad replicativa con los codones bajo presión selectiva positiva identificados por razón dN/dS con el servidor DataMonkey (**Tabla 9**). En la Tabla 9 se muestra la variación de aminoácidos con respecto a la referencia, en las posiciones antes mencionadas. Los virus utilizados fueron aquéllos para los cuales se conocía el valor de capacidad replicativa. De esta manera, se pudieron localizar

mutaciones putativas con efecto en la capacidad replicativa que no han sido reportadas previamente.

Se observaron posiciones en las cuales un polimorfismo específico estaba presente en la mayoría de las muestras: codón 94, presencia de V en un 69% de las muestras; codón 420, presencia de K en un 64% de las muestras; codón 459, presencia de H en un 76% de las muestras y codón 539, presencia de K en un 60% de las muestras. Así, se puede decir que estas mutaciones están conservadas y acumuladas en las secuencias analizadas, sin un efecto evidente en la capacidad replicativa viral. Adicionalmente, se observaron 7 codones (de los 40 seleccionados por DataMonkey) que presentaron un aminoácido específico diferente a la secuencia referencia, representado mayoritariamente en los virus con mayor capacidad replicativa. Así, estos cambios específicos, podrían actuar como mutaciones compensatorias (presencia preferencial en muestras con alta capacidad replicativa): E12Q, Q69K, H230Q, 234V, N263S, R297K Y S547L. Adicionalmente, se encontraron 4 codones que podrían actuar como mutaciones de escape al estar representados preferentemente en virus con baja capacidad replicativa: T84V, A390N, T445N y S470N (**Fig. 27**).

Tabla 9. Relación de posiciones bajo presión selectiva (DataMonkey) y capacidad replicativa.

Fitness	Codón	7	12	54	62	69	76	82	84	94	95	115	127	184	230	234	259	263	291	297	321	351	384	390	391	406	420	445	459	470	472	482	505	508	509	512	532	537	539	544	547			
		V	E	S	G	Q	R	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.4671184656	#1140	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <td>H</td> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.4901061382	#1087	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <td>H</td> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.5146699952	#1056	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <td>H</td> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.6700415888	#1703	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>I</td> <td>R</td> <td>V</td> <td>E</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	I	R	V	E	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S
0.6932408300	#0623	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>E</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	E	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.6971369363	#0425	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8022121077	#C121	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>?</td> <td>K</td> <td>A</td> <td>T</td> <td>?</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8372065379	#1001	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.837914874	#0729	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8495395688	#C174	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.882888594	#1062	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8680811924	#1054	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8730392359	#0967	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8785476201	#1084	?	E	S	G	Q	R <td>V</td> <td>T</td> <td>?</td> <td>K</td> <td>A</td> <td>T</td> <td>?</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8809860366	#1299	?	E	S	G	Q	R <td>V</td> <td>T</td> <td>?</td> <td>K</td> <td>A</td> <td>T</td> <td>?</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8897875128	#1694	?	E	S	G	Q	R <td>V</td> <td>T</td> <td>?</td> <td>K</td> <td>A</td> <td>T</td> <td>?</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8901322678	#1209	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.8924621015	#0538	V	?	A	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9165017397	#0949	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9283394952	#1031	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9393860734	#1149	V	Q	A	I	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9393851983	#1698	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9391814853	#1143	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9576282516	#0807	V	Q	A	E	R	?	?	?	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9589901387	#0760	V	Q	A	E	R	?	?	?	?	K	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.9821225058	#1469	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
0.990381462	#0412	V	Q	A	I	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0078605655	#0355	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0211336146	#1714	V	E	S	E	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0222363663	#0322	V	Q	P	R	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0393814028	#1374	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0610429924	#1126	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0953969340	#1203	V	E	R	G	Q	R <td>V</td> <td>T</td> <td>?</td> <td>R</td> <td>A</td> <td>T</td> <td>?</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	?	R	A	T	?	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.0966997106	#0986	V	E	R	G	H	K <td>V</td> <td>V</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	V	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.1000565485	#1166	V	Q	S	G	R	K <td>V</td> <td>V</td> <td>V</td> <td>V</td> <td>V</td> <td>V</td> <td>V</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	V	V	V	V	V	V	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.128561707	#0586	V	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.1268956941	#1130	V	K	A	T	K	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.172426487	#1130	V	Q	A	V	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.180435158	#1010	V	Q	A	V	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.190335158	#1129	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.1950532600	#1129	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.1983917864	#1550	?	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.208947961	#0687	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.2568451518	#0196	V	Q	A	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.2634865017	#0651	V	E	K	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S
1.265943994	#1377	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.2668813225	#0884	V	?	S	E	K	K <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.2918248184	#1535	V	Q	S	E	Q	K <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th> <th>G</th> <th>N</th> <th>T</th> <th>R</th> <th>S</th> <th>A</th> <th>S</th> <th>A</th> <th>A</th> <th>-</th> <th>-</th> <th>I</th> <th>R</th> <th>T</th> <th>H</th> <th>S</th> <th>P</th> <th>T</th> <th>E</th> <th>T</th> <th>P</th> <th>K</th> <th>L</th> <th>S</th> <th>R</th> <th>N</th> <th>S</th>	V	T	V	R	A	T	S	H	I	G	N	T	R	S	A	S	A	A	-	-	I	R	T	H	S	P	T	E	T	P	K	L	S	R	N	S	
1.326556047	#0526	V	E	S	G	Q	R <td>V</td> <td>T</td> <td>V</td> <td>R</td> <td>A</td> <td>T</td> <td>S</td> <th>H</th> <th>I</th>	V	T	V	R	A	T	S	H	I																												

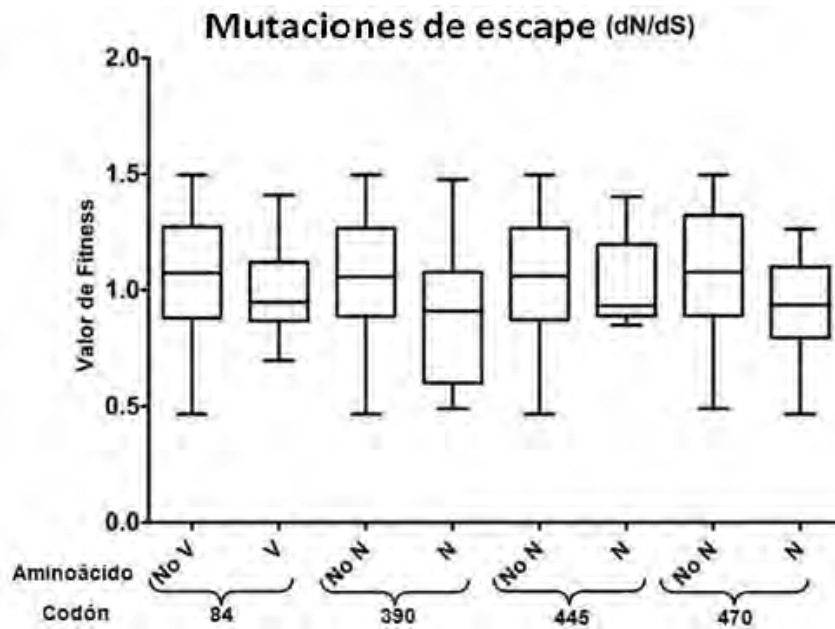
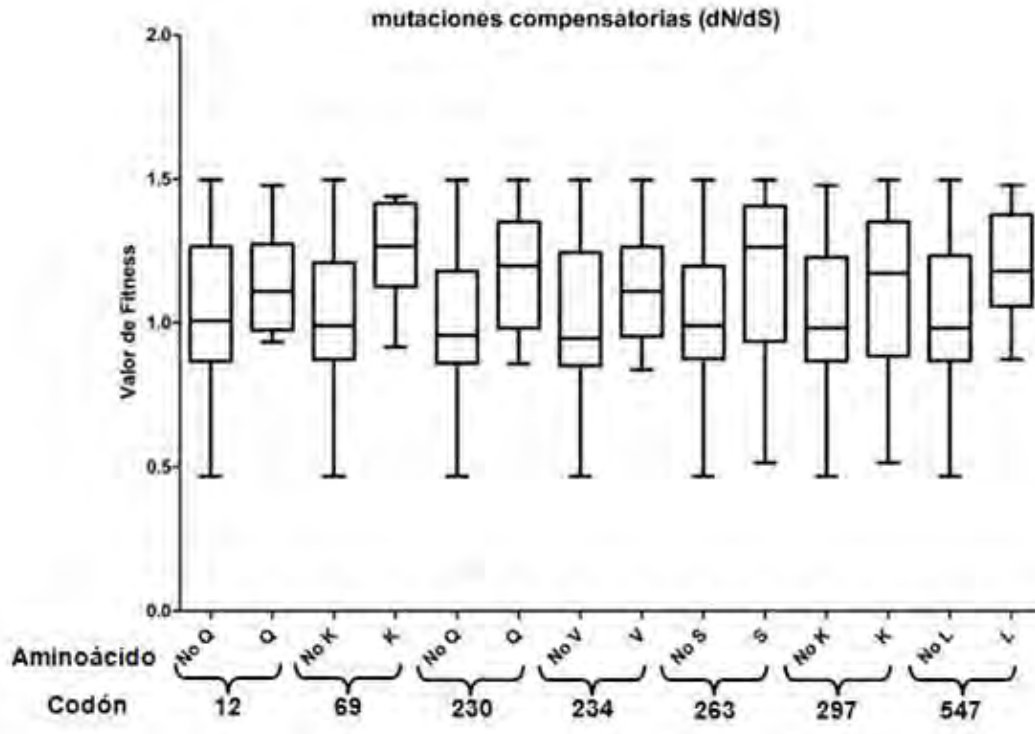


Figura 27. Gráfico de caja y bigotes de las mutaciones compensatorias y de escape encontradas por el análisis de razones de sustituciones no sinónimas / Sinónimas (DataMonkey). Ninguna fue estadísticamente representativa. En cada diagrama la caja sólida indica el rango entre los percentiles 25 y 75, la línea indica la mediana, los bigotes indican los valores máximos y mínimos no atípicos.

Para corroborar si las tendencias descritas anteriormente eran significativas estadísticamente, se realizaron pruebas exactas de Fisher, considerando los valores de



capacidad replicativa asociados a secuencias con y sin las mutaciones en los 11 codones mencionados (**Tabla 10**). No se observaron diferencias significativas en ninguna de las probables mutaciones con impacto en capacidad replicativa. Sin embargo, el impacto de estas mutaciones en la capacidad replicativa viral no puede ser descartado y se requerirá incrementar el número de virus con valores de capacidad replicativa asociados para confirmar estas observaciones.

Tabla 10. Codones bajo presión selectiva con posibilidad de impacto en la capacidad replicativa.

Codón*	Mutación	Valor de p	Estadísticamente Significativo ( $p < 0.05$ )
12	E --> Q	0.1443	NO
69	Q --> K	0.3578	NO
84	T --> V	0.1789	NO
230	H --> Q	0.1001	NO
234	I --> V	0.0570	NO
263	N --> S	0.0700	NO
297	R --> K	0.4703	NO
390	A --> N	0.2070	NO
445	T --> N	0.1020	NO
470	S --> N	0.4077	NO
547	S --> L	0.2529	NO

\*Codones con posibilidad de impacto en la capacidad replicativa y sus cambios de aminoácidos. Los 11 codones fueron determinados bajo presión selectiva positiva con el servidor en línea DataMonkey.

Con respecto a los datos del modelo PDN, se analizó el posible impacto de las 12 posiciones no reportadas en literatura y no encontradas por DataMonkey en la capacidad replicativa viral (**Tabla 11**). Las posiciones 35, 120, 128, 168 y 335 no mostraron ninguna relación específica con la capacidad replicativa pues se presentaron en prácticamente todas las muestras. Las posiciones 215 y 385 tampoco mostraron relación específica con la capacidad replicativa pues no presentaron frecuencias distintas en los virus con alta y baja capacidad replicativa. Interesantemente, las posiciones 18, 130, 203, 378 y 390, se presentaron preferencialmente en los virus con valores mayores o menores de capacidad replicativa, por lo que podrían representar mutaciones compensatorias o de escape (**Fig. 28**). Sin embargo, al aplicar la prueba exacta de Fisher para buscar comprobar la significancia de las asociaciones sólo el codón 390 fue estadísticamente significativo (**Tabla 12**).

Tabla 11. Relación de codones bajo presión selectiva (PDN) y capacidad replicativa.

Fitness	Codón	18	35	120	128	130	168	203	215	335	378	385	390
		HXB2	K	V	A	V	Q	V	E	V	K	M	N
	B023804	R	V	A	V	Q	V	D	L	R	L	G	V
0.4671184656	#1140	K	V	A	V	Q	V	E	L	K	M	N	V
0.4301061382	#1087	K	V	A	V	Q	V	E	L	K	M	N	V
0.5148693952	#1056	K	V	A	V	Q	V	E	L	K	M	N	I
0.6704015088	#1703	K	V	A	V	Q	V	E	L	K	M	N	V
0.6932408300	#0623	K	V	A	V	Q	V	E	L	K	M	F	V
0.6971363963	#0425	K	V	A	V	Q	V	E	L	K	M	N	V
0.8022170177	#CT21	K	V	?	V	Q	V	E	L	K	M	N	V
0.8372045579	#1001	K	V	-	T	Q	V	E	L	K	M	N	V
0.8377914874	#0769	K	V	A	?	Q	V	E	L	K	M	N	I
0.8495395638	#CT24	K	V	A	V	Q	V	E	M	K	M	N	V
0.8582883504	#1062	K	V	A	V	Q	V	E	L	K	M	N	V
0.8608011924	#1054	K	V	A	V	Q	V	E	L	K	M	N	V
0.8733092339	#0967	K	V	A	V	Q	V	E	?	K	M	N	I
0.8785476201	#1239	K	V	A	V	Q	V	E	L	K	M	N	V
0.8809860356	#1084	K	V	A	V	Q	V	E	L	K	M	N	V
0.8897875128	#1634	K	V	A	V	Q	V	E	L	K	M	N	V
0.8901322678	#1209	?	V	A	V	Q	V	E	L	K	M	N	V
0.8924621015	#0538	K	V	A	V	Q	V	E	L	K	M	N	V
0.9165017397	#0949	K	V	A	V	Q	V	E	L	K	M	N	I
0.9283394952	#1031	K	V	A	V	Q	V	?	?	K	M	N	I
0.9339680734	#1149	K	V	A	V	Q	V	E	?	K	I	N	?
0.9339851983	#1698	K	V	A	N	Q	V	E	V	K	M	N	V
0.9391814853	#1143	K	V	A	V	Q	V	E	L	K	M	N	V
0.9576282516	#0807	K	V	A	V	Q	V	E	L	K	M	N	V
0.9588901387	#0760	K	V	A	V	Q	V	E	V	K	M	N	I
0.9821225058	#1469	K	V	A	V	Q	V	E	M	K	M	N	V
0.9903811462	#0412	K	M	A	V	Q	V	E	L	K	M	N	V
1.0076805055	#0355	K	V	A	V	Q	V	D	L	K	M	N	I
1.0211356146	#1714	K	V	A	V	Q	V	E	L	K	M	N	V
1.0222363653	#0322	K	V	A	V	Q	V	E	L	K	M	N	V
1.0539814038	#1374	K	V	A	V	Q	V	E	L	K	M	N	V
1.0610429324	#1126	K	V	A	V	Q	V	E	L	K	M	N	V
1.0953969340	#1203	K	I	A	V	Q	V	E	L	K	M	N	V
1.0966997106	#0986	K	V	A	V	Q	V	E	L	K	M	N	I
1.1100056185	#1186	K	V	A	V	Q	V	E	L	K	M	N	I
1.1236561707	#0586	K	V	A	V	Q	V	E	L	K	M	N	P
1.1263936541	#1301	G	V	A	V	Q	V	E	L	K	M	N	V
1.1724236487	#1130	K	V	A	V	Q	V	E	L	K	M	N	I
1.1804355158	#1010	K	V	A	V	Q	V	E	L	K	M	N	V
1.1959532690	#1129	K	V	?	V	Q	V	E	?	K	M	N	I
1.1982917564	#1550	K	V	A	V	Q	V	E	L	K	M	N	I
1.2089477961	#0687	K	V	A	V	Q	V	E	M	K	M	N	I
1.2568451518	#0196	K	V	A	V	Q	V	E	L	K	M	N	I
1.2634865017	#0651	K	V	A	V	Q	V	E	?	K	M	N	V
1.2662943994	#1377	K	V	A	V	Q	V	E	L	K	M	N	I
1.2668813225	#0884	K	V	A	V	Q	V	E	L	K	M	N	I
1.2918248184	#1535	R	V	A	V	Q	V	E	L	K	M	N	V
1.3326585047	#0526	K	V	A	V	Q	V	E	L	K	M	N	V
1.3426531944	#0577	K	V	A	V	Q	V	E	V	K	M	N	I
1.3508598846	#0393	K	V	A	V	Q	V	E	L	K	M	N	V
1.4028194771	#0259	K	V	A	V	Q	V	E	L	K	M	N	I
1.4093949407	#0403	K	V	A	G	Q	V	E	L	K	M	N	I
1.4101487149	#0422	K	V	A	V	H	V	D	M	K	M	N	I
1.4154759400	#0390	K	V	A	V	Q	V	E	L	K	M	N	A
1.4400817593	#0413	K	V	A	V	Q	V	D	L	K	M	N	V
1.4772371712	#1640	K	V	A	V	H	V	E	V	K	M	N	V
1.4965384479	#0989	?	V	A	A	Q	?	E	L	K	M	N	I
1.4968904114	#0941	R	V	A	V	Q	V	E	L	K	M	F	I

Se muestran los virus ordenados en orden creciente de capacidad replicativa. En colores (según el aminoácido) se muestran las mutaciones importantes para la capacidad replicativa. Con una línea negra se separan los virus de alta y baja capacidad replicativa.

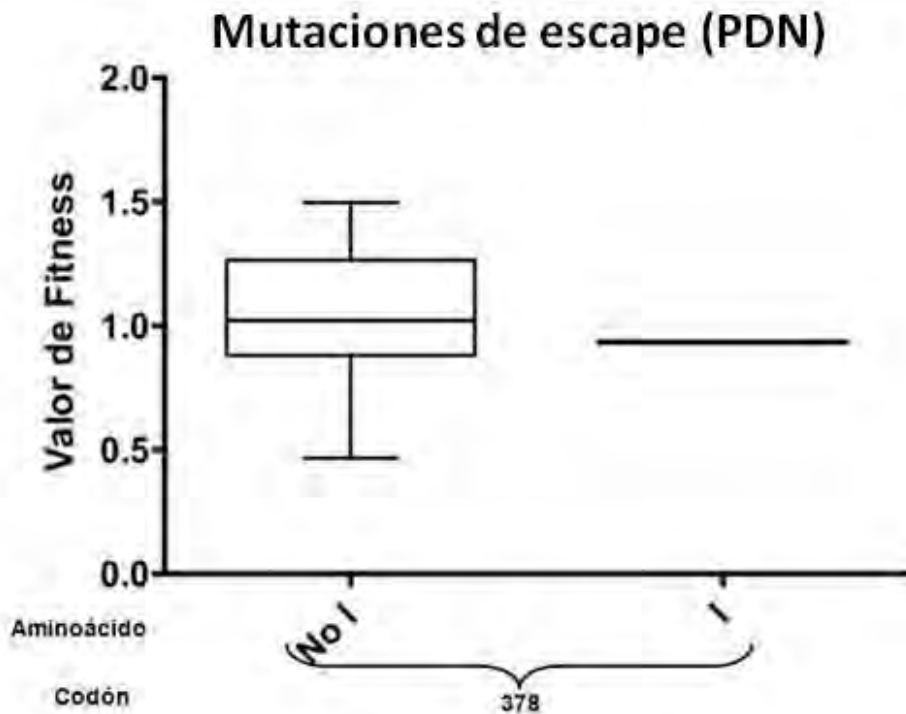
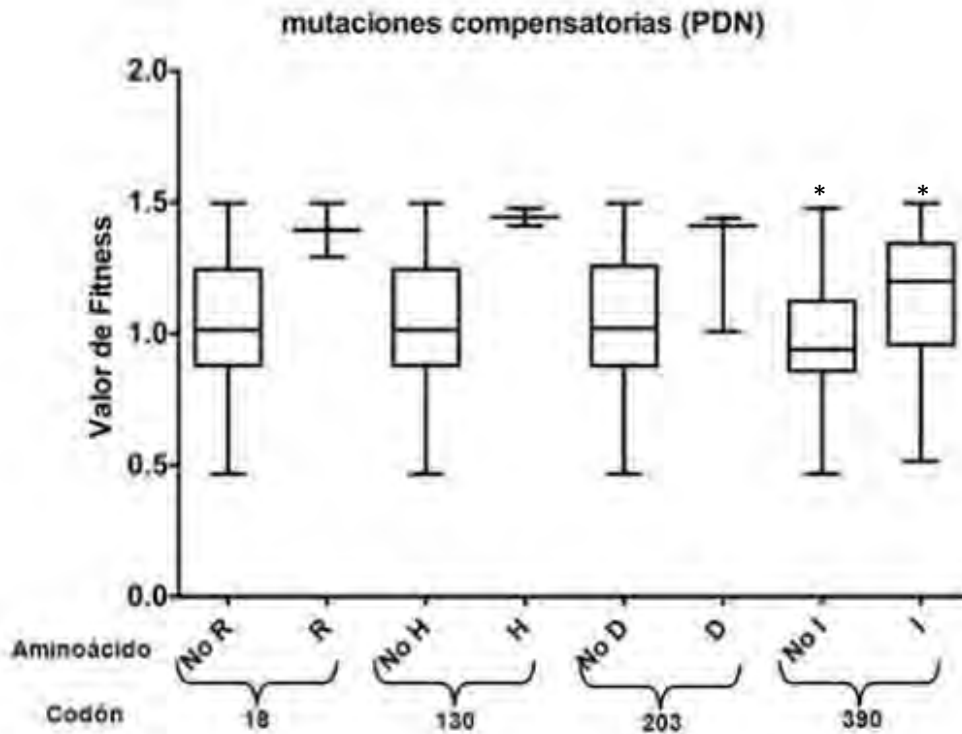


Figura 28. Gráfico de caja y bigotes de las mutaciones compensatorias y de escape encontradas por el análisis de Redes de dependencia Filogenética (PDN). Sólo la posición 390 fue estadísticamente representativa. En cada diagrama la caja sólida indica el rango entre los percentiles 25 y 75, la línea indica la mediana, los bigotes indican los valores máximos y mínimos no atípicos.

Tabla 12. Codones con asociación a HLA en la cohorte mexicana por el modelo PDN, no reportadas en la literatura, con posibilidad de impacto en la capacidad replicativa.

Codón*	Mutación	Valor de p	Estadísticamente Significativo (p < 0.05)
18	K --> R	0.4912	NO
35	V --> V	-	-
120	A --> A	-	-
128	V --> V	-	-
130	Q --> H	0.4912	NO
168	V --> V	-	-
203	E --> D	0.3816	NO
215	V --> M	-	-
335	K --> K	-	-
378	M --> I	1.0000	NO
385	N --> P	-	-
390	V --> I	0.0307	SI

\*Se muestran los codones con asociaciones a HLA por el modelo PDN, no reportados previamente en cohortes caucásicas, y sus respectivos cambios de aminoácidos. De de las 12 mutaciones observadas 7 no presentaron una tendencia de aparición preferente en muestras de alta o baja capacidad replicativa y por tanto no se calculó su p. Sólo la muestra 390 fue estadísticamente significativa.

## 7. Obtención de un mapa de evolución asociada a HLA del gen gag en VIH circulante en México.

Conjuntando los diversos análisis realizados, se logró obtener un mapa de escape inmunológico de *gag* para el VIH circulante en la cohorte mexicana, en el cual se incluyeron las posiciones encontradas en la literatura (población caucásica), las posiciones determinadas como bajo presión selectiva positiva por el servidor DataMonkey (200 secuencias mexicanas) y las posiciones asociadas a HLA con el modelo PDN (500 secuencias mexicanas). Se incluyeron algunos alelos HLA asociados a las posiciones, así como, algunos epítomos reportados previamente sobrelapados con posiciones de interés (Fig. 29).



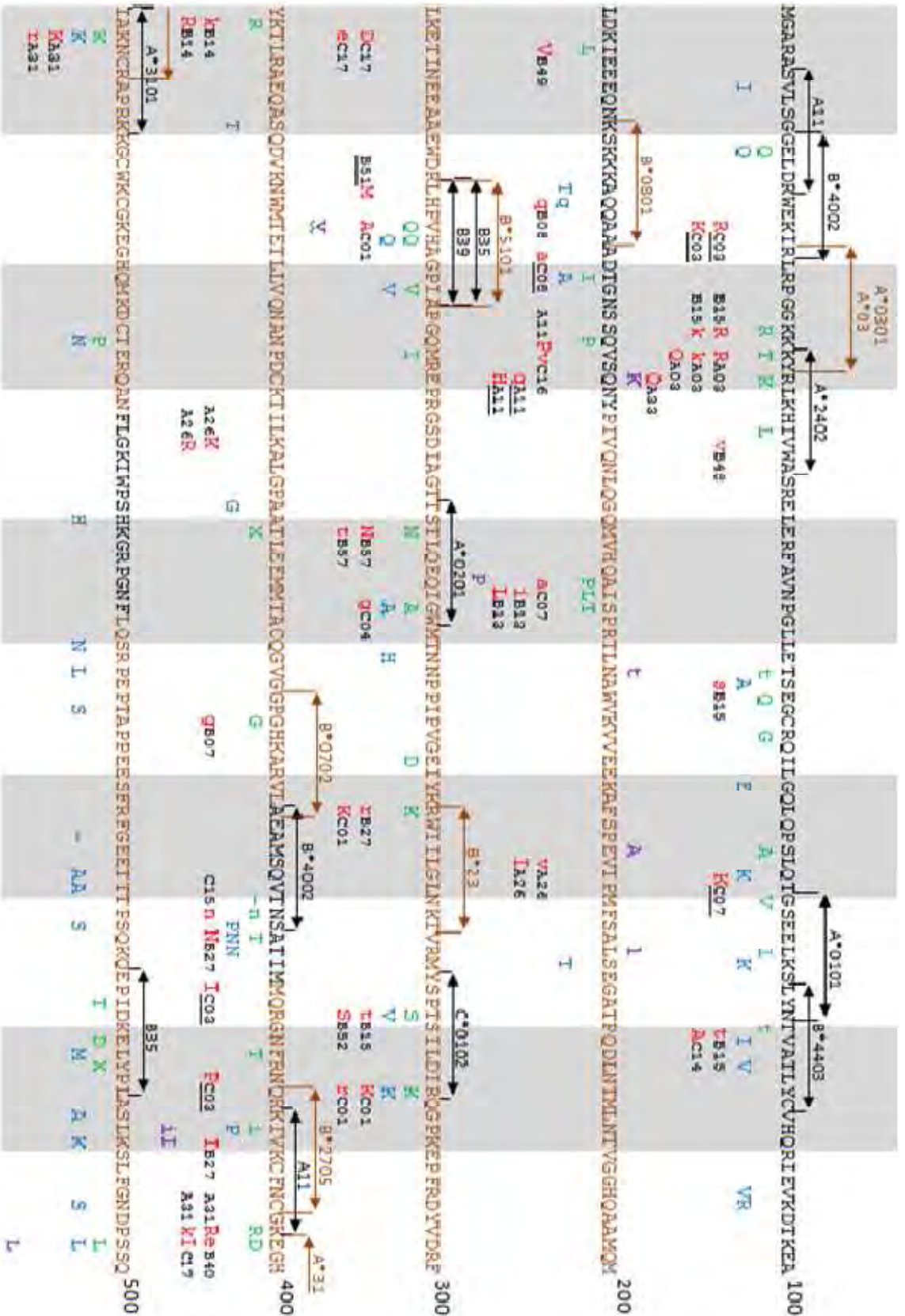


Figura 29. Mapa de escape inmunológico de *gag* en VIH circulante en México. Como referencia se muestra la secuencia consenso obtenida para las 200 muestras mexicanas [31]. Se muestran 100 aminoácidos por línea y las barras grises y blancas alternadas separan bloques de 10 aminoácidos. Se alternan letras negras y café en el consenso para identificar cada una de las zonas que codifican para las distintas proteínas de *gag* (p17, p24, p2, p7, p1 y p6). Los polimorfismos en verde son los reportados previamente en la literatura [7, 47], los azules son los determinados como bajo presión selectiva positiva por el servidor DataMonkey y los rojos son los asociados a HLA por el análisis de PDN [10]. Además, los polimorfismos marcados en morado y negritas son aquellos relacionados con efectos en la capacidad replicativa. Aunado a esto se muestran las relaciones polimorfismo-HLA determinados por PDN y se encuentran subrayados los alelos más frecuentes en la población mexicana. En letras mayúsculas se muestran las asociaciones que tienden a cambiar el aminoácido y en letras minúsculas aquellas que tienden a conservarlo. Finalmente, se muestran algunos epítotos confirmados experimentalmente, reportados previamente en la base de datos de los Álamos [31] y relacionados con alelos HLA frecuentes en la población mexicana (en negro) y algunos epítotos (café) obtenidos mediante la herramienta de predicción de epítotos de la base de datos de los Álamos.

## CAPÍTULO VI. DISCUSIÓN.

La distribución geográfica de los pacientes analizados permite observar que a pesar de que faltarían tomar en cuenta algunos estados, sobre todo en el sureste y norte del país, la cohorte analizada es altamente representativa de la población infectada por VIH en México, debido a que entre los estados seleccionados se encuentran aquellos que presentan más casos de pacientes infectados por VIH de acuerdo a los reportes oficiales (Distrito Federal, Estado de México, Veracruz y Jalisco) [12]. Con respecto a la proporción de género, la cohorte refleja lo observado previamente para la mayoría de las cohortes latinoamericanas, pues se observa que se mantiene una mayor proporción de hombres infectados con VIH, con respecto a las mujeres. Además la proporción reportada oficialmente [12], 72.8% hombres y 27.2% mujeres, es muy semejante a la encontrada en la cohorte de estudio, 75.5% varones y 24.5% mujeres. En cuanto a la edad se tiene un amplio rango que cubre desde adultos jóvenes hasta personas de la tercera edad, observándose que la media tiende a dirigirse hacia las personas jóvenes.

Los valores de carga viral y de linfocitos T CD4+ en conjunto, son útiles para determinar la etapa de la infección en la que se encuentra el paciente. Así mismo, estos parámetros permiten determinar cuándo es probable que se presenten enfermedades oportunistas, son útiles para el monitoreo del tratamiento antirretroviral y para estimar el riesgo de transmisión, incluyendo la materno-fetal. La cohorte en estudio estuvo enriquecida en pacientes en etapas tempranas de la infección con una mediana de conteo de linfocitos T CD4+ más alta que la esperada para la población mexicana en general, la cual se caracteriza por una detección tardía de la infección [12]. Este sesgo en la selección de pacientes en la cohorte del presente estudio se debió a que previamente se han reportado diferencias en la capacidad replicativa entre virus en etapas tempranas y tardías de la infección, con una ganancia de capacidad replicativa a lo largo del tiempo [7]. Así, para el presente estudio se seleccionaron preferentemente pacientes en etapas tempranas, con el objeto de estimar el impacto de la presencia de mutaciones seleccionadas por HLA en la capacidad replicativa viral en virus en etapas semejantes de la infección.

Dada la importancia de las proteínas Gag para la estructura y función de los viriones, se esperaba una baja variabilidad en esta región genética del virus, asociada a la conservación de la función de las proteínas de la región [20]. Sin embargo, al comparar árboles filogenéticos de los virus de la cohorte mexicana basados en el gen *gag* con árboles de los mismos virus basados en el gen *pol* se observó una mayor divergencia en las secuencias del gen *gag*, sugiriendo un mayor grado de evolución en esta región del virus. Es probable que la alta inmunogenicidad de *gag* contribuya a esta diversificación, con la inclusión de múltiples sitios bajo presión selectiva inmune en los virus circulantes en la región. Es interesante que solo el árbol filogenético basado en *gag* mostrara agrupamientos regionales característicos que podrían estar reflejando variaciones en la presión selectiva inmune a nivel regional, con distribuciones distintas de alelos HLA en diferentes regiones del país impactando diferencialmente la adaptación viral. Estos clusters representan subgrupos del virus que se organizan debido a la similitud particular que presentan algunas de las secuencias, las cuales bien podrían estar determinadas por presión selectiva por HLA [44, 45]. Alternativamente, posibles efectos fundadores en la diversificación regional del virus circulante en el país podrían estar reflejándose en los análisis filogenéticos realizados. En el caso de las 6 secuencias encontradas en el árbol, cuya cercanía era muy evidente, se observó que pertenecían al mismo estado y que podrían ser posibles pares de transmisión.

Los estudios de variabilidad del gen *gag* de VIH circulante en México mostraron resultados esperados. Las regiones p17 y p2p7p1p6 presentaron mayor variabilidad, teniendo una mayor cantidad de codones con una frecuencia de polimorfismos alta. Por el contrario, la zona correspondiente a p24 resultó ser la más conservada, al presentar una menor frecuencia de polimorfismos. Esto concuerda con lo reportado por Brumme, et al [8], acerca de la alta conservación de p24 y la mayor variación en p17 y p2p7p1p6. Esta observación fue consistente en el análisis de frecuencia de polimorfismos y el análisis de entropía de Shannon. La alta conservación de p24 podría explicarse debido a que, dada su importancia funcional y estructural para el virus, prácticamente cualquier variación en la esta zona podría afectar la capacidad replicativa del VIH. Aparentemente, la zonas

correspondientes a p17 y a p2p7p1p6 permitirían un poco más de flexibilidad en cuanto a la presencia de polimorfismos, manteniendo las funciones de las proteínas correspondientes.

Posteriormente, nos preguntamos si la variabilidad observada en Gag podría estar asociada a presión selectiva inmune. Por tanto, se procedió a realizar un análisis de presión selectiva por razones de sustituciones sinónimas y no sinónimas (dN/dS) en la cohorte mexicana. Al comparar la cantidad de zonas seleccionadas negativamente (301) y las seleccionadas positivamente (40) se observa una tendencia hacia la conservación del gen. Sin embargo, es notable la existencia de una gran cantidad de posiciones bajo presión selectiva positiva, que muy probablemente podrían estar asociadas a selección por HLA, confirmando la alta inmunogenicidad de Gag y la alta densidad de epítomos de CTLs encontrados en la región [31].

Como era esperado, p24 resultó ser la proteína de la región con menor número de posiciones bajo presión selectiva positiva, pero con mayor número de posiciones bajo presión selectiva negativa, comprobándose la tendencia a la conservación de la mayoría de las posiciones de la proteína y la poca tolerancia en variabilidad para asegurar la funcionalidad de la misma. Más aun, observaciones de mayor variabilidad en p2p7p1p6 y p17, concuerdan también con la presencia de una mayor proporción de posiciones bajo presión selectiva positiva.

Para evaluar la importancia de los genes HLA en la evolución del gen *gag* del VIH circulante en México se utilizaron los mapas de epítomos confirmados de la base de datos de los Álamos. Esto permitió observar que todas las posiciones bajo presión selectiva positiva podrían estar asociadas a selección por diversos alelos HLA. Es importante mencionar, que aunque se espera que los epítomos en *gag* reconocidos en población mexicana varíen con respecto a los reportados en otras poblaciones dadas las distribuciones de frecuencias HLA características de los mexicanos, la presencia de las posiciones bajo presión selectivas en zonas altamente inmunogénicas de las distintas proteínas Gag harían altamente factible su asociación con selección inmune.



Interesantemente, un 25% de las posiciones se asociaron a alelos HLA con altas frecuencias en la población mexicana, incrementándose la factibilidad de las asociaciones.

Como un acercamiento alternativo, se exploró la factibilidad de asociación de las posiciones bajo presión selectiva positiva a selección inmune, mediante comparación con posiciones en *gag* previamente asociadas con alelos HLA específicos. No hay que olvidar que los mapas de asociaciones HLA-polimorfismo de VIH han sido determinados en cohortes principalmente caucásicas y que sería factible encontrar asociaciones previamente no descritas en dichas poblaciones, al estudiar la evolución dependiente de HLA del VIH en contextos inmunogenéticos diferentes como el presentado en este estudio para la cohorte mexicana. Interesantemente, dentro de las posiciones bajo presión selectiva positiva se encontraron algunas previamente ya reportadas y que podrían representar rasgos comunes en la evolución del virus en diferentes poblaciones. Sin embargo, también se observaron posiciones previamente no reportadas, que podrían representar nuevas asociaciones en el contexto inmunogenético de los mexicanos, que será necesario estudiar más a fondo.

Para tener un punto de referencia más específico de la población mexicana, se compararon las posiciones bajo presión selectiva positiva con posiciones con asociaciones a alelos HLA identificadas mediante el modelo PDN. Al hacer las comparaciones con el análisis de PDN se observó que este método podría ser más sensible para discriminar asociaciones reales, al corregir para los principales factores que podrían confundir la identificación de asociaciones HLA-polimorfismo de VIH: efectos de la filogenia de VIH, desequilibrio de ligamiento de alelos HLA y covariación de codones de VIH, evitando así falsos positivos. Interesantemente, solo 3 de las posiciones asociadas a presión selectiva inmune por el modelo PDN coincidieron con el análisis de DataMonkey y se encontraron 12 posiciones adicionales no reportadas previamente. Se requieren estudios con mayor número de pacientes para establecer si la falta de asociación de estas posiciones se debe a una baja potencia del análisis por el modelo PDN para identificar asociaciones con HLA en ella; o alternativamente a que las posiciones representen asociaciones indirectas con HLA, presiones selectivas alternas o simples falsos positivos.

Conjuntando las observaciones anteriores, hemos mostrado evidencia de un impacto importante de la respuesta inmune restringida por HLA en la determinación de la variabilidad del gen *gag* del VIH circulante en México. Parte de la variabilidad asociada a HLA de los virus mexicanos se comparte muy probablemente con la de virus circulantes en poblaciones caucásicas; sin embargo, nuestros estudios sugieren fuertemente también la existencia de asociaciones HLA-polimorfismo de VIH específicas de la población mexicana.

Una vez comprobado el impacto de la presión selectiva inmune sobre la evolución del VIH en México, se exploró la influencia de posibles mutaciones seleccionadas por HLA en capacidad replicativa del virus en el contexto inmunogenético de la población mexicana. Aunque las diferencias no fueron significativas en la mayoría de los casos, la clara tendencia de agrupación de secuencias ya sean de alta o baja capacidad replicativa mostradas por el árbol filogenético (fig. 25) evidencia la presencia de cambios particulares en la secuencias de estos virus que están afectando la capacidad replicativa. Aunado a esto, se observaron polimorfismos específicos con tendencias interesantes en la influencia sobre la capacidad replicativa viral. Las posiciones analizadas incluyeron codones identificados como bajo presión selectiva positiva, posiciones previamente reportadas como asociadas a cambios en capacidad replicativa en poblaciones caucásicas y posiciones asociadas a HLA en la cohorte mexicana por el modelo PDN. Se requiere incrementar el número de pacientes con mediciones de capacidad replicativa para confirmar las observaciones de este estudio.

Un dato interesante del estudio fue la alta frecuencia de mutaciones compensatorias en la zona de *gag* correspondiente a p24, esto es de esperarse dadas las restricciones funcionales de esta proteína, ya que pequeñas variaciones en su secuencia podrían alterar el ensamblaje de las subunidades de la cápside viral [30]. Así, es lógico que mutaciones de escape con costos en la capacidad replicativa tiendan a compensarse en esta zona, específicamente. Como era esperado, los virus con valores más altos de capacidad replicativa tendieron a presentar mayores frecuencias de mutaciones identificadas previamente como compensatorias.

Al conjuntar todos los resultados de los diversos análisis realizados en las muestras se pudo obtener un mapa del gen *gag* que resume las asociaciones con presión selectiva positiva, con presión selectiva negativa y con selección por HLA en el contexto inmunogenético específico de la población mexicana. En este mapa se muestra que hay codones que presentan variantes que se comparten con otras poblaciones inmunogenéticamente distintas y que hay codones con variantes que se seleccionan por alelos específicos de nuestra población. Cabe mencionar, que aunque será necesario aumentar el número de pacientes participantes para incrementar la potencia de encontrar asociaciones HLA-polimorfismo de VIH en población mexicana, el presente estudio ya muestra características únicas de la evolución del gen *gag* del VIH en México y siendo el primer trabajo que trata de describir y explicar la variabilidad de este gen en virus circulantes en el país, guiará estudios futuros para comprender mejor la adaptación del VIH a nivel poblacional. En específico, será importante continuar con el análisis de posiciones con asociaciones de interés, incluyendo los codones identificados como bajo presión selectiva y como asociados a HLA simultáneamente por el modelo PDN, confirmando su relevancia mediante dos metodologías distintas; así como los codones encontrados por el análisis de PDN que no han sido reportados, que podrían representar características únicas de adaptación del VIH en nuestra población.

Aunque es necesario también aumentar el número de virus cuyos valores de capacidad replicativa relativa estén disponibles, el presente estudio sugiere un impacto de la selección por HLA en cambios en la capacidad replicativa viral en México. La identificación de posiciones con impacto en capacidad replicativa como V390I, que previamente no se han identificado en otras poblaciones, apoya la hipótesis de que la selección por alelos HLA característicos en nuestra población puede impactar de manera diferencial la capacidad replicativa del virus en pacientes mexicanos, con respecto a otras poblaciones.

## CAPÍTULO VII. CONCLUSIÓN Y PERSPECTIVAS.

### 1. Conclusión.

A través de diversas metodologías se lograron obtener las secuencias del gen *gag* de VIH-I subtipo B presentes en 200 pacientes, formando una cohorte representativa de la población mexicana y determinándose las relaciones filogenéticas existentes entre ellos. Los polimorfismos que se determinaron permitieron demostrar que la alta variabilidad presentada, sobre todo en las zonas de p17 y p2p7p1p6, fue en gran parte influenciada debido a la presión selectiva ejercida por el sistema inmune del hospedero a través de los alelos HLA-I; mientras que p24 tiende a conservarse debido a su importancia en la correcta formación de la estructura del virus.

Con los resultados obtenidos y la comparación con la teoría se pudo corroborar que las mutaciones en p24 tuvieron un mayor costo en la capacidad replicativa y que la aparición de mutaciones compensatorias permitirá el restablecimiento de la capacidad replicativa. Sin embargo, la relación de nuevas mutaciones con su efecto en la disminución de la capacidad replicativa no mostró una relación contundente pero si mostró una tendencia que podría convertirse en estadísticamente significativa si se aumenta el número de muestras analizadas.

Al ser el presente estudio el primero que realizó un análisis evolutivo en *gag* en la población mexicana permitió comparar lo observado en estudios con otras poblaciones, determinándose que la variabilidad específica que presentan el grupo de secuencias de cada población está relacionada con el contexto inmunogenético de la población en sí misma. Compartiéndose ciertas mutaciones y relaciones HLA-polimorfismo, pero teniendo otras características propias de la población de la que se trate.

Finalmente, al conjuntar todos los resultados obtenidos se logró obtener el primer mapa de escape inmunológico específico de la población mexicana, considerando mutaciones previamente compartidas con otras poblaciones y aquellas encontradas específicamente para nuestra población.

## 2. Perspectivas.

A pesar de que el presente trabajo demuestra una tendencia en la aparición de ciertas mutaciones y de selecciones por HLA-I sobre el VIH específica de la población mexicana, esta serie de análisis deberá ser llevada a cabo con una cantidad de muestras que sea mucho mayor y de la misma forma, deberán aplicarse más análisis estadísticos. Como se muestra en el estudio realizado por Carlson et al [10], la capacidad de predicción de los análisis evolutivos aumenta conforme se tienen en cuenta más factores que puedan afectar los mismos tales como, efectos de la filogenia de VIH, desequilibrio de ligamiento de alelos HLA, covariación de codones, etcétera. Así mismo, esta capacidad predictiva se verá enriquecida también al aumentar el número de miembros que forman la cohorte, pues esto dará una muestra más representativa de la población estudiada. Considerando lo anterior, los resultados que fueran obtenidos posteriormente podrían extrapolarse a la población de una manera más representativa, asegurando una significancia estadística de cada uno de los análisis realizados.

Estos estudios se deberán realizar en más poblaciones para cotejar de qué manera el contexto inmunogenético de cada una de ellas está influenciando al VIH y así poder hacer comparaciones entre ellas. De esta manera, se podrán mejorar las estrategias de combate contra la infección por este virus. El continuar con estos estudios en diversas poblaciones podría ayudar a encontrar blancos específicos que servirían como posibles zonas diana para ser atacadas por fármacos en el tratamiento o por una posible vacuna para la prevención de la enfermedad.

## CAPÍTULO VIII. BIBLIOGRAFÍA.

1. Abbas, A.K.; Lichtman, A.H.; Pillai, S. (2011). *Cellular and Molecular Immunology E-Book: With STUDENT CONSULT Online Access*. Elsevier, Cap. 6. [Disponible en: <http://books.google.com.mx>].
2. Allen, T.M.; Altfeld, M.; Geer, S.C.; Kalife, E.T.; Moore, C.; O'Sullivan, K.M.; DeSouza, I.; Feeney, M.E.; Eldridge, R.L.; Maier, E.L.; Kaufmann, D.E.; Lahaie, M.P.; Reyor, L.; Tanzi, G.; Johnston, M.N.; Brander, C.; Draenert, R.; Rockstroh, J.K.; Jessen, H.; Rosenerg, E.S.; Mallal, S.A.; Walker, B.D. (2005). *Selective escape from CD8<sup>+</sup> T-cell responses represents a mayor driving force of Human Immunodeficiency Virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution*. J. Virol., 79 (21): 13239-13249.
3. Allen, T.M.; Altfeld, M.; Yu, X.G.; O'Sullivan, K.M.; Lichterfeld, M.; Le Gall, S.; John, M.; Mothe, B.R.; Lee, P.K.; Kalife, E.T.; Cohen, D.E.; Freedberg, K.A.; Strick, D.A.; Johnston, M.N.; Sette, A.; Rosenberg, E.S.; Mallal, S.A.; Goulder, P.J.R., Brander, C.; Walker, B.D. ( 2004). *Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in Human Immunodeficiency Virus type 1 infection*. J. Virol., 78 (13): 7069-7078.
4. AVERT (2011). *History of AIDS up to 1986*. UK: Autor [Disponible en: <http://www.avert.org/aids-history-86.htm>].
5. Ávila-Ríos, S.; Ormsby, C.E.; Carlson, J.M.; Valenzuela-Ponce, H.; Blanco-Heredia, J.; Garrido-Rodríguez, D.; Garcia-Morales, C.; John, M.; Espinosa, E.; Reyes-Terán, G. (2009). *Unique features of HLA-mediated HIV evolution in a Mexican cohort: a comparative study*. Retrovirology, 6: 72.
6. Ávila-Ríos, S.; Reyes-Terán, G.; Espinosa, E. (2007). *Cornering HIV: taking advantage of interactions between selective pressures*. Med. Hypoth., 69: 422-431.
7. Brockman, M.A.; Brumme, Z.L.; Brumme, C.J.; Miura, T.; Sela, J.; Rosato, P.C.; Kadie, C.M.; Carlson, J.M.; Markle, T.J.; Streeck, H.; Kelleher, A.D.; Markowitz, M.; Jessen, H.; Rosenberg, E.; Altfeld, M.; Harrigan, P.R.; Heckerman, D.; Walker, B.D.; Allen, T.M. (2010). *Early selection in Gag by protective HLA alleles contributes to*

- reduced HIV-1 replication capacity that may be largely compensated in chronic infection.* J. Virol., 84 (22): 11937-11949.
8. Brumme, Z. L.; Tao, I.; Szeto, S.; Brumme, C. J.; Carlson, J.; Chan, D.; Kadie, C.; Frahm, N.; Brander, C.; Walker, B.; Heckerman, D.; Harrigan, P. R. (2008). *Human leukocyte antigen-specific polymorphisms in HIV-1 Gag and their association with viral load in chronic untreated infection.* AIDS, 22 (11): 1277-1286.
  9. Burwitz, B.J.; Sacha, J.B.; Reed, J.S.; Newman, L.P.; Norante, F.A.; Bimber, B.N.; Wilson, N.A.; Watkins, D.I.; O'Connor D.H. (2011). *Pyrosequencing reveals restricted patterns of CD8+ T cell escape-associated compensatory mutations in simian immunodeficiency virus.* J Virol., 85(24):13088-96.
  10. Carlson, J.M., Brumme, Z.L., Rousseau, C.M., Brumme, C.J., Matthews, P., Kadie, C., Mullins, J.I., Walker, B.D., Harrigan, P.R., Goulder, P.J.R., Heckerman, D. (2008). *Phylogenetic Dependency Networks: Inferring patterns of CTL escape and codon covariation in HIV-1 Gag.* PLoS Comp. Biol., 4 (11): 1-23.
  11. Centro de Investigación en Enfermedades Infecciosas (2011). 30 años del VIH-SIDA perspectivas desde México. México: Autor. 1-146.
  12. Centro Nacional para la Prevención y Control del VIH SIDA (2011). *El VIH/SIDA en México 2011 Numeralia epidemiológica.* México: Autor. [Disponible en: [http://www.censida.salud.gob.mx/descargas/2011/NUMERALIA\\_SIDA\\_2011.pdf](http://www.censida.salud.gob.mx/descargas/2011/NUMERALIA_SIDA_2011.pdf)].
  13. Centro Nacional para la Prevención y Control del VIH SIDA (2011). *Vigilancia Epidemiológica de casos de VIH/SIDA en México.* México: Autor. [Disponible en: [http://www.censida.salud.gob.mx/descargas/2011/vih\\_sidapub31dic2011.pdf](http://www.censida.salud.gob.mx/descargas/2011/vih_sidapub31dic2011.pdf)].
  14. Córdoba Villalobos, J. A.; P.d.L.R.S., Valdespino, J. L. (2009). *25 años de SIDA en México. Logros, desaciertos y retos.* México: Instituto Nacional de Salud Pública. (2ª Ed.). 135-191.
  15. Crawford, H.; Prado, J.G.; Leslie, A.; Hué, S.; Honeyborne, I.; Reddy, S.; Stok, M.; Mncube, Z.; Brander, C.; Rousseau, C.; Mullins, J.I.; Kaslow, R.; Goepfert, P.; Allen, S.; Hunter, E.; Mulenga, J.; Kiepiela, P.; Walker, B.D.; Goulder, P.J.R. (2007). *Compensatory mutation partially restores fitness and delays reversion of escape*

- mutations within the immunodominant HLA-B\*5703-restricted Gag epitope in chronic Human Immunodeficiency Virus type 1 infection. J. Virol., 81 (15): 8346.*
16. Espinosa, E. (2006). *Encontrando las huellas de la selección en el VIH. Rev. Latinoam. Microbiol., 48 (2): 84-89.*
  17. Fauci, A.S; Kasper, D.L.; Braunvald, E.; Hauser S.L.; Longo, D.L.; Jameson, J.L.; Loscalzo, J. (2006). *Harrison principios de medicina interna. (Edición 17).* [Disponible en: [www.harrisonmedicina.com](http://www.harrisonmedicina.com)].
  18. Frahm, N.; Korber, B.T.; Adams, C.M.; Szinger, J.J.; Draenert, R.; Addo, M.M.; Feeney, M.E.; Yusim, K.; Sango, K.; Brown, N.V.; SenGupta, D.; Piechocka-Trocha, A.; Simonis, T.; Marincola, F.M.; Wurcel, A.G.; Stone, D.R.; Russell, C.J.; Adolf, P.; Cohen, D.; Roach, T.; StJohn, A.; Khatri, A.; Davis, K.; Mullins, J.; Goulder, P.J.R.; Walker, B.D.; Brander, C. (2004). *Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in Human Immunodeficiency Virus across multiple ethnicities. J. Virol., 78 (5): 2187-2200.*
  19. Frankel AD, Young JA. (1998). *HIV-1: fifteen proteins and an RNA. Annu. Rev. Biochem., 67: 1-25.*
  20. Freed, E.O. (1998). HIV-1 Gag proteins: Diverse functions in the virus life cycle. *Virology, 251: 1-15.*
  21. Freed, E.O.; Martin, M.A. (2007). *HIVs and their replication. Philadelphia: Krieger DM, Vol. II. (5a Ed.) 2108-2183.*
  22. Gómez-Carrillo, M.; Salomon, H.; Pand, M.A.; Kijak, G.; Avila, M.M. (2001). *Distribución de subtipos y recombinantes del HIV situación en la Argentina. MEDICINA (Buenos Aires), 61: 881-889.*
  23. Goulder, P.J.R.; Watkins, D.I. (2008). *Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat. Rev. Immunol., 8(8):619-30.*
  24. InfoMed (2011). Resumen mundial de la epidemia de VIH/sida 2010. [Disponible en: <http://articulos.sld.cu/sida/?cat=106>].
  25. Kawashima, Y.; Pfafferott, K.; Frater, J.; Matthews, P.; Payne, R.; Addo, M.; Gatanaga, H.; Fujiwara, M.; Hachiya, A.; Koizumi, H.; Kuse, N.; Oka, S.; Duda, A.; Prendergast, A.; Crawford, H.; Leslie, A.; Brumme, z.; Brumme, C.; Allen, T.;



- Brander, C.; Kaslow, R.; Tang, J.; Hunter, E.; Allen, S.; Mulenga, J.; Branch, S.; Roach, T.; John, M.; Mallal, S.; Ogwu, A.; Shapiro, R.; Prado, J.G.; Fidler, S.; Weber, J.; Pybus, O.G.; Klenerman, P.; Ndung'u, T.; Phillips, R.; Heckerman, D.; Harrigan, P.R.; Walker, B.D.; Takiguchi, M.; Goulder, P. (2009). *Adaption of HIV-1 to human leukocyte antigen class I*. *Nature*, 458: 641-645.
26. Kiepiela, P.; Leslie, A.J.; Honeyborne, I.; Ramduth, D.; Thobakgale, C.; Chetty, S.; Rathnavalu, P.; Moore, C.; Pfafferott, K.J.; Hilton, L.; Zimbwa, P.; Moore, S.; Allen, T.; Brander, C.; Addo, M.M.; Altfeld, M.; James, I.; Mallal, S.; Bunce, M.; Barber, L.D.; Szinger, J.; Day, C.; Klenerman, P.; Mullins, J.; Korber, B.; Coovadia, H.M.; Walker, B.D.; Goulder, P.J.R. (2004). *Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA*. *Nature*, 432: 769-775.
27. Kiepiela, P.; Ngumbela, K.; Thobakgale, C.; Ramdduth, D.; Honeyborne, I.; Moodley, E.; Reddy, S.; Pierres, C.; Mncube, Z.; Mkhwanazi, N.; Bishop, K.; Stok, M.; Nair, K.; Khan N. Crawford, H.; Payne, R.; Leslie, A.; Prado, J.; Prendergast, A.; Frater, J.; McCarthy, N.; Brander, C.; Learn, G.H.; Nickle, D.; Rousseau, C.; Coovadia, H.; Mullins, J.I.; Heckerman, D.; Walker, B.D.; Goulder P. (2007). *CD8<sup>+</sup> T-cell responses to different HIV proteins have discordant associations with viral load*. *Nat. Med.*, 13 (1): 46-52.
28. Kuritzkes, D.R.; Walker, B.D. (2007). *HIV-1: pathogenesis, clinical manifestations and treatment*. Philadelphia: Knipe DM, Vol. II. (5a Ed.) 2187-2214.
29. Leslie, A.; Kavanagh, D.; Honeyborne, I.; Pfafferott, K.; Edwards, C.; Pillay, T.; Hilton, L., Thobakgale, C.; Ramduth, D.; Draenert, R.; Le Gall, S.; Luzzi, G.; Edwards, A.; Brander, C.; Sewell, A.K.; Moore, S.; Mullins, J.; Moore, C.; Mallal, S.; Bhardawj, N.; Yusim, K.; Phillips, R.; Klenerman, P.; Korber, B.; Kiepiela, P.; Walker, B.; Goulder, P. (2005). *Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA*. *J. Exp. Med.*, 201 (6): 891-902.
30. Leslie, A.; Matthews, P.C.; Listgarten, J.; Carlson, J.M.; Kadie, C.; Ndung'u, T.; Brander, C.; Coovadia, H.; Walker, B.D.; Heckerman, D.; Goulder, P.J.R. (2010).

- Additive contribution of HLA class I alleles in the immune control of HIV-1 infection.* J. Virol., 84 (19): 9879-9888.
31. Los Álamos Base de Datos de Inmunología (2012). [Disponible en <http://www.hiv.lanl.gov/content/index>].
  32. Martínez-Picado, J.; Prado, J.G.; Fry, E.E.; Pfafferott, K.; Leslie, A.; Chetty, S.; Thobakgale, C.; Honeyborne, I.; Crawford, H.; Matthews, P.; Pillay, T.; Rousseau, C.; Mullins, J.I.; Brander, C.; Walker, B.D.; Stuart, D.I.; Kiepiela, P.; Goulder, P. (2006). *Fitness cost of escape mutations in p24 Gag in association with control of Human Immunodeficiency Virus type 1.* J. Virol., 80 (7): 3616-3623.
  33. McCutchan, F.E., (2003). *Global distribution of HIV-1 recombinants and subtypes.* IAVA report. [Disponible en: [www.iavireport.org](http://www.iavireport.org)].
  34. Morikawa, Y. (2003). *HIV capsid assembly.* Current HIV Research, 1:1-14.
  35. Mudd, P.A.; Ericson, A.J.; Walsh, A.D.; León, E.J.; Wilson, N.A.; Maness, N.J.; Friedrich, T.C.; Watkins, D.I. (2011). *CD8+ T cell escape mutations in simian immunodeficiency virus SIVmac239 cause fitness defects in vivo and many revert after transmission.* J Virol., 85(23):12804-10.
  36. Newman, J.L.; Butcher, E.W.; Patel, D.T.; Ylenko, Y.M.; Summers, M.F. (2004). *Flexibility in the P2 domain of the HIV-1 Gag polyprotein.* Protein Sci., 13: 2101-2107.
  37. Noviello, C.M.; López, C.S.; Kukull, B.; McNett, H.; Still, A.; Eccles J.; Sloan, R.; Narklis, E. (2011). *Second-site compensatory mutations of HIV-1 capsid mutations.* J. Virol., 85 (10):4730-4738.
  38. Overbaugh, J.; Morris, L. (2012). *The Antibody Response against HIV-1.* Cold Spring Harbor Perspectives in Medicine, [Disponible en: [www.perspectivesinmedicine.org](http://www.perspectivesinmedicine.org)].
  39. Piana, S.; Carloni, P.; Rothlisberger, U. (2002). *Drug resistance in HIV-1 protease: flexibility-assisted mechanism of compensatory mutations.* Protein Sci., 11: 2393-2402.

40. Roshni, M.; Sarvjeet, S.; Ashok, K. (2003). *The cytotoxic T lymphocyte (CTL) response to tumour cells*. Expert Rev. Mol. Med., 5: 28. [Disponibile en: <http://www.expertreviews.org/>].
41. Srinivasan, A.; Ayyavoo, V.; Mahalingam, S.; Kannan, A.; Boyd, A.; Datta, D.; Kalyanaraman, V.S.; Cristillo, A.; Collman, R.G.; Morellet, N.; Sawaya, B.E.; Murali, R. (2008). *A comprehensive analysis of the naturally occurring polymorphisms in HIV-1 Vpr: potential impact on CTL epitopes*. Virol. J., 5: 99.
42. The Kaiser Family Foundation (2008). *The HIV/AIDS Epidemic in Mexico*. USA: Autor. [Disponibile en: <http://www.kff.org/hivaids/upload/7797.pdf>].
43. WHO (2012). *Antiretroviral therapy*. USA: Autor [Disponibile en: <http://www.who.int/hiv/topics/treatment/en/index.html>].
44. Woese, C. R. (1998). *The Universal Ancestor*. Proc. Nat. Acad. Sci., 95: 6854-6859.
45. Woese, C. R. (2002). *On the evolution of cells*. Proc. Nat. Acad. Sci., 99: 8742-8747.
46. Wright, J.K.; Brumme, Z.L.; Carlson, J.M.; Heckerman, D.; Kadie, C.M.; Brumme, C.J.; Wang, B.; Losina, E.; Miura, T.; Chonco, F.; Stok, M.; Mncube, Z.; Bishop, K.; Goulder, P.J.R.; Walker, B.D.; Brockman, M.A.; Ndung'u, T. (2010). *Gag-protease mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters*. J. Virol., 84 (20): 10820-10831.
47. Zabrina, B.; Mina, J.; Jonathan, C.; Chanson J.; Dennison, C.; Mark, A.; Luke, C.; Iris, T.; Sharon, S.; Pamela, R.; Jennifer, S.; Carl, M.; Nicole, F.; Christian, B.; David, W.; Sharon, A.; Richard, H.; Bruce, D.; Richard, H.; David, H.; Simon, M. (2009). *HLA-Associated Immune Escape Pathways in HIV-1 Subtype B Gag, Pol and Nef Proteins*. PLoS ONE, 4 (8): e6687.