



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

ESCUELA DE BIOLOGÍA

**Evolución asociada al antígeno leucocitario humano del
gen *gag* del virus de inmunodeficiencia humana circulante
en Centroamérica**

Tesis para obtener el título de

BIÓLOGA

PRESENTA:

Sánchez León Anahi Tania

TUTOR: Dr. Santiago Ávila Ríos

Mayo 2016



¿Qué es el futuro? ¿Qué es el pasado? ¿Qué somos? ¿Cuál es el fluido mágico que nos rodea y oculta las cosas que más necesitamos saber? Vivimos y morimos en medio de maravillas

Napoleón Bonaparte.

AGRADECIMIENTOS

A mis padres que siempre han sido mi mayor motivación para seguir adelante. Por hacer innumerable cantidad de sacrificios para que jamás me faltara nada. Por consentirme y brindarme la oportunidad de seguir estudiando, y transmitirme su fortaleza. Han sido, son y serán siempre un gran motor en mi vida.

A mi hermana que me enseñó todo lo que no se debe hacer, a ella que siempre se ha preocupado por mi y porque sin sus pasos no tendría dirección.

A mi familia por disculpar mis ausencias físicas y mentales.

A mis amigos que a pesar de haber optado por caminos y objetivos distintos siempre han estado para hacerme reír, para compartir conmigo viajes, fiestas, juegos, enojos, tristezas y nuevas experiencias que me han formado como persona. Tantos momentos vividos y tan poco espacio para escribirlos.

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 a la Dr. Rosalina Reyes un gran apoyo en estos años de licenciatura.

Al Dr. Santiago Ávila Ríos quien me incluyó en este gran proyecto y ha sido un excelente tutor, y además es una excelente persona.

En especial gracias a Maribel Soto Nava, quien fue un gran apoyo en este proyecto y que sin su paciencia y dedicación no habría podido aprender todo de una forma tan fácil.

En especial gracias al CIENI por abrirme sus puertas y permitir que no sólo yo sino más estudiantes sigan aprendiendo; a todas las personas que me encontré en sus instalaciones durante el tiempo que duro esta tesis, por compartir sus conocimientos conmigo y por tener la paciencia de enseñarme tantas cosas.

A la Escuela de Biología y a la BUAP por abrirme sus puertas y permitirme ser de los privilegiados que pueden terminar una carrera en una de las mejores universidades del país. Gracias a la vida por cada una de las personas que me he topado en el camino, por el conocimiento compartido y cada una de las experiencias que me he vivido, pues cada una ha influido para que yo llegara a donde estoy ahora.

Este trabajo fue financiado por la Comisión de Equidad y Género de las Legislaturas LX-LXI y la Comisión de Igualdad de Género de la Legislatura LXII de la H. Cámara de Diputados de la República Mexicana, así como por el Consejo Nacional de Ciencia y Tecnología (CONACyT SALUD-20-2013-01-202475).

Dedicatoria

Dedico esta tesis a mis padres Héctor Sánchez Poblano y Alba León Tenorio quienes me han dado su apoyo incondicional no sólo económico sino también moral y sentimental para poder llegar a ser la persona que soy ahora.

A ellos los pilares de mi vida quienes siempre han respetado mis decisiones aun cuando no estan de acuerdo con ellas.

Este logro también es suyo.

ÍNDICE

AGRADECIMIENTOS	3
Dedicatoria	5
RESUMEN	8
INTRODUCCIÓN	10
Origen y epidemiología molecular del VIH	13
Estructura del VIH: genoma y proteínas	15
Proteínas estructurales	17
Proteínas reguladoras.....	18
Proteínas accesorias	18
Ciclo de replicación	19
Historia natural de la enfermedad	21
El VIH y la respuesta de linfocitos T citotóxicos	22
Presión selectiva por HLA	27
Planteamiento del problema	30
HIPOTESIS	32
OBJETIVOS	32
METODOLOGÍA.....	33
Pacientes	33
Extracción del RNA viral.	35
Amplificación del gen <i>gag</i>	35
Electroforesis	36
Purificación del producto de PCR	37
Secuenciación.....	37
Subtipificación de VIH	41

Análisis filogenético.....	41
Análisis de posiciones bajo presión selectiva	41
Análisis de variabilidad genética	42
Análisis de la distribución de alelos HLA.....	42
Análisis de asociación HLA-polimorfismo de VIH	42
RESULTADOS.....	44
Descripción del grupo de pacientes participantes.....	44
Predominancia de clusters específicos de país en el VIH circulante en Centroamérica	45
Distintos patrones de posiciones bajo presión selectiva positiva en las secuencias de <i>gag</i> de los virus de los diferentes países	47
Diferentes patrones de variabilidad entre las secuencias de los virus circulantes en los diferentes países.	49
Alta diversidad en distribución de frecuencias HLA entre las poblaciones de los países centroamericanos	52
Asociaciones HLA-polimorfismo en <i>gag</i>	56
Integración de datos.....	62
DISCUSION.	65
CONCLUSIONES y PERSPECTIVAS	69
BIBLIOGRAFÍA	71
ANEXOS:	74
ABREVIATURAS	74
GLOSARIO	74
ÍNDICE DE TABLAS	75
ÍNDICE DE FIGURAS.....	76

RESUMEN

El Virus de inmunodeficiencia Humana (VIH) continua siendo un problema de salud pública a nivel mundial con más de 25 millones de muertes en las pasadas 3 décadas y aproximadamente 34 millones de personas infectadas en el mundo. En Centroamérica, el VIH continúa representando un grave problema de salud pública al ser la región más afectada en América, después del Caribe e incluyendo países con altas prevalencias de infección como Belice, Guatemala, Honduras y Panamá. Se requieren estudios sobre la diversidad genética y evolución del virus a nivel regional con el objeto de contribuir al entendimiento de la adaptación del VIH a nivel poblacional y al diseño de estrategias antirretrovirales. Es bien sabido que las respuestas de linfocitos T citotóxicos (CTLs), definidas y restringidas por los genes HLA de clase I, controlan parcialmente la viremia en las etapas tempranas de infección, ejerciendo una fuerte presión selectiva sobre el virus. Péptidos derivados de las proteínas virales son presentados en la superficie de la célula infectada en el contexto de moléculas HLA de clase I. Los CTLs reconocen el complejo HLA-péptido viral y destruyen a la célula infectada. Sin embargo, dada la alta capacidad de variación del VIH, el virus puede escapar de esta respuesta asociada a HLA, mediante la selección de mutaciones que evitan la unión del péptido viral a la molécula HLA o su reconocimiento por los CTLs. Interesantemente, la adaptación del VIH a las respuestas restringidas por HLA es altamente predecible para una molécula HLA específica. Además, las mutaciones de escape restringidas al HLA muchas veces persisten en el virus al ser transmitido, alterando la inmunidad del hospedero y la progresión de la enfermedad. Dado el notable polimorfismo de los genes HLA y su distribución heterogénea en diferentes poblaciones del mundo se espera que la distribución y prevalencia de mutaciones de escape inmune asociadas al HLA sea característica de cada población, de acuerdo a las frecuencias HLA observadas en ella. Existen estudios previos que han iniciado la descripción de la adaptación viral a nivel poblacional a esta presión evolutiva, principalmente en poblaciones caucásicas, mediante la determinación de asociaciones HLA-polimorfismo de VIH. Sin embargo, la evolución asociada a HLA del VIH circulante en Centroamérica es

desconocida. El objetivo del presente estudio fue describir por primera vez el impacto de la presión selectiva ejercida por respuestas inmunes restringidas por HLA en el contexto inmunogenético propio de Centroamérica, en la variabilidad de la proteína Gag del VIH-1. Se secuenció el gen *gag* de 997 muestras de plasma de cinco cohortes de individuos infectados por VIH sin tratamiento antirretroviral previo, procedentes de Belice, Panamá, Nicaragua, Honduras y Guatemala. Se realizó la extracción de RNA y la amplificación de *gag* utilizando una PCR anidada. La secuenciación se realizó con el método de Sanger. Las secuencias se alinearon y se realizaron análisis filogenéticos y de variabilidad genética con estimaciones de entropía de Shannon. Así mismo, se estimaron aminoácidos bajo presión selectiva por razón dN/dS. Finalmente, utilizando los datos de frecuencias HLA-I previamente obtenidas en el laboratorio, se investigaron asociaciones HLA-polimorfismo de VIH. Los análisis filogenéticos mostraron clusters de virus altamente específicos de país. Los estudios de variabilidad del gen *gag* mostraron mayor variabilidad en las regiones p17 y p2p7p1p6 y una alta conservación en p24, posiblemente debida a su importancia funcional y estructural para el virus. Se encontraron asociaciones de HLA-polimorfismo de VIH previamente reportadas en 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 Centroamérica. Mostrando evidencia de un impacto importante de las respuestas inmunes restringidas por HLA en la determinación de la variabilidad del gen *gag* del VIH-1 circulante en poblaciones de Centroamérica. Siendo el primer trabajo que describe la variabilidad de *gag* en el VIH circulante en Centroamérica, el presente estudio constituye una base importante para el entendimiento de la adaptación viral en el contexto inmunogenético local, y establece las bases para continuar con el análisis de posiciones con asociaciones de interés. Este estudio contribuye con información importante para el diseño de vacunas contra VIH-1 y posibles nuevos mecanismos para el control del VIH-1.

INTRODUCCIÓN

A 36 años de la aparición del virus de inmunodeficiencia humana (VIH) la evolución de la epidemia del VIH se puede fragmentar en regiones, países, grupos étnicos hasta en estratos socio-económicos. Sin embargo, la epidemia es, en esencia, la suma de individuos infectados, en tanto que el VIH se transforma continuamente, infecta diversas poblaciones celulares, con características biológicas de cambio continuo [1]. Desde que el virus de inmunodeficiencia humana fue identificado en 1983, ha mostrado ser la causa para el síndrome de inmunodeficiencia adquirida (SIDA) en 1984, por lo que ha sido el mayor foco de atención en la creación de una vacuna, y las proteínas virales pueden ser el blanco de potenciales anticuerpos para la prevención de la infección [1].

Globalmente, la pandemia de VIH/SIDA constituye un problema de salud pública de primera magnitud en el mundo, con grandes repercusiones demográficas, sociales y económicas en muchos países. En el 2014 36.9 millones [34.3–41.4 millones] de personas vivían con el VIH en el mundo. En 2014, alrededor de 2 millones de personas se infectaron con el VIH y 1.2 millones de personas murieron de enfermedades relacionadas con el sida. [2].


De acuerdo a datos de la ONUSIDA desde el comienzo de la epidemia, aproximadamente 78 millones [71–87 millones] de personas contrajeron la infección por el VIH y 39 millones [35–43 millones] de personas han fallecido a causa de enfermedades relacionadas con el VIH. El 66 % de las nuevas infecciones por el VIH se dieron en el África subsahariana, alrededor de 600 en niños de menos de 15 años y aproximadamente 5,000 se dan en adultos de más de 15 años, de los cuales, casi un 48 % son mujeres y alrededor de un 30 % son jóvenes de 15 a 24 años.

Según datos de junio de 2015, había 15.8 millones de personas en tratamiento. Ello representa un 37% [35%-39%] de todas las personas que viven con el VIH. Un 38% [36%-40%] de los adultos que viven con el VIH reciben actualmente el tratamiento que necesitan, mientras que tan solo un 24% [22%-26%] de los niños que viven con el VIH tienen acceso al mismo.

Para América Latina, las cifras no son más alentadoras ya que en 2014, **1.7 millones [1.4-2.4 millones]** de personas vivían con el VIH 87,000 [71,000-170,000] nuevas infecciones por el virus durante el año de las cuales 2,000 fueron en niños [2].

Según estimaciones de la Organización Panamericana de la Salud (OPS), la Organización Mundial de la Salud (OMS) y el ONUSIDA, la infección por VIH afectó a 1 de cada 200 adultos de 15 a 49 años de edad en América del Norte y América Latina a finales del Siglo XX [3]. De todos los países de América Latina continental, los del norte de América Central parecen ser los más afectados por el VIH. Sin embargo, la situación es muy diversa, aún dentro del istmo [2].

Tabla1. Datos para Latinoamérica sobre VIH.



Región	Personas que viven con el VIH 2014		Nuevas infecciones por el VIH en 2014			Muertes relacionadas con el sida en 2014 (total)
	total	niños	total	adultos	niños	
América Latina	1,7 millones [1,4-2,0 millones]	33 000 [29 000–40 000]	87 000 [70 000–100 000]	85 000 [68 000–100 000]	2000 [1300-2900]	41 000 [30 000–82 000]
MUNDIAL	36,9 millones [34,3–41,4 millones]	2,6 millones [2,4-2,8 millones]	2,0 millones [1,9–2,2 millones]	1,8 millones [1,7–2,0 millones]	220 000 [190 000-260 000]	1,2 millones [980 000-1,6 millones]

Tomado de [2].

Cuatro de los países con prevalencias más altas en América Latina están en Centroamérica (2.5% de la población, en Belice; 1.6% en Honduras, y 0.9% en Guatemala y el Salvador). EL SIDA es una de las diez causas de muerte más frecuentes entre los países de Centroamérica (Honduras, Guatemala y Panamá) y

se estima que existen más de doscientas mil personas VIH/SIDA en Centroamérica [4].

En Centroamérica, la epidemia está concentrada en determinadas áreas geográficas, como las zonas urbanas, las rutas comerciales y los puertos, en grupos de alta vulnerabilidad como los trabajadores comerciales del sexo y sus clientes, los hombres que tienen relaciones sexuales con hombres, los privados de libertad y la población garífuna (un grupo de población afro-caribeño de la costa Atlántica de Honduras). Sin embargo hay una preocupación creciente por los grupos asociados con las poblaciones móviles, trabajadoras comerciales de sexo (TCS) y los hombres que tienen sexo con otro hombre (HSH), que tienen tasas de prevalencia de VIH más altas, y representan la población puente para la transmisión de la epidemia a la población en general [3].

La transmisión de VIH en Centroamérica está principalmente asociada al sexo heterosexual, lo que se asemeja más al patrón caribeño que al de América del Sur. Y aunque hay más hombres que mujeres con SIDA en Centroamérica, la brecha de género se está cerrando. Otros factores que contribuyen a la epidemia en la región son la inmigración, el turismo y la proximidad al Caribe donde la prevalencia es aún mayor [4].

La epidemia se está feminizando y se ha empeorado debido a factores culturales, incluyendo que las mujeres con frecuencia son forzadas a tener sexo sin consentimiento y sin poder negociar el uso de condones. Un ejemplo claro de esto es Guatemala con un total de 26,113 mujeres guatemaltecas mayores de 15 años viven con VIH, según estimaciones de 2011. La infección por VIH en mujeres se ha triplicado en 22 años, y la relación de masculinidad, que antes de 1996 era de entre 5 y 7 hombres por cada mujer, ahora es de dos hombres por cada mujer. Actualmente, los datos señalan que la epidemia está aumentando en todos los países de Centroamérica (Tabla 2) [4].

Tabla 2. Casos de VIH y Sida acumulados hasta 2014, en países de la Región Centroamericana.

País	VIH/SIDA - habitantes infectados con VIH/SIDA	Fallecimientos por SIDA
Guatemala	49,000 [34000-69000]	1700 [1100 - 2600]
Honduras	23,000 [19,000-27, 000]	1200 [1000 - 1600]
Panamá	17,000 [13,000 - 24,000]	<1000 [<500 - <1000]
Nicaragua	10,000 [7,600-15,000]	<500 [<500 - <1000]
Belice	2,700 [2,200 - 7,700]	<100 [<100 - <200]

Tomado de [4].

Independientemente de la situación de la epidemia del VIH/SIDA en cualquier país del mundo, el conocimiento universal de la virología del VIH indudablemente tendrá un impacto en la generación de nuevas estrategias de tratamiento que ayuden a controlar la enorme carga humana, social, económica y política que representa el VIH [1].

Origen y epidemiología molecular del VIH

A mediados del siglo pasado se observó el inicio de una enfermedad inmune que empezó silenciosamente en África Central y Subsahariana, y que poco a poco fue creciendo e invadió gran parte de la población mundial.

El VIH es el resultado de saltos zoonóticos de retrovirus entre monos y humanos. Ahora es claro que el VIH-1 se originó a partir del virus de inmunodeficiencia simiana del chimpancé (SIVcpz), mientras que el VIH-2 se generó a partir del SIV del mono mangabey ahumado (sooty mangabey en inglés, SIVsmm) [5].

Como resultado de múltiples eventos de transmisión zoonótica, se han generado diversos linajes de VIH: VIH-1 con el grupo M (“main”), el O (“outlier”), el N (“new”) y el P y el VIH-2 con grupos de la A-H. El grupo M del VIH-1 es el responsable de

la epidemia a nivel global y presenta nueve subtipos (A, B, C, D, F, G, H, J, K) (Figura 1), los cuales presentan entre sí algunas diferencias en su comportamiento biológico y en su distribución geográfica [5].

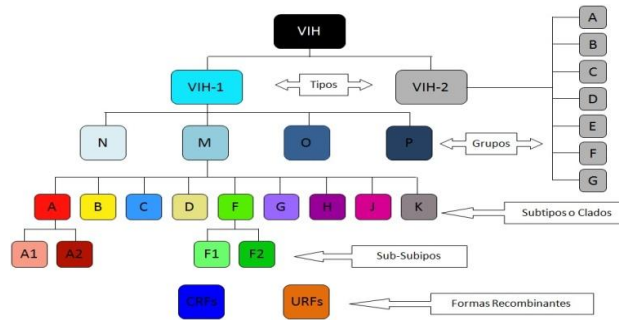


Figura 1. Grupos, tipos y subtipos del VIH. Descripción de la diversificación del virus en diferentes linajes.

Actualmente se han caracterizado 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 transcriptasa inversa viral (RT) enzima que puede cambiar el templado usado durante su actividad, debido a la presencia de dos copias del genoma del virus en cada partícula viral. Algunas de estas variantes se han diseminado dando lugar a las formas recombinantes circulantes (CRF). 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 [1].

La distribución de los subtipos de VIH-1 en los individuos infectados depende de la proporción de personas infectadas en cada país, que genera una estimación regional y global sobre la distribución de subtipos de VIH-1

Entre el 2004 y el 2007 el subtipo C se encontró en el 48% de la población mundial, los subtipos A y B causaron el 12% y 11% de infecciones respectivamente, seguidos de CRFC02_AG con un 8%, CRF01_AE en 5%, subtipo G 5% y el D 2%. Subtipos como el F, H, J y el K causan juntos cerca de 1% a nivel mundial. El subtipo B es dominante en el continente Americano (Norte

de América, el Caribe, América Latina), Australia y en Europa occidental y central (Figura 2) [6].

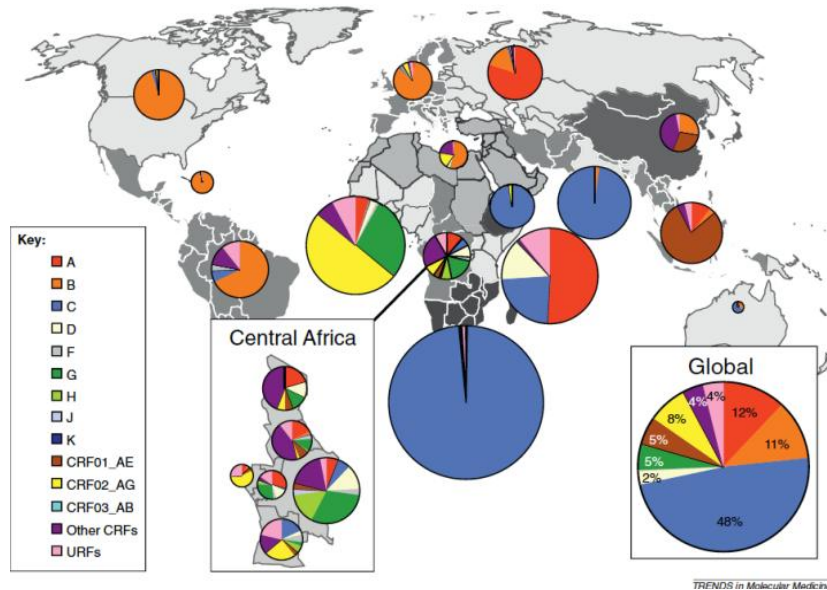


Figura 2. Distribución global de subtipos de VIH-I y sus recombinantes (CRF). El color de cada figura representa el subtipo y el tamaño de los pasteles es proporcional al número de personas que viven con VIH-1 en cada región representada [5]. Se observa que el subtipo C predomina a nivel mundial por su presencia en poblaciones con gran prevalencia a la infección (África sub-Sahariana e India). Por otro lado, en América el subtipo predominante es el B [6].

Estructura del VIH: genoma y proteínas

El VIH se ubica taxonómicamente en la subfamilia *Lentiviridae*, perteneciente a la Familia *Retroviridae*. Los miembros de esta familia se caracterizan por presentar un genoma compuesto por dos hebras sencillas de RNA+, de aproximadamente 10 kb contenidas dentro de un virión, por lo que se les consideran pseudodiploides. Los retrovirus son característicos de los vertebrados, y su ciclo biológico depende de una polimerasa cuya capacidad para leer el genoma viral de RNA y sintetizar una hebra doble de ADN, la llevó a ser denominada reverso transcriptasa (RT) o transcriptasa inversa. El ADN viral se integra al genoma de la célula hospedera recibiendo el nombre de provirus. El provirus está flanqueado por secuencias largas repetidas (o LTR, por sus iniciales en inglés) que son generadas por reverso transcripción [1,7].

El genoma del VIH 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* (Figura 3).

Las LTRs se localizan en cada extremo del genoma viral 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 [1, 9].

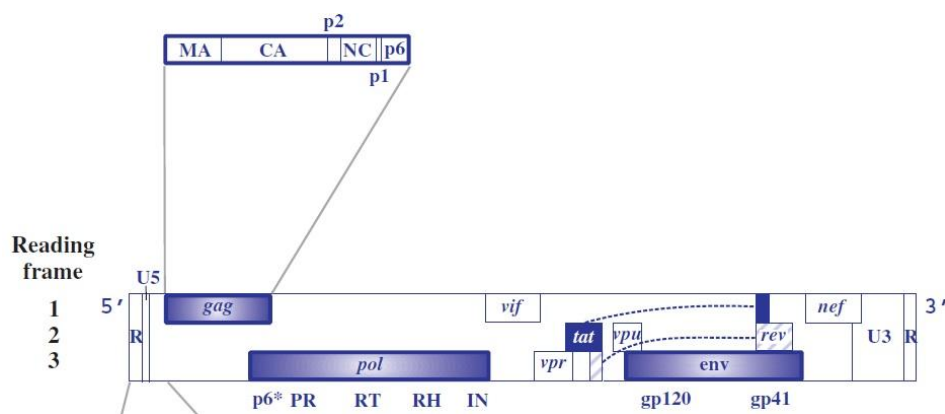


Figura 3. Organización del genoma del VIH-I. Se observan tres marcos de lectura abiertos (ORF) principales que comprenden a los genes *gag*, *pol* y *env*. Además, se observan seis ORF más que codifican para las proteínas accesorias/reguladoras. El fragmento de *gag* está ampliado. Además existen dos secuencias de repetición terminal que flanquean el genoma, LTR5' y LTR3' que contienen secuencias promotoras y reguladoras de transcripción [9].

En total, los genes del VIH codifican para nueve proteínas clasificadas en tres clases: proteínas estructurales: Gag, Pol y Env, que forman 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* (Figura 4) [7,8].

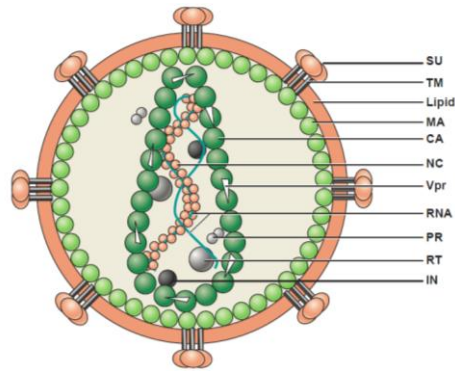


Figura 4: Diagrama de la estructura del VIH. Descripción gráfica de las proteínas del VIH en su estructura. SU: proteína de superficie (gp120), TM: proteína transmembranal (gp41), MA: matriz (p17), CA: cápside (p24), NC: núcleo cápside (p7), Vpr: proteína viral R (p23), PR: proteasa (p10), RT: Transcriptasa inversa (p66/p51), IN: integrasa (p32) [8].

Proteínas estructurales

El gen *gag* codifica para la poliproteína precursora Gag (Pr55Gag) cuyo orden de proteínas maduras es el siguiente: p17/p24/p2/p7/p1/p6 [1]. 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). MA (p17) se localiza en la superficie interna de la bicapa lipídica y es responsable de la estabilidad del virus [10]. CA (p24) es la cápside del virus, y está compuesta por 2 dominios: el dominio N-terminal “nuclear” y el dominio C-terminal de “dimerización”, que favorece las interacciones Gag-Gag. NC (p7) facilita la transcripción inversa (protege de nucleasas e incorpora un RNA para el inicio de la transcripción) y es necesaria para el empaquetamiento del genoma del virus, constituyendo la nucleocápside del virus. La proteína p6 favorece la liberación de los viriones [9].

El gen *pol* produce una poliproteína conocida como Gag-Pol (Pr160Gag-Pol) y deriva en tres enzimas: proteasa, transcriptasa inversa e integrasa [10]. PR es una proteasa encargada de cortar los polipeptidos durante la gemación del nuevo virión y produciendo cambios estructurales y morfológicos conocidos como maduración viral [9]. 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 (3.4×10^{-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 [7].

IN permite la inserción del genoma viral dentro del material genético del hospedero [8].

Env es un polipeptido de 160 kDa, 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 forman trímeros que producen espículas en la envoltura del virus. SU (de mayor variabilidad) 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 la 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 [10].

Proteínas reguladoras

Tat es requerida para el inicio de la transcripción inversa por su actividad de transactivador transcripcional. Rev induce la transición de la fase temprana a la tardía en la expresión de los genes del VIH. [7].

Proteínas accesorias

Nef, de aproximadamente 27kDa. 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. Vpr de 14 kDa 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. Vpu 16 kDa exclusiva de VIH-I, reduce la expresión de la molécula CD4 y favorece la liberación de los viriones de las células infectadas. Vif de 23 kDa, fundamental para la replicación y el adecuado empaquetamiento de las proteínas virales [1,7].

Ciclo de replicación

Reconocimiento y adhesión: El VIH infecta a aquellas células que poseen el receptor de membrana CD4. La replicación del VIH inicia con la unión de gp120 y la molécula CD4 de la célula blanco, la cual funciona como receptor principal del virus y está presente en linfocitos T cooperadores, macrófagos CD4+ y algunas células dendríticas [8].

Entrada: requiere de la unión de la región V3 de gp120 a un correceptor. Los principales correceptores del virus funcionan como receptores de quimiocinas, destacando CXCR4 y CCR5 [8].

La unión de gp120 con el correceptor produce cambios conformacionales exponiendo el dominio de fusión de gp41 y permite la fusión de las membranas viral y celular. La cápside viral entra al citoplasma de la célula diana y el RNA viral se transcribe a DNA de doble cadena por medio de la RT, produciéndose la desencapsidación viral. Posteriormente se forma el llamado complejo de preintegración (PIC) que contiene a las proteínas virales con señales de localización nuclear: IN, NC, Vpr y el DNA viral recién sintetizado. Una vez que el PIC se encuentra dentro del núcleo, la IN 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 [8].

Después de la integración comienza la transcripción de genes virales, para lo cual son necesarios factores de transcripción celulares y proteínas virales reguladoras como Tat. Los primeros RNAm virales sufren splicing, formando las proteínas Tat, Rev, Nef, Vif, Vpr y Vpu. A medida que la concentración de Rev aumenta, se favorece la salida de los RNAm virales del núcleo, evitando así los eventos de splicing. En consecuencia, los RNAm virales de la etapa tardía se caracterizan por tener uno o ningún evento de splicing, lo que conlleva a la producción de las poliproteínas Pol, Env y Gag y los RNAs que se empaquetarán en las nuevas partículas virales [11].

Los precursores Gag, Pol, Env y el nuevo genoma de RNA se dirigen a la membrana celular para ser ensamblados. Residuos miristilados en el extremo MA

de los precursores Gag y Gag-Pol son responsables de su localización membranal. Al gemar los nuevos viriones, estos adquieren parte de la membrana lipídica de la célula hospedera y los productos del gen *env* (gp120 y gp41). Durante la maduración del nuevo virión, la enzima proteasa corta cada poliproteína para formar las proteínas virales funcionales, las cuales se ensamblan para formar la matriz, cápside y nucleocápside virales. El VIH maduro será capaz de infectar a otras células e iniciar de nuevo el ciclo de replicación (Figura 5) [7].

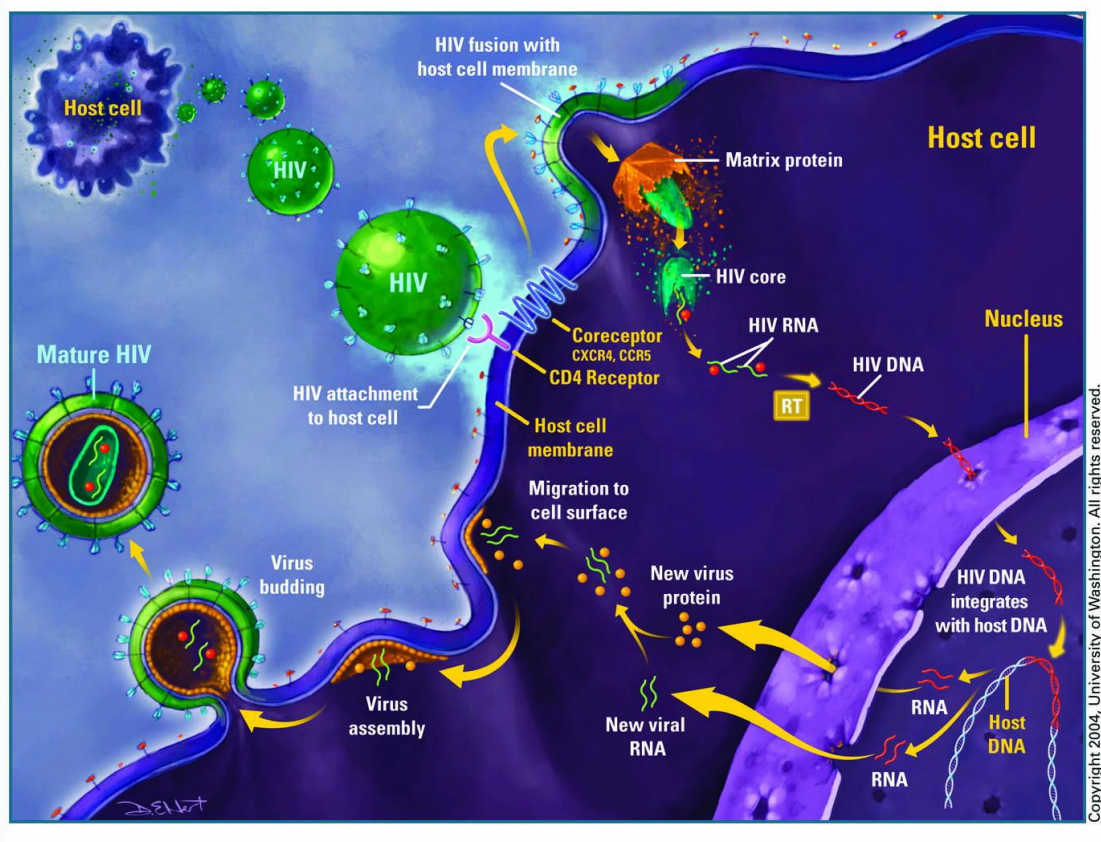


Figura 5. Descripción esquemática de los eventos que ocurren durante el ciclo de replicación del VIH. Interacción entre gp120, CD4 y el receptor de quimiocinas (CCR5 o CXCR4). Fusión de membranas gracias a gp41, entrada de la CA, desnudamiento del virión, transcripción inversa del genoma viral, formación 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

virion por el corte de precursores gracias a la PR. Inicio de nuevo ciclo de replicación. Tomado de HIV.websttudy.org

Historia natural de la enfermedad

La infección por VIH se divide clínicamente en etapas que son diferenciadas por la interacción entre el virus y la respuesta inmune del hospedero.

- Infección primaria o aguda: se caracteriza por altas cargas virales, asociadas a una replicación descontrolada del virus, así como una caída temporal de la cuenta de Linfocitos T CD4+ en sangre.
- Etapa crónica o de latencia clínica: caracterizada por una estabilización de la carga viral y la cuenta de linfocitos T CD4+ en sangre, con concentraciones entre 200 a 500 células/ μ l
- Etapa tardía o de SIDA: donde los pacientes presentan concentraciones por debajo de 200 células T CD4+/ μ l, perdiendo el control parcial de la replicación viral con aumentos en la carga viral en sangre e incrementando el riesgo de adquisición de infecciones oportunistas. Resultando eventualmente en la muerte del hospedero [1, 8].

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 (Figura 6).

La habilidad del sistema inmune para montar una respuesta específica (CTL's) contra el VIH en las primeras semanas es un factor clave en el subsecuente avance de la enfermedad [12].

La fase de SIDA se caracteriza por una inmunodeficiencia grave, lo que conlleva a una mayor susceptibilidad de contraer enfermedades oportunistas [1].

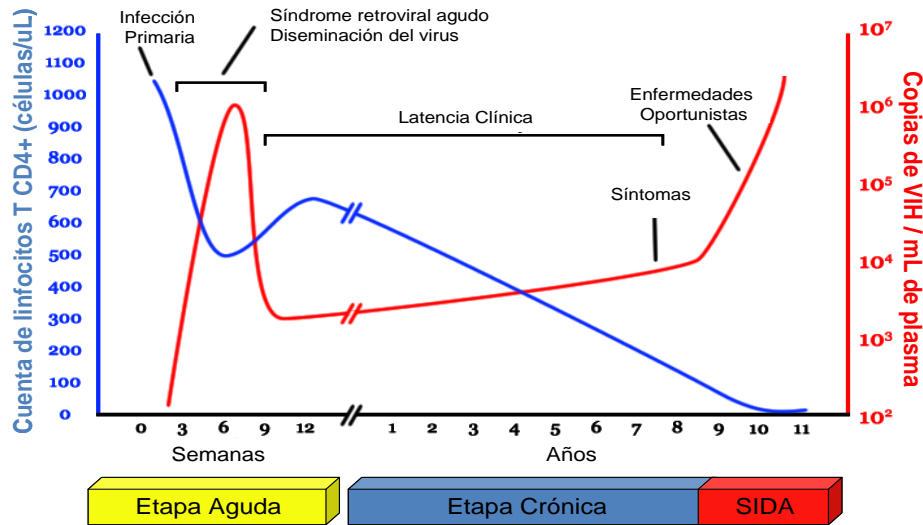


Figura 6. Historia natural de la enfermedad causada por VIH. Se observa la variación de los linfocitos T CD4+ y del número de copias de RNA del VIH a través de las diversas etapas en las que se divide la enfermedad.

El VIH y la respuesta de linfocitos T citotóxicos

El sistema inmunitario debe enfrentar todo tipo de agentes patógenos y ha desarrollado por evolución múltiples estrategias para combatir la invasión. La inmunidad adaptativa es capaz de reconocer y eliminar de manera selectiva microorganismos y moléculas extrañas específicos (es decir, antígenos ajenos)[8]. Las células T expresan en su membrana una molécula de unión a antígeno única, denominada receptor de célula T (TCR, *T-cell receptor*). Existen dos subpoblaciones de células T bien definidas: células T colaboradoras (TH) y células T citotóxicas (TC). Ambos tipos pueden distinguirse entre sí por la presencia de las glucoproteínas de membrana CD4 o CD8 respectivamente. Las células T maduras CD4+ suelen funcionar como células TH, en tanto que las CD8+ lo hacen casi siempre como células TC [13].

La mayoría de los receptores de célula T únicamente puede identificar antígenos unidos a proteínas de membrana celular llamadas moléculas del complejo principal de histocompatibilidad (MHC, *major histocompatibility complex*). Las moléculas del MHC, también llamadas HLA (*human leucocyte antigen*), son glucoproteínas polimórficas, que se encuentran en las membranas celulares. Existen dos tipos principales de moléculas MHC: moléculas MHC clase I, que casi

todas las células nucleadas de las especies de vertebrados expresan, y moléculas MHC clase II, que sólo las células presentadoras de antígeno expresan (APC, *antigen-presenting cells*). Las moléculas MHC clásicas ligan péptidos antigénicos para que éstos sean presentados a diversas subpoblaciones de células del sistema inmune.

En los seres humanos, el MHC se localiza en el brazo corto del cromosoma 6 y ocupa un gran segmento de ADN. Los genes del MHC se expresan de forma codominante en cada sujeto. En otras palabras, para un gen del MHC dado, cada sujeto expresa los alelos que se heredan de cada uno de los dos progenitores. Para el sujeto, esto maximiza el número de moléculas del MHC disponibles para unirse a péptidos que presentan a los linfocitos T. De los tres genes clásicos del MHC clase I (HLA*A, HLA*B y HLA*C), el alelo HLA*B es el más polimórfico con 4179 moléculas diferentes descritas, comparado con 3356 del HLA*A y 2902 del HLA*C [13,14,15]. Específicamente el alelo HLA*B es el locus más polimórfico en el genoma (cromosoma 6) de la población entera. Existen diferencias en las variaciones de los alelos de HLA*B, con una extraordinaria diversidad, en las distintas frecuencias en poblaciones de todo el mundo. (Figura 7) [14].

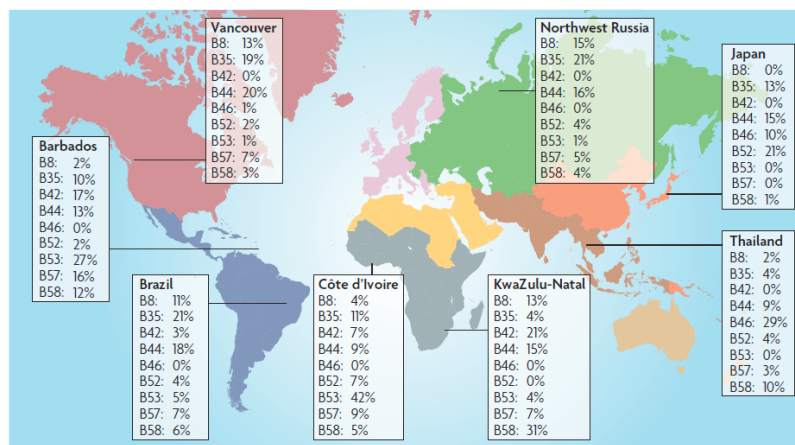


Figure 7. Diversidad y distribución de los alelos HLA*B en el mundo. La diversidad del HLA clase I es ilustrada con la prevalencia de nueve moléculas de HLA*B en ocho poblaciones. Por simplicidad se muestran solo dos dígitos del alelo [13].

La distribución de los alelos en Centroamérica ha sido poco estudiada en relación al VIH-1 teniendo mayores datos para poblaciones caucásicas.

Las variaciones en las moléculas del MHC entre los sujetos influyen en el péptido que ligan y en su reconocimiento por el linfocito T. Linfocitos T con TCR específicos pueden reconocer péptidos específicos que son capaces de unirse solo a un número determinado de moléculas MHC específicas. Este fenómeno se llama restricción por el MHC.

Los linfocitos T citotóxicos (CTL) tiene una función vital en la vigilancia de las células del cuerpo y la eliminación de cualquiera que exhiba antígeno en el contexto de moléculas HLA de clase I, como las células infectadas por virus, las células tumorales y las células de un injerto de tejido extraño [12]. En el caso del control de infecciones virales, los CTLs se estimulan por antígenos presentados a su TCR, en el contexto de moléculas HLA clase I expresadas por células infectadas. Las células infectadas procesan las proteínas virales y las presentan en forma de péptidos de 8 a 10 aminoácidos para que éstos puedan unirse a 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].

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].

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+ reaparece [14].

Existen estudios sobre análisis comparativos del genoma (GWAS, *genome-wide association studies*) en individuos Caucásicos con VIH-1, que presentaron por primera vez un impacto en la variación genética humana con respecto a la carga viral y la progresión de la enfermedad. Estos estudios han confirmado el papel fundamental del polimorfismo del MHC en el control del VIH, específicamente alelos HLA-I [16-17]. EL primer estudio de GWAS en VIH-1 reveló tres asociaciones de polimorfismos en un solo nucleótido (SNP) con bajos niveles de carga viral individuos con infecciones crónicas. Interesantemente, estas tres asociaciones se localizan en la región HLA del cromosoma 6, incluyendo: un SNP ligado a HLA*B57, un SNP localizado cerca del HLA*C y un SNP en la subunidad de la RNA polimerasa. El segundo SNP fue previamente asociado con una alta transcripción de RNAm. Se ha especulado que este SNP podría estar relacionado con el incremento en la expresión de HLA*C en la superficie de la células [18].

Las moléculas HLA promueven el mecanismo por el cual el sistema inmune genera una respuesta específica al patógeno lo que genera una fuerte presión selectiva sobre el VIH. Sin embargo, no todos los alelos HLA son igualmente eficientes para controlar la replicación del virus y se ha demostrado que la expresión de alelos HLA específicos puede asociarse con una progresión más lenta o más rápida de la enfermedad por VIH (Tablas 3 y 4) [17]. Consistentemente en estudios en diversas poblaciones, varios alelos HLA*B35 se han asociado con una progresión acelerada a SIDA. Por otra parte el alelo HLA*B57 presente en el 11% de la población EUA y en el 11% de individuos VIH positivos, ha sido identificado con una progresión lenta, baja carga viral y nulos síntomas de la infección primaria por VIH [19]. Este fenómeno constituye otra

fuerte evidencia del papel de la respuesta de CTL en el control de la infección por VIH [20].

Tabla 3. Alelos HLA-I protectores hacia la progresión de la infección por VIH.

HLA-B	References	Cohorts
Protective HLA I Alleles		
B*13:02	Honeyborne et al., 2007; Fellay et al., 2009; Pereyra et al., 2010	C clade South Africa; B clade Caucasian
B*14/14:02	Pereyra et al., 2010; Lazaryan et al., 2011	B clade Caucasian; African American
B*27/27:05	Kaslow et al., 1996; O'Brien et al., 2001, Fellay et al., 2009; Pereyra et al., 2010	B clade Caucasian
B*42:01	Carlson et al., 2012	C clade South Africa/Botswana/Zimbabwe
B*44:03	Leslie et al., 2010; Carlson et al., 2012	C clade, South Africa; South Africa/Botswana/Zimbabwe
B*51	Kaslow et al., 1996; O'Brien et al., 2001	B clade Caucasian
B*52:01	Pereyra et al., 2010	B clade Caucasian
B*57/B*57:01	Kaslow et al., 1996; Migueles et al., 2000; O'Brien et al., 2001; Pereyra et al., 2010	B clade Caucasian; Caucasian/Hispanic
B*57:02	Leslie et al., 2010	C clade South Africa
B*57:03	Costello et al., 1999; Leslie et al., 2010	A clade Rwanda; C clade South Africa
	Tang et al., 2010; Carlson et al., 2012	C clade Zambia; South Africa/Botswana/Zimbabwe
B*57:03	Pereyra et al., 2010; Lazaryan et al., 2011	B clade African American
B*58:01	Kiepiela et al., 2004; Lazaryan et al., 2006	C clade South Africa; Zambia
	Carlson et al., 2012	C clade South Africa/Botswana/Zimbabwe
B*81:01	Kiepiela et al., 2004; Tang et al., 2010	C clade, South Africa; Zambia; South Africa/Botswana/Zimbabwe
	Carlson et al., 2012; Pereyra et al., 2010	B clade African American; C clade
A*25/25:01	Kaslow et al., 1996; Pereyra et al., 2010	B clade Caucasian
A*32/32:01	Kaslow et al., 1996; Fellay et al., 2009	B clade Caucasian;
	Lazaryan et al., 2011	B clade North American African American
A*74/74:01	Tang et al., 2010; Matthews et al., 2011; Lazaryan et al., 2011; Koehler et al., 2010	C clade Zambia; South Africa/Botswana/Zimbabwe; B clade African American;

Tomado de [17].

Tabla 4. Alelos HLA-I susceptibles a la progresión de la infección por VIH.

Disease-Susceptible HLA I Alleles		
B*07:02	Pereyra et al., 2010	B clade Caucasian
B*08/08:01	Steel et al., 1988; Pereyra et al., 2010; Carlson et al., 2012	B clade Caucasian C clade South Africa/ Botswana/Zimbabwe
B*18/18:01	Leslie et al., 2010; Carlson et al., 2012	C clade South Africa/ Botswana/Zimbabwe
B*35	Carrington et al., 1999; O'Brien et al., 2001	B clade Caucasian
B*35:01	Pereyra et al., 2010	B clade North American European
B*35:02/35:03	Gao et al., 2001; Fellay et al., 2009	B clade Caucasian
B*45/45:01	Tang et al., 2010; Carlson et al., 2012	C clade Zambia; South Africa/Botswana/ Zimbabwe
B*51:01	Carlson et al., 2012	C clade South Africa/ Botswana/Zimbabwe
B*53:01	Gao et al., 2001; Pereyra et al., 2010	B clade Caucasian; African American
B*54/55/56	Hendel et al., 1999; Dorak et al., 2003	B clade Caucasian
B*58:02	Kiepiela et al., 2004; Tang et al., 2010; Carlson et al., 2012	C clade South Africa; Zambia South Africa/Botswana/ Zimbabwe
A*36:01	Tang et al., 2010; Carlson et al., 2012	C clade Zambia; South Africa/Botswana/ Zimbabwe
A*36:01	Tang et al., 2010; Carlson et al., 2012	C clade Zambia; South Africa/Botswana/ Zimbabwe
B*07:02	Pereyra et al., 2010	B clade Caucasian
B*08/08:01	Steel et al., 1988; Pereyra et al., 2010;	B clade Caucasian

Tomado de [17].

Presión selectiva por HLA

Como se ha descrito anteriormente, las respuestas de CTL, definidas y restringidas por las moléculas HLA de clase I, son altamente eficientes para el control de la replicación viral, al menos en etapas tempranas de la infección. Por esta razón, estas respuestas representan una importante fuerza evolutiva para el VIH a nivel individual y poblacional [17]. Dada la gran capacidad de variación del VIH, es posible que se seleccionen variantes con mutaciones en fragmentos de las proteínas virales que son presentadas por las moléculas HLA clase I a los CTLs. Estas mutaciones de escape provocan que el péptido viral se vuelva incapaz de

unirse a la molécula HLA o que el epítotope en cuestión ya no sea reconocido por el CTL. 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 [21]. Así, 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 epitopes virales presentados por moléculas HLA clase I, siendo esta adaptación viral altamente conservada y predecible [23].

De acuerdo con la historia evolutiva de las poblaciones y el hecho de que cada sujeto expresa los alelos HLA que se heredan de cada uno de los dos progenitores, las frecuencias de alelos HLA varían ampliamente en las poblaciones, en algunas ocasiones presentándose alelos HLA únicos de poblaciones específicas. Consecuentemente, la presión selectiva asociada al HLA y la frecuencia de mutaciones de escape en el virus dependerá de las frecuencias alélicas HLA de cada población [23].

Estudios previos han demostrado que los genes HLA están dejando una huella poblacional característica en la evolución del virus al seleccionarse mutaciones de escape que se ven reflejadas en los virus circulantes, de acuerdo a la frecuencia alélica HLA de la población [24]. Las mutaciones de escape observadas en diferentes poblaciones son altamente variables, de acuerdo a la distribución de frecuencias alélicas HLA característica de cada región del mundo, creándose un contexto inmunogenético único para cada población. Las secuencias circulantes de VIH en las poblaciones tienen adaptaciones específicas a los alelos expresados en cada población a pesar de que en algunas de estas poblaciones como en México, EUA y Japón compartan el mismo subtipo de virus (Subtipo B) (Figura 8). Además, estas asociaciones HLA-polimorfismo viral son altamente conservadas y predecibles[22].

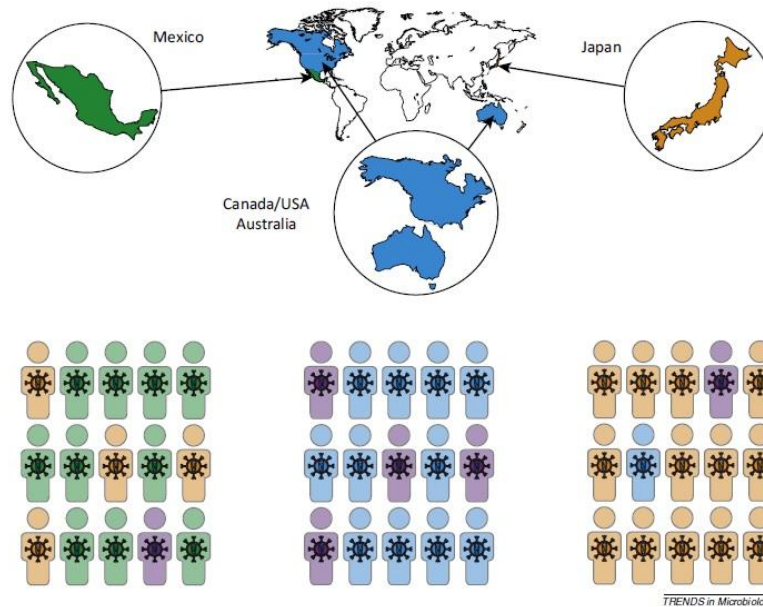


Figura 8. El HLA es el mayor control global de la diversidad del VIH en el mundo. Existe una adaptación específica del virus a cada población. Se muestran poblaciones de México, Canadá/EUA/Australia y Japón con sus distribuciones de HLA representados en colores Verde, Azul y Naranja, respetivamente, así como la adaptación de las secuencias del virus a cada uno de ellos, mostrado con los diferentes colores del virus en cada hospedero.

Aunque las mutaciones de escape a CTLs confieren una ventaja al virus 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, el cual influye en su capacidad de dispersión a nivel poblacional [12, 14, 19, 22].

La identificación de polimorfismos asociados a HLA ha servido como una guía para descubrir nuevos epítopes en el VIH. Particularmente *gag* es un gen altamente inmunogénico, con zonas que requieren una alta conservación para el correcto ensamblaje de las estructuras del virión. Por esta razón, mutaciones de escape en Gag frecuentemente se asocian a costos en la capacidad replicativa del virus. El costo en la capacidad replicativa del virus debido a mutaciones de escape se ha relacionado a variaciones en la velocidad de avance de la enfermedad [16]. Alelos HLA protectores en el contexto de la infección por VIH como B*57:01 y B*27:05, relacionados con una progresión lenta a SIDA, frecuentemente se

asocian con selección de mutaciones de escape en *gag*, con altos costos en la capacidad replicativa [1, 22, 23].

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 provocara que las mutantes con alto costo reviertan y las de bajo costo permanezcan, cambiando el consenso poblacional con el avance de la epidemia [14,21].

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 y granzimas, liberadas por los CTLs [1]. Por tanto, la respuesta de los CTLs 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 [21]. 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 epitopos antes de su presentación o que disminuyen la afinidad de la unión HLA/péptido [20]. La presencia de estas mutaciones de escape al sistema inmune están muchas veces asociadas a un costo en la capacidad replicativa del virus, lo que conduce a una reversión al transmitirse a un individuo carente del HLA selector [13].

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 [14,20,22]. 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 diversos contextos inmunogenéticos poblacionales. 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 [13,17].

Debido al alto polimorfismo de los genes HLA, la frecuencia de alelos en distintas poblaciones es diferente, pudiendo impactar de forma distinta en la evolución del VIH según el contexto inmunogenético de la población que se trate. 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, observando que el patrón de aparición de estas mutaciones varía según las combinaciones específicas de HLA presentes en cada una de las cohortes estudiadas [14,20,22,23]. Sin embargo, el conocimiento de la adaptación viral a respuestas restringidas por HLA en poblaciones latinoamericanas es limitado.

En Centroamérica no solamente se desconoce si existen características particulares de la evolución asociada a HLA del VIH circulante, sino que no existen datos sobre la caracterización inmunogenética de las poblaciones de la región. Estudiar la evolución del VIH en el gen *gag* dependiente de HLA en poblaciones de diferentes partes del mundo es importante dada la alta diversidad de la distribución alélica HLA, que puede causar diferencias importantes en la evolución viral en diferentes regiones con impacto directo en el diseño de estrategias de control y posibles vacunas efectivas a nivel regional. Estos análisis ayudan a entender de qué manera está evolucionando el VIH en un contexto inmunogenético característico, pudiéndose encontrar nuevos alelos HLA asociados con un mejor o peor control de la replicación viral, así como sus correspondientes sitios blanco en el genoma viral para el entendimiento de mecanismos naturales de control viral y su posible aplicación en estrategias terapéuticas. El conocimiento de la evolución del VIH mediada por HLA en Centroamérica tiene aplicación en la investigación en vacunas basadas en CTL, particularmente en la búsqueda de epítopes capaces de inducir respuestas eficientes de CTL contra sitios en proteínas virales que exhiben variabilidad

limitada, hacia los cuales existan restricciones de escape para el virus por implicar costos en la capacidad replicativa [14,21].

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 poblaciones de Centroamérica. A partir de secuencias del gen *gag* de VIH de una cohorte de pacientes centroamericanos, se evaluó la diversidad poblacional del VIH circulante en cada país, así como la posibilidad de que esta diversidad esté asociada a presión selectiva mediada por HLA a nivel poblacional, en el contexto inmunogenético de las poblaciones de la región.

De acuerdo con la importancia previamente descrita de las respuestas de CTLs dirigidas a Gag para el control inmune del VIH, los resultados de este estudio contribuyen a la investigación sobre vacunas en el contexto global, y a la identificación de posibles sitios blanco para la generación de inmunógenos que pudieran ser útiles para controlar el virus en estas personas que viven con el virus en Centroamérica.

HIPOTESIS

Se espera que la adaptación evolutiva asociadas a HLA en del gen *gag* del VIH circulante en Centroamérica tenga características únicas debido al contexto inmunogenético estudiado.

OBJETIVOS

General

Estudiar la evolución asociada a HLA en el gen *gag* del VIH circulante en Centroamérica.

Específicos:

- ❖ Obtener secuencias de *gag* representativas del VIH circulante en Centroamérica incluyendo Guatemala, Belice, Honduras, Nicaragua y Panamá.

- ❖ Comparar la diversidad genética del virus circulante en cada país a través de análisis filogenéticos y determinación de entropía.
- ❖ Obtener la subtipificación de los virus circulantes en la región a partir de secuenciación y análisis filogenéticos del gen *gag*.
- ❖ Buscar sitios bajo presión selectiva por análisis de tasa de sustituciones sinónimas vs. no sinónimas en las secuencias de los virus circulantes en la región
- ❖ Recopilar los datos de tipificación de los genes HLA de clase I A, B y C de la cohorte de individuos infectados por VIH en Centroamérica.
- ❖ Asociar la presencia de polimorfismos en *gag* con la presencia de alelos HLA específicos.
- ❖ Comparar los patrones de asociaciones HLA-polimorfismo de VIH observados en cada país de la región.

METODOLOGÍA

Pacientes

Se seleccionaron pacientes participantes en el proyecto de “Epidemiología Molecular y Evolución del VIH en la Región Mesoamericana”, conducido por del Centro de Investigación en Enfermedades Infecciosas (CIENI) del Instituto Nacional de Enfermedades Respiratorias (INER). El CIENI, INER ha establecido colaboraciones con centros de salud y clínicas de atención a individuos que viven con VIH en la región con el objetivo de estudiar la variabilidad genética, prevalencia de farmacorresistencia y adaptación del virus en el contexto inmunogenético local. Las clínicas participantes invitan a los pacientes a participar en el estudio y colectan muestras de sangre, las cuales son enviadas y procesadas en el CIENI, INER. Este proyecto fue revisado y aprobado por el Comité de Ética en Investigación del INER y es conducido de acuerdo a la Declaración de Helsinki. Todos los participantes firmaron consentimiento escrito para el uso de las muestras de sangre donadas en investigación. Para todos los participantes se realizaron las pruebas de carga viral, conteo de linfocitos T CD4+

y genotipo del gen *pol* viral. Los resultados de estas pruebas fueron enviados a los centros participantes para su uso en el seguimiento clínico de los pacientes.

El grupo de estudio del presente trabajo se conformó con individuos infectados por VIH provenientes de 5 países: Belice (54), Guatemala (342), Honduras (285), Nicaragua (205) y Panamá (111), quienes se encontraban vírgenes a tratamiento en el momento de adquirir la muestra. A partir de cada muestra de sangre se obtuvo plasma para la extracción y secuenciación del RNA viral y células mononucleares de sangre periférica (PBMCs) para la extracción de ADN genómico y genotipificación de los genes A, B y C del HLA de clase I. (Figura 9):

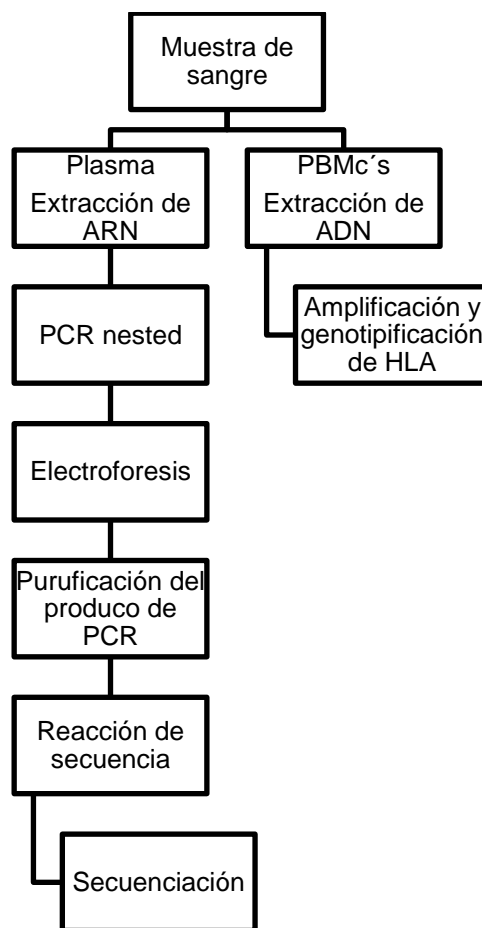


Figura 9: Descripción del trabajo experimental.

Extracción del RNA viral.

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). 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 ionica.

Amplificación del gen gag

Nested PCR o PCR anidada es una variante de la PCR convencional con la excepción de que esta comprende dos fases de amplificación con distintos primers en cada etapa. Esto con el fin de incrementar la sensibilidad y o la especificidad. La primera reacción permite amplificar regiones más extensas de ADN que contienen el segmento de interés. Después con el producto de la primera amplificación se realiza la segunda PCR para obtener la región específica. En este caso fue la mejor opción ya que se amplificó ADNc de la obtención de RNA vírico presente en una baja abundancia por una población heterogénea en la sangre. Este método reduce al mínimo la probabilidad de amplificar un segmento erróneo.

La primera reacción para la secuenciación del gen gag se realizó con los primers “forward” G23Fi universal (5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3') y “reverse” 2cRx (5'-TAACCCTGCGGGATGTGGTATTCC-3') previamente diseñados y la enzima SuperScriptR III One-Step RT-PCR PlatinumR Taq High Fidelity (Invitrogen, Carlsbad, CA) (Tabla 5).

Tabla 5: Descripción de los reactivos y cantidad de cada uno en la primera reacción de PCR.

Reactivos	Cantidad por muestra
dH ₂ O	8.5µl
Buffer 2x	12.5µl
PF (20 pmol/µl)	0.5µl
PR (20 pmol/µl)	0.5µl
SSIII RT/Platinum Taq Pol	0.5µl
RNA viral	2.5µl
Total	25µl

La reacción de amplificación fue llevada al termociclador con el siguiente programa:

Número de ciclos	T (C°)	Tiempo
1	55	30min
	94	2min
35	94	15s
	55	30s
	68	2min
1	68	5min
	4	∞

Para la segunda reacción se utilizaron los primers “forward” G1 (5'-GCAGGACTCGGCTTGCTGAA-3') y “reverse” G10 (5'-TATCATCTGCTCCTGTATC-3') previamente diseñados, y la enzima Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA) (Tabla 6).

Tabla 6: Descripción de los reactivos y cantidad de cada uno en la segunda reacción de PCR

Reactivos	Cantidad por muestra
dH ₂ O	8.5µl
Buffer 10x	12.5µl
MgCl ₂ (50µM)	1µl
dNTPs (1.25µM)	4µl
PF (20 pmol/µl)	0.1µl
PR (20 pmol/µl)	0.1µl
Platinum Taq Pol	0.25µl
cDNA 1ra vuelta	2.5µl
Total	25µl

Con el siguiente programa en el termociclador:

Número de ciclos	T (C°)	Tiempo
1	94	3min
35	94	30s
	56	30s
	72	2min
1	72	5min
	4	∞

Electroforesis

Posteriormente se procedió a la visualización de los productos de PCR por medio de electroforesis en gel de agarosa al 1% (Figura 10). El gel se preparó con 180ml de TBE (Tris/Borate/EDTA) 1X y 1.80g de agarosa, ocupando 18µl de gel red para revelar el DNA.

Se agregaron 3µl de ADN de la reacción de PCR obtenida previamente y 3µl de buffer de carga, también se agregó de forma independiente el marcador de peso

molecular ϕ x 174 RF DNA/*Hae* III Fragments (Invitrogen, Carlsbad, CA). Se dejó correr a 100volts y 400mA seleccionando como positivas aquellas bandas que se encuentren aproximadamente en las 1500 pb.

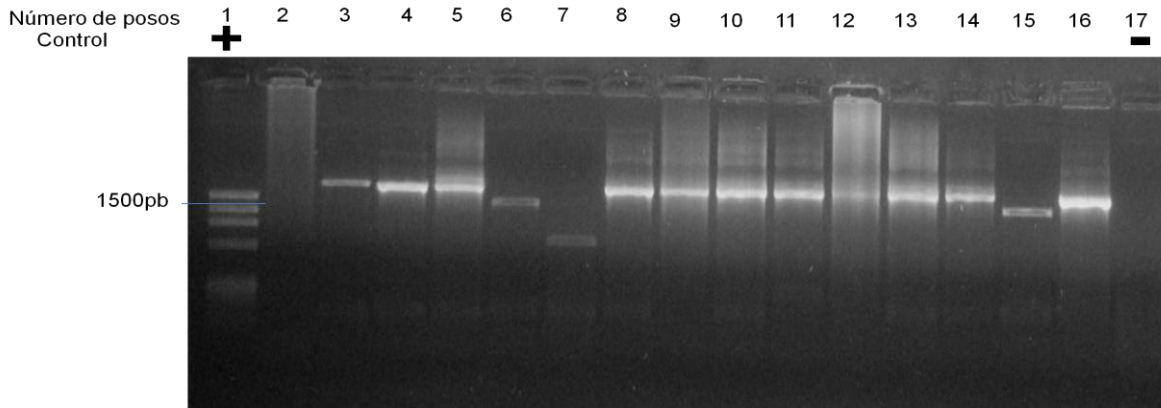


Figura 10: Muestra de un gel de agarosa al 1%. El marcador de peso molecular en la primera columna (izquierda). Se observan las muestras que corresponden al producto de amplificación del gen *Gag* (aproximadamente 1500 pb) y el control negativo al final (ala derecha).

Purificación del producto de PCR

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

Los procedimientos están diseñados para asegurar que no hay contaminación cruzada de una muestra a otra y permite la manipulación segura de muestras potencialmente infecciosas.

Secuenciación

Para la secuenciación del gen *gag* se empleó la química Big Dye Terminator v.3.1 cycle-sequencing (Life Technologies, Carlsbad, CA). Se utilizaron 8 primers

(obtenidos previamente) de secuenciación para cubrir toda la región de interés con al menos dos secuencias.

Cuatro primers “reverse”:

GS1 (5'-TTATCTAAAGCTTCCTTGGTGTCT-3')

GAS4 (5'-GGTTCTCTCATCTGGCCTGG-3')

GAS6 (5'AAAATAGTCTTACAATCTGG-3')

GA2274R (5'-TCTTTATTGTGACGAGGGGTCTG-3')

Cuatro primers “forward”:

G2 (5'-GCGGCGACTGGTGAGTA-3')

GAS3 (5'-CATCAATGAGGAAGCTGCAG-3')

GAS5 (5'-CTCTAAGAGCCGAGCAAGCT-3')

HPR1977 (5'-GTTAAGTGTTCATTGTGG-3')

Se empleó la siguiente mezcla de reactivos para cada una de las reacciones de secuencia, distribuyendo las reacciones en placas de 96 pozos:

Reactivos	Cantidad por muestra
dH2O	2µl
Buffer 5x	3.5µl
Primer (2pmol/µl)	2µl
ADN (10 a 40 ng)	2µl
Big Dye v3.1	0.5 µl
Total	10 µl

Con el siguiente programa en el termociclador:

Número de ciclos	T (C°)	Tiempo
1	96	1min
25	96	10s
	50	5s
	60	1min
1	4	∞

Posteriormente se purificó la placa con los productos de secuenciación. Brevemente, se agregó 1µL de acetato de sodio y 50 µL de etanol al 100% por pozo, se mezcló con un vortex y se procedió a realizar una centrifugación a

20000rpm durante 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 µL de etanol al 70%, mezclando con vortex y centrifugando a 20000 rpm durante 5 minutos. A continuación se llevo a cabo la eliminación del etanol por una inversión en papel y se centrifugo a 1500rpm 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 µL de formamida, se mezcló con el uso del vortex y se centrifugó a 1500 rpm 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 fue realizado con el software Sequencing Analysis v5.4 (Applied Biosystems) (Figura 11).

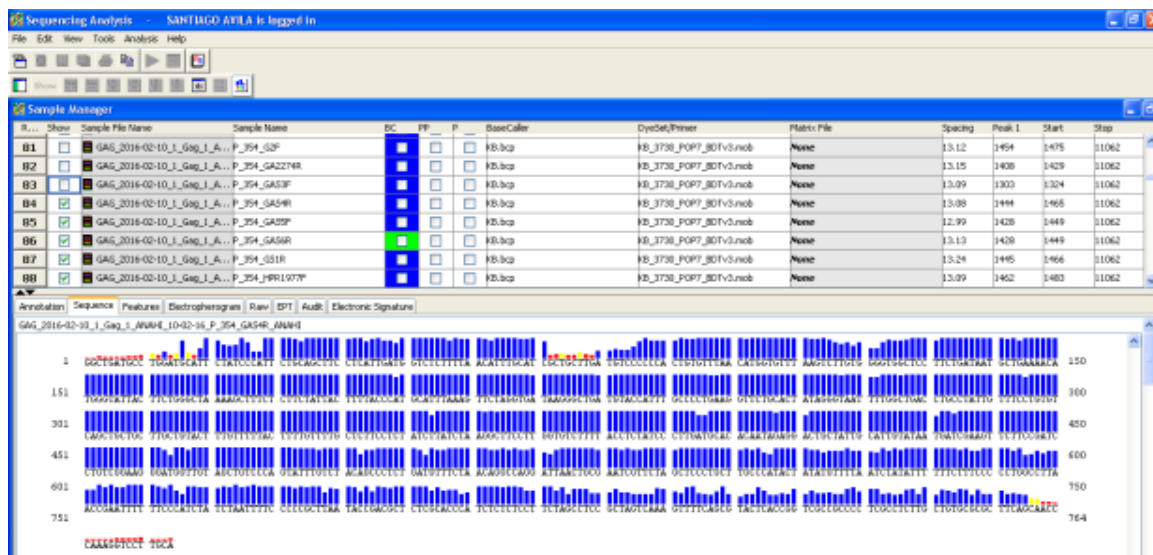


Figura 11. Sequencing Analysis v5.4. Se muestra pantalla representativa del análisis de una secuencia.

Posteriormente, las 8 secuencias obtenidas para cada paciente fueron ensambladas con el software Geneious (versión 5.6.7, Biomatters, Nueva Zelanda), obteniéndose una secuencia consenso para cada paciente (Figura 12).

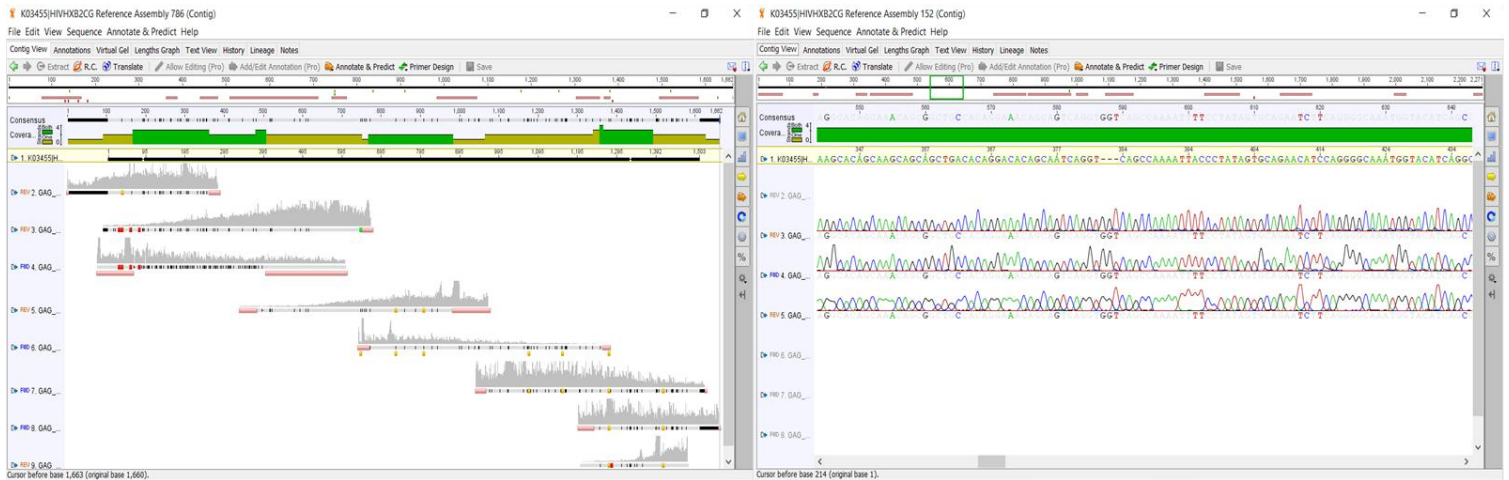


Figura 12. Software Geneious (versión 5.6.7). Se muestra pantalla representativa del ensamble de una muestra.

Las secuencias virales de todos los pacientes participantes fueron manualmente alineadas con el programa MEGA 5.05 (Molecular Evolutionary Genetics Analysis) (Figura 13), para posteriormente emplear el alineamiento para diversos análisis filogenéticos y evolutivos.

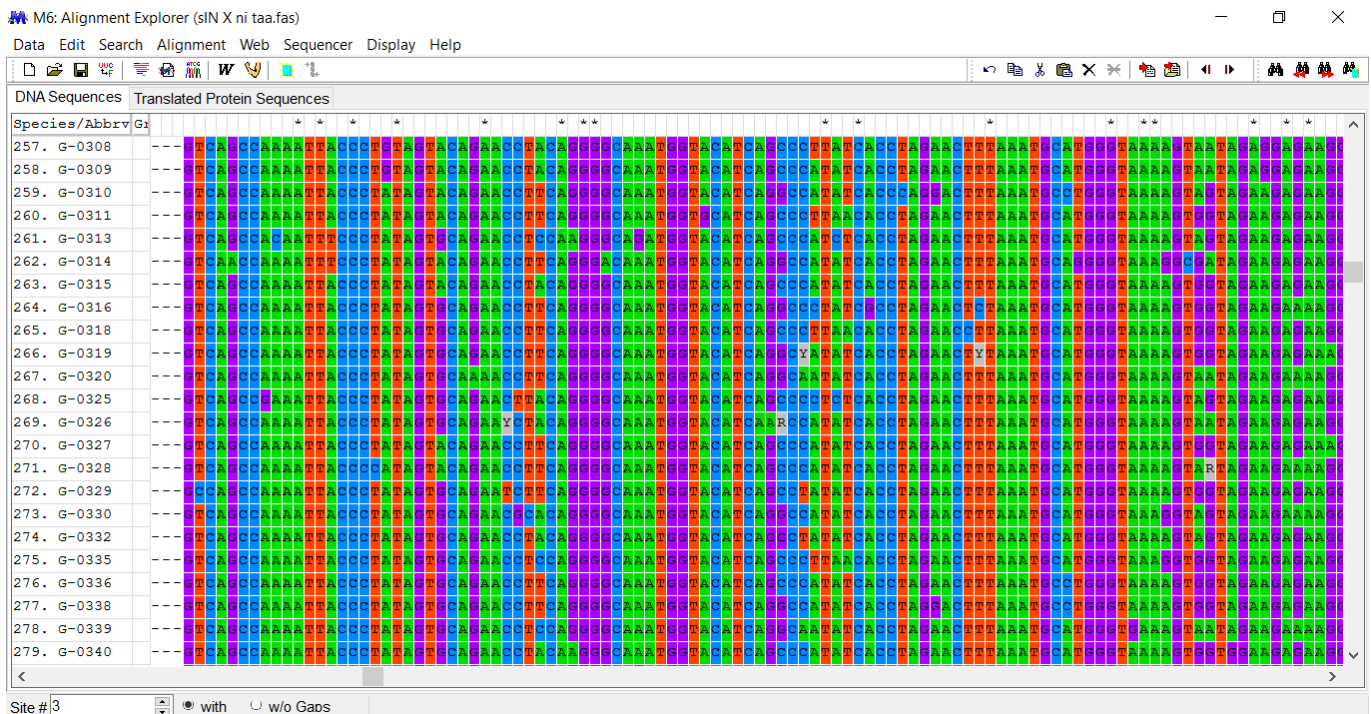


Figura 13. Programa MEGA 5.05 (Molecular Evolutionary Genetics Analysis). Se muestra una porción significativa del alineamiento de las secuencias.

Subtipificación de VIH

Para obtener la subtipificación del virus en cada una de las secuencias de *gag* en Centroamérica se recurrió a los programas RIP (Recombinant Identification Program) [25] y REGA HIV-1 Subtyping Tool - Version 3.0 [26], ambas disponibles en línea. Estas herramientas fueron diseñadas para identificar el subtipo específico de cada secuencia.

Análisis filogenético

A través del software MEGA 5.05 se realizó el alineamiento de las secuencias consenso para el gen *gag* de los virus de cada uno de los individuos participantes en el estudio. El alineamiento fue editado manualmente, y posteriormente se construyó un árbol filogenético usando como referencias:

- B.US.86.pNL4-3
- B.FR.83.HXB2
- B.CO.01.PCM074
- B.AR.00.ARMS008
- C.BR.92.BR025-d
- C.ZA.04.SK164B1
- D.CD.83.ELI
- F1.BR.93.93BR020_1
- F1.FR.96.MP411
-

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 replicas de bootstrap.

Análisis de posiciones bajo presión selectiva

Se identificaron las posiciones bajo presión selectiva positiva en el alineamiento de secuencias *gag* de los virus circulantes en Centroamérica. El modelo base-codón en evolución molecular permite inferir el sitio seleccionado en un alineamiento de secuencias homólogas para estimar el índice relativo de sustituciones sinónimas (dS) y no sinónimas (dN). Utilizando el servidor DataMonkey

(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 positiva.

Análisis de variabilidad genética

Se calculó la variabilidad de los aminoácidos en las secuencias, estimando la entropía de Shannon, con la herramienta Entropy de la base de datos de VIH los Alamos

(<http://www.VIH.lanl.gov/content/sequence/ENTROPY/entropy/entropy.html>).

Esta herramienta estima la entropía (cambio) de cada posición de cada aminoácido del alineamiento, mostrando la variación de la población. La entropía de Shannon se considera una medida que estima el cambio de un aminoácido dentro de un grupo de secuencias. Se realizaron comparaciones entre los países con la finalidad de observar el grado de variabilidad entre éstos.

Análisis de la distribución de alelos HLA

Se tomaron 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 Mesoamérica del CIENI, INER (Valenzuela-Ponce, et al, datos no publicados). Las frecuencias alélicas HLA fueron estimadas manualmente y comparadas entre los países centroamericanos mediante pruebas exactas de Fisher. Se realizó corrección para múltiples comparaciones mediante el método de razón de falsos positivos, empleando el estadístico “q”, ocupando la herramienta False Discovery Rate (FDR) de Microsoft (<http://research.microsoft.com/en-us/um/redmond/projects/MSCompBio/FalseDiscoveryRate/>). El nivel de significancia fue establecido como $p < 0.05$, $q < 0.2$.

Análisis de asociación HLA-polimorfismo de VIH

Para determinar asociaciones entre la expresión de alelos HLA específicos y la presencia de polimorfismos en posiciones determinadas del gen *gag* del VIH en el contexto inmunogenético de las poblaciones en Centroamericana, se utilizó un

modelo evolutivo denominado, redes de dependencia filogenética o PDN, desarrollado por Carlson *et al.* [20]. Este modelo arroja asociaciones HLA-polimorfismo del gen *gag*, corrigiendo para factores que frecuentemente confunden estas asociaciones, incluyendo el desequilibrio de ligamiento de los alelos HLA, coevolución de codones dentro de la secuencia de Gag y factores filogenéticos. El modelo se basa en el análisis 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. También se compara la verosimilitud de un modelo nulo que supone que la secuencia blanco no está bajo presión del atributo predictor, contra la de un modelo alternativo que supone que la variable blanco sí 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.

Para un mejor manejo en el análisis de las secuencias de *gag* se dividió el gen en tres, de acuerdo a las proteínas codificadas: Matriz (p 17; aminoácidos 1-132), Cápside(p24; 133-363) y región terminal (364-500) incluyendo p2 (364-377), Nucleocapside (p7; 378-432), p1 (433-448) y p6 (449-500).

RESULTADOS

Descripción del grupo de pacientes participantes

Se obtuvieron 997 muestras de pacientes infectados con VIH-1, vírgenes a tratamiento con antirretrovirales y en la etapa crónica de la infección. Las muestras provinieron de 5 países diferentes de Centroamérica: Belice, Panamá, Honduras, Guatemala y Nicaragua (Figura 14).

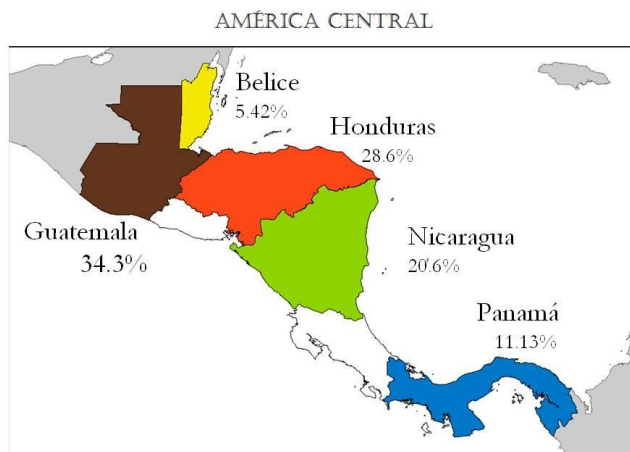


Figura 14: Distribución geográfica de la cohorte estudiada. Se muestra la proporción de pacientes por cada país. Cada uno con un código de color. 997 individuos centroamericanos infectados con VIH-1, vírgenes a tratamiento antirretroviral, procedentes de 5 países.

En los análisis de subtipo viral se encontró que de acuerdo a lo reportado por Hemelaar *et. al* [6] el subtipo B predomina en Centroamérica, con un 75% para Belice, 92.8% en Panamá, 92.3% Nicaragua, 89.7% en Honduras y 92.5% para Guatemala. Con respecto a los virus no-B se encontró en Belice pacientes con virus B/F1, B/C y C; en Panamá: C, D y C/D; Honduras fue el país que presentó mayor diversidad encontrándose cinco subtipos no-B: B/D/F1, B/F, B/D, D y A1; en Nicaragua sólo se encontró un subtipo adicional (subtipo D) y en Guatemala se identificaron los subtipos: B/D, C y C/D. Los subtipos C y D estuvieron presentes en la mayoría de los países de la región.

En cuanto a las proporciones por género se encontró mayor porcentaje de hombres infectados en todos los países comparados. La proporción de mujeres en todos los casos fue menor al 50%, pero con diferencias importantes entre países

con 42.5% en Belice, 38.4% en Nicaragua, 41.3% en Honduras, 43.5% en Guatemala, pero solo 25% en Panamá (Tabla 7). La variación con respecto a edad no fue significativa entre los países encontrando una media para Belice, Panamá, Nicaragua, Honduras y Guatemala en los 31-34 años (Tabla 7).

En cuanto a los valores de carga viral (viral load: VL) se presentaron medianas muy variables entre países; los individuos de Nicaragua con 65162 copias/ml fueron los más altos, mientras que los individuos de Panamá, con 21929 copias/ml, los más bajos. Los valores para el conteo de linfocitos T CD4+ variaron también entre los países, con una mediana de 417 células/ μ l en Belice, hasta una mediana de 210 células/ μ l en Guatemala (Tabla 7).

Tabla 7: Características clínicas de las cohorte

	Belice	Panamá	Nicaragua	Honduras	Guatemala
% de pacientes	5.42%	11.13%	20.6%	28.6%	34.3%
% de subtipo B	75%	92.80%	92.30%	89.70%	92.50%
Edad (mediana IQR)	33 (27-43)	31 (24-38)	32 (25-40)	34 (27-42)	34 (28-45)
% de Mujeres	42.5%	25%	38.4%	41.3%	43.5%
Carga viral (mediana IQR)	37189 (89604-8455)	21929 (7612-92179)	65162 (4470-257260)	38787 (6324-176309)	54803 (12907-191225)
CD4 (mediana IQR)	416.5 (265-564)	340 (223-505)	300 (306-358)	292 (96-494)	210 (102-314)

Predominancia de clusters específicos de país en el VIH circulante en Centroamérica

A partir del alineamiento de las 997 secuencias del gen *gag* de los virus provenientes de todos los países participantes, se realizó un árbol filogenético para evaluar las relaciones de distancia genética y los patrones de agrupamiento entre las secuencias. La topología del árbol (Figura 15) nos habla de las relaciones de parentesco entre las secuencias de los cinco países estudiados y la longitud de las ramas habla de la distancia genética que hay entre las secuencias. En primera instancia se observó un patrón de agrupamiento

fuertemente país-específico, pudiéndose identificar mayoritariamente clusters (grupos) compuestos principalmente por secuencias del mismo país. Así mismo, se observaron algunos clusters internacionales, conteniendo secuencias de dos o más países. En particular, se observó un alto número de secuencias intercaladas de Guatemala y Honduras en varios clusters.

Como era esperado, secuencias de subtipo no-B se agruparon con las secuencias de referencia correspondientes, como es el caso de Belice que mostró un grupo cercano a las referencias de subtipos C y F1. De cualquier manera, la mayoría de las secuencias se agruparon con referencias de subtipo B, confirmando que este subtipo es el dominante en Centroamérica.

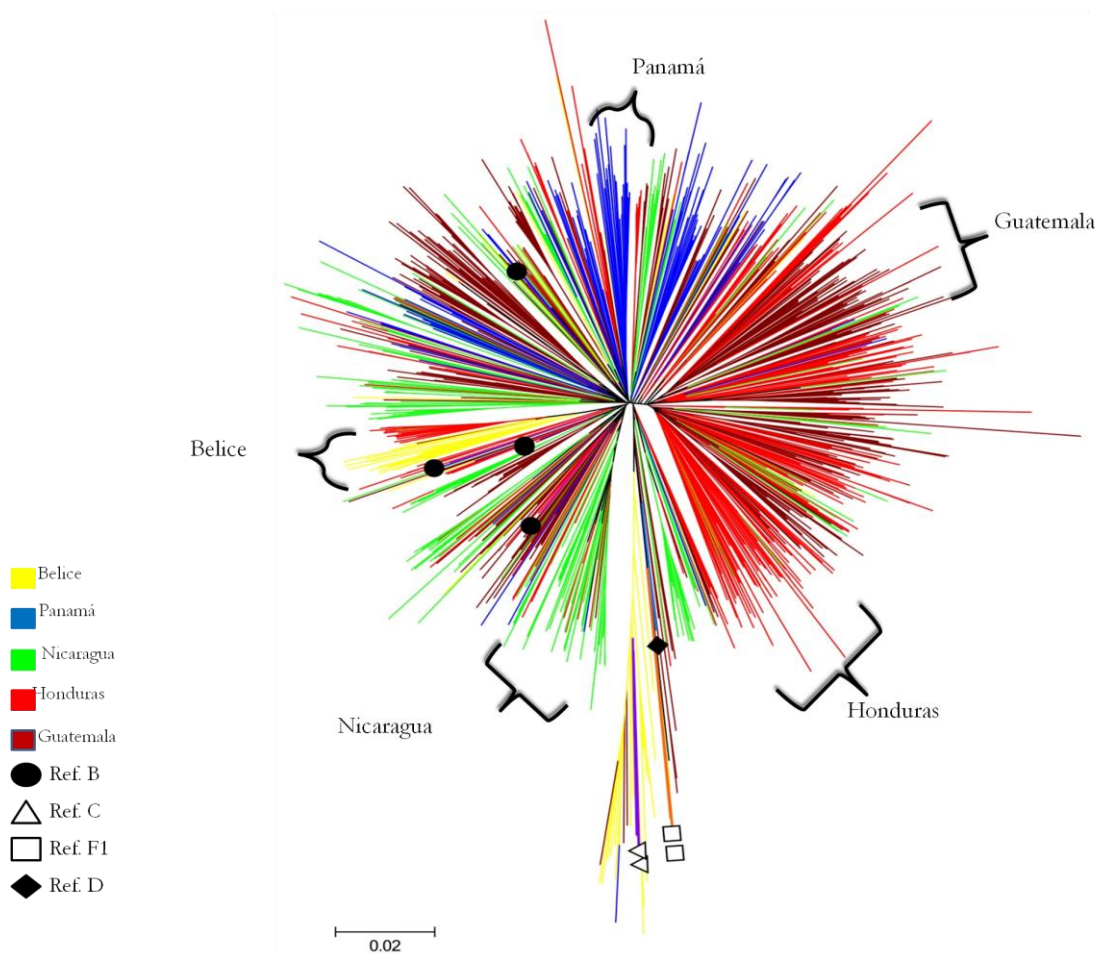


Figura 15. Relaciones filogenéticas de las secuencias de *gag* en individuos vírgenes a tratamiento antirretroviral de cinco países de Centroamérica. El árbol se construyó por el método de Neighbor Joining con el modelo de dos parametros de Kimura a partir del alineamiento de 997 secuencias del gen *gag* del VIH-1 de los individuos participantes en el estudio. La confiabilidad de los arboles

se evaluó mediante 1000 repeticiones de bootstrap. Se utilizó un código de colores para marcar cada país así como íconos que identifican la posición de las nueve referencias usadas. Se señalan algunos clusters notables país-específicos.

Distintos patrones de posiciones bajo presión selectiva positiva en las secuencias de gag de los virus de los diferentes países

Para analizar el papel de la selección en la diversidad genética de *gag* del VIH circulante en Centroamérica y evaluar posibles diferencias regionales se identificaron posiciones bajo presión selectiva positiva por comparación de la razón de sustituciones sinónimas y no sinónimas. Se obtuvo una gráfica por país mostrando los sitios de aminoácidos bajo presión de selección positiva y negativa (Figura 16).

Se pudieron localizar en cada país sitios específicos bajo presión selectiva positiva (tendencia al cambio con respecto al consenso), donde la razón de sustituciones no sinónimas fue significativamente mayor que la de sustituciones sinónimas ($dN>dS$). En el total de los cinco países analizados se obtuvieron 222 sitios bajo selección positiva, con un nivel de significancia de $p<0.01$ (Tabla 8).

Se observó que hay más posiciones con tendencia a conservar el aminoácido presente en la secuencia consenso que aquéllas que tienden a cambiar. Esto confirma lo 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 del VIH-1.

En general, la capacidad de identificación de sitios bajo presión selectiva fue proporcional al número de secuencias analizadas por país. Así, en Honduras y Guatemala se encontraron más sitios bajo presión positiva, 60 y 57 sitios respectivamente (Figura 16). Interesantemente, en Belice se localizaron 10 sitios únicos (que no presentó ningún otro país), seis de estos sitios en la región terminal de Gag (p2p7p1p6). Al comparar todas las proteínas de gag, p2p7p1p6 se identificó como la sección con más sitios bajo presión selectiva positiva. Guatemala y Honduras también presentaron sitios únicos bajo presión positiva, los cuales fueron localizados a lo largo de toda la secuencia, concentrándose al final

(p2p7p1p6) y al inicio (p17) de *gag*, en contraste con Panamá y Nicaragua, donde todos los sitios de selección bajo presión positiva fueron compartidos con otros países.

Al ordenar las proteínas por el número de sitios bajo presión positiva observados, éstas quedaron de la siguiente manera: p2p7p1p6>p17>p24.

El patrón que se genera en los sitios de selección bajo presión positiva indica sitios más polimórficos que otros como lo son p17 y p2p7p1p6, mientras que p24 se sigue caracterizando por ser una de las zonas más conservadas.

Tabla 8. Sitios bajo presión selectiva positiva por país, en cada posición de *gag*.

	Número de sitios bajo presión selectiva positiva*																		
	P17			P24			P2			P7			P1			P6			Total
Centroamérica	27			28			7			16			3			27			108**
País	C	U	T	C	U	T	C	U	T	C	U	T	C	U	T	C	U	T	
Belice	9	0	9	1	4	5	0	0	2	4	4	6	0	1	1	1	1	2	24
Panamá	13	0	13	7	0	7	2	0	2	4	0	4	1	0	1	8	0	8	35
Nicaragua	18	0	18	8	0	8	5	0	5	5	0	5	1	0	0	11	0	11	46
Honduras	16	4	20	11	1	12	4	1	5	2	4	6	1	1	6	9	15	60	
Guatemala	18	2	20	11	2	13	4	0	4	6	2	8	1	0	1	10	4	14	57

*Obtenidos por razón de sustituciones sinónimas (dS) vs. no sinónimas (dN) con el servidor DataMonkey (dN>dS, p<0.01). C, sitios que se comparten con al menos uno de los otros países; U, sitios únicos del país; T, Total.

**Número de sitios totales en la secuencia consenso para Centroamérica.

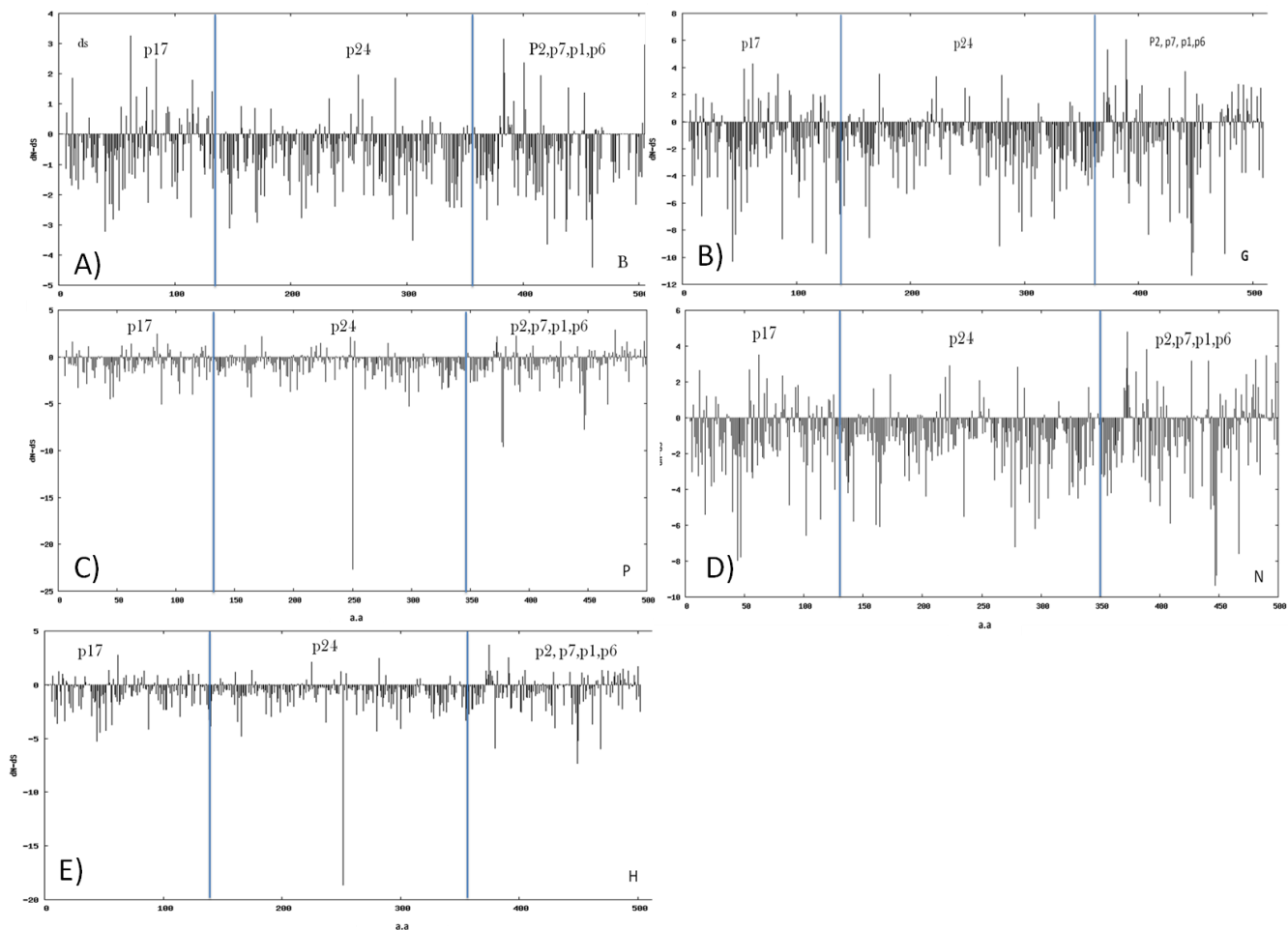


Figura 16. Mutaciones sinónimas y no sinónimas en la secuencia consenso para cada país de Centroamérica. Se muestran las mutaciones sinónimas con dirección hacia arriba y las no sinónimas con dirección hacia abajo en cada uno de los aminoácidos. La secuencia de *gag* está dividida por líneas azules que delimitan las proteínas agrupando las últimas cuatro en un solo bloque. A)Belize, B)Guatemala, C)Panamá, D)Nicaragua, E)Honduras

Diferentes patrones de variabilidad entre las secuencias de los virus circulantes en los diferentes países

La entropía de Shannon es una medida cuantitativa de incertidumbre aplicable a alineamientos de secuencias de DNA. Esta medida puede usarse para estimar la variabilidad en cada posición del alineamiento tomando en cuenta tanto la frecuencia, como el número de bases presentes.

Para explorar la variabilidad del VIH circulante en los países participantes, se realizaron comparaciones de entropía de Shannon entre los alineamientos del gen *gag* de los virus circulantes en cada país con la finalidad de observar el grado de variabilidad entre estos.

En los gráficos obtenidos para cada país, se observó que consistentemente la zona correspondiente a p24 fue la que presentó relativamente menor cantidad de sitios con alta entropía y por el contrario las zonas de p17 y p2p7p1p6 presentaron mayor cantidad de sitios con alta entropía (Figura 17). De esta manera se corroboran observaciones previas que sugieren una mayor conservación de p24 con respecto a otras proteínas de *gag*.

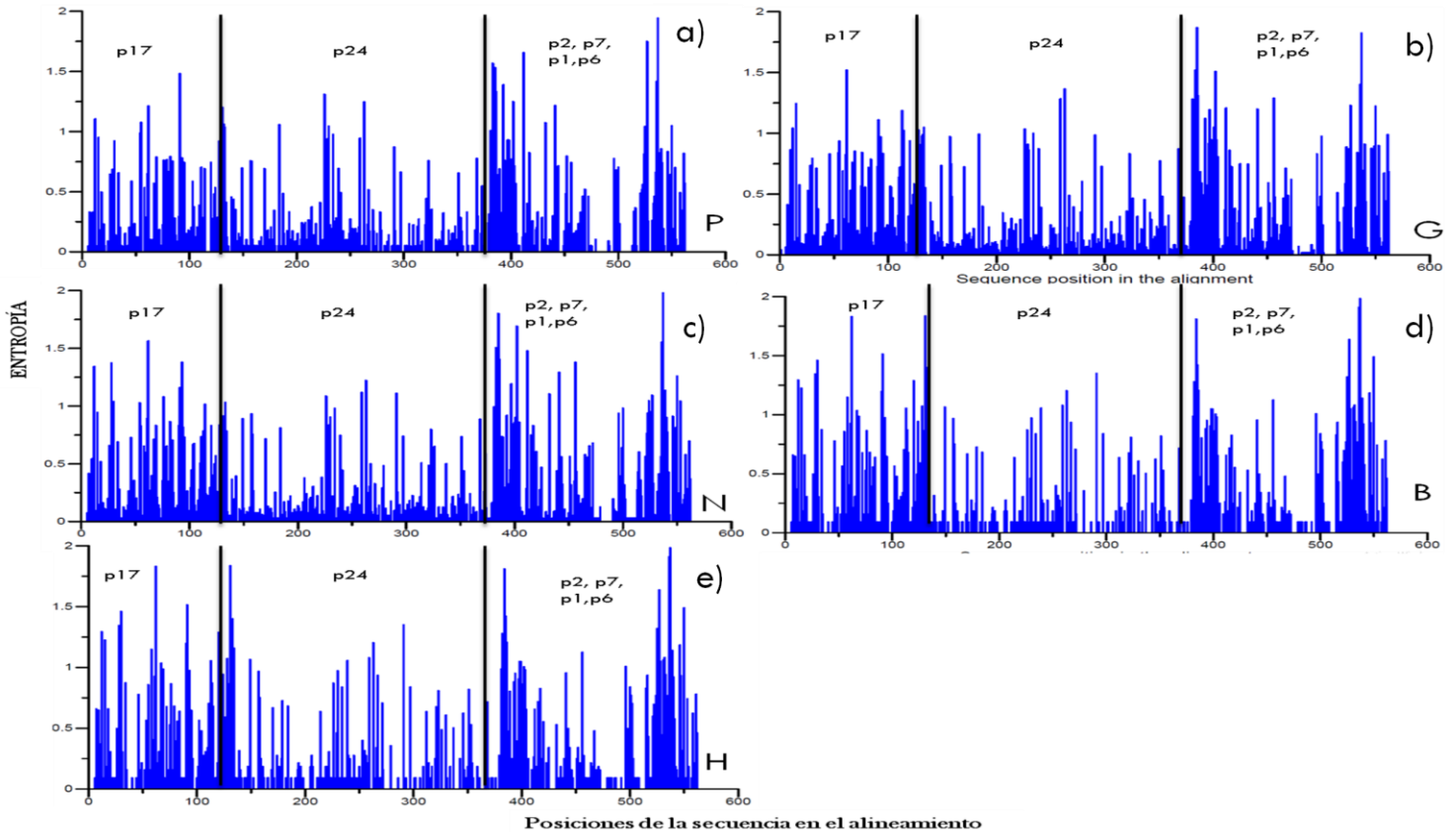


Figura17. Entropía de Shannon en la secuencia del gen *gag* de VIH circulante en los países centroamericanos. Se muestran los valores de entropía de Shannon por nucleótido en el alineamiento del país correspondiente, estimados con la herramienta Entropy-One de la Base de Datos de VIH de Los Alamos. Las proteínas de Gag se encuentran divididas con líneas naranjas en 3 zonas correspondientes a p17, p24 y p2p7p1p6. a) Panamá, b) Guatemala, c) Nicaragua, d) Belice, e) Honduras.

Con el fin de evaluar la variabilidad de las secuencias entre países se realizaron comparaciones de entropía de Shannon por posición entre ellos. En general se observaron mayores diferencias de entropía en la zona terminal de *gag*

(p2p7p1p6) y en p17, mientras que la región de p24 mostró un menor número de diferencias entre países (Figura 18).

Interesantemente, Nicaragua presentó mayor número de posiciones con mayor entropía ($p < 0.05$) en p17 que Panamá y Guatemala. Así mismo, Panamá mostró un bajo número de posiciones con alta entropía en p17, al compararse con Guatemala y Nicaragua. Adicionalmente en las comparaciones de Belice con los otros países muestra claras diferencias en la zona final de *gag*, con alta entropía.

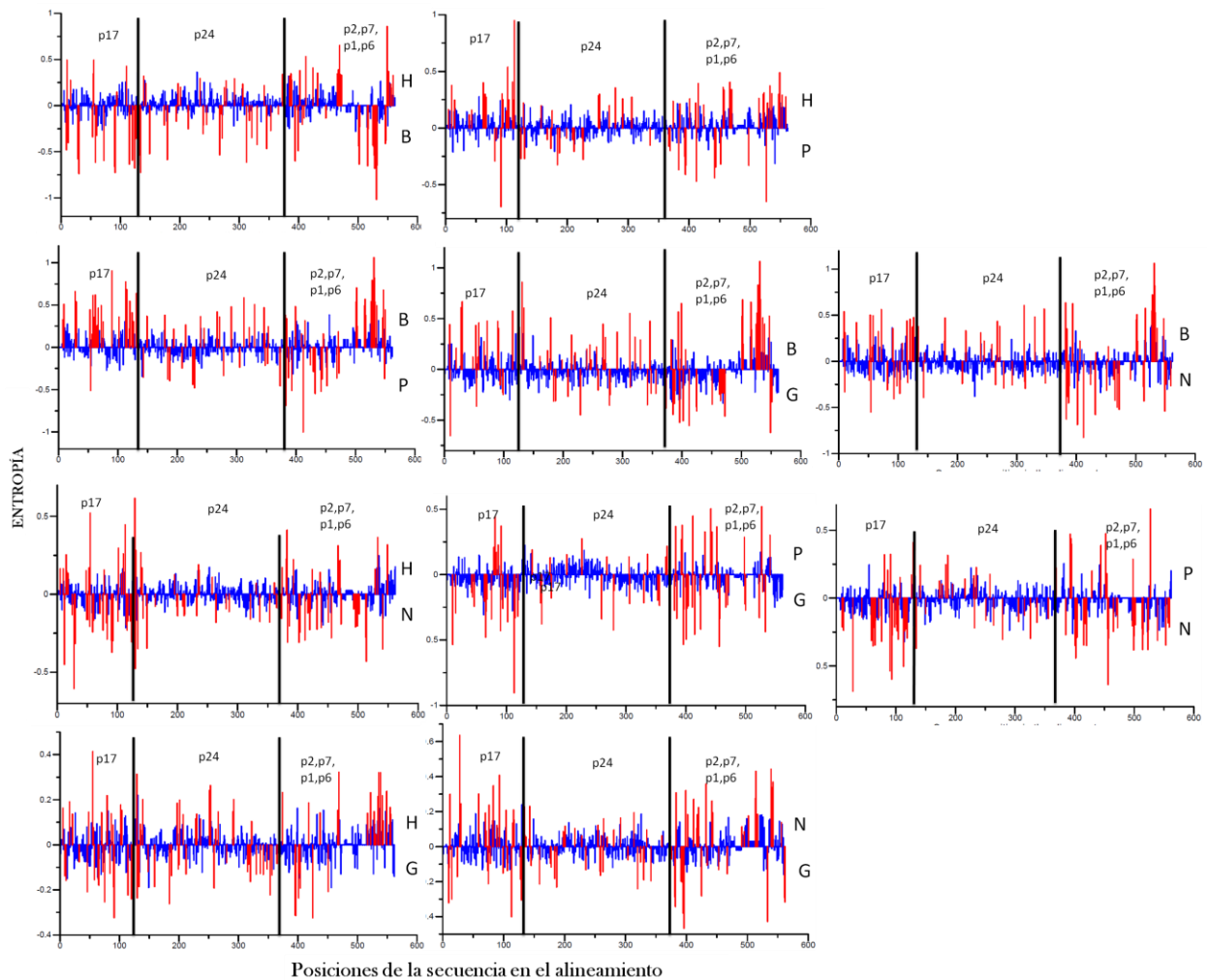


Figura 18. Comparación de entropía de Shannon por posición de *gag* del VIH entre los países centroamericanos. Se muestran las diferencias de entropía de Shannon por posición de nucleótido entre pares de países, estimadas con la herramienta Entropy-Two, disponible en la Base de Datos de VIH de Los Alamos. a) Panamá vs Guatemala; b) Panamá vs Nicaragua; c) Nicaragua vs Guatemala vs Nicaragua.

Alta diversidad en distribución de frecuencias HLA entre las poblaciones de los países centroamericanos

Con el objeto de evaluar el impacto de la selección por HLA en el VIH circulante en los distintos países centroamericanos, se estimaron y compararon las distribuciones de frecuencias de los genes *A*, *B* y *C* del HLA de clase I en cada país (Figuras 19-21). Las distribuciones de frecuencias se compararon con las observadas en una cohorte mayoritariamente caucásica de Canadá (cohorte HOMER, HAART Observational Medical Evaluation and Research) [27]. Los alelos HLA observados en mayor frecuencia en cada país se muestran en la Tabla 9.

Se identificaron 54 alelos HLA-A en diferentes frecuencias para los países de Centroamérica. El alelo A*02:01 fue el más frecuente en la mayoría de los países, excepto en Panamá, donde ocupó el sitio número seis. El segundo alelo más frecuente fue A*24:02 en todos los países analizados. Se notó menor variabilidad entre Honduras, Guatemala y Nicaragua con respecto a Belice y Panamá. Característicamente Panamá mostró a los alelos A*33:01, A*24:01 y A*24:03 con mayor frecuencia que otros países, donde estos alelos HLA presentaron una frecuencia muy baja o nula como en el caso del último alelo mencionado que no se encontró en los otros países estudiados presentándose en HOMER pero con una baja frecuencia. Al comparar la cohorte de HOMER se observaron frecuencias significativamente diferentes. Para esta cohorte, los alelos de mayor frecuencia fueron el grupo A*02, localizado también en todos los países centroamericanos con una baja frecuencia, A*01:01 y A*03:01, se comparten sólo con Honduras (Figura 19).

En cuanto a HLA*B se identificaron 99 alelos diferentes. Los alelos de mayor frecuencia para *B* fueron B*53:01, en Belice y Panamá con una frecuencia muy baja para Honduras; B*35:01, teniendo una proporción muy baja en Honduras mientras que en HOMER no se observó; B*40:02, en mayor frecuencia para Nicaragua y Guatemala y casi nulo en Honduras, entre otros. Panamá presentó a B*07:02 con mayor frecuencia que en cualquier otro país. Honduras presentó al alelo B*27:05 con una frecuencia alta compartiendo este sólo con Belice que lo

presentó en una frecuencia muy baja. Belice mostró diferentes frecuencias con respecto a los otros países analizados de Centroamérica.

Para HLA*C se obtuvieron 37 alelos diferentes. Este alelo presentó menor variación en los países analizados. Belice destacó con los alelos C*03:04 siendo este el segundo alelo con mayor frecuencia en su población, C*16:01, con frecuencias similares a las de Panamá y Honduras, y C*17:01 localizado los otros países con frecuencias muy bajas. Los alelos con mayor frecuencia en C fueron: C*04:01 y C*07:02 para todas las poblaciones incluyendo HOMER. En la tabla 9 se pueden apreciar los 5 primeros alelos (A, B y C) con mayor frecuencia para cada población.

Tabla 9. Alelos presentes en mayor frecuencia en cada país.

Belice	Panamá	Nicaragua	Honduras	Guatemala	Homer
A*02:01 A*24:02 A*23:01 A*30:02 A*31:01	A*33:01 A*24:02 A*24:03 A*24:01 A*29:02	A*02:01 A*24:02 A*68:63 A*03:01 A*02:06	A*02:01 A*24:02 A*03:01 A*68:01 A*01:01	A*02:01 A*24:02 A*02:06 A*68:03 A*31:01	A*01:01 A*03:01 A*24:02 A*11:01
B*53:01 B*44:03 B*35:01 B*15:10 B*58:01	B*35:01 B*51:01 B*07:02 B*40:02 B*35:43	B*35:01 B*40:02 B*07:02 B*35:43 B*14:02	B*27:05 B*07:02 B*40:01 B*50:01 B*08:01	B*35:01 B*40:02 B*39:05 B*35:12 B*39:08	B*07:02 B*08:01 B*44:02 B*15:01
C*04:01 C*03:04 C*07:01 C*07:02 C*17:01	C*04:01 C*07:02 C*01:02 C*07:01 C*03:05	C*04:01 C*07:02 C*01:02 C*07:01 C*03:05	C*07:02 C*04:01 C*07:01 C*03:05 C*01:02	C*07:02 C*04:01 C*01:02 C*03:04 C*07:01	C*07:02 C*07:01 C*03:04 C*05:01 C*04:01

Diversidad de Alelos A

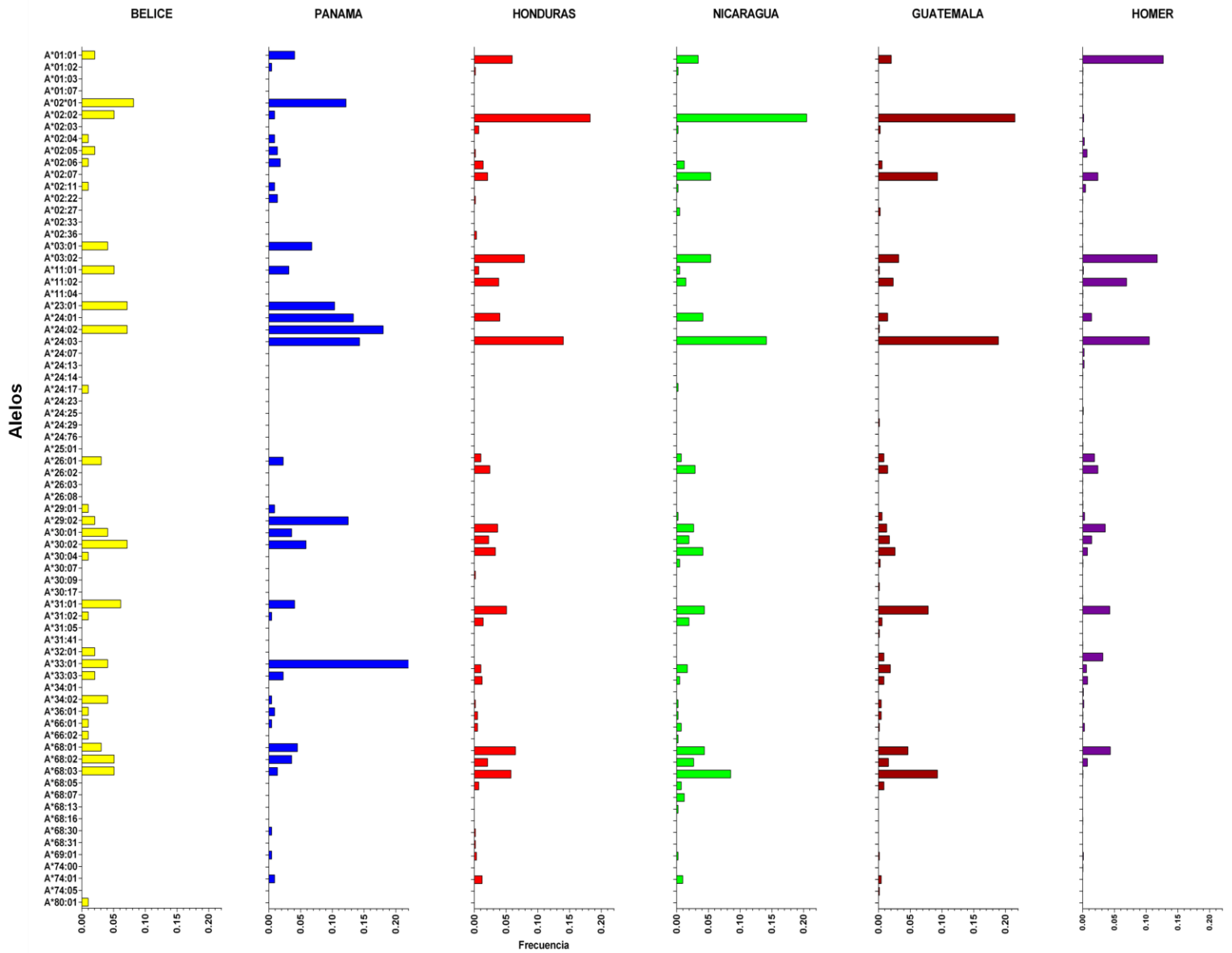


Figura 19. Distribución de la frecuencia de HLA-A en las poblaciones de Centroamérica. Se comparan las frecuencias con las observadas en una cohorte Canadiense mayoritariamente caucásica (HOMER) [27].

Diversidad de Alelos B

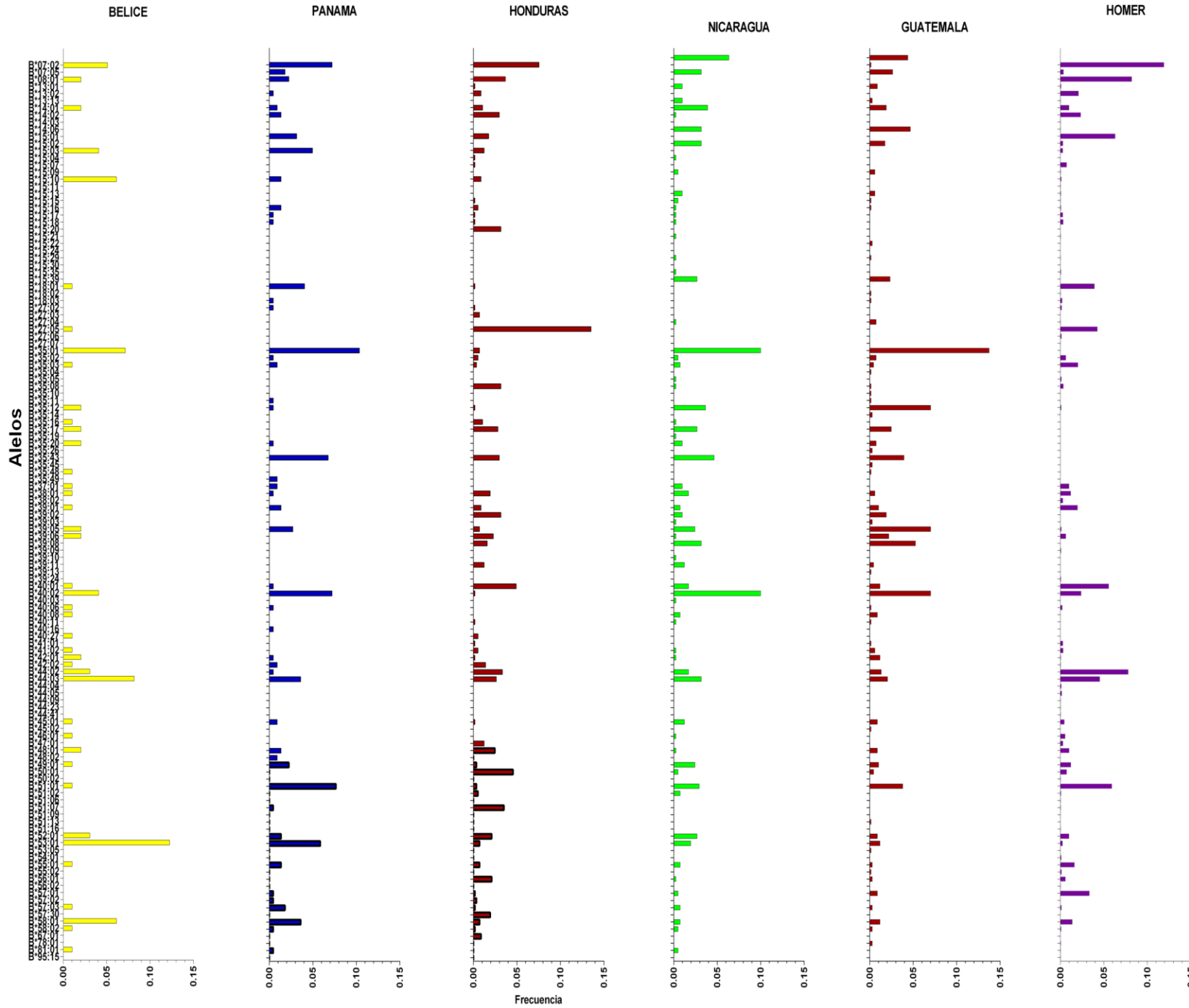


Figura 20. Distribución de la frecuencia de HLA-B en las poblaciones de Centroamérica. Se comparan las frecuencias con las observadas en una cohorte Canadiense mayoritariamente caucásica (HOMER).

Diversidad de Alelos C

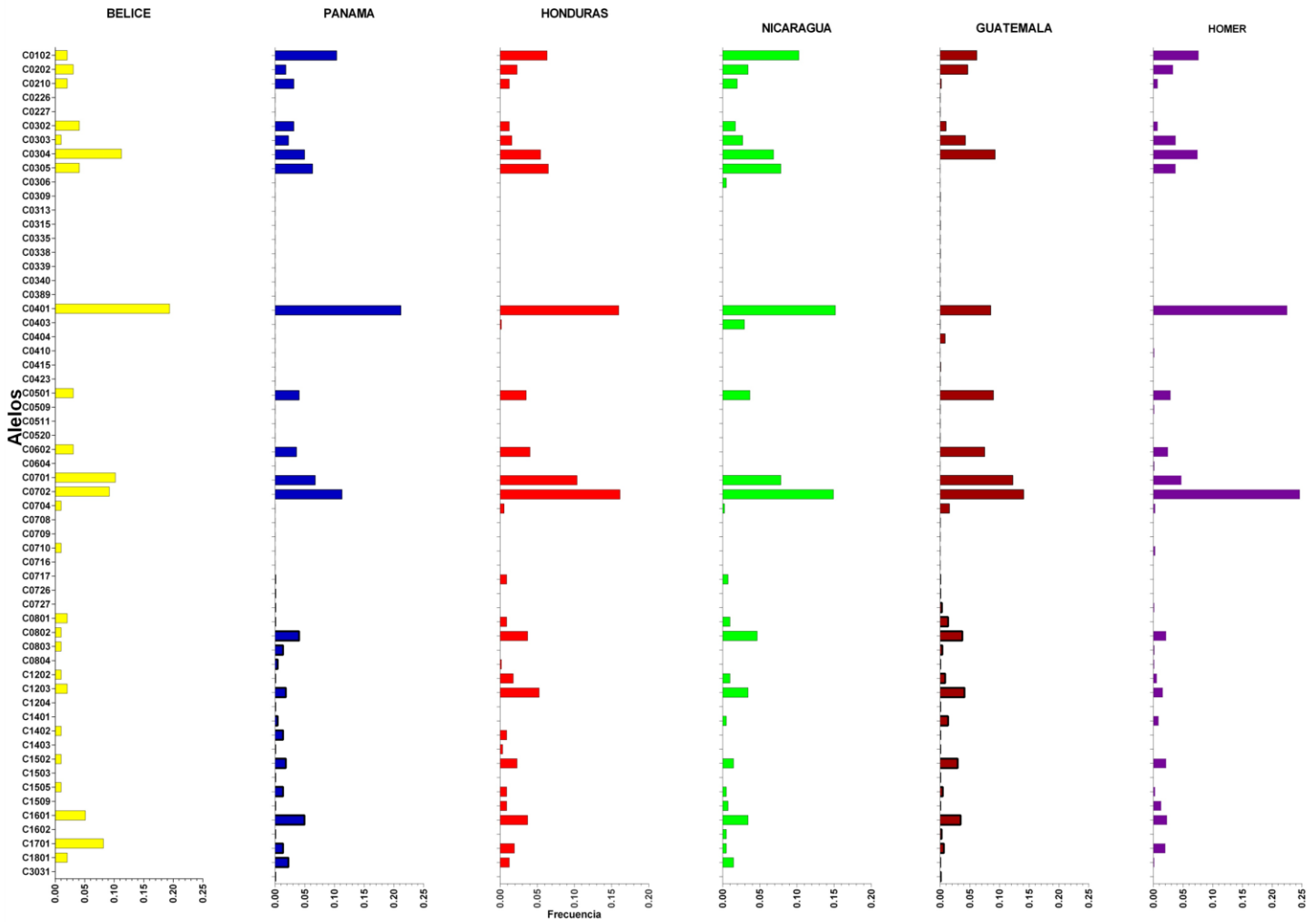


Figura 21: Distribución de la frecuencia de HLA-C en las poblaciones de Centroamérica. Se comparan las frecuencias con las observadas en una cohorte Canadiense mayoritariamente caucásica (HOMER).

Asociaciones HLA-polimorfismo en *gag*

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 inmunogénético único de las poblaciones de Centroamérica, se utilizaron datos de asociaciones entre la expresión de alelos HLA específicos y la presencia de polimorfismos en el virus usando un modelo evolutivo llamado Redes de Dependencia Filogenética (PDN).

Con esta herramienta fue posible identificar asociaciones HLA-polimorfismo de VIH corrigiendo para factores que típicamente pueden confundir las asociaciones,

como efectos filogenéticos, coevolución de codones del VIH y desequilibrio de ligamiento de los alelos HLA.

Se obtuvieron 41 asociaciones significativamente relevantes, en 26 posiciones dentro de la secuencia consenso para las poblaciones estudiadas de las cuales 19 fueron entre el alelo HLA selector y el aminoácido adaptado, es decir, aquél que permite que el VIH escape a la respuesta restringida por el alelo HLA en cuestión (Tabla 10). El resto de las asociaciones (22) fueron entre el alelo HLA relevante y el aminoácido no adaptado, es decir, aquél cuya presencia hace que el virus sea susceptible a la respuesta restringida por el alelo HLA en cuestión.

De éstas se localizaron 10 asociaciones HLA-polimorfismo en ocho sitios de p17, de las cuales 4 fueron con aminoácidos adaptados; 13 asociaciones en diez sitios de p24 con 7 con aminoácidos adaptados; cero en p2; 9 asociaciones en cuatro sitios de p7 de las cuales 4 fueron con aminoácidos adaptados; 2 asociaciones en un sitio de p1, 1 con aminoácido adaptado y 2 asociaciones en dos sitios p6, una con aminoácido adaptado.

Al hacer la comparación entre las 41 asociaciones HLA-polimorfismo de VIH obtenidas por el modelo de PDN en la cohorte de Centroamérica y asociaciones reportadas previamente en la literatura en pacientes caucásicos [28], se encontraron 18 asociaciones en 8 posiciones que coincidían en las dos cohortes: 30, 242, 280, 397, 398, 403, 427, 436. En 5 asociaciones ubicadas en 4 posiciones diferentes, se identificó una diferencia en el alelo HLA o el aminoácido asociado entre poblaciones. Por ejemplo, en el caso de la posición 147 se observaron dos asociaciones con los mismos aminoácidos adaptados y no adaptados para las dos cohortes comparadas, pero con el alelo A*25:01 asociado en Centroamérica y el alelo B*13 asociado en caucásicos. Otro ejemplo es, en las posiciones 68 y 119 con asociaciones con los alelos B*52:01 y B*35:01 respectivamente, y un aminoácido no adaptado para Centroamérica, en contraste con caucásicos que presentan un aminoácido adaptado diferente asociado con A*02 en ambas posiciones [20].

Del análisis se identificarón 17 posiciones en *gag* que mostraron 22 asociaciones con HLA en la cohorte centroamericana que no han sido reportadas en otras

cohortes, mostrando características particulares de la evolución asociada al HLA-1 del VIH-1 en Centroamérica (Tabla 11).

De estas asociaciones se encontraron 23 posibles nuevos epítopes que hasta ahora, no han sido publicados con anterioridad en otras cohortes. Estos resultados en conjunto, muestran características particulares de la evolución asociada a HLA del VIH en el contexto inmunogenético particular, previamente no estudiado, de las poblaciones centroamericanas.

Tabla 10: Asociaciones HLA-polimorfismo de VIH en el gen *gag* significativas en una cohorte de 997 individuos centroamericanos, de acuerdo al modelo de redes de dependencia filogenética.

Posición y aminoácido	Alelo predictor ₁	seleccionado ₂	Dirección ₃	Epitope Adaptado ₄	EpitopesPublicados ₅	Valor de P ₆	Valor de Q ₇
A*02:06	342@T	No Adaptado	SLEEMMTA			6.87E-06	0.02270214
A*02:06	342@S	Adaptado	SLEEMMTA			3.41E-05	0.0911271
A*03:01	28@K	No Adaptado	RLRPGGKKQYK	(KIRLRPGGK),(RLRPGGKKK),(RLRPGGKKKY)		4.81E-15	2.70E-10
A*03:01	28@Q	Adaptado	RLRPGGKK	(KIRLRPGGK),(RLRPGGKKK),(RLRPGGKKKY)		1.34E-09	1.25E-05
A*03:01	252@G	Adaptado	PIPVGEIYK			0.00011347	0.19852104
A*11:01	91@K	No Adaptado	ATLYCVHQR	(TLYCVHQQ)		5.77E-05	0.12961112
A*11:01	15@R	Adaptado	SVLSGGELDR	(KIRLRPGGK)		9.99E-05	0.18711308
A*24:02	30@R	Adaptado	KYRLKHIVW	(KYKLBHIVW)		3.97E-05	0.09685205
A*24:02	206@N	No Adaptado	AMQMLKETI			8.10E-05	0.15693499
A*25:01	147@L	Adaptado	SPRTLNAW	(QAISRTLNAW)		5.35E-07	0.0025057
A*25:01	147@I	No Adaptado	SPRTLNAW	(QAISRTLNAW)		1.89E-06	0.008172
A*29:02	472@P	Adaptado	KQEPIDKELY			0.00011807	0.19852104
A*30:01	28@K	No Adaptado	RLRPGGKKQYK	(KIRLRPGGK),(RLRPGGKKK),(RLRPGGKKKY)		8.07E-05	0.15693499
A*30:01	28@R	Adaptado	RLRPGGKKRY	(KIRLRPGGK),(RLRPGGKKK),(RLRPGGKKKY)		4.07E-05	0.09515189
A*31:01	397@R	Adaptado	KIVKCFNCGR			4.52E-09	3.63E-05
A*31:01	397@K	No Adaptado	KIVKCFNCGR			6.27E-09	4.40E-05
A*31:01	403@K	Adaptado	KNCRAPRK			3.65E-08	0.00022809
A*31:01	403@R	No Adaptado	KNCRAPRK			6.13E-08	0.00034406
A*33:01	221@G	No Adaptado	RPVAPGQMR			7.90E-05	0.15693499
B*13:02	436@R	Adaptado	RQANFLGRI	(RQANFLGKI)		7.75E-05	0.17417417
B*13:02	436@K	No Adaptado	RQANFLGRI	(RQANFLGKI)		9.42E-05	0.18896253
B*15:01	26@R	Adaptado	RLRPGGRKKY			7.47E-10	8.39E-06
B*15:01	215@I	Adaptado	AEWDRIHPVHA			8.10E-05	0.15693499
B*15:01	26@K	No Adaptado	RLRPGGRKKY			2.07E-05	0.0553326
B*39:08	238@G	No Adaptado	SDIAETTSTL			5.21E-05	0.12190525
B*40:02	427@T	No Adaptado	NERQANFL	(TERQANFL)		4.58E-10	6.43E-06
B*40:02	427@N	Adaptado	NERQANFL	(TERQANFL)		2.33E-06	0.00912008
B*40:02	398@D	Adaptado	KDGHIAKNC			0.00012371	0.19852104
B*40:02	398@E	No Adaptado	GHIAKNCRA			4.85E-09	3.41E-05
B*49:01	12@E	No Adaptado	GKLDLWEKI			3.90E-05	0.09685205
B*52:01	280@S	Adaptado	RMYPSSILDI	(GLNKIVRMY),(RMYSPTSI)		3.34E-05	0.0911271
B*52:01	68@L	No Adaptado	RQILGQLQPSI			0.00012108	0.19852104
B*53:01	119@A	No Adaptado	TGNSSQVSQNY			8.67E-05	0.18044165
B*57:01	242@N	Adaptado	TSNLQEQIAW	(TSTLQEIQIW,TSTLQEQIAW)		2.44E-06	0.00912008
B*57:01	242@T	No Adaptado	TSNLQEQIAW	(TSTLQEIQIW,TSTLQEQIAW)		1.35E-05	0.04004287
B*58:01	242@N	Adaptado	TSNLQEQIAW	(TSTLQEIQIW,TSTLQEQIAW)		3.29E-12	9.24E-08
B*58:01	242@T	No Adaptado	TSNLQEQIAW	(TSTLQEIQIW,TSTLQEQIAW)		3.06E-10	5.74E-06
C*02:10	485@P	No Adaptado	LYHLTSLRSLF			4.74E-06	0.01665045
C*03:05	398@E	No Adaptado	KVGHIAKNCRA			2.49E-07	0.00127101
C*07:01	378@V	No Adaptado	VTNSATIMMQR			1.14E-05	0.03550032
C*08:02	303@V	Adaptado	DYVDRFYKVL			0.00010905	0.19757877

1. Alelo HLA predictor asociado a la respuesta inmune ejerciendo presión selectiva sobre la posición indicada...
2. Sitio en la secuencia de *gag* dónde se localiza el aminoácido asociado a selección por el alelo HLA, empleando como referencia de numeración la secuencia HXB2.
3. Dirección que marca el sentido de la asociación, si es adaptado el alelo siempre va a localizar a ese aminoácido en particular, si es no adaptado el alelo nunca se va a encontrar con ese aminoácido. Los aminoácidos adaptados son los que permiten al virus escapar, los enriquecidos en presencia del alelo HLA asociado; representan la adaptación/escape del virus a la respuesta restringida por dicho HLA. Los aminoácidos no adaptados, son los susceptibles al reconocimiento inmunológico, los enriquecidos en ausencia del alelo HLA.
4. Epítotope adaptado refleja el péptido con la mutación de escape asociada a la respuesta de linfocitos T citotóxicos restringida por el alelo HLA en cuestión.
5. Epítotope publicados son aquéllos que se han descrito y comprobados experimentalmente con anterioridad en la región donde se encuentra la posición.
6. Valor de $p < 0.05$
7. Valor de $q < 0.2$

Tabla 11: Asociaciones HLA-polimorfismo únicas en Centroamérica.

Alelo predictor ₁	Posición y aminoácido seleccionado ₂		Epitope nuevo		
	Dirección ₃	Adaptado ₄	Valor de P ₅	Valor de Q ₆	
A*11:01	15@R	Adaptado	SVLSGGELDR	9.99E-05	0.18711308
B*15:01	26@K	No Adaptado	RLRPGGRKKY	2.07E-05	0.0553326
B*15:01	26@R	Adaptado	RLRPGGRKKY	7.47E-10	8.39E-06
A*30:01	28@K	No Adaptado	RLRPGGKKQYK	8.07E-05	0.15693499
A*30:01	28@R	Adaptado	RLRPGGKKRY	4.07E-05	0.09515189
A*24:02	30@R	Adaptado	KYRLKHIVW	3.97E-05	0.09685205
A*24:02	206@N	No Adaptado	AMQMLKETI	8.10E-05	0.15693499
B*15:01	215@I	Adaptado	AEWDRIHPVHA	8.10E-05	0.15693499
A*33:01	221@G	No Adaptado	RPVAPGQMR	7.90E-05	0.15693499
B*39:08	238@G	No Adaptado	SDIAETTSTL	5.21E-05	0.12190525
B*57:01	242@N	Adaptado	TSNLQEQUIAW	2.44E-06	0.00912008
B*58:01	242@N	Adaptado	TSNLQEQUIAW	3.29E-12	9.24E-08
B*57:01	242@T	No Adaptado	TSNLQEQUIAW	1.35E-05	0.04004287
B*58:01	242@T	No Adaptado	TSNLQEQUIAW	3.06E-10	5.74E-06
A*03:01	252@G	Adaptado	PIPVGEIYK	0.00011347	0.19852104
C*08:02	303@V	Adaptado	DYVDRFYKVL	0.00010905	0.19757877
A*02:06	342@S	Adaptado	SLEEMMTA	3.41E-05	0.0911271
A*02:06	342@T	No Adaptado	SLEEMMTA	6.87E-06	0.02270214
C*07:01	378@V	No Adaptado	VTNSATIMMQR	1.14E-05	0.03550032
C*03:05	398@E	No Adaptado	KVGHIAKNCRA	2.49E-07	0.00127101
A*31:01	403@K	Adaptado	KNCRAPRK	3.65E-08	0.00022809
A*31:01	403@R	No Adaptado	KNCRAPRK	6.13E-08	0.00034406
A*29:02	472@P	Adaptado	KQEPIDKELY	0.00011807	0.19852104
C*02:10	485@P	No Adaptado	LYHLTSLRSLF	4.74E-06	0.01665045

Igual que arriba.

Para explorar la influencia característica del contexto inmunogenético de Centroamérica en la evolución del VIH, se estudió la frecuencia de los alelos HLA predictores involucrados en asociaciones previamente no descritas en la cohorte centroamericana. Las posiciones asociadas a dichos alelos se compararon con las asociadas a presión selectiva positiva en cada país. Además, las frecuencias de estos alelos HLA fueron comparadas con las de los alelos más frecuentes para HLA-A, HLA-B y HLA-C en cada país (Tabla 12). Interesantemente, los 17 sitios con asociaciones únicas en Centroamérica incluyeron alelos como A*02:06, A*03:01, C*03:05, C*07:01, y C*08:02 en diferentes proporciones en los países centroamericanos. La posición 206 con el aminoácido N (no adaptado) fue

seleccionado por el alelo A*24:02 que es el segundo más frecuente en las cinco poblaciones de Centroamérica. Al buscar los alelos en los datos de Goulder *et al.*(2012) se localizaron dos alelos protectores B*52:01 y A*2501 mientras que B*35:01 fue caracterizado como un alelo de progresión de la infección.

TABLA 12. Frecuencias de alelos por país en relación a las asociaciones de HLA-polimorfismos únicos de Centroamérica.*

Alelo	Aminoácido	Sitio en gag	Frecuencia del Alelo por país					
			Belice	Panamá	Nicaragua	Honduras	Guatemala	Homer
A*02:06	<u>S</u> -T	342	0.01020408	0.01801802	0.05365854	0.02105263*	0.09302326*	0.024
A*03:01	<u>G</u>	252	0.04081633	0.06756757*	0.05365854	0.07894737	0.03197674*	0.1174
A*11:01	<u>R</u>	15	0.05102041	0.03153153	0.01463415*	0.03859649	0.02325581	0.0691
A*24:02	<u>R</u> -N	30-206	0.07142857	0.18018018	0.14146341	0.14035088	0.18895349	0.1051
A*29:02	<u>P</u>	472	0.02040816	0.125	0.02682927	0.03684211	0.0130814	0.0355
A*31:01	<u>K</u> -R	403	0.06122449	0.04054054	0.04390244	0.05087719	0.07848837	0.0428
A*33:01	G	221	0.04081633	0.44444	0.0707317	0.01052632	0.01889535	0.0059
B*15:01	<u>K</u> - <u>R</u> - <u>I</u>	26-215	0	0.03153153	0.0370732	0.01754386	0.04671533	0.0629
B*39:08	G	238	0	0	0.03170732	0.01578947	0.05255474	0
B*57:01	<u>N</u> -T	242	0	0.0045045	0.00487805	0.00175439	0.00875912	0.0333
B*58:01	<u>N</u> -T	242	0.06122449	0.03603604	0.00731707	0.00701754	0.01167883	0.0136
C*02:10	<u>P</u>	485	0.02040816	0.03153153	0.0195599	0.0122807*	0.00716332	0.0018
C*03:05	E	398	0.04081633	0.0630606	0.07823961	0.06491228	0.03724928	0
C*07:01	<u>V</u>	378	0.10204082	0.06756757	0.07823961	0.10350877	0.04727794	0.1224
C*08:02	<u>V</u>	303	0.01020408	0.04054054	0.04645477	0.03684211	0.02148997	0.0373

*Sitios que coinciden con selección bajo presión selectiva positiva en el país marcado.

Aminoácidos subrayados representan aminoácidos adaptados, los no subrayados son aminoácidos no adaptados.

Integración de datos

Al comparar las posiciones asociadas a HLA por el modelo PDN con las posiciones determinadas bajo presión selectiva positiva con el servidor DataMonkey, se observó que diez coincidían como posiciones identificadas como bajo presión selectiva positiva y a la vez asociadas a HLA (Tabla 13).

Tabla 13. Asociaciones HLA-polimorfismo por país

Posición en gag ¹	Aminoácido ²	Alelo asociado ³	Países con asociación ⁴
12	E	B*49:01	B-P-N-H y G
15	R	A*11:01	H
26	R	K*B1:501	B-P-N y H
252	G	A*03:01	P y G
280	S	B*52:01	P, N y G
342	S-T	A*02:06	H y G
398	E E-D	C*03:05 B*40:02	N y G
403	K-R	A*31:01	B-P-N-H y G
427	T-N	B*40:02	P, N y G
485	P	C*02:10	H

1. Posición ubicada dentro de la secuencia de *gag* donde se localizaron asociaciones aminoácido-HLA marcado con presión positiva por algún país.
2. Aminoácido seleccionado por algún aminoácido
3. Alelo predictor, asociado a un aminoácido en una posición específica
4. B=Belice, P=Panamá, N=Nicaragua, H=Honduras, G=Guatemala.

De las 42 asociaciones HLA-polimorfismo en 26 posiciones se detectó que nueve posiciones coinciden con presión selectiva positiva. Su distribución varía en toda la proteína de Gag, en p17 el número de sitios fue del 6%, para p24 del 4.3% mientras que para la última sección (p2,p7,p1,p6) fue del 5%. Lo cual confirma a p24 como la zona más conservada. Al comparar con el nivel de entropía que presentaron los países, se notan coincidencias en la última zona siendo la de mayor entropía seguida de p17, encontrando un pico marcado en la parte intermedia de p24, coincidiendo con lo encontrado en el análisis de asociaciones HLA-polimorfismos. Las posiciones exclusivas de Centroamérica se distribuyen en toda la proteína notando un mayor número de asociaciones al inicio de p17 y en la parte intermedia de p24.

De las posiciones identificadas solo para Centroamérica en el gen *gag* bajo presión selectiva positiva se localizaron seis de las cuales cinco coinciden con Guatemala y la restante solo para Honduras. Confirmando en los análisis de Entropía para Guatemala que presentó mayor número de sitios con alta entropía en comparación de otros países.

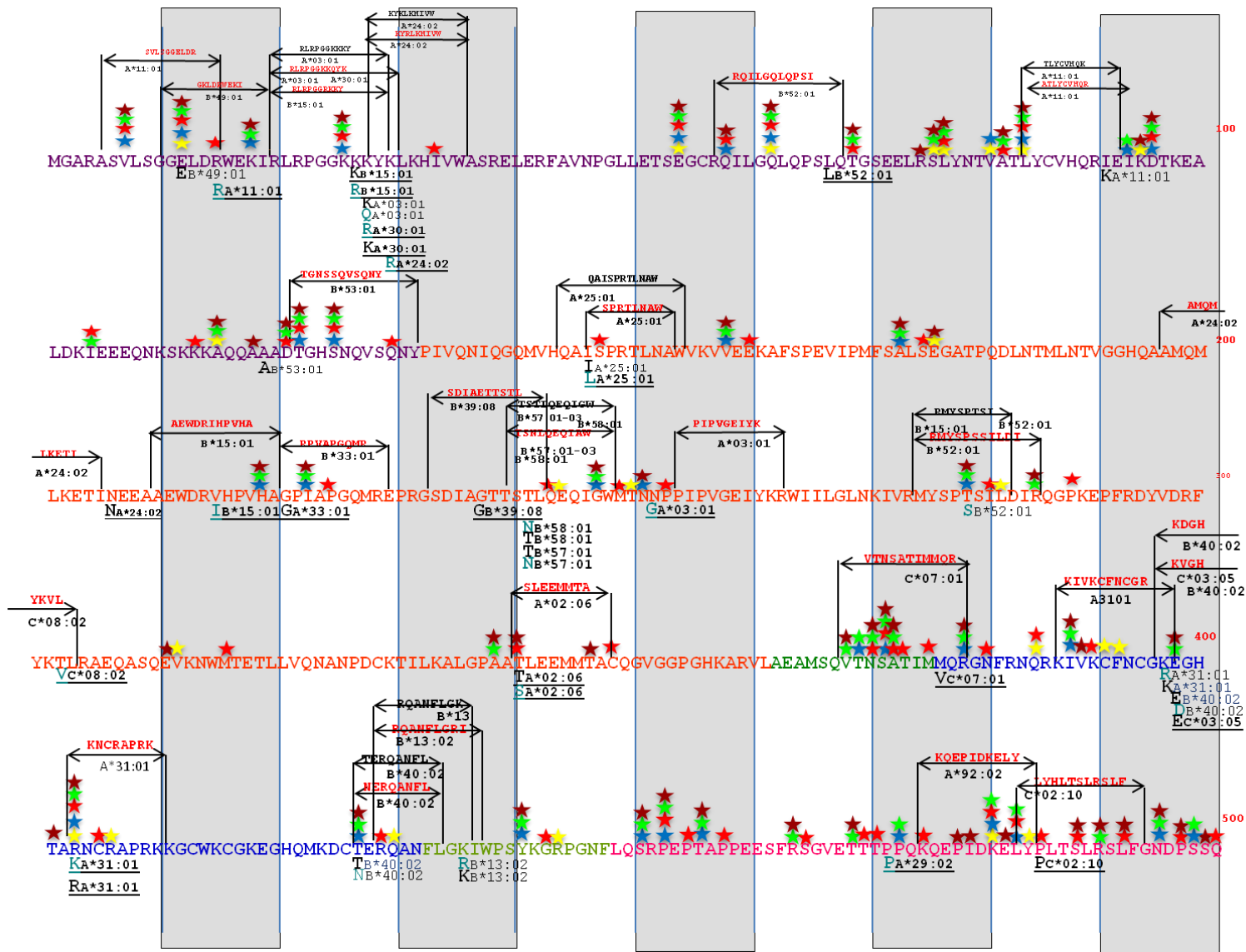


Figura 22. Mapa de Gag (997 secuencias en a.a) para el VIH-1 circulante en las poblaciones de Centroamérica. Se muestra la secuencia consenso. Se muestran 100 aminoácidos por línea divididos en bloques de diez por barras verticales. Las proteínas de Gag se marcan con colores alternados en el orden p17, p24, p2, p7, p1, p6. Se incluyen los sitios bajo presión selectiva positiva para cada país obtenidos con el servidor DataMonkey y se muestran como estrellas sobre la posición relevante con código de color de acuerdo al país donde se detectaron: Belice-Amarillo, Panamá-Azul, Nicaragua-Verde, Honduras-Rojo y Guatemala-Café. Los aminoácidos adaptados (aquéllos enriquecidos en presencia del alelo HLA selector) se muestran de color azul, los no adaptados (aquéllos enriquecidos en ausencia del alelo HLA selector) se muestran en color negro. Se ubicaron los 42 sitios de asociación de HLA-polimorfismo Subrayando las asociaciones únicas para la cohorte de Centroamérica. Los alelos y aminoácidos con una línea debajo representan sitios únicos para las poblaciones de Centroamérica.

DISCUSION

En el presente estudio se describió por primera vez la variabilidad del VIH circulando en distintos países de Centroamérica, con especial énfasis en el papel de los genes HLA de clase I en la evolución del virus en la región. La distribución geográfica de los individuos analizados fue un punto clave para definir contextos inmunogenéticos únicos, previamente no estudiados, que contribuyen de manera importante a la comprensión de la evolución asociada a HLA del VIH a nivel global. La cohorte analizada es altamente representativa de la población infectada por VIH-1 en Centroamérica, compartiendo características demográficas y clínicas con los datos reportados oficialmente en la región. Con respecto a la proporción de género, la cohorte refleja lo observado previamente para la mayoría de las cohortes latinoamericanas. Las proporciones de hombres-mujeres (1.8:1) infectados reportadas por la Organización Panamericana de la Salud para Centroamérica fueron semejantes a las encontradas en las cohortes de estudio con porcentajes de 42.5%, 25%, 38.4%, 41.3% y 43.5% para Belice, Panamá, Nicaragua, Honduras y Guatemala respectivamente [3].

En cuanto a la edad en su mayoría son adultos jóvenes para todos los países, observándose que la media tiende entre los 31 y 34 años para las cinco cohortes. Los valores de carga viral y de linfocitos T CD4+ en conjunto, son útiles para determinar la etapa de la infección en los individuos en el estudio. Estos parámetros son útiles para el monitoreo del tratamiento antirretroviral, para estimar el riesgo de transmisión, incluyendo la materno-fetal y permiten determinar cuándo es probable que se presenten enfermedades oportunistas [1,2]. En nuestra cohorte los valores variaron, teniendo a países como Guatemala, Honduras y Nicaragua con menor número de CD4+ y una carga viral elevada mencionados en orden de progresión a SIDA. Para Panamá y Belice los valores presentaron un mejor pronóstico los cuales mostraron mayor número de CD4 y menor carga viral. Al estudiar el árbol filogenético de los virus de la cohorte de Centroamérica basados en el gen *gag* se observaron claramente clusters por país, Esto sugiere una importante diseminación del virus a nivel nacional con diversos grados de entrecruzamiento de las epidemias a nivel regional. Nuestros resultados sugieren

que también podría existir una evolución localizada del virus, con agrupamientos que podrían estar reflejando variaciones en la presión selectiva inmune a nivel regional, resultado de distribuciones distintas de alelos HLA en diferentes regiones de los países impactando diferencialmente la adaptación de virus al hospedero. 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 [13,15]. Alternativamente, posibles efectos fundadores en la diversificación regional del virus circulante en cada país podrían estar reflejando en los análisis filogenéticos realizados. Así mismo, efectos epidemiológicos con influencia de barreras culturales y políticas, además de considerar posibles contactos comerciales y turísticos, muy probablemente tienen influencia en la diversificación del virus en la región.

Los análisis de variabilidad del gen *gag* de VIH-1 circulante en los países de Centroamérica mostraron resultados esperados. Las regiones p17 y p2p7p1p6 presentaron mayor variabilidad, teniendo una mayor cantidad de aminoácidos 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, y colaboradores [8], acerca de la alta conservación de p24 y la mayor variación en p17 y p2p7p1p6. Esta observación se confirmó en el análisis de frecuencia de polimorfismos y el análisis de entropía de Shannon. La conservación de p24 podría explicarse dada su importancia funcional y estructural para el virus, ya que prácticamente cualquier variación en la esta zona puede afectar la capacidad replicativa del VIH-1. Aparentemente, las zonas correspondientes a p17 y a p2p7p1p6 permiten mayor flexibilidad en cuanto a la presencia de polimorfismos, manteniendo las funciones de las proteínas correspondientes.

Por otro lado se notó la existencia de un número considerable de posiciones bajo presión selectiva positiva en el virus circulante en cada país. Interesantemente, algunos países presentaron posiciones bajo presión selectiva únicas y otras que compartían con los demás países analizados. Algunas de estas posiciones bajo

presión selectiva se encontraron también asociadas a selección por HLA, confirmando la alta inmunogenicidad de Gag [14], así como la importancia de las respuestas restringidas por HLA en la evolución del virus.

En el análisis de frecuencias alélicas HLA por país se demostró que existe diversidad en la distribución de frecuencias en los países analizados, incluso tratándose de una región geográficamente pequeña y con poblaciones en común. Se observaron alelos HLA en que coincidían en los diferentes países estudiados como el caso de C*03:05 que se encontró entre los 10 primeros alelos en los individuos de la cohorte estudiada. Algunos alelos fueron característicos únicos de cada población, claro ejemplo Panamá con los alelos: B*07:02, B*35:49, B*40:16 y B*48:02 que ningún otro país presentó. Estas observaciones definen un contexto inmunogenético único en el cual estudiar la evolución asociada a HLA del VIH.

Aunque se esperaba que los epítomos en Gag reconocidos en poblaciones de Centroamérica variaran con respecto a los reportados en otras poblaciones dadas las distribuciones de frecuencias HLA características de estas cohortes debido a un contexto inmunogenético único, también se observaron posiciones bajo presión selectiva en zonas previamente predichas como altamente inmunogénicas, con epítomos para alelos HLA distintos previamente descritos.

Parte de la variabilidad asociada a HLA de los virus en estas poblaciones 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 las poblaciones en Centroamérica.

Dentro de las posiciones bajo presión selectiva positiva en el modelo de PDN 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 países de Centroamérica, que será necesario estudiar a fondo.

Interesantemente, diez de las posiciones bajo presión selectiva positiva coincidieron con sitios de presión selectiva inmune asociados a HLA por el modelo

PDN, mientras que 13 sitios asociados a HLA no mostraron coincidencia con sitios identificados como bajo presión selectiva positiva. Adicionalmente también se encontraron 17 sitios con 25 aminoácidos que no han sido reportados previamente en la literatura, que se encuentran bajo presión selectiva mediada por HLA y que podrían pertenecer específicamente a la poblaciones de Centroamérica. De éstos, cinco sitios también presentaron presión selectiva positiva para Guatemala en tres casos y Honduras en dos.

Adicionalmente se identificaron 23 epitopos no descritos anteriormente para otras poblaciones lo que sugiere que estarían en zonas específicas para la cohorte de Centroamérica, ameritando una revisión más exhaustiva de estos.

Es bien sabido que la identificación de asociaciones HLA-polimorfismo de VIH requiere de cohortes grandes (Carlson 2008), por lo que es posible, que en el presente estudio no se estén identificando todos los sitios existentes. Sin embargo, el tamaño de la cohorte del presente estudio es comparable con estudios previamente reportados en cohortes caucásicas y asiáticas (14-17). Se requieren otros estudios para establecer si la falta de asociación con HLA de posiciones bajo presión selectiva se debe a una baja potencia del análisis en el modelo PDN para identificar asociaciones con el número de participantes disponible, o si las posiciones representan asociaciones indirectas con HLA, presiones selectivas alternas o asociaciones negativas. De cualquier manera, con este estudio se ha 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 poblaciones de Centroamérica.

Los resultados de los diversos análisis realizados se compilaron en un mapa del gen *gag* que resume las asociaciones con presión selectiva positiva, y con selección por HLA en el contexto inmunogenético específico de las poblaciones de Centroamérica. En este mapa se mostró que hay aminoácidos que presentan variantes que se comparten con otras poblaciones inmunogenéticamente distintas y que hay aminoácidos con variantes que se seleccionan por alelos específicos de nuestra cohorte. El presente estudio muestra características únicas de la evolución del gen *gag* del VIH-1 en Centroamérica siendo el primer trabajo que

trata de describir y explicar la variabilidad de este gen en virus circulantes de esta región, lo cual guiará estudios futuros para comprender mejor la adaptación del VIH-1 a nivel poblacional.

CONCLUSIONES y PERSPECTIVAS

En este estudio se logró secuenciar el gen *gag* del VIH-1 en 997 individuos, de cinco países que conformaron una cohorte altamente representativa de los virus circulantes en Centroamérica. Los polimorfismos que se identificaron mostraron alta variabilidad a lo largo del gen, sobre todo en las zonas de p17 y p2p7p1p6, que en gran parte puede ser influenciada por la presión selectiva ejercida por el sistema inmune del hospedero por medio de los alelos HLA-I. Se observó que la proteína p24 tiende a conservarse debido a su importancia en la correcta formación de la cápside del virus.

Nuestros resultados muestran por primera vez un panorama de la influencia de la selección por HLA en la evolución del VIH en Centroamérica, enfocándose en la importancia de esta presión selectiva sobre la diversidad del virus circulante en la región.

Este trabajo representa la primera caracterización de las asociaciones de HLA-polimorfismo de VIH en una cohorte de Centroamérica. El análisis evolutivo de las secuencias de *gag* de virus circulante en Centroamérica permitió identificar posiciones asociadas a HLA únicas en Centroamérica. Estos resultados podrían revelar nuevas asociaciones con el control inmune del VIH-1, las cuales podrían ser de gran relevancia en la a portación de conocimiento para el diseño de una vacuna basada en CTLs.

El continuar con estos estudios para Centroamérica no sólo incluyendo al Salvador y Costa Rica sino también para otras poblaciones en América Latina permitiría tener una perspectiva más completa de la evolución del virus en poblaciones latinoamericanas y ayudaría a localizar blancos específicos no identificados para otras poblaciones que pudieran ser relevantes en el diseño de una vacuna basada en CTLs o que pudieran funcionar como blancos para otras estrategias

antirretrovirales. Adicionalmente, la realización de estudios sobre el impacto de la selección por HLA en *gag* con respecto a la capacidad replicativa viral en estudios *in vitro* así como valorar el impacto de la adaptación viral asociada a HLA en la progresión de la enfermedad por VIH, permite estudiar más a fondo la evolución poblacional de este patógeno y contribuir con estrategias relevantes para su control

BIBLIOGRAFÍA

1. Cordoba Villalobos, J. A.; De León Rosales, S.P., Valdespino, J. L.(Eds.) 2009 . **25 años de SIDA en México. Logros, desaciertos y retos.** (2ª Ed.). Mexico: Instituto Nacional de Salud Publica. 135-191.
2. ONUSIDA 2015
3. Organización Panamericana de la Salud (OPS) 2014.
4. Plan estratégico regional de VIH y sida de Centroamérica y república dominicana 2010 – 2015., Centroamérica 2010
5. Hemelaar, J. 2011. **The origin and diversity of the VIH-1 pandemic.** Trends in Molecular Medicine, March 2012, Vol. 18, No. 3
6. Hemelaar,J.; Gouws,E.; Ghys,P.D.; Osmanov, S.; WHO-UNAIDS Network for HIV Isolation and Characterisation. 2011. **Global trends in molecular epidemiology of HIV-1 during 2000– 2007.** AIDS. 2011 Mar 13;25(5):679-89.
7. Freed, E.O.; Martin, M.A. 2007. **VIHs and their replication.** Philadelphia: Knipe DM, Vol. II. (5a Ed.) 2108-2183.
8. Nicholas H. Acheson. 2011. **Fundamentals of molecular virology.**2nd ed.United States. ISBN 978-0-470-90059-8
9. Freed, E.O. (1998). **VIH-1 Gag proteins: Diverse functions in the virus life cycle.** Virology, 251: 1-15
- 10.Heinrich G. Gottlinger. 2011. **VIH-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release.** Reviews. Harvard Medical School, Boston, MA. 02115.
- 11.Neil, M. Bell and Andrew M.L.2012. **VIH Gag polyprotein: processing and early viral particle assembly.** Trends in Microbiology, March 2013. Vol 2. No.3.
- 12.Langford E.S., Ananworanich J., Cooper D.A.2007. **Predictors of disease progression in HIV infection: a review.** AIDS Research and Therapy, 4:11
- 13.Kindt, T.J. Goldsby, R. A. Osborne B. A.2007. INMUNOLOGÍA de Kuby. 6ta ed. McGRAW-HILL INTERAMERICANA EDITORES, S.A. de C.V.

14. Goulder, P. J. R., & Watkins, D. I. (2008). ***Impact of MHC class I diversity on immune control of immunodeficiency virus replication.*** *Nature Reviews. Immunology*, 8(8), 619–630. <http://doi.org/10.1038/nri2357>
15. IPD - IMGT/HLA 2015 <https://www.ebi.ac.uk/ipd/imgt/hla/stats.html>
16. McLaren PJ, Coulonges C, Ripke S, van den Berg L, Buchbinder S, et al. 2013. ***Association Study of Common Genetic Variants and HIV-1 Acquisition in 6,300 Infected Cases and 7,200 Controls.*** *PLoS Pathog* 9(7): e1003515. doi: 10.1371/journal.ppat.1003515
17. Goulder, P. J. R., & Walker, B. D. (2012). ***HIV and HLA Class I: an evolving relationship.*** *Immunity*, 37(3), 426–440. <http://doi.org/10.1016/j.immuni.2012.09.005>
18. Alter, G., & Altfeld, M. (2009). ***NK cells in HIV-1 infection: Evidence for their role in the control of HIV-1 infection.*** *Journal of Internal Medicine*, 265(1), 29–42. <http://doi.org/10.1111/j.1365-2796.2008.02045.x>
19. Mascola JR.(2013). ***VIH 1 neutralizing antibodies understanding nature's pathways.***
20. 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 VIH-1 Gag.*** *PLoS Comp. Biol.*, 4 (11): 1-23.
21. Carlson, J. M., Listgarten, J., Pfeifer, N., Tan, V., Kadie, C., Walker, B. D., Heckerman, D. (2012). ***Widespread Impact of HLA Restriction on Immune Control and Escape Pathways of HIV-1.*** *Journal of Virology*, 86(9), 5230–5243. <http://doi.org/10.1128/JVI.06728-11>
22. Carlson J.M.; Le A.Q.; Shahid A.; Brumme Z.L. (2014). ***VIH adpitation to HLA: a window into virus-host inmune interactions.*** *Trends in Micribiology* XX(2015) 1-13
23. Crawford, H.; Prado, J.G.; Leslie, A.; Hue, 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.
24. Avila-Rios, S.; Ormsby, C.E.; Carlson, J.M.; Valenzuela-Ponce, H.; Blanco-Heredia, J.; Garrido-Rodriguez, D.; Garcia-Morales, C.; John, M.; Espinosa, E.; Reyes-Teran, G.(2009). **Unique features of HLA-mediated HIV evolution in a Mexican cohort: a comparative study.** Retrovirology, 6: 72
25. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. Siepel AC, Halpern AL, Macken C, Korber BT.
26. *AIDS Res Hum Retroviruses*. 1995 Nov;11(11):1413-6. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. Peña ACP, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, Gomez A, Camacho RJ, de Oliveira T, Vandamme A-M. *Infectious Genetics and Evolution* 2013; 19:337-48. doi: 10.1016/j.meegid.2013.04.03
27. Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, et al. (2007) **Evidence of Differential HLA Class I-Mediated Viral Evolution in Functional and Accessory/Regulatory Genes of HIV-1.** PLoS Pathog 3(7): e94. doi: 10.1371/journal.ppat.0030094
28. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, et al. (2009) **HLA-Associated Immune Escape Pathways in HIV-1 Subtype B Gag, Pol and Nef Proteins.** PLoS ONE 4(8): e6687. doi: 10.1371/journal.pone.0006687

ANEXOS:

ABREVIATURAS

a.a: Aminoácido

APC: Celula Presentadora de Antigeno, en ingles.

CRF: Formas Recombinantes Circulantes, en ingles.

CTLs: Linfocitos T Citotoxicos, en ingles.

dN: Sustituciones No Sinonimas, en ingles.

DNA: Acido Desoxirribonucleico, en ingles.

dS: Sustituciones Sinonimas, en ingles.

GAG: Gen-Antígeno Grupo-Específico.

GWAS: Estudios sobre el Análisis Comparativos del Genoma

HLA: Antigeno Leucocitario Humano, en ingles.

MHC: Complejo Mayor de Histocompatibilidad, en ingles.

OMS: Organización Mundial de la Salud.

ORF: Fragmentos de Lectura Abierta, en ingles.

PDN: Redes de Dependencia Filogenetica, en ingles.

RNA: Acido Ribonucleico, en ingles.

RT: Transcriptasa Inversa, en ingles.

SIV: Virus de Inmunodeficiencia Simiana, en ingles.

SIDA: Sindrome de Inmunodeficiencia Adquirida.

TCR: Receptor de celulas T.

VIH: Virus de Inmunodeficiencia Humana.

GLOSARIO

Alelo: Es cada una de las posibles formas alternativas de un gen dado, que difiere en su secuencia de ADN y afecta a su función. Un organismo diploide tiene siempre dos alelos de cada gen, que pueden ser iguales (homocigosis) o diferentes (heterocigosis).

Epítopo: (también conocido como determinante antigénico) es un sitio de un antígeno reconocido por un anticuerpo o por un receptor de antígeno. Un epítopo

de célula T es un péptido corto derivado de un antígeno proteínico. Se une a una molécula del MHC y es reconocido por una célula T particular. Los epítomos de célula B son determinantes antigénicos reconocidos por células B y por lo general son motivos estructurales ubicados sobre la superficie del antígeno.

LTR: del inglés *Long Terminal Repeat*) es una secuencia de nucleótidos característica que se encuentra en cada extremo de un elementoretroviral que ha sido integrado en el genoma hospedador. Está implicada en el proceso de integración.

ORF: del inglés *reading frame* o Marco de lectura abierto; porción del genoma que contiene una secuencias de bases que potencialmente puede codificar una proteína. En un gen, el ORF se localiza en el codón de inicio y el codón de paro (stop).

Splicing: Es un proceso co-transcripcional de corte y empalme de RNA.

Set point: Concentración vírica estabilizada.

TAXA. Plural de taxón, lo mismo que taxones.

TAXÓN: Especie

ÍNDICE DE TABLAS

- ❖ Tabla 1. Datos para Latinoamérica sobre VIH.
- ❖ Tabla 2. Casos de VIH y Sida acumulados hasta 2014, en países de la Región Centroamericana.
- ❖ Tablas 3 y 4. Alelos HLA-I (A y B) protectores y susceptibles a la progresión de la infección.
- ❖ Tabla 5: Descripción de los reactivos y cantidad de cada uno en la primera reacción de PCR
- ❖ Tabla 6: Descripción de los reactivos y cantidad de cada uno en la segunda reacción de PCR
- ❖ Tabla 7: Características clínicas de las cohorte
- ❖ Tabla 8: Sitios bajo presión selectiva positiva por país, en cada posición de Gag.
- ❖ Tabla 9. Alelos presentes en mayor frecuencia en cada población.

- ❖ Tabla 10: Asociaciones HLA-polimorfismo de VIH en el gen *gag* en una cohorte de 997 individuos centroamericanos, de acuerdo al modelo de PDN.
- ❖ Tabla 11. Asociaciones HLA-polimorfismo por país
- ❖ TABLA 12: Frecuencias de alelos por país en relación a las asociaciones de HLA-polimorfismos únicos de Centroamérica. Mostrando en rojo los que muestran adaptación.

ÍNDICE DE FIGURAS

- ❖ Figura 1. Grupos, subtipos del VIH.
- ❖ Figura 2. Distribución global de subtipos de VIH-I y sus recombinantes (CRF)
- ❖ Figura 3: Diagrama de la estructura del VIH
- ❖ Figura 4. Organización del genoma del VIH-I.
- ❖ Figura 5. Descripción esquemática de los eventos que ocurren durante el ciclo de replicación del VIH.
- ❖ Figura 6. Historia natural de la enfermedad causada por VIH
- ❖ Figure 7: Diversidad y distribución de los alelos HLA- B en el mundo
- ❖ Figura 8. El HLA es el mayor control global de la diversidad del VIH en el mundo el trabajo experimental.
- ❖ Figura 9: Descripción del trabajo experimental.
- ❖ Figura 10: Muestra de un gel de agarosa al 1%.
- ❖ Figura 11. Sequencing Analysis v5.4
- ❖ Figura 12. Software Geneios (versión 5.6.7).
- ❖ Figura 13. Programa MEGA 5.05
- ❖ Figura 14: Distribución geográfica de la cohorte estudiada
- ❖ Figura 15: Relaciones filogenéticas de las secuencias de Gag (VIH-1) en individuos vírgenes a tratamiento antirretroviral de cinco países de Centroamérica.
- ❖ Figura 16: Mutaciones sinónimas y no sinónimas en la secuencia consenso de cada país de Centroamérica
- ❖ Figura17. Entropía en la secuencia de gag.

- ❖ Figura 18. Entropía, variabilidad de un país con respecto a otro.
- ❖ Figura 19: Distribución de la frecuencia de HLA-A en las poblaciones de Centroamérica y de un grupo externo de Canadá
- ❖ Figura 20. Distribución de la frecuencia de HLA-B en las poblaciones de Centroamérica y de un grupo externo de Canadá
- ❖ Figura 21: Distribución de la frecuencia de HLA-B en las poblaciones de Centroamérica y de un grupo externo de Canadá.
- ❖ Figura 22. Mapa de Gag (997 secuencias) para el VIH-1 circulante en las poblaciones de Centroamérica