



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**POSGRADO EN CIENCIAS BIOLÓGICAS**  
FACULTAD DE MEDICINA  
BIOMEDICINA

**IMPACTO DE MUTACIONES SELECCIONADAS POR HLA EN LA CAPACIDAD REPLICATIVA  
DEL VIRUS DE INMUNODEFICIENCIA HUMANA CIRCULANTE EN LA REGIÓN  
MESOAMERICANA**

**TESIS**

QUE PARA OPTAR POR EL GRADO DE:  
**MAESTRA EN CIENCIAS BIOLÓGICAS**

PRESENTA:

**THALIA ALEJANDRA GARCÍA TÉLLEZ**

**TUTOR(A) PRINCIPAL DE TESIS: DR. SANTIAGO ÁVILA RÍOS**

INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS, INER

**COMITÉ TUTOR: DR. LUIS GUILLERMO LLORENTE PETERS**

INSTITUTO NACIONAL DE CIENCIAS MÉDICAS Y NUTRICIÓN SALVADOR ZUBIRÁN, INCMNSZ

**DRA. LEONOR HUERTA HERNÁNDEZ**

INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

**MÉXICO, D.F. AGOSTO, 2015**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

Dr. Isidro Ávila Martínez  
Director General de Administración Escolar, UNAM  
Presente

Me permito informar a usted que el Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, en su sesión ordinaria del día 08 de junio de 2015, aprobó el jurado para la presentación de su examen para obtener el grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** de la alumna **GARCÍA TÉLLEZ THALÍA ALEJANDRA** con número de cuenta **304139529**, con la tesis titulada **"IMPACTO DE MUTACIONES SELECCIONADAS POR HLA EN LA CAPACIDAD REPLICATIVA DEL VIRUS DE INMUNODEFICIENCIA HUMANA CIRCULANTE EN LA REGIÓN MESOAMERICANA"**, realizada bajo la dirección del **DR. SANTIAGO ÁVILA RIOS**:

Presidente: DR. ANTONIO EUSEBIO LAZCANO ARAUJO REYES  
Vocal: DRA. MARTHA ESTHELA PÉREZ RODRÍGUEZ  
Secretario: DR. LUIS GUILLERMO LLORENTE PETERS  
Suplente: DRA. LETICIA ROCHA ZAVALETA  
Suplente: DRA. LEONOR HUERTA HERNÁNDEZ

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE  
"POR MI RAZA HABLARÁ EL ESPÍRITU"  
Cd. Universitaria, D.F., a 03 de agosto de 2015

*M. del Coro Ariz*

DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA  
COORDINADORA DEL PROGRAMA



## **AGRADECIMIENTOS**

- a) Primeramente, agradezco al Posgrado en Ciencias Biológicas, UNAM.
- b) La realización de esta tesis fue gracias al Consejo Nacional de Ciencia y Tecnología.  
Beca CONACYT. No registro becario: 280492. No. Apoyo: 345468
- c) Agradezco La Comisión de Equidad y Género de las legislaturas LX y LXI y la comisión de Igualdad de Género de la Legislatura LXII de la H. Cámara de Diputados de la República Mexicana.

## **AGRADECIMIENTOS PERSONALES**

Mi agradecimiento total va para quienes me dieron herramientas y oportunidades.

Al Dr. Gustavo Reyes Terán y al Dr. Santiago Ávila Ríos por la confianza que me brindaron día con día. Por la oportunidad de unirme al CIENI y confiarme importantes tareas. Por la motivación constante para lograr grandes cosas, por la libertad para compartir mis metas y el respaldo para seguir construyendo un camino exitoso haciendo algo que disfruto.

Al equipo CIENI por la oportunidad de ser parte de sus vidas y hacerlos parte de la mía. A Humberto, Daniela, Olivia y Sandra por conformar una pequeña familia que me ayudó a soñar en grande y a caminar segura. A Paloma, Lety, Dany Tapia, Tania, Selma, Vero, Clau, Maggie, Zeidy, Berenice y todos aquellos que le dan vida a los pasillos y laboratorios y que me hicieron sentir en casa.

Gracias a mis tutores, colaboradores y colegas por compartir su conocimiento y experiencia. A Maribel Soto por su paciencia y apoyo, a la Dra. Zabrinna Brumme, el Dr. Mark Brockman, y el Dr. Jonathan Carlson por proveerme de herramientas para alcanzar los objetivos de mi proyecto. Gracias al equipo del Dr. Carlos Mejía por la confianza y respaldo para generar información para su país. Y a todos los expertos nacionales e internacionales que se involucraron de una u otra forma con el éxito de mi investigación.

Gracias a mi familia, que durante 26 años me ha dado los medios para alcanzar mis sueños. A mis padres por enseñarme a ser tenaz, dedicada y apasionada. Por demostrarme que el único límite está en mi cabeza, y el arma más grande está en mi corazón. A mi hermana por repetirme incansablemente “tu puedes” y ser ejemplo de fortaleza. A mis seres queridos por estar en las buenas y no tan buenas, siempre de la mano. A mis amigas por recordarme lo feliz que soy al compartir lo que amo con personas insustituibles en mi vida.

A la vida, por la oportunidad de demostrarme que soy capaz de aprender, conocer y comprender tanto como mi curiosidad me exija. Que soy capaz de fijarme objetivos y alcanzarlos con trabajo arduo. Que vale la pena el cansancio, la frustración y los errores, porque siempre se obtiene algo bueno de ello.

Gracias a los pacientes del CIENI, quienes nos dan un motivo de lucha y trabajo cada día y depositan su confianza y esperanza en nuestro trabajo.

**Gracias por ser parte de lo que me hace feliz.**

The most important thing in life is to be happy... It depends for what you are highly determined and motivated.. It is not a question about high salary, low salary, whatever, is a question of motivation.

If you are happy with what you are working on, that is the most important thing in your life.

Françoise Barré-Sinoussi

## Índice

ÍNDICE DE FIGURAS .....	i
RESUMEN .....	1
ABSTRACT .....	3
CAPITULO 1. INTRODUCCIÓN.....	5
I. EPIDEMIOLOGÍA DE LA INFECCIÓN POR VIH y SIDA.....	5
I.1 DESCUBRIMIENTO DEL VIRUS DE INMUNODEFICIENCIA HUMANA ....	5
I.2 LA EPIDEMIA EN EL MUNDO .....	7
I.3 LA INFECCIÓN POR VIH EN LATINOAMÉRICA.....	11
I.4 LA INFECCIÓN EN MÉXICO Y GUATEMALA .....	13
II. GENERALIDADES DEL VIRUS DE INMUNODEFICIENCIA HUMANA .....	16
II.1 CLASIFICACIÓN DEL VIRUS .....	16
II.2 TROPISMO, RECEPTORES Y CO-RECEPTORES.....	19
II.3 ORGANIZACIÓN GENÓMICA Y PROTEÍNAS VIRALES.....	20
II.4 CICLO DE REPLICACIÓN DEL VIH.....	25
II.5 HISTORIA NATURAL DE LA INFECCIÓN .....	32
III. RESPUESTA CITOTÓXICA Y EVOLUCIÓN VIRAL MEDIADA POR HLA... 37	
III.1 ANTIGENO LEUCOCITARIO HUMANO (HLA) .....	37
III.2 PRESENTACIÓN ANTIGÉNICA MEDIADA POR HLA-I.....	40
III.3 MADURACIÓN Y ACTIVACIÓN DE LOS LINFOCITOS CD8 <sup>+</sup> .....	43
III.4 RESPUESTA CITOTÓXICA MEDIADA POR LINFOCITOS T CD8 <sup>+</sup> .....	47
III.5 INTERACCIÓN ENTRE EL VIH Y EL HOSPEDERO: EVOLUCIÓN MEDIADA POR HLA.....	49
CAPITULO 2. OBJETIVOS .....	54
OBJETIVO GENERAL .....	54

OBJETIVOS PARTICULARES .....	54
CAPITULO 3. ANTECEDENTES .....	55
CAPITULO 4. METODOLOGÍA.....	57
4.1 SELECCIÓN DE PACIENTES .....	57
4.2 PROCESAMIENTO DE MUESTRAS.....	57
4.2.1 EXTRACCIÓN DE ADN.....	57
4.2.2 EXTRACCIÓN DE ARN VIRAL.....	58
4.3 SECUENCIACIÓN .....	58
4.3.1 AMPLIFICACIÓN .....	58
4.3.2 SECUENCIACIÓN AUTOMÁTICA.....	60
4.3.3 ALINEAMIENTO Y ANÁLISIS FILOGENÉTICO .....	61
4.3.4 ANÁLISIS DE PRESIÓN SELECTIVA POR MODELO DE CODONES. ....	61
4.4 ENSAYO DE CAPACIDAD REPLICATIVA (ECR) .....	62
4.4.1 AMPLIFICACIÓN DEL FRAGMENTO GAG-PRO .....	62
4.4.2 PREPARACIÓN DEL VECTOR.....	63
4.4.3 CULTIVO CELULAR.....	64
4.4.4 CO-TRANSFECCIÓN .....	64
4.4.5 PROPAGACIÓN DE VIRUS QUIMÉRICOS .....	65
4.4.6 TITULACIÓN .....	66
4.4.7 CÁLCULO DE CAPACIDAD REPLICATIVA VIRAL (CRV).....	68
4.5 TIPIFICACIÓN HLA .....	69
4.6 ANÁLISIS ESTADÍSTICOS.....	72
CAPITULO 5. RESULTADOS .....	74
5.1 DESCRIPCIÓN DE LA COHORTE DE ESTUDIO .....	74
5.2 FRECUENCIAS ALÉLICAS HLA-I .....	74



5.3 ANÁLISIS DE CAPACIDAD REPLICATIVA VIRAL (CRV).....	77
5.4 ANÁLISIS MOLECULAR DE GAG-VIH-1: VARIABILIDAD Y PRESIÓN SELECTIVA .....	81
5.5 IMPACTO DE MUTACIONES EN GAG SOBRE LA CAPACIDAD REPLICATIVA DEL VIH.....	86
5.6 EFECTO DEL HLA EN LA SELECCIÓN DE MUTACIONES Y ALTERACIÓN DE LA CAPACIDAD REPLICATIVA VIRAL .....	90
CAPITULO 6. DISCUSIÓN.....	95
CAPITULO 7. CONCLUSIONES.....	106
LITERATURA CITADA.....	107
APENDICE I.....	111

## ÍNDICE DE FIGURAS

FIGURA 1. DISTRIBUCIÓN MUNDIAL DE LA INFECCIÓN POR VIH.....	7
FIGURA 2. DISTRIBUCIÓN POR PAÍS DE CASOS DE VIH EN EL MUNDO.....	8
FIGURA 3. PREVALENCIA DE VIH POR POBLACIÓN DE RIESGO. ....	10
FIGURA 4. PERSONAS VIVIENDO CON VIH EN LATINOAMÉRICA, 2013. ....	11
FIGURA 5. NUEVAS INFECCIONES POR VIH EN LATINOAMÉRICA, 2013. ....	12
FIGURA 6. INCIDENCIA ACUMULADA DE CASOS DE VIH Y SIDA SEGÚN ENTIDAD FEDERATIVA. ....	14
FIGURA 7. MORTALIDAD EN GUATEMALA POR CAUSAS ASOCIADAS A VIH/SIDA SEGÚN OCURRENCIA Y PROCEDENCIA.. ....	16
FIGURA 8. ÁRBOL FILOGENÉTICO DE SIV Y VIH.. ....	18
FIGURA 9. DISTRIBUCIÓN MUNDIAL DE LOS SUBTIPOS DE VIH-1. ....	18
FIGURA 10. REPRESENTACIÓN ESQUEMÁTICA DE UN VIRUS DE INMUNODEFICIENCIA HUMANA.....	20
FIGURA 11. GENOMA Y PROTEÍNAS DEL VIH.....	23
FIGURA 12. FUSIÓN DE MEMBRANAS MEDIADA POR GP120 Y GP41.....	26
FIGURA 13 MADURACIÓN DEL VIH.....	30
FIGURA 14 ESQUEMA CLÁSICO DEL CICLO DE REPLICACIÓN DEL VIH.....	31
FIGURA 15 PATOGÉNESIS DE LA INFECCIÓN POR VIH. ....	34
FIGURA 16 RESPUESTA INMUNE EN LA INFECCIÓN POR VIH. ....	36
FIGURA 17 PATOLOGÍAS ASOCIADAS A SIDA. ....	37
FIGURA 18 MAPA CROMOSÓMICO DEL ANTÍGENO LEUCOCITARIO HUMANO.....	39
FIGURA 19 ESTRUCTURA DE UNA MOLÉCULA HLA DE CLASE I.....	40
FIGURA 20 PROCESAMIENTO Y PRESENTACIÓN ANTIGÉNICA RESTRINGIDA POR MOLÉCULAS HLA DE CLASE I.....	42
FIGURA 21 MADURACIÓN DE LOS LINFOCITOS T. ....	44
FIGURA 22 RECEPTOR DE LINFOCITO T.....	47
FIGURA 23 MECANISMOS EFECTORES DE LA RESPUESTA LÍTICA MEDIADA POR LINFOCITOS T CD8+.....	48
FIGURA 24 PRESIÓN SELECTIVA MEDIADA POR HLA.. ....	50
FIGURA 25 REVERSIÓN Y PERSISTENCIA DE MUTACIONES DE ESCAPE. ....	51
FIGURA 26 EJEMPLO DE DISTRIBUCIÓN POBLACIONAL DE ALGUNOS ALELOS HLA-B EN EL MUNDO.....	52
FIGURA 27 ESTRATEGIA DE ANÁLISIS DE CITOMETRÍA DE FLUJO USADA DURANTE EL ENSAYO DE CAPACIDAD REPLICATIVA VIRAL. ..	67
FIGURA 28 DIAGRAMA DE ENSAYO DE CAPACIDAD REPLICATIVA VIRAL.....	73
FIGURA 29 DISTRIBUCIÓN GEOGRÁFICA DE LA COHORTE DE ESTUDIO. ....	75
FIGURA 30 FRECUENCIAS ALÉLICAS DE HLA-I EN EL GRUPO DE ESTUDIO. ....	76
FIGURA 31 CORRELACIÓN DE PARÁMETROS CLÍNICOS Y CAPACIDAD REPLICATIVA VIRAL.....	77
FIGURA 32 ASOCIACIÓN ENTRE CARGA VIRAL Y ALELOS HLA-I. ....	78
FIGURA 33 ASOCIACIÓN ENTRE CUENTA DE LINFOCITOS T-CD4+ Y HLA-I.....	79

FIGURA 34 ASOCIACIÓN ENTRE CAPACIDAD REPLICATIVA VIRAL (CRV) Y ALELOS HLA-I.....	80
FIGURA 35 COMPARACIÓN FILOGENÉTICA DE SECUENCIAS DE GAG ANTES Y DESPUÉS DEL ENSAYO DE CAPACIDAD REPLICATIVA....	81
FIGURA 36 GRÁFICA DE VARIABILIDAD EN LA SECUENCIA DE GAG. ....	82
FIGURA 37 EFECTO DE LA PRESENCIA DE MUTACIONES EN P17 SOBRE LA CAPACIDAD REPLICATIVA VIRAL.....	86
FIGURA 38 EFECTO DE LA PRESENCIA DE MUTACIONES EN P24 SOBRE LA CAPACIDAD REPLICATIVA VIRAL.....	87
FIGURA 39 EFECTO DE LA PRESENCIA DE MUTACIONES EN P2P7P1P6 SOBRE LA CAPACIDAD REPLICATIVA VIRAL.....	87
FIGURA 40 ANÁLISIS FILOGENÉTICA DE LA MUTACIÓN T84V.. ....	88
FIGURA 41 ANÁLISIS FILOGENÉTICO DE LA MUTACIÓN A115K.....	89
FIGURA 42 ANÁLISIS FILOGENÉTICO DE LA MUTACIÓN S373T. ....	90
FIGURA 43 MAPA DE PREISÓN SELECTIVA DE HLA SOBRE GAG DE VIH-1. ....	93

## RESUMEN

**ANTECEDENTES** El control de la infección por VIH ha sido asociado a la respuesta citotóxica mediada por los alelos HLA-I. El efecto es consecuencia de la unión de péptidos virales, particularmente derivados del polipéptido Gag, a moléculas HLA de clase I que pueden ser reconocidas por linfocitos T CD8+ citotóxicos (CTL, inglés), con la consecuente eliminación de las células infectadas. Sin embargo, el VIH es capaz de evadir rápida y eficazmente esta presión inmune a través de la adquisición de mutaciones que disminuyen la eficacia de la respuesta citotóxica. Se ha observado que algunas de las mutaciones de escape afectan la capacidad replicativa viral (CRV) y pueden representar blancos interesantes de control del VIH. Por otro lado, algunas mutaciones pueden ser compensadas o seleccionadas por alelos frecuentes en la población y representar una vía de adaptación del VIH y progresión de la infección. La distribución distinta de alelos HLA-I entre poblaciones alrededor del mundo, sugiere la existencia de mecanismos de progresión población-específicos. Dado que las estrategias actuales de vacunación buscan inducir la respuesta mediada por CTLs, complementar la imagen de los sitios, las vías y la cinética del escape viral no solo ayudara a entender en qué medida la inmunidad del hospedero favorece la evolución del VIH, sino que dirigirá el diseño racional de futuros inmunógenos vacunales. En el presente estudio, se analizó el efecto de mutaciones seleccionadas por alelos HLA-I sobre la CRV en una cohorte guatemalteca no caracterizada previamente.

**METODOS** Se tipificaron a alta resolución los genes HLA-I A, B y C de 160 individuos Guatemaltecos infectados con VIH-1 subtipo B. Se construyeron virus recombinantes que expresaran la región genética gag-proteasa en el esqueleto del virus de referencia NL43. La CRV fue evaluada en una línea celular linfocitaria capaz de expresar el gen de la proteína verde fluorescente al ser infectada por VIH. Las secuencias de Gag fueron obtenidas y analizadas para posiciones bajo presión selectiva, estimando el cociente de mutaciones sinónimas y no sinónimas. Se usó la prueba de Spearman para evaluar la correlación entre CRV y progresión. El estadístico U Mann-Whitney se utilizó para identificar alelos HLA-I y polimorfismos asociados a cambios en la CRV. Se construyeron árboles filogenéticos de las mutaciones con impacto en la CRV usando el modelo de

Maximun Likelihood y 1000 repeticiones de Bootstrap. Se construyó la secuencia consenso viral y se buscaron epítomos de interés con la herramienta Motif Scan y la base de datos de Los Alamos.

**RESULTADOS** Se observó una correlación negativa ( $r=-0.1810$ ) entre CRV y cuenta de CD4+ ( $p=0.0220$ ) y una correlación positiva entre CRV y Carga Viral ( $r=0.1833$ ,  $p=0.0204$ ). Los alelos B\*18:01 ( $p=0.0355$ ), C\*05:01 ( $p=0.0276$ ) y C\*12:02 ( $p=0.0462$ ) se asociaron a altos valores de copias de carga viral. Los alelos A\*24:02 ( $p=0.0241$ ), C0501 ( $p=0.0295$ ) y C\*12:02 ( $p=0.0084$ ) fueron asociados a altas cuentas de células CD4+; mientras que B\*35:01 ( $p=0.0506$ ) y C\*04:01 ( $p=0.0028$ ) se asociaron a bajas cuentas de estas células. Los alelos A\*31:01 ( $p=0.0507$ ), A\*68:01 ( $p=0.0164$ ), B\*49:01 ( $p=0.0291$ ), C\*04:01 ( $p=0.016$ ) y C\*05:01 ( $p=0.0285$ ) se vieron asociados a altos valores de CRV. Se encontraron 18 posiciones bajo presión selectiva, 4 de las cuales fueron asociadas a CRV elevada, T84V ( $p=0.0593$ ), A115K ( $p=0.0410$ ), I91N ( $p=0.0091$ ) y S373T ( $p=0.055$ ). Se identificaron los epítomos restringidos por A\*31:01, A\*68:01, B\*35:01 y C\*04:01 por mostrar un impacto directo en la CRV y se mapearon los epítomos restringidos en Gag por los alelos más frecuentes en la población.

**CONCLUSION** Se describió por primera vez el contexto inmunogenético de los individuos infectados con VIH-1 de Guatemala identificándose alelos frecuentes de origen amerindio y alelos no caracterizados extensivamente en otras poblaciones. Se identificó un papel controversial de los alelos en desequilibrio de ligamiento B\*18:01 y C\*05:01. Las observaciones sugieren que la presencia de A\*68:01 selecciona mutaciones compensatorias dentro o fuera de Gag. Se identificó una posible vía de progresión a la infección basada en la selección simultanea de mutaciones de escape por parte de los alelos B\*35:01 y C\*04:01. Se observó que las mutaciones de escape T84V y A115K, seleccionadas por B\*35:01 y C\*04:01, podrían representar un evento de co-variación o compensación. Se construyó el primer mapa de selección inmune mediada por HLA en la secuencia de Gag circulante en la población guatemalteca.

## ABSTRACT

**BACKGROUND** Immune control of HIV infection has been associated with HLA-I alleles. This effect results from the binding of viral peptides, particularly derived from the Gag polyprotein, to HLA molecules that can be recognized by CD8+ cytotoxic T cells (CTL), with the consequent elimination of infected cells. However, HIV can readily evade this immune selective pressure through the acquisition of mutations that diminish the efficacy of cytotoxic response. Some of these escape mutations have been shown to affect viral replicative capacity (VRC) and could represent interesting targets for HIV control. On the other hand, some scape mutations may be compensated or selected by highly frequent alleles in a population, what could represent an adaptation pathway of the virus linked to disease progression. Different HLA-I allele frequency distributions in populations around the world suggest the existence of population-specific mechanisms of disease progression. The current vaccine development strategies aimed to induce a CTL response. So, to obtain the complete information about sites, pathway and dynamic of viral scape will improve the understanding of how host immunity shaped the viral evolution and may contribute to the vaccine design. We present results on the effect of particular HLA alleles and specific mutations on HIV VRC in a previously uncharacterized Guatemalan cohort.

**METHODS** HLA class I A, B and C genes were typed at high resolution from 160 Guatemalan HIV-1 subtype B-infected individuals. Chimeric viruses expressing the gag-protease genetic region in a NL4-3 backbone were obtained. VRC was evaluated with a reporter CD4+ T cell assay. Gag sequences were obtained and analyzed for positions under selective pressure, estimating rates of synonymous/non-synonymous substitutions. Spearman correlation was used to evaluate the effect of VRC on disease progression. Mann-Whitney tests were used to identify HLA-I alleles and polymorphisms associated with changes in VRC. Phylogenetic trees were built for each mutations linked to change in VRC using Maximum Likelihood method and 1000 Bootstrap. A consensus sequence was made and epitopes restricted by important alleles where mapped.

**RESULTS** Positive correlation between plasma viral load and VRC, and negative between CD4+ T cell count and VRC were observed in the Guatemalan cohort ( $r=0.1833$ ,  $p=0.0204$

and  $r=-0.1810$ ,  $p=0.0220$  respectively). B\*18:01 ( $p=0.0355$ ), C\*05:01 ( $p=0.0276$ ) and C\*12:02 ( $p=0.0462$ ) alleles were associated with high values of viral load. HLA alleles that showed a positive impact on CD4<sup>+</sup> T cell count included A\*24:02 ( $p=0.0241$ ), C\*05:01 ( $p=0.0295$ ) and C\*12:02 ( $p=0.0084$ ), while B\*35:01 ( $p=0.0506$ ) and C\*04:01 ( $p=0.0028$ ) were associated with decreased T cell counts. On the other hand, A\*31:01 ( $p=0.0507$ ), A\*68:01 ( $p=0.0164$ ), B\*49:01 ( $p=0.0291$ ), C\*04:01 ( $p=0.016$ ) and C\*05:01 ( $p=0.0285$ ) were associated with higher values of VRC. A total of 18 codons were found to be under positive selective pressure. Four of them were associated with increased VRC, T84V ( $p=0.0593$ ), A115K ( $p=0.0410$ ), I91N ( $p=0.0091$ ) and S373T ( $p=0.055$ ). Several epitopes restricted by A\*31:01, A\*68:01, B\*35:01 and C\*04:01 were identified. It was observed the selective pressure of most prevalent HLA alleles.

**CONCLUSION** The immunogenetic context of Guatemalan HIV-1-infected individuals was described for the first time. The most prevalent alleles were Amerindian and there were observed some alleles that have not been well characterized in other populations. A controversial role of B\*18:01 and C\*05:01 which are in linkage disequilibrium, was found. The results suggest A\*68:01 selects compensatory mutations within Gag or within other HIV genes, because this allele recognized a mutated epitope. It was identified a possible escape pathway dependent of simultaneous restriction of B\*35:01 y C\*04:01. Escape mutations T84V and A115K, which are selected by B\*35:01 and C\*04:01, might represent a co-variation or compensation event. The first epitope map of HLA selective pressure on Gag was built for Guatemalan population.

# “IMPACTO DE MUTACIONES SELECCIONADAS POR HLA EN LA CAPACIDAD REPLICATIVA DEL VIRUS DE INMUNODEFICIENCIA HUMANA CIRCULANTE EN LA REGIÓN MESOAMERICANA”

---

## **CAPITULO 1. INTRODUCCIÓN**

### **I. EPIDEMIOLOGÍA DE LA INFECCIÓN POR VIH y SIDA**

#### **I.1 DESCUBRIMIENTO DEL VIRUS DE INMUNODEFICIENCIA HUMANA**

El 5 de junio de 1981 el Centro de Control y Prevención de Enfermedades de los Estados Unidos (CDC) publicó un artículo en el *Reporte Semanal de Morbilidad y Mortalidad* sobre cinco pacientes hospitalizados por neumonía atípica causada por *Pneumocystis jiroveci*. Un mes después, el reporte había ascendido a veintiséis individuos diagnosticados con la misma enfermedad, aunada a una marcada depleción de Linfocitos T CD4+, adenopatía y variadas infecciones oportunistas, lo que representaba un franco estado de inmunosupresión y deterioro que posteriormente sería llamado Síndrome de Inmunodeficiencia Adquirida (SIDA).

La creciente detección de pacientes aparentemente con el mismo síndrome se volvió preocupante; lo que había iniciado como un padecimiento “único” de hombres homosexuales se había convertido ya en una enfermedad claramente contagiosa, mortal y desconocida también presente en personas hemofílicas, individuos con previas transfusiones y parejas de éstas. Surgió, entonces, la sospecha de que la rara enfermedad se transmitía por vía sanguínea y sexual.

Diversos grupos de investigación tenían toda su atención puesta en descubrir el agente causal de este padecimiento. En Estados Unidos, el doctor Robert C. Gallo había descubierto a los únicos retrovirus capaces de causar enfermedad en los humanos, nombrados Virus Linfotrópico T Humano (HLTV, por sus siglas en inglés) 1 y 2. Estos agentes eran responsables de un tipo particular de leucemias asociadas a un estado de debilidad y desgaste similar al observado en los pacientes con SIDA, despertando la sospecha de que un virus similar podía ser la etiología buscada.

En 1982 el grupo liderado por el Doctor Luc Montagnier del Instituto Pasteur en Francia



comenzó a trabajar en la misteriosa enfermedad, que aunque no se había detectado en un número notable de franceses, resultaba un problema que ya había rebasado la frontera estadounidense. Su interés inicial fue descubrir en una etapa temprana del síndrome, caracterizada por la aparición de adenopatía, la presencia de un virus; si esta búsqueda era exitosa comprobaría que el agente propuesto era causal del SIDA y no consecuencia de la inmunodepresión. Fue así como en enero de 1983, con la colaboración de los médicos Willy Rozenbaum y Françoise Brun-Vézinet, obtuvo una biopsia de ganglio. Junto con Jean-Claude Cherman y la joven científica Françoise Barré-Sinoussi realizaron cultivos celulares de linfocitos T, población celular que se veía claramente disminuida en las personas afectadas, y midieron la actividad de la transcriptasa reversa, una enzima característica de los retrovirus. Los resultados positivos en dicha prueba confirmaron que un retrovirus era el agente causal del SIDA. Múltiples trabajos colaborativos llevaron a este grupo a descubrir, en junio de 1983, que el virus en cuestión era estructural e inmunológicamente distinto del HLTV <sup>1</sup>.

Inicialmente se nombró al nuevo retrovirus LAV (Lymphadenopathy Associated Virus, Virus asociado a linfadenopatía) y su caracterización continuó a pesar del escepticismo de la sociedad científica. Hasta que en 1985 el grupo del Doctor Gallo en Bethesda e investigadores liderados por Jay Levy en San Francisco, ambos en Estados Unidos, confirmaron los resultados obtenidos por el grupo francés. La caracterización molecular de los agentes encontrados en los tres laboratorios culminó en la clasificación de este nuevo retrovirus dentro del grupo *Lentiviridae* y su identificación como el Virus de Inmunodeficiencia Humana (VIH)

Así como las investigaciones sobre el VIH avanzaban, la enfermedad también lo hacía y a pasos agigantados. Para esos años, la cantidad de infecciones reportadas en más de 15 estados de la Unión Americana y algunos países de Europa ya rebasaba los doce mil casos, convirtiéndose en una epidemia de atención urgente.

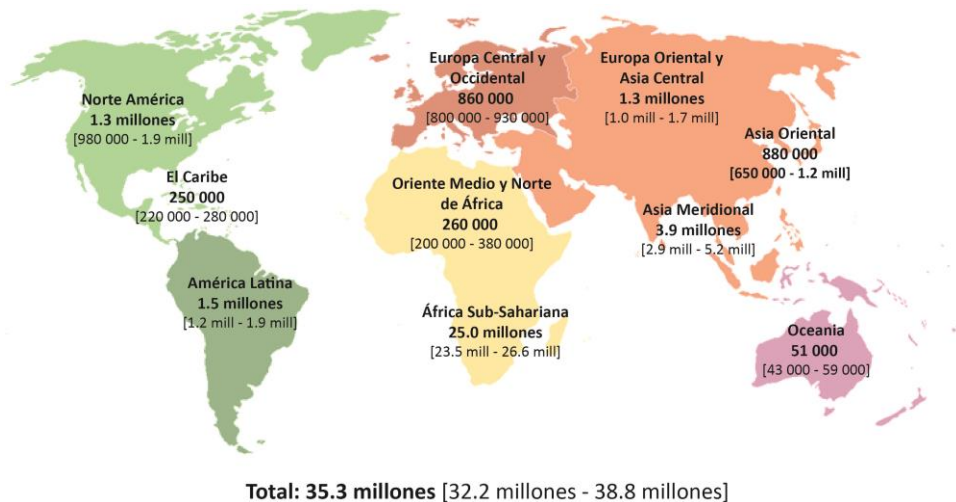
Mientras los intentos por desarrollar una vacuna o encontrar una terapia efectiva continuaban, en 1986 se descubre un virus estrechamente relacionado con el VIH. Éste fue encontrado únicamente en pacientes africanos y presentaba una progresión clínica más lenta, pero por su probada relación evolutiva con el agente etiológico del SIDA se le denominó VIH-2.

## I.2 LA EPIDEMIA EN EL MUNDO

En 2011 la Asamblea General de las Naciones Unidas estableció objetivos específicos relacionados con la transmisión, atención y control de la infección por VIH. Dentro de las llamadas “Metas del Milenio”, se propusieron diversas medidas para ampliar y mejorar la atención para personas que viven con VIH. Pero sin duda, el objetivo más ambicioso fue la reducción de la transmisión sexual y por usuarios de drogas intravenosas (UDI) al 50% y la erradicación de la transmisión a niños. A pocos meses de que el plazo propuesto se cumpla, los avances han sido notables, pero los retos continúan en todo el mundo.

Desde el registro de los primeros casos de VIH en 1981 hasta el año 2013 se han contabilizado 78 (71 - 87) millones de infecciones y 39 (35 - 43) millones de muertes relacionadas con SIDA<sup>2</sup>.

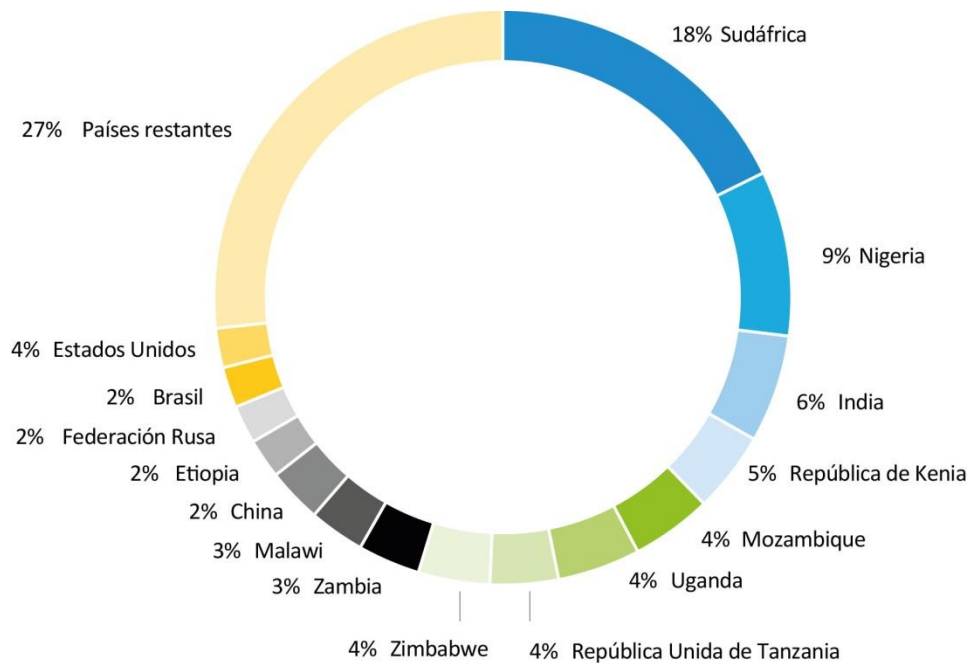
De acuerdo al último reporte emitido por el programa conjunto de las naciones unidas sobre VIH/SIDA (ONUSIDA), se estima que en todo el mundo existen 35.3 (32.2 - 38.8) millones de personas viviendo con VIH; de las cuales, 19 millones desconocen su estado de seropositividad. De las personas estimadas que viven con el virus, 24.7 (23.5 - 26.1) millones, es decir, aproximadamente el 70% del total viven en África subsahariana<sup>3</sup> (Figura 1).



**Figura 1. Distribución mundial de la infección por VIH.** Casos registrados de infección por VIH en las 10 regiones determinadas por la Organización Mundial de la Salud. África sub-Sahariana contabiliza cerca del 70% de las infecciones totales.

Solo en el año 2013, se registraron 2.1 (1.9 - 2.4) millones de nuevas infecciones y 1.5 (1.4 -1.7) millones de muertes relacionadas con SIDA. Se estimaron 3.2 (2.9 - 3.5)

millones de niños menores de 15 años viviendo con VIH y 4 (3.6 - 4.6) millones de jóvenes entre 15 y 24 años que han sido infectados alrededor del mundo <sup>3</sup>. Aunque se han registrado casos en prácticamente todo el mundo, en el año 2013 se observó que el 75% de las personas que viven con VIH se distribuyen entre 15 países, la mayoría de ellos del continente Africano en donde la atención en salud sigue siendo insuficiente (Figura 2).



**Figura 2. Distribución por país de casos de VIH en el mundo.** Cerca del 75% de los casos de VIH contabilizados por la OMS se concentran en los 15 países detallados en el esquema. Sud-África es la región de mayor incidencia, aportando el 18% de ellos. Modificada de <sup>3</sup>.

El sector de la población que registra más casos de VIH varía de país a país, mientras que en África subsahariana las mujeres jóvenes representan la población de mayor riesgo, en el Continente Americano los hombres que tienen sexo con hombres (HSH) se asocian 19 veces más con casos de VIH que la población en general (Figura 3).

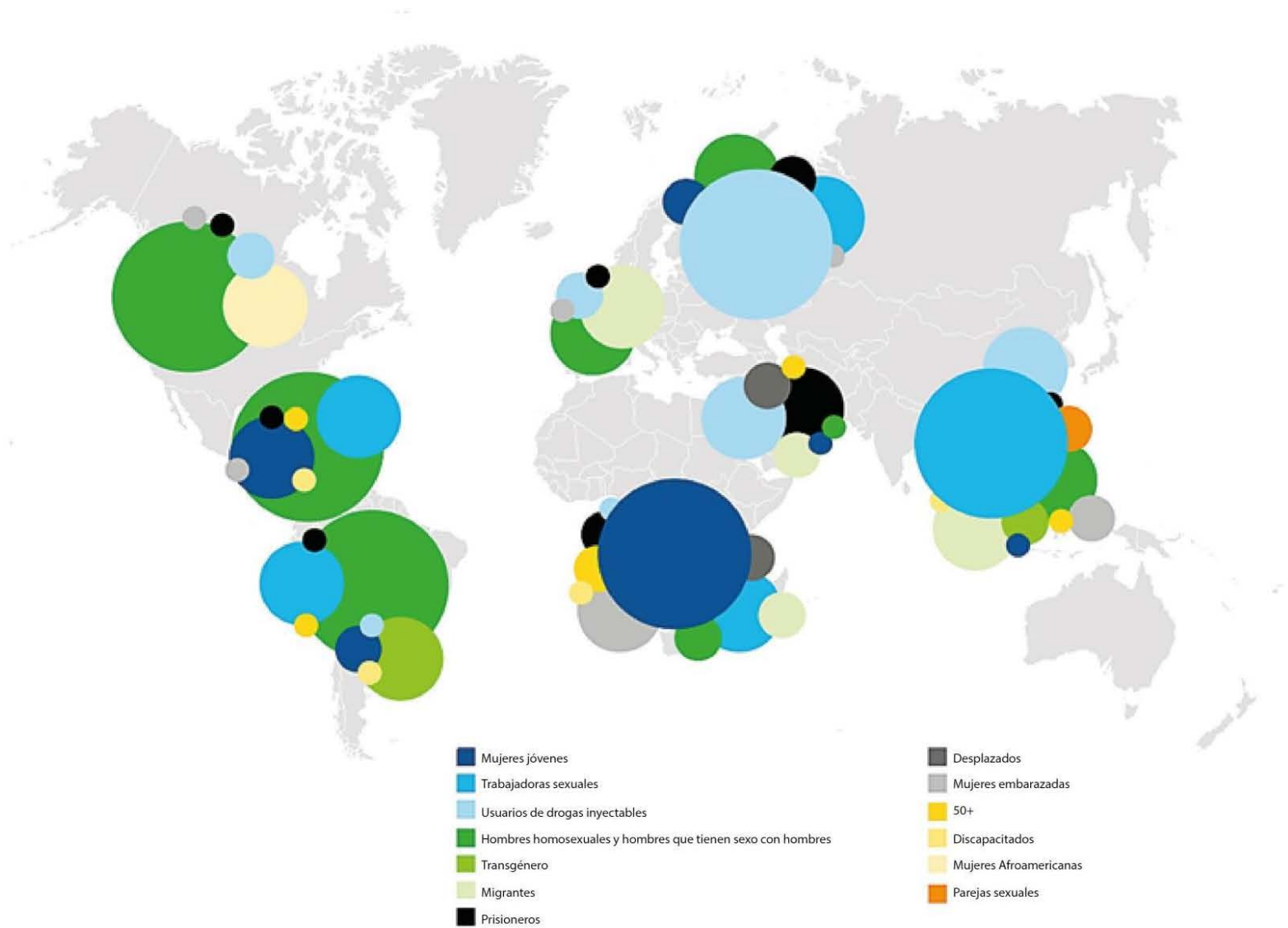
A pesar del panorama sombrío que aún prevalece en ciertas regiones, la respuesta ante la infección ha permitido que de forma global los nuevos casos decaigan en un 38% (de 2001 a 2013). Además, a finales del 2013 se estimó una inversión anual de 19.1 millones de dólares en la respuesta al VIH/SIDA en países de medianos y bajos ingresos, lo que permitió que cerca del 37% (12.9 millones) de las personas que viven con el virus tengan

acceso a Tratamiento Antirretroviral (TAR) y se eviten millones de muertes asociadas a SIDA.

Para incrementar las cifras de éxito, los esfuerzos globales se han dirigido a extender la cobertura en países de medianos y bajos ingresos, basándose principalmente en:

- Establecimiento de un momento óptimo de inicio de tratamiento mediante el uso de indicadores clínicos eficaces (Carga viral y cuenta de linfocitos T CD4+)
- Reducción del costo de antirretrovirales (ARVs) y simplificación de los regímenes actuales.
- Reducción de la incidencia de eventos adversos asociados a toxicidad de los medicamentos actuales.

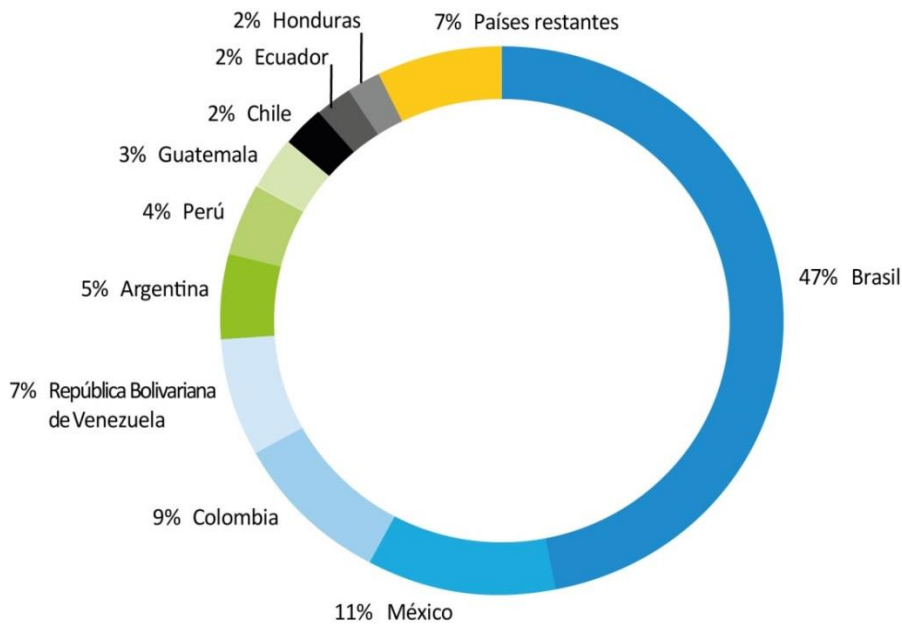
Dada la magnitud que ha adquirido esta epidemia, los recursos destinados a su investigación, prevención y tratamiento han ido en aumento. Sin embargo, una estrategia prometedora que pueda solucionar este creciente problema aún se ve lejana, por lo que profundizar en el conocimiento del virus, su comportamiento a nivel poblacional y sus interacciones con el hospedero, dirigen a la comunidad científica en la búsqueda de un manejo retroviral adecuado, e incluso, el desarrollo de una vacuna efectiva.



**Figura 3. Prevalencia de VIH por población de riesgo.** Identificado por colores se muestran los sectores poblacionales afectados en cada país. El tamaño de cada círculo representa la proporción de pacientes pertenecientes a determinado sector poblacional de riesgo. En el continente Americano, los hombres que tienen sexo con hombres son el grupo más afectado por VIH. Modificada de <sup>3</sup>

### I.3 LA INFECCIÓN POR VIH EN LATINOAMÉRICA

América Latina es considerada una región de medianos a bajos ingresos. Con sistemas de salud desiguales, la ONUSIDA estima que al 2013, 1.6 (1.4 – 2.4) millones de personas viven con VIH y en el último año las nuevas infecciones se contabilizaron en 94,000 (71,000 – 170,000). La epidemia en esta región está concentrada en hombres jóvenes; el 75% de los casos se localizan en Brasil, Colombia, México y Venezuela, mientras que el 9% corresponde a personas de Centro América (Figura 4).

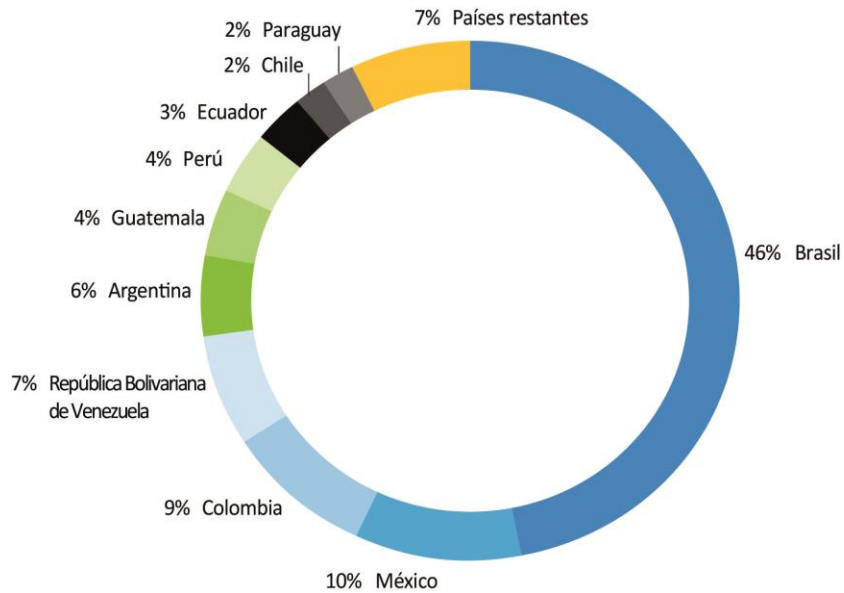


**Figura 4. Personas viviendo con VIH en Latinoamérica, 2013.** De acuerdo al último estimado de ONUSIDA, Brasil y México son los países más afectados por VIH en la región latinoamericana, mientras que Honduras, Ecuador y Chile son las regiones con menos casos contabilizados. Modificada de <sup>3</sup>

Gracias a los esfuerzos de cada país junto con el apoyo internacional, se ha logrado proporcionar tratamiento antirretroviral a cerca de la mitad de personas infectadas en esta área 45% (33-51%). Sin embargo, el reparto de medicamentos no es equivalente entre los países que componen la región, México se encuentra entre aquellos con >40% de cobertura, mientras que Guatemala otorga TAR a entre el 20 y 40% de sus pacientes.

De las personas estimadas con VIH en América Latina, 70% han sido diagnosticadas, 45% han iniciado TAR y 66% han logrado tener supresión de la replicación del virus. El inicio del tratamiento varía de acuerdo a las políticas de salud de cada entidad y aunque cada vez es más temprano (>500 CD4 en algunos países), el diagnóstico de la infección se sigue dando en etapas avanzadas de la enfermedad.

En México, de 2005 a 2013, los nuevos casos mostraron un decremento del 39% mientras que en Brasil aumentaron 11% en el mismo periodo; estas diferencias han hecho que la disminución generalizada de los nuevos registros de infección sea lenta en toda la región (Figura 5). Los sectores más vulnerables de la población lo componen personas transgénero, hombres homosexuales y hombres que tienen sexo con hombres.



**Figura 5. Nuevas infecciones por VIH en Latinoamérica, 2013.** Los nuevos casos de VIH en Latinoamérica se presentan de forma heterogénea entre los países de la región. Brasil y México contabilizaron en 2013 la mayoría de nuevas infecciones, mientras que en Paraguay y Chile el incremento en pacientes infectados fue mínimo. Tomada de <sup>3</sup>.

Los estudios epidemiológicos en esta región son insuficientes, lo cual incrementa el estimado de sub-registros de nuevas infecciones y muertes asociadas a VIH. La dispensación de medicamentos, aun en países vecinos, es alarmantemente distinta; así como la infraestructura hospitalaria y personal médico capacitado y dispuesto para atender a los pacientes, debido a la disparidad de recursos con los que cuenta cada nación. Se sabe que diez países en la región son capaces de cubrir el 90% o más del gasto en la atención a la infección con recursos propios; sin embargo, Bolivia, Guatemala, Honduras y Nicaragua dependen de recursos internacionales para financiar el 40% o más de la atención a la pandemia.

Por ello, el registro adecuado de casos, la detección temprana y la dispensación de TAR son algunos de los mayores retos que enfrenta Latinoamérica de cara a la pandemia.

El proyecto ambicioso de 2015, aunado al objetivo de detener la pandemia en 2030, hacen imperativa una acción colectiva del gobierno, la sociedad civil y la comunidad médica y científica. La Organización Mundial de la Salud ha establecido ocho planes de acción para alcanzar dichas metas. Entre ellos, se puntualiza la necesidad de estudiar la epidemia a nivel local. La identificación de áreas de alta incidencia de VIH, las actividades de riesgo que favorecen el incremento del número de infecciones, así como el registro adecuado y completo de casos en cada distrito, contribuirán a adecuar las políticas de acción a las necesidades de cada comunidad.

#### **I.4 LA INFECCIÓN EN MÉXICO Y GUATEMALA**

El Centro Nacional para la Prevención y el Control del VIH y SIDA (CENSIDA) en México, reportó 223,995 casos acumulados de VIH o SIDA desde la aparición del primer caso en México, en 1983, hasta el 30 de septiembre del 2014. La tasa en este mismo periodo (1983- 2014) se ha calculado en 186.7 casos/100,000 hab. Siendo, Distrito Federal (414.7), Yucatán (294.5), Veracruz (280.1), Baja California (272.2) y Campeche (271.7) <sup>4,5</sup> las entidades federativas con más reportes (Figura 6).

A finales de 2013, se informó de 180,000 (140,000 – 230,000) nuevas infecciones, a las que se sumaron 2,894 en el último trimestre de 2014 y 2,785 nuevos reportes de SIDA. Sin embargo, dentro del estimado global, se calcula que cerca del 50% de las personas que están infectadas desconocen su estado serológico <sup>5</sup>.

De acuerdo a ONUSIDA, en México la epidemia está concentrada entre HSH, usuarios de drogas intravenosas (UDIs), trabajadoras y trabajadores sexuales, personas transgénero, transexuales y travestis. No es de sorprender, entonces, que el 80.2% de los casos correspondan a hombres, la mayoría de ellos alrededor de los 30 años <sup>5</sup>. Es importante mencionar, que la vía de transmisión sexual sigue siendo la predominante en el país, siendo asociada al 95.1% de los casos registrados.

En todo el país, 89,410 personas reciben tratamiento antirretroviral de algún sistema de salud público o privado; de ellas, 57,073 lo reciben directamente de la Secretaría de Salud<sup>5</sup>. No obstante, el país aún enfrenta dos desafíos en la atención a personas con VIH y SIDA:



- ⚡ Solo el 33% de las personas con VIH tienen carga viral indetectable.
- ⚡ La tendencia de mortalidad varía por condición de derechohabiencia, sexo, edad y entre entidades federativas.

Por lo que la actualización del conocimiento y el esfuerzo constante para mejorar la atención, infraestructura y servicios dirigidos a estos pacientes, permitirá el éxito en la detección y tratamiento oportuno.

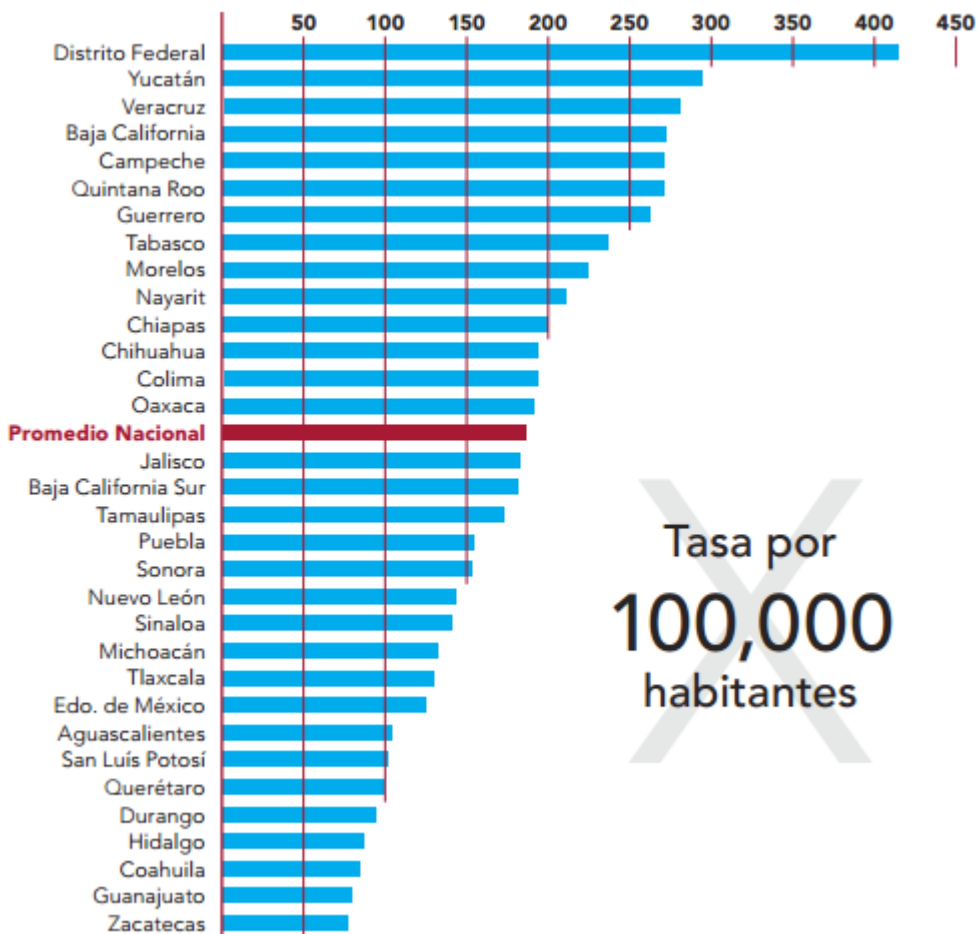


Figura 6. Incidencia acumulada de casos de VIH y SIDA según Entidad Federativa. Se muestran los casos acumulados por cada 100,000 habitantes en los 32 estados de la República Mexicana en el periodo 1983 a 2014. Catorce entidades se encuentran por encima de la media nacional, siendo el Distrito Federal el que presenta la mayor incidencia de casos de VIH. Tomada de <sup>4</sup>.

Por otro lado, la infección por VIH en Guatemala fue detectada en 1984 en dos pacientes masculinos, dos años después tres mujeres más recibieron el mismo diagnóstico. Estas cinco personas fueron remitidas a servicios de salud en estadíos avanzados de la infección, como era normal en aquellos años en todo el mundo. A partir de esos casos y

hasta 2003, en este país solo se realizaba vigilancia sobre pacientes en etapas avanzadas o SIDA; y es hasta el año 2004 cuando se incorpora en los datos globales de esta infección a pacientes en cualquier otro estadio clínico.

De acuerdo al Centro Nacional de Epidemiología (CNE), de 1984 hasta septiembre del 2014, se han acumulado 32, 858 personas afectadas, de estas, 42% con VIH y 58% con SIDA <sup>6</sup>. El mayor porcentaje de las personas afectadas tienen entre 30 y 39 años de edad y se concentran en 8 departamentos del país. El grupo de mayor riesgo lo representan hombres que tiene sexo con hombres, mujeres transgénero y clientes de servicios sexuales. La vía de transmisión sexual es responsable del 95% de los casos contabilizados en la nación.

En cuanto a la mortalidad por afecciones asociadas a VIH y SIDA, de enero del 2005 a diciembre 2013 se contabilizaron 5,529 muertes. De acuerdo al INE, se observó una tendencia descendente en las muertes durante este periodo, en 2005 la tasa de mortalidad por causas asociadas a VIH se estimaba en 5.6 (por cada 100,000 hab.), mientras que en 2013 alcanzó 2.9, sin embargo la razón de población se mantuvo estable, siendo 2.4 hombres afectados por cada mujer. El grupo de edad con mayor mortalidad corresponde a individuos entre 40 y 59 años<sup>7</sup>.

De las defunciones registradas, 85.56% fueron de personas provenientes de zonas urbanas y cerca del 20% de personas identificadas como indígenas. Los departamentos que reportaron la mayor ocurrencia de mortalidad en 2013 fueron Escuintla, Guatemala Izabal y Quetzaltenango (Figura 7A). Sin embargo, los pacientes que murieron por causas asociadas a VIH no eran originarios necesariamente de estas provincias. De acuerdo al origen de los individuos, los departamentos que contabilizan mayor cantidad de defunciones por VIH/SIDA son Escuintla, Izabal, Retalhuleu y Suchitepequez (Figura 7B).

A pesar de las cifras elevadas y el sub-registro de morbi-mortalidad en Guatemala, se han obtenido logros significativos a través del establecimiento y mejoramiento de las 18 Unidades de Atención Integral a pacientes con VIH y las unidades centinela encargadas de la detección y canalización de pacientes seropositivos a servicios de salud de mayor nivel. Uno de estos logros ha sido la reducción de casos de transmisión vertical, previniéndose alrededor de 400 casos de transmisión materno-fetal en 2013 <sup>8</sup>. Así mismo,

cerca de 16,386 personas ya reciben TAR. La inversión en prevención, detección y atención del VIH y SIDA ha ido en incremento en el último lustro, y se estima que se gasten 86.8 millones de dólares en 2015, provenientes de presupuesto público, privado y apoyo internacional <sup>8</sup>.

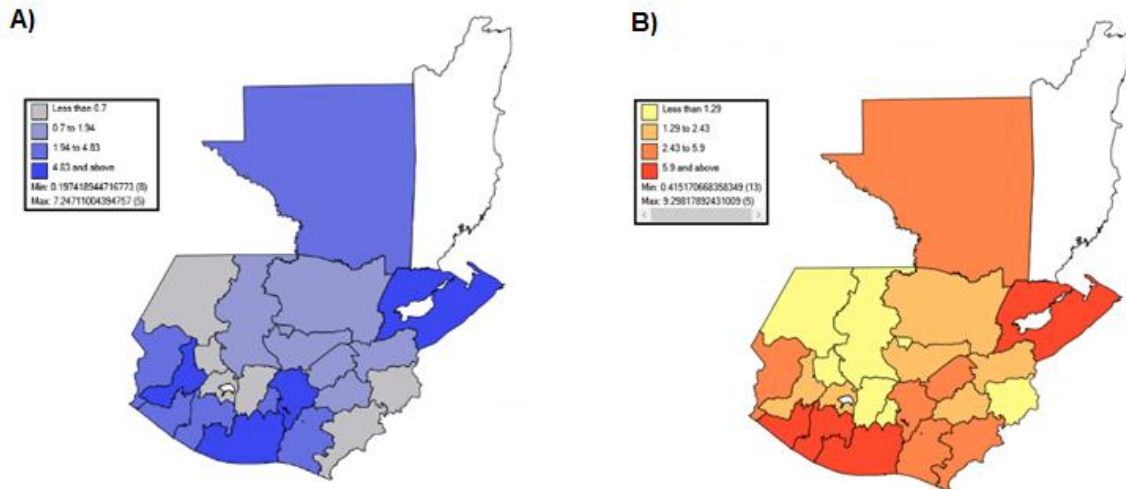


Figura 7. Mortalidad en Guatemala por causas asociadas a VIH/SIDA según ocurrencia y procedencia. A. Se muestra la prevalencia de decesos en pacientes con VIH de acuerdo al lugar de ocurrencia y registro de la defunción. B. Se muestra la prevalencia de defunciones por causas asociadas a VIH/SIDA por lugar de origen de los individuos.

## II. GENERALIDADES DEL VIRUS DE INMUNODEFICIENCIA HUMANA

### II.1 CLASIFICACIÓN DEL VIRUS

El VIH pertenece a la familia *Retroviridae*, que posee una enzima capaz de transcribir su genoma en sentido inverso, es decir, convierten su genoma de RNA en ADN para poder ser insertado en el genoma de la célula hospedera. Esta familia incluye 7 géneros, de los cuales, únicamente 2 causan enfermedades en humanos. Los *Retrovirus BLV-HTLV* que originan un tipo particular de leucemia y los *Lentivirus*, género al cual pertenecen el VIH y un virus similar que causa inmunodeficiencia en simios y del que se cree que proviene el primero.

Una de las teorías mayormente estudiadas ubica el origen de la pandemia en un evento de zoonosis entre simios y humanos en el Centro-Occidente de África. El hallazgo de virus que causan una enfermedad equivalente en casi 40 especies distintas de primates no humanos fortalece esta proposición. Análisis del genoma de VIH y de distintos Virus de

Inmunodeficiencia en Simios (SIV, en inglés) permitieron establecer con mayor certeza esa relación filogenética y una clasificación genética de los virus circulantes alrededor del mundo.

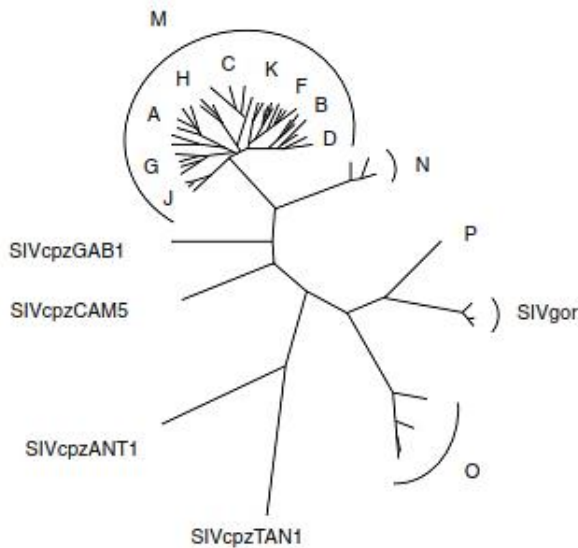
La primera división se estableció entre dos tipos principales de VIH, VIH-1 que es el más ampliamente distribuido a nivel mundial y VIH-2 que es un virus estructuralmente homólogo al VIH-1, pero con diferencias notables en su secuencia. La infección por este último se centra principalmente en algunos países de África y aunque se han presentado casos fuera de dicho continente, la dispersión de este tipo viral es lenta.

Al analizar filogenéticamente la secuencia de los distintos tipos de VIH y SIV, fue posible crear una subclasificación más detallada, así como establecer posibles fechas de origen de cada subtipo. Dentro del VIH-1 se identifican cuatro linajes: M (Main), O (Outlier), N (Non-M, Non-O) y P (Pending identification). En el grupo M se ubica cerca del 90% de las infecciones de todo el mundo, el grupo O está concentrado en Camerún, Gabón y Guinea Ecuatorial; por otro lado, el grupo N ha sido encontrado únicamente en pacientes de África occidental, mientras que el grupo P fue identificado en 2009 en dos individuos de Yaounde, Camerún<sup>9,10</sup>.

De acuerdo a estudios filogenéticos, se sabe que los grupos M y N se originaron de manera directa del SIVcpz, el virus responsable de la inmunodeficiencia en los primates *Pan troglodytes troglodytes* que habitan en África Central y África Occidental. El grupo P y O, por otro lado, se encuentran más cercanos evolutivamente al SIVgor<sup>9,11</sup> que infecta a gorilas originarios de las tierras bajas del occidente de África (Figura 8).

Por estudios moleculares evolutivos se cree que el origen del grupo M puede situarse alrededor de 1900, siendo el linaje más antiguo en los humanos, seguido por el grupo O, que pudo haber aparecido alrededor de 1920 (1840-1920); y finalmente, la transmisión inter especies que derivó en el grupo N se cree que sucedió en 1963 (1984-1977)<sup>9,10</sup>.

Dentro de cada uno de los linajes mencionados se encuentran agrupados distintos subtipos virales o clados, entre los cuales hay una variación del 17 al 35% de su secuencia proteica. Por otro lado, la diferencia en la secuencia de aminoácidos entre virus pertenecientes al mismo subtipo se estima entre el 8 y el 17%<sup>9</sup>.



**Figura 8. Árbol filogenético de SIV y VIH.** Se ha establecido la relación evolutiva entre los Virus de Inmunodeficiencia en simios (SIV, en inglés) y el Virus de Inmunodeficiencia Humana (VIH). Los linajes M y N de VIH-1 se ubican cercanos al virus que infecta primates de la especie *Pan troglodytes troglodytes* (SIVcpz); mientras que los linajes P y O se relacionan con el virus de gorila (SIVgor). Tomada de <sup>9</sup>.

Los subtipos que han sido identificados dentro del grupo M son A, B, C, D, F, G, H, J y K; mientras que el grupo O ha sido subdividido en los clados I al V. Los clados del grupo M son los que se han detectado en todo el mundo en diferentes proporciones en cada población. El subtipo C tiene las cifras más altas de prevalencia global por ser predominante en África sub-Sahariana y la India; mientras que el subtipo B es el de mayor incidencia en el mundo, incluido el continente Americano (Figura 9).



**Figura 9. Distribución mundial de los subtipos de VIH-1.** Se muestra la distribución de los subtipos de VIH-1 en cada continente. En mayúsculas se indica el clado predominante. Tomada de <sup>12</sup>

En las regiones donde existe una prevalencia considerable de varios subtipos virales es posible la infección de un individuo por más de un clado; por ejemplo, en África han sido detectadas infecciones duales en 10-20% de los pacientes. Este fenómeno favorece la recombinación entre subtipos virales o cepas de diferentes grupos, siendo más frecuente la generación de virus híbridos del grupo M.

A los virus recombinantes de grupo M se les denomina formas recombinantes circulantes (CRFs) si son encontrados en tres o más individuos epidemiológicamente no relacionados, o como URFs (formas recombinantes únicas) cuando no cumplen este criterio. Hasta el momento, se han identificado 66 CRF<sup>13</sup> y se estima que el 20% de las infecciones en el mundo, se dan por virus recombinantes de ambas clases<sup>11</sup>.

## **II.2 TROPISMO, RECEPTORES Y CO-RECEPTORES**

El tropismo viral se define como la atracción altamente específica de un virus hacia un tejido del hospedero, determinado por los marcadores de superficie de las células que lo componen.

Diversos estudios demostraron que el VIH tiene una afinidad marcada por células en cuya superficie esté expresada la molécula CD4. Es por ello que las principales células blanco son Linfocitos T ayudadores (CD4+), cuya depleción es característica de la infección, y macrófagos, que han sido identificados como reservorios virales. Aunque las células dendríticas expresan en cantidades reducidas CD4, también son susceptibles de ser infectadas.

El receptor CD4 es miembro de la superfamilia de las inmunoglobulinas que de forma natural estabiliza la interacción entre el receptor del linfocito T y las moléculas del complejo mayor de histocompatibilidad de clase II (MHC-II) para el proceso de presentación antigénica<sup>14</sup>. En el contexto de la infección, se ha identificado un sitio de unión de alta afinidad hacia la proteína viral gp120 en el fragmento N-terminal de su dominio extracelular. Esta interacción es el primer paso en el ciclo replicativo del VIH.

Se sabe que la simple unión de gp120 y CD4<sup>+</sup> no es suficiente para permitir la entrada del virus a la célula, es decir, es necesaria la participación de una serie de correceptores que refuercen la respuesta. Se sabe que miembros de la superfamilia de proteínas de siete asas transmembranales acoplados a proteína G cumplen con esta función.

Los dos correceptores principales para VIH son CXCR4 y CCR5. Ambos funcionan normalmente en la célula como receptores de quimiocinas. El uso preferencial de alguno de ellos define el tropismo viral, clasificando al virus en M-trópicos o R5 si el correceptor usado es CCR5 y T-trópicos o X4 si el correceptor CXCR4 es el utilizado. Se han aislado

algunos virus cuya afinidad es intermedia, definiéndose como virus de tropismo dual <sup>15</sup>. Se ha observado que el uso de CCR5 se da preferencialmente en etapas tempranas de la infección, mientras que los virus que usan CXCR4 aparecen típicamente en etapas avanzadas y se asocian con un aumento en la patogenicidad<sup>16</sup>.

Otros receptores de quimiocinas que se han propuesto como posibles correceptores para el VIH incluyen CCR2b, CCR3, CCR8, CCR9, APJ entre otros. Sin embargo, su importancia biológica aun no es clara <sup>17</sup>.

### II.3 ORGANIZACIÓN GENÓMICA Y PROTEÍNAS VIRALES

El VIH es un virus esférico, con un diámetro aproximado de 100 a 120 nm. Posee un genoma de ARN lineal de cadena sencilla y sentido positivo, con una longitud cercana a las 9.3 kb. Cada partícula viral o virión posee dos copias, no necesariamente idénticas, de su genoma empaquetado en una estructura llamada cápside, de geometría cónica. Además, está envuelto por una bicapa lipídica que es un fragmento de la membrana celular de la célula hospedera a partir de la cual gemó y en la que se hallan ancladas proteínas virales (gp120/gp41) (Figura 10) y en ocasiones proteínas celulares<sup>18</sup>.

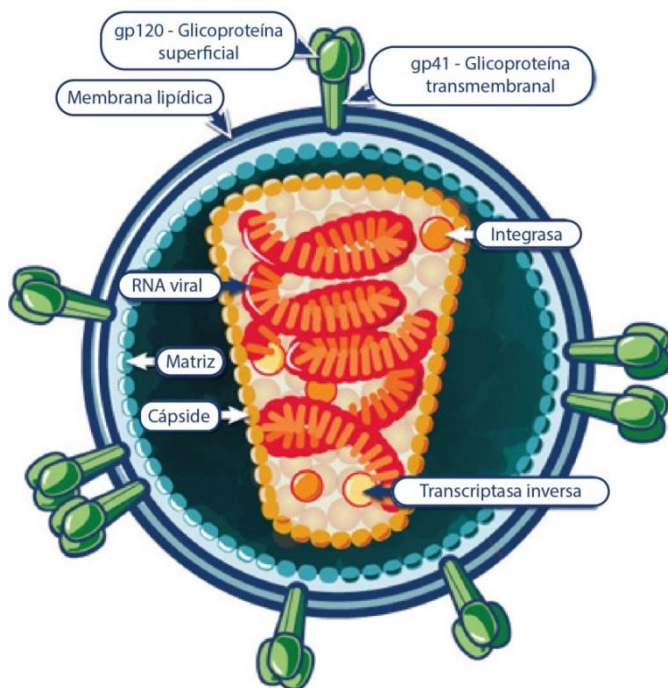


Figura 10. Representación esquemática de un Virus de Inmunodeficiencia Humana. Cada partícula viral se encuentra envuelta en una bicapa lipídica sobre la que se expresan espículas conformadas por gp120 y gp41. Al interior está conformada la matriz que protege la cápside cónica que a su vez, aloja proteínas y enzimas virales, así como dos copias de RNA. Modificada de: <http://www.niaid.nih.gov/topics/hiv/aids/understanding/howhivcausesaids/pages/howhiv.aspx>

Como en todos los retrovirus, el genoma del VIH está compuesto por 3 genes básicos que codifican para proteínas estructurales, enzimas y espículas. Además, posee seis genes codificantes de proteínas que favorecen la patogenicidad y la replicación (Figura 11). En total, en el ARN viral están contenidos 9 genes y dos LTRs (extremos). Dichos genes son traducidos en tres marcos de lectura distintos para producir 15 proteínas que se detallan a continuación.

Las **proteínas estructurales** están codificadas en el gen *gag* (group specific antigen). Éste se traduce en una poliproteína de 55 kDa (Pr55<sub>Gag</sub>) que debe sufrir procesamiento proteolítico para originar polipéptidos funcionales capaces de autoensamblarse para formar al virión. Los fragmentos originados a partir de dicho procesamiento son **p17** (matriz, MA), **p24** (cápside, CA), **p7** (nucleocápside, NC) y **p6**; además de **p1** y **p2** que funcionan como péptidos espaciadores (Figura 11).

La proteína p17 consta de 5 alfa hélices, 2 hélices cortas y 1  $\beta$  plegada <sup>18</sup>. Alrededor de 200 copias de este péptido se ensamblan en trímeros que posteriormente se asocian para formar una superficie básica que se une a la parte interna de la envoltura viral (Figura 10). Además, esta proteína se conjuga con ácido mirístico en los residuos de glicina de su extremo amino terminal, lo que le confiere afinidad hacia membranas lipídicas.

Por otra parte, p24 es una proteína compuesta por dos dominios de  $\alpha$ -hélices y un asa flexible que funciona de manera multimérica. Esta proteína es capaz de ensamblar anillos de 5 o 6 unidades. Los anillos posteriormente se agregan para formar una estructura similar a un fullereno. La conformación inicial contiene en su base hexámeros de p24 y hacia la parte terminal pentámeros de la proteína, lo que produce una geometría cónica <sup>19</sup>. Para el ensamblaje de dicha estructura, es necesaria la interacción con una prolina-isomerasa humana conocida como ciclofilina A (CypA) cuyo dominio de reconocimiento se encuentra en el asa de p24 <sup>18</sup>.

El genoma viral está protegido por la nucleocápside. Las proteínas que conforman esta estructura, p7, tienen dos motivos del tipo dedos de zinc que permiten la interacción con el ARN. Aunada a esta función estructural, NC se une a cuatro asas o loops en el ácido nucléico conocidos como “señal de encapsidación” para promover su dimerización y empaquetamiento en los viriones nacientes, actúa como chaperona para la transcripción



reversa del material genético protegiéndolo de nucleasas y colocando un tRNA para el inicio de este proceso identificado como tRNA<sub>Lys3</sub> y coadyuva en la integración del genoma viral <sup>18</sup>.

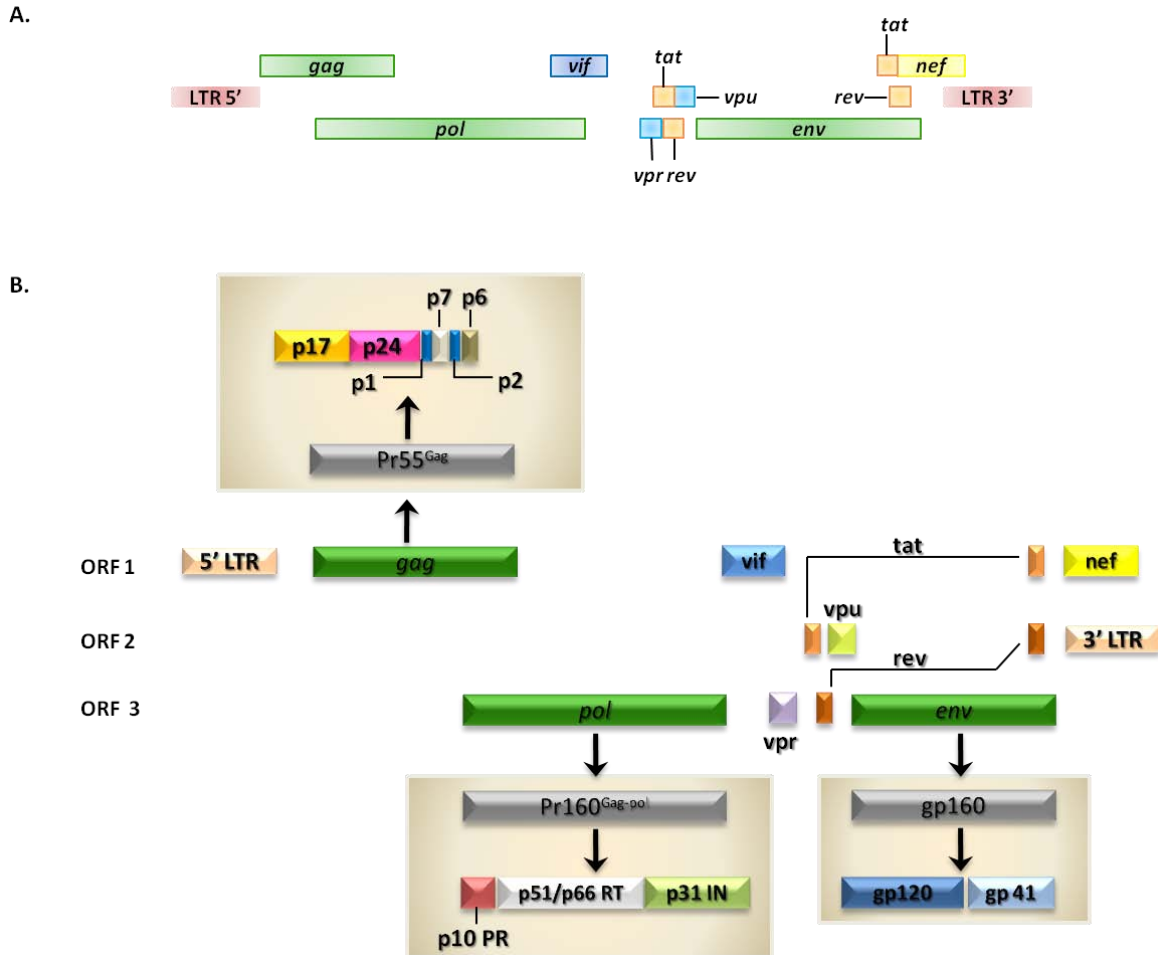
A partir del extremo C-terminal de *gag* se genera p6, una proteína indispensable para la liberación de partículas virales que interactúa con factores celulares de la familia ESCRT.

El gen *pol* produce una poliproteína conocida como Pr160<sup>Gag-pol</sup> a partir de la cual derivan las tres **enzimas virales** (Figura 11): **Proteasa** (PR, p10), **Transcriptasa inversa** (RT, p66/p51) e **Integrasa** (IN, p32).

La proteasa es una aspartil-proteasa localizada fuera de la cápside y es activa en forma de dímero. Esta enzima es responsable de cortar los precursores Pr55 y autocatalizar el corte de Pr160 en diversos polipéptidos justo después de la gemación del virión. Al generarse distintas proteínas se presentan cambios morfológicos conocidos como maduración del virus, confiriéndole infectividad.

La transcriptasa reversa es una DNA polimerasa RNA-dependiente que produce una copia de ADN de doble cadena a partir del ARN viral para producir más partículas virales. Alrededor de 50 copias de RT son empaquetadas en la cápside junto al genoma viral. Funciona como un heterodímero compuesto de las subunidades p66 y p51. En p66 se encuentran dos dominios catalíticos, uno encargado de la polimerización de ADN y el otro que actúa como ARNasa H, que digiere el ARN del dímero transitorio ARN-ADN durante el proceso de transcripción reversa. Sin embargo, ninguna de las subunidades tiene actividad de exonucleasa 3'-5', es decir, la polimerasa no es capaz de corregir errores de transcripción, lo que aumenta la variabilidad génica del VIH. De hecho, se estima una tasa de mutación de  $3.4 \times 10^{-5}$  mutaciones/nucleótido/ciclo <sup>20</sup>.

Por su parte, La Integrasa es la enzima que cataliza la inserción del genoma viral en el ADN de la célula hospedera. Este proceso lo realiza mediante la catálisis de dos reacciones, el procesamiento del extremo 3' de la doble cadena de ADN (provirus) que genera extremos cohesivos y la integración de éste en el genoma celular.



**Figura 11. Genoma y proteínas del VIH.** A. El genoma del VIH de 9.3kb posee tres marcos de lectura (ORF) codificantes para proteínas estructurales (*gag*, *pol* y *env*), proteínas accesorias (*vif*, *vpu* y *vpr*) y proteínas reguladoras (*tat* y *rev*). Además, existen dos secuencias de repetición flanqueantes, LTR5' y LTR3' que contienen secuencias promotoras y reguladoras de transcripción. B. Los genes *gag* y *pol* se transcriben en los precursores Pr55 y Pr160 respectivamente. A partir de Pr55 se generan matriz (p17, MA), cápside (p24, CA), nucleocápside (p7, NC), p6 y dos péptidos espaciadores (p1 y p2). De Pr160 se originan las enzimas proteasa (p10, PR), transcriptasa inversa (p51, RT) e Integrasa (p32, IN). *Env* se traduce en gp160 que al escindirse origina a gp120 (Superficie, SU) y gp41 (Transmembranal, TM).

En cuanto a las **proteínas de membrana** del VIH, éstas se originan a partir del gen *env*, el cual se traduce en la glicoproteína gp160. A partir de esta poliproteína, se produce un péptido superficial conocido como **gp120** (SU) y uno transmembranal llamado **gp41** (TM).

Gp120 es un polipéptido hidrofílico que se encuentra en la superficie de la envoltura y posee cinco asas (V1-V5) cuya secuencia es altamente variable. Su función primordial es unirse a la molécula de superficie CD4 presente en las células blanco del virus e interactuar con receptores de quimiocinas (CCR5 y CXCR4) para promover la infección.

La proteína gp41 atraviesa en siete ocasiones la membrana y se ensambla con gp120 en 72 espículas que se localizan ancladas a la membrana lipídica que rodea al virus y que además se encuentran “decoradas” con carbohidratos para evitar su reconocimiento por parte del sistema inmune <sup>21</sup>.

La interacción del dominio superficial con los receptores celulares, induce cambios conformacionales en gp41 que promueven la fusión de la membrana celular con la viral, proceso que constituye el primer paso del ciclo de replicación del VIH.

Finalmente, las **proteínas accesorias** se encuentran en menor cantidad pero no por ello tienen funciones poco importantes. Se encuentran codificadas en marcos de lectura sobrelapados y algunas son producto del procesamiento de ARNm.

**Tat** (trans-activator of transcription) y **Rev** (Regulator of virion) son proteínas de regulación transcripcional. La primera de ellas puede ser encontrada en dos formas, un péptido de 71 aminoácidos llamado Tat-1 exon y otro de 86 a.a. o Tat-2 exon, esta última es la mayormente observada en el núcleo de la célula infectada, aunque ambas formas se pueden indentificar en individuos con infecciones persistentes. Su función consiste en potenciar la procesividad de la ARN polimerasa aumentando la producción de todas las proteínas virales. Por su parte, Rev es una proteína de 19 kD que se une al mRNA de las proteínas estructurales para regular su producción y procesamiento (splicing). Es una proteína sintetizada en la etapa temprana de la infección y aumenta su concentración durante la replicación viral para regular la exportación de proteínas desde el núcleo hasta el citoplasma y estabilizarlas antes de ser ensambladas.

**Vif** (Viral infectivity factor) es una proteína asociada con la infectividad de los viriones así como con la producción de los mismos, pues se ha observado que cuando esta molécula está ausente, las nuevas partículas virales son defectuosas. Además, está relacionada con la supresión de mecanismos anti-virales montados por el hospedero, principalmente al inhibir la enzima APOBEC3G, una citidin-desaminasa que funciona como factor antirretroviral al causar hipermutación en el VIH<sup>22</sup>.

**Vpr** (Viral protein r) es una proteína de 96 aminoácidos que se asocia con factores virales y celulares para conformar el complejo que dirige al provirus hasta el sitio de integración.

Así mismo, se ha sugerido que causa arresto celular en G2 inhibiendo la apoptosis de la célula infectada, pero favoreciendo su diferenciación. Por otro lado, podría actuar como un transactivador de genes celulares favorables para la replicación viral principalmente en macrófagos. En VIH-2 y algunos SIV esta proteína es reemplazada por Vpx, que se cree que surgió por un evento de duplicación del gen *vpr*.

**Vpu** (Viral protein u) es una proteína exclusiva de VIH-1 y los SIV más cercanos a él. Presenta un dominio transmembranal y otro citoplasmático. Se le han adjudicado dos funciones principales a) favorecer la degradación de moléculas CD4 por vía proteosomal y b) facilitar la liberación de los virus de la membrana plasmática a través del bloqueo de la proteína celular enlacina (Tetherina, inglés)<sup>22</sup>.

**Nef** (Negative regulatory factor) es la proteína más inmunogénica del VIH que es producida de forma temprana en la infección. Es capaz de inhibir la expresión de CD4 y moléculas del complejo mayor de histocompatibilidad (MHC) para favorecer la persistencia del virus y evasión del sistema inmune. Además interactúa con proteínas celulares para modificar algunas vías de señalización a favor de la replicación viral<sup>22</sup>.

Mención aparte merecen algunos elementos estructurales del genoma del VIH, como dos regiones que flanquean el RNA conocidas como repeticiones terminales largas (LTR) que son secuencias idénticas divididas en tres dominios: U3, R y U5. U3 del LTR 5' contiene gran parte de los elementos de control transcripcional, incluido el promotor. R forma estructuras híbridas de ARN-ADN para permitir a RT la transcripción de hebras completas de ADN. En U5 se encuentran sitios de unión a factores de transcripción que regulan la síntesis de ARNm durante la infección. Los LTRs participan también en la integración del provirus en el ADN hospedero<sup>13</sup>.

## **II.4 CICLO DE REPLICACIÓN DEL VIH**

A pesar del tamaño relativamente pequeño del genoma del VIH, sus proteínas son capaces de reclutar algunas moléculas celulares y actuar de manera conjunta para la replicación eficaz y continua del virus.

El ciclo de replicación viral es un proceso complejo y su duración depende de la célula en

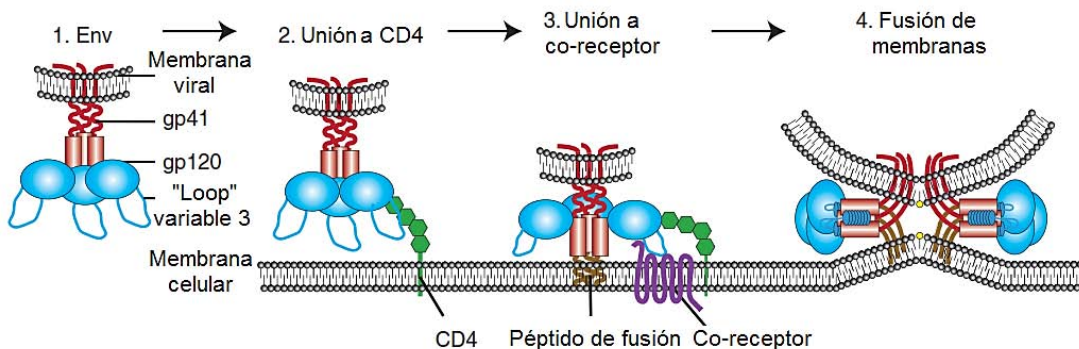
la que ocurra y el estado de activación de la misma. Los eventos que ocurren en la etapa temprana, son:

### 1. Unión y entrada

La entrada del virus a la célula es un proceso que no causa daño inmediato pero desencadena una cascada de señales que facilitan la replicación viral. En este proceso participan la proteína gp120 y el trímero de gp41 asociado a ella, que en conjunto conforman una espícula viral.

El proceso de entrada, esquematizado en la Figura 12, requiere primeramente de la formación de un puente entre la proteína viral gp120 y el receptor CD4. Esta interacción expone dominios de gp120, primordialmente el asa V3, que se unen al correceptor (CXCR4 o CCR5). La unión a ambas proteínas celulares promueve la inserción de un péptido de fusión localizado en el extremo amino terminal de gp41 en la membrana celular. La inserción de dicho péptido por parte de dos unidades de gp41 forman una bisagra que facilita el acercamiento de la membrana viral a la membrana celular<sup>11,19</sup>.

El acercamiento de ambas membranas es seguido por el agrupamiento de dominios helicoidales presentes en los extremos de cada unidad de gp41; formando un “poro de fusión” de 6 hélices que culmina con la fusión de membranas y la liberación en la cápside al citoplasma de la célula<sup>23</sup>.



**Figura 12. Fusión de membranas mediada por gp120 y gp41.** La proteína viral gp120 reconoce al receptor CD4 en la célula blanco. La subsecuente exposición de dominios en el asa V3 permite la interacción con CXCR4/CCR5 que facilita la inserción del péptido de fusión de gp41 en la membrana celular para provocar el acercamiento y posterior fusión de las bicapas lipídicas. Modificada de <sup>24</sup>

## 2. Descapsidación

Una vez dentro de la célula, es necesaria la disolución de la cápside para la liberación de proteínas y ARN viral. Por mucho tiempo, se creyó que la descapsidación era un proceso que ocurría de forma parcial en el citoplasma para permitir el inicio de la transcripción reversa del RNA viral y una vez completada ésta, el resto de las unidades de p7 se disociaban para permitir el traslado del ADN viral al núcleo. Sin embargo, se desconocían los factores que protegían al ARN de las nucleasas celulares cuando éste quedaba expuesto.

Hay evidencia reciente que demuestra que, en realidad, la descapsidación se da en una zona adyacente al poro nuclear y solo después de que el ARN viral ha sido retro-transcrito al interior; de este modo, el genoma recién sintetizado se mantiene protegido. Durante este proceso, es necesaria la interacción de dos secuencias ubicadas en el LTR viral, el tracto de polipurinas y la secuencia central de terminación. Estas secuencias interactúan con la maquinaria enzimática y forman una estructura de ADN de tres hebras conocido como “DNA flap” que favorece el movimiento de la cápside hasta las cercanías del núcleo para la translocación gura y efectiva del ADN viral<sup>25</sup>. Cuando el ARN ha sido retro transcrito casi en su totalidad, la cápside empieza a disolverse por la parte más angosta. Esto permite la asociación de proteínas celulares que favorecen el paso hacia el núcleo del genoma viral.

Diversos estudios han demostrado que la descapsidación es un proceso susceptible. El estudio de la replicación viral en macacos permitió la identificación de TRIM5 $\alpha$ , una proteína celular que actúa como factor de restricción. Ésta, acelera la decapsidación mediante el reconocimiento de p24 dentro de la estructura de cápside, lo cual expone al genoma viral a degradación. En los humanos, la proteína TRIM21 actúa de forma análoga a TRIM5 $\alpha$ <sup>19</sup>.

## 3. Transcripción reversa

El ARN viral es transcrito en una doble cadena de ADN gracias a la actividad de la RT. Durante esta etapa diversas variantes pueden ser generadas por la incapacidad de la enzima para corregir errores. Además de esta enzima, se ha demostrado que NC juega

un papel importante al posicionar el ARNt que marca el punto de inicio de la transcripción. Para el funcionamiento de RT, son requeridos dos cationes de  $Mg^{2+}$  que activan la función de polimerización de ADN y estabilizan a la cadena naciente y el nucleótido a incorporar.

Dado que el ARN viral se usa como molde de transcripción, un híbrido entre éste y una hebra de ADN es formado durante ese paso de la replicación. Para complementar el ADN de doble cadena capaz de insertarse en el genoma celular, es necesaria la degradación de la porción de ARN, la cual se lleva a cabo por la actividad del dominio ARNasa H de RT.

La generación del ADN viral o provirus puede ser inhibida por el factor de restricción APOBEC3G. Esta desaminasa convierte las citidinas del provirus en uracilo, pudiendo causar hipermutación por modificación en el apareamiento de bases en la hebra complementaria. Como respuesta, la proteína Vif del VIH se une a APOBEC3G e induce su degradación<sup>19</sup>.

Cuando la transcripción reversa se logra de forma exitosa, el ADN viral se asocia con las proteínas IN, NC y Vpr, formando lo que se conoce como Complejo de Pre-Integración (PIC).

#### **4. Integración**

Para conformar el PIC, la enzima Integrasa se une a ambos LTRs virales y corta un dinucleótido altamente conservado de GT en cada extremo (procesamiento 3'). Este corte deja extremos cohesivos que permitirán la inserción del provirus en el cromosoma hospedero<sup>26,27</sup>.

Una vez procesado el provirus, el PIC es dirigido hacia el núcleo. Una proteína celular de unión a cromatina conocida como p75 o LEDGF (lens epithelium-derived growth factor) dirige al provirus hacia regiones transcripcionalmente activas para su inserción<sup>28</sup>. En el sitio de inserción, la IN usa los 3-OH libres generados en el procesamiento del ADN viral para cortar el ADN cromosómico y posteriormente inserta el provirus. A este proceso se le conoce como transferencia de hebra y es terminado por enzimas celulares que reparan los

espacios producidos en el corte<sup>26,27</sup>. Con la integración finaliza la fase temprana del ciclo de replicación.

## **5. Transcripción y traducción de genes virales**

La transcripción de los genes virales inicia en el promotor de la región U3. Para que la elongación sea correcta, es necesaria la proteína Tat. Esta proteína viral recluta un factor celular conocido como factor b, el cual está constituido por la ciclina cinasa y la ciclina T1. El factor b se une a los transcritos virales y favorece su elongación<sup>19</sup>.

La transcripción se lleva a cabo en dos fases, en la primera se generan ARNm cortos por procesamiento múltiple que son rápidamente exportados del núcleo para generar las proteínas Tat, Rev, Vif, Vpr y Vpu. En la segunda fase, se generan ARN mensajeros que sufrieron un solo corte y ARNm no procesados que requieren unirse a Rev<sup>19</sup>. Rev sirve como adaptador para la exportación de esos templados que se traducirán en proteínas estructurales o bien, serán las cadenas de ARN que se empaquen en las nuevas partículas virales.

## **6. Ensamblaje**

Los transcritos monoprocesados son rápidamente traducidos en poliproteínas que proceden de los genes *gag*, *pol* y *env*. El precursor poliprotéico Gag es suficiente para la conformación de partículas similares a virus y para unirse a la membrana celular a través de ácido mirístico y aminoácidos básicos localizados en el fragmento de MA<sup>19</sup>. En cada partícula o virión son empaquetados los precursores proteicos, proteínas de expresión temprana y dos hebras de ARN no procesado. Una vez conformados los viriones y unidos a la membrana celular para ser envueltos, se lleva a cabo el proceso de salida por gemación.

## **7. Gemación**

Este paso en la replicación viral es orquestado por interacción de motivos ricos en prolina localizados en Gag, mejor conocidos como dominios tardíos (L), y proteínas vacuolares de la célula hospedera conocidas como VPS. De forma natural, las VPS participan en la formación de cuerpos vesiculares y la escisión de membranas durante la división celular.



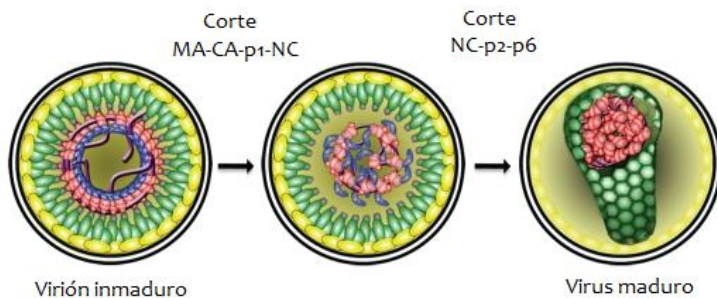
Las proteínas de este grupo que participa, además, en la gemación de los virus son identificadas como ESCRT (endosomal sorting complex required for transport, en inglés). ESCRT-I y ECRT-II participan en la formación de los cuerpos de gemación y ESCRT-III en la escisión de la membrana para liberar al virión <sup>19</sup>. La región terminal de Gag, que corresponde a p6, se une a la proteína Tsg101, parte de los complejos ESCRT para dirigir a los viriones y facilitar su salida.

La gemación de las partículas virales puede ser inhibida por una proteína celular llamada teterina (también conocida como CD317 o BST2), que retiene a los virus unidos a la membrana celular <sup>19</sup>. La proteína Vpu contrarresta la acción de teterina al inhibir su unión a los viriones.

## 8. Maduración

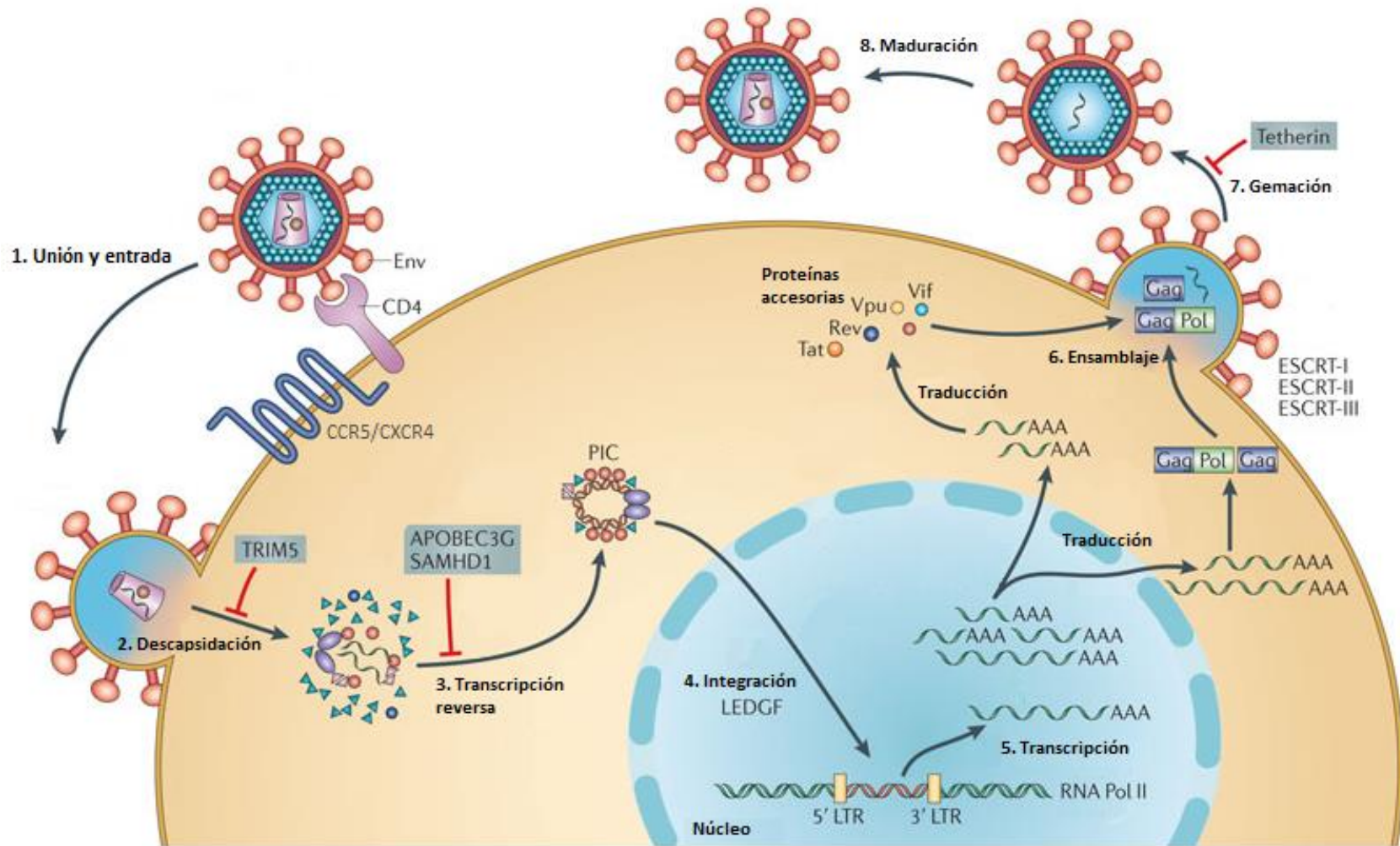
Los viriones correctamente conformados y envueltos no son infectivos, es por ello que requieren el procesamiento de los polipéptidos empaquetados. Este procesamiento constituye el paso final de la replicación y es llevado a cabo por la PR de forma casi inmediata a la gemación.

PR corta los precursores Gag y Gag-pol para producir las proteínas estructurales independientes MA, CA y NC, así como las enzimas funcionales PR, RT e IN. El corte de MA-CA permite la estructuración de la cápside cónica que resguardara a las proteínas y el ARN (Figura 13). Al madurar, el virus queda listo para poder unirse a otra célula y comenzar un nuevo ciclo de replicación.



**Figura 13 Maduración del VIH.** Una vez que el virus gema de la célula, múltiples copias del polipéptido Gag-Pro se encuentran organizadas de forma radial, al ser escindidas por la proteasa, se liberan las proteínas estructurales individuales que le dan forma a un virus infectivo. Traducida de <http://www.idibaps.org/aidsresearch/teams/team7/gallery.html>

El ciclo de replicación se esquematiza, en su totalidad, en la Figura 14.



**Figura 14 Esquema clásico del ciclo de replicación del VIH.** 1) Se da la unión entre el linfocito T blanco y el virus a través de la interacción de gp120 y gp41 con el receptor CD4 y algún co-receptor CCR5 o CXCR4, respectivamente. 2) La cápsida viral es liberada en el interior de la célula hospedera y se desintegra permitiendo la salida del material genético y las enzimas virales. 3) El ARN es transcrito en ADN de doble cadena por acción de la TR. 4) El ADN proviral junto con otras proteínas celulares y virales conforman el complejo de preintegración (PIC) que transloca al núcleo. El provirus es integrado en el genoma de la célula hospedera gracias a IN. 5) La transcripción del ADN viral origina ARN no procesados y con eventos sencillos o dobles de corte y empalme para empaquetarse, traducirse el polipéptido Gag-Pro y proteínas accesorias Tat, Rev, Vpr y Vpu. 6) El ARN viral y las proteínas son transportadas a la membrana plasmática para conformar nuevas partículas virales. 7) Los viriones salen de la célula mediante gemación, envolviéndose en una porción de la membrana plasmática. 8) La proteasa viral corta los polipéptidos virales para conferir infectividad al virión. Modificada de <sup>19</sup>

## **II.5 HISTORIA NATURAL DE LA INFECCIÓN**

La mayoría de las infecciones por VIH ocurren por contacto sexual, lo cual facilita el acceso del virus a la mucosa genital y rectal, donde la replicación inicial es facilitada por la abundancia de blancos celulares.

Parte de la respuesta inmune primaria a dicha infección se da por las células dendríticas que actúan como presentadoras de antígenos. El desplazamiento de estas células hacia los nódulos linfáticos para llevar a cabo su función, facilita el encuentro entre el virus y nuevas células susceptibles. La posterior diseminación del virus a través de la sangre, permite la llegada a otros tejidos y facilita su detección experimental. Generalmente los primeros eventos no causan síntomas, por lo que la identificación del virus se da varios meses e incluso años después de la adquisición del mismo.

A través del estudio de la dinámica viral y celular en los individuos infectados, se han logrado establecer distintas etapas que constituyen la historia natural de la infección y se detallan a continuación:

### **ECLIPSE**

Es un periodo de aproximadamente 10 días posteriores a la transmisión y durante los cuales el virus es indetectable en plasma. La replicación viral está restringida a las mucosas y principalmente se da en linfocitos T CD4<sup>+</sup>, pocos macrófago y monocitos son infectados durante esta etapa.

El periodo de eclipse antecede un pico de viremia que surge alrededor de 21-28 días post-infección<sup>29</sup>.

### **ETAPA AGUDA**

Es reconocida como las primeras semanas entre la transmisión del virus y antes de la seroconversión, es decir, antes de que se alcance un título de anticuerpos VIH-específicos detectable en plasma.

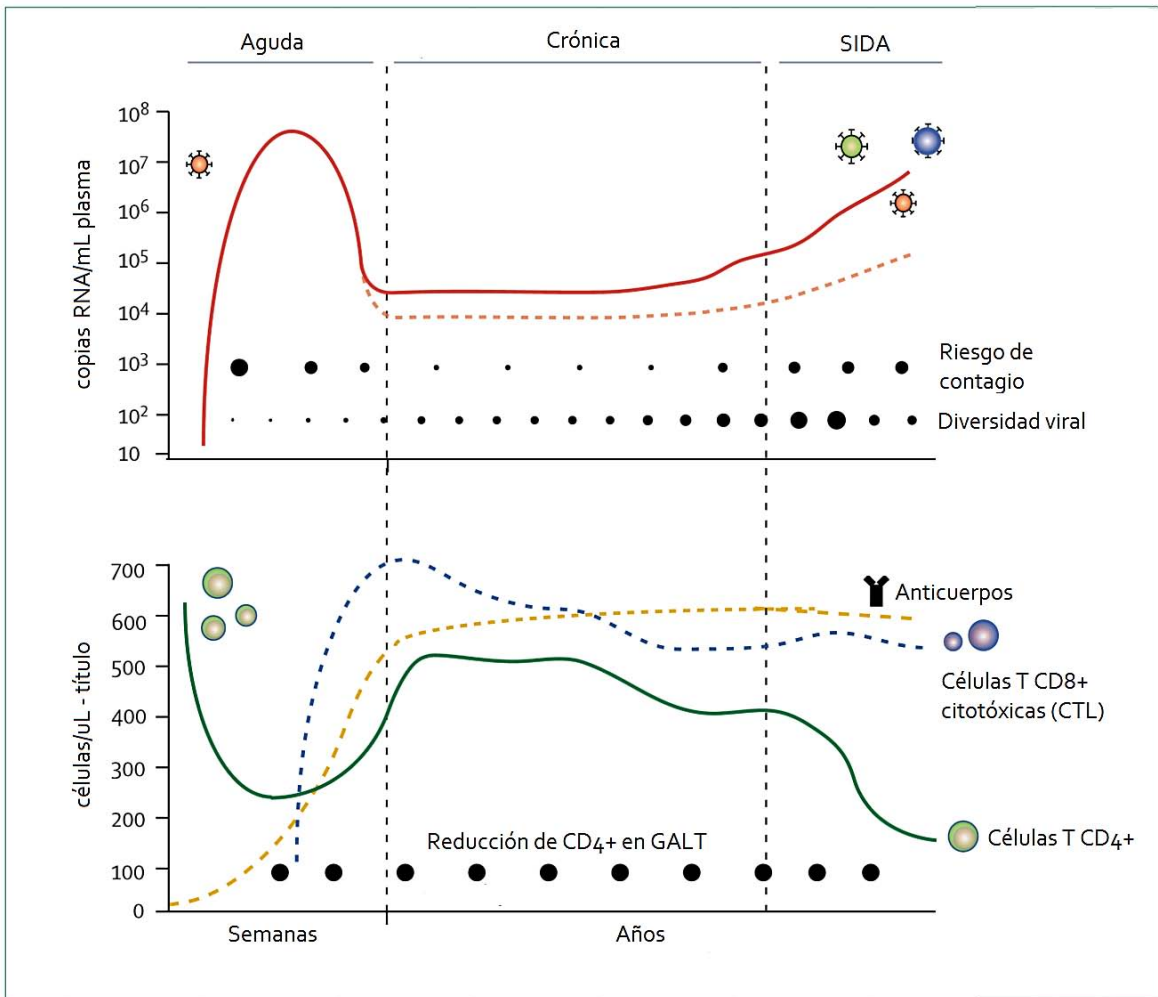
Este periodo se caracteriza por una elevación súbita de la replicación viral (pico de viremia) después de 21-28 días de infección, que permite el establecimiento de una infección secundaria en nódulos linfáticos y su posterior diseminación a través del torrente sanguíneo<sup>30</sup>. El tejido linfoide asociado a intestino (GALT, en inglés) se ve severamente afectado, pues en él residen grandes cantidades de células T CD4+ CCR5+ que constituyen el principal blanco viral durante este periodo. Alrededor del 80% de los linfocitos residentes en GALT son infectados o eliminados por apoptosis durante las primeras 3 semanas de la infección<sup>29</sup>.

Dada la alta tasa de replicación, la diversidad viral aumenta notoriamente y la viremia puede incrementarse hasta  $10^6 - 10^7$  copias de RNA por mL de sangre; mientras que la cuenta plasmática de linfocitos T CD4+ disminuye de manera significativa, llevando al paciente a un estado de inmunosupresión transitorio<sup>11</sup>.

A pesar de que este periodo generalmente es asintomático, existen algunos pacientes que pueden presentar el llamado síndrome retroviral agudo, una enfermedad parecida a la mononucleosis infecciosa que causa fiebre, erupciones, dolor articular e inflamación de los nódulos linfáticos mientras se producen anticuerpos específicos contra el virus<sup>31</sup>.

Independientemente de la presencia de síntomas, el individuo infectado es altamente contagioso (Figura 15). Las manifestaciones que aparecen tras el ingreso del VIH en el organismo pueden guardar relación con la dosis infectante, la virulencia de la cepa y la capacidad de respuesta del sujeto infectado.

De 12 a 20 semanas posteriores al contagio la replicación viral se controla y estabiliza, suceso conocido como "set point" ( $10^3 - 10^5$  copias/mL plasma)<sup>32</sup>. Diversos estudios han permitido identificar el papel trascendental de los linfocitos T CD8 en este proceso y han permitido definir al set point como un parámetro parcialmente predictivo de la progresión de la infección, pues refleja el estado inmunológico del individuo al momento de establecer una respuesta efectiva contra el virus.



**Figura 15 Patogénesis de la infección por VIH. A) Variación de factores virales.** Durante la etapa aguda el virus replica de manera progresiva hasta alcanzar un punto máximo. La replicación es controlada y se mantiene estable durante la etapa crónica. En ausencia de tratamiento antirretroviral, la viremia puede incrementarse nuevamente favoreciendo la progresión a SIDA. El riesgo de transmisión y la diversidad viral mantienen una relación inversa a lo largo de la infección. **B) Respuesta inmune.** Como blanco principal del VIH los linfocitos T CD4+ circulantes se ven severamente disminuidos durante la etapa aguda. La etapa crónica se caracteriza por la aparición de la respuesta citotóxica mediada por linfocitos T CD8+, el máximo título de anticuerpos y una recuperación parcial de células T CD4+. Los linfocitos T CD4+ circulantes disminuyen de forma progresiva durante la infección hasta un mínimo de 200 cel/uL que marca el estadio de SIDA. Linfocitos T CD4+ residentes en intestino (GALT) desaparecen al inicio de la infección sin recuperación posterior. Modificada de <sup>11</sup>.

## ETAPA CRÓNICA (Latencia clínica)

De 3 a 4 semanas posteriores a la transmisión, aparece la respuesta mediada por anticuerpos; aunque existen individuos a quienes les puede tomar hasta 6 meses. Algunos autores definen el periodo posterior a la seroconversión y precedente de la etapa crónica como **infección primaria**; sin embargo, de forma general, la respuesta humoral

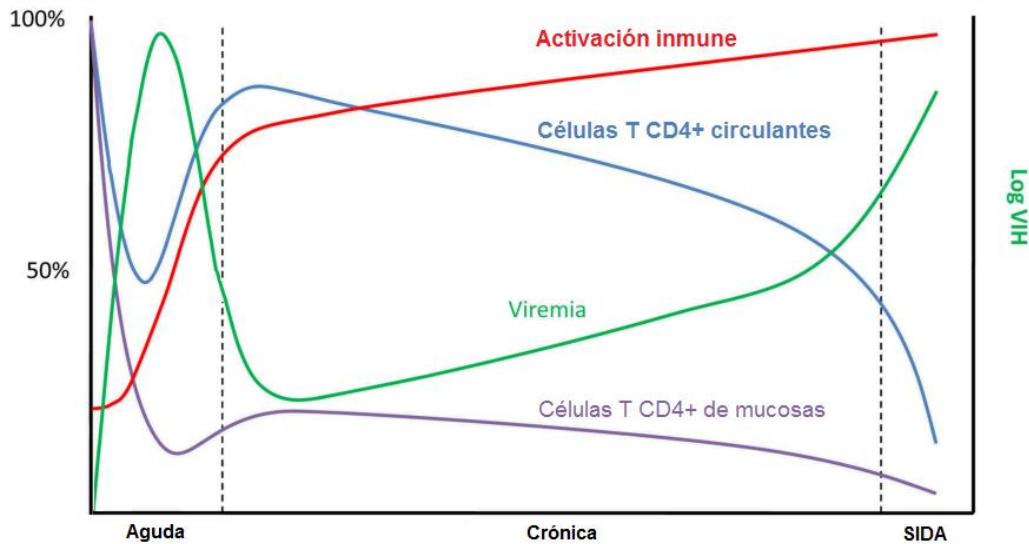
se considera parte de una fase típicamente asintomática y de duración variable conocida como etapa crónica.

Durante este periodo existe una hiper-activación del sistema inmune, que aumenta por la translocación bacteriana favorecida por el daño a la mucosa intestinal y la replicación viral permanente. La carga viral suele oscilar entre  $<50$  a  $>1,000,000$  copias ARN/mL entre individuos y se pueden llegar a producir hasta  $10^8$  viriones/día<sup>30</sup>. Aunado a esto, la aparición de las variantes de escape a la respuesta inmune se ve favorecida.

La disminución de linfocitos T-CD4+ en GALT es irreversible. Sin embargo, se puede observar un aumento parcial de estas células en sangre una vez que se establece el control sobre la replicación viral (Figura 16). Dado que la proporción de linfocitos infectados en comparación con el total de células mononucleares circulantes es bajo (1:50,000), se cree que dicha disminución no se da por un efecto lítico directo del virus sobre las células infectadas, sino por la inducción de apoptosis de linfocitos no infectados, la eliminación de células infectadas por parte del sistema inmune, daño celular producto de la activación sistémica y una renovación fallida de estas células<sup>30</sup>.

La duración de la fase crónica es altamente variable entre individuos, aunque se estima que el tiempo de progresión a estadios avanzados oscila entre 8 y 13 años. En personas con infección crónica, pueden o no presentarse adenopatías, cifras baja de plaquetas o trastornos neurológicos mínimos.

La mayoría de detecciones clínicas de la infección se da durante esta etapa. La intervención terapéutica oportuna disminuye la carga viral hasta niveles indetectables y permite una reconstitución parcial de las células inmunes. Existe una pequeña proporción de individuos infectados ( $<1\%$ ) capaces de permanecer en este estadio por décadas sin necesidad de tratamiento antirretroviral; denominados “controladores”. El motivo del control sobresaliente de la replicación viral es controversial, ya que depende de características intrínsecas de la respuesta inmune del individuo, sus características inmunogenéticas como el haplotipo HLA y KIR, modificaciones en receptores celulares, factores virales, entre otros (ver Sección III).

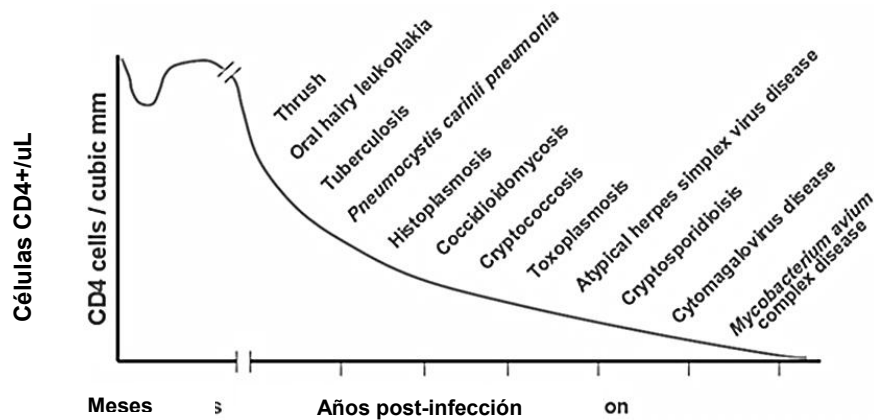


**Figura 16 Respuesta inmune en la infección por VIH.** La fase aguda se caracteriza por un pico de viremia y una disminución drástica de linfocitos T CD4 de mucosas y sangre. La pérdida de células inmunes en mucosas es irreversible, los linfocitos CD4+ en sangre se recuperan parcialmente cuando se monta una respuesta efectiva. A lo largo de la infección se pierden de forma progresiva células CD4+ hasta el quiebre del sistema inmune que favorece el repunte de la viremia y constituye la fase de SIDA. Modificada de <sup>33</sup>.

## SIDA

En ausencia de tratamiento antirretroviral, el control natural de la replicación viral puede fallar. El repunte de la viremia se acompaña con una franca disminución de linfocitos T-CD4+ hasta por debajo de 200 cel/uL; ambas son características del estadio final de la infección conocido como SIDA. Aunado a ello, existe una profunda alteración del estado general del paciente (síndrome de desgaste), aparecen graves infecciones oportunistas, ciertas neoplasias o alteraciones neurológicas.

Durante la fase de SIDA, se ha observado que existe una correlación directa entre la aparición de enfermedades oportunistas como candidiasis oral, pneumococcosis, tuberculosis y varicela zoster; y la disminución de células T CD4+ por debajo de 500 células/uL. El riesgo de tener complicaciones como neumonía por *Pneumocystis jiroveci*, *Candida esophagitis*, histoplasmosis diseminada y otras infecciones fúngicas sistémicas se incrementa sustancialmente cuando el nivel de linfocitos T CD4+ cae a valores cercanos a 200 células/uL. La aparición de otros trastornos como reactivación de infecciones por citomegalovirus (CMV), criptosporidiosis, entre otras, son indicativas de una franca falla del sistema inmune con conteos de ~50 células T CD4+ /uL (Figura 17).



**Figura 17 Patologías asociadas a SIDA.** Distintos trastornos en los pacientes con infección por VIH aparecen de forma progresiva y con una correlación directa con el descenso en la cuenta de células T CD4+. La acumulación de infecciones y trastornos inmunes derivan en el desarrollo de Síndrome de Inmunodeficiencia Adquirida (SIDA), generalmente cuando los linfocitos T CD4+ se encuentran por debajo de 500 células/uL. Tomada de <http://www.microbiologybook.org/Spanish-Virology/spanish-chapter7-3.htm>

Entre las neoplasias típicas de SIDA se incluye el sarcoma de Kaposi causado por un virus de herpes tipo 8, linfoma asociado a infección por virus Eipsten- Barr y carcinoma anal o cervical asociado al virus de papiloma humano (VPH).

El pronóstico de supervivencia de pacientes con SIDA es variable. La edad, el mecanismo de contagio, el tratamiento antirretroviral y la capacidad de respuesta del organismo, parecen influir en la sobrevivencia del individuo.

Aunque existen amplias variaciones individuales, se estima que en 10 años el 50% de los adultos y el 80% de los niños habrán evolucionado a estadios avanzados, aunque la progresión de la enfermedad puede verse influida por numerosos factores. Únicamente de 1 a 2% de los pacientes progresan a SIDA antes de 2 años<sup>2</sup>.

### III. RESPUESTA CITOTÓXICA Y EVOLUCIÓN VIRAL MEDIADA POR HLA

#### III.1 ANTIGENO LEUCOCITARIO HUMANO (HLA)

En 1940 los experimentos de George Snell permitieron la identificación de un grupo de genes asociado al rechazo o la aceptación de un injerto en ratones. Cuando la información genética era idéntica, el tejido era reconocido como propio, favoreciendo su funcionalidad; en cambio, si existían versiones (alelos) distintas de estos genes entre el donador y el receptor, se montaba una respuesta inmune en contra del injerto. A dichos



genes se les agrupó y definió como Complejo Principal de Histocompatibilidad (MHC, por sus siglas en inglés)<sup>34</sup>.

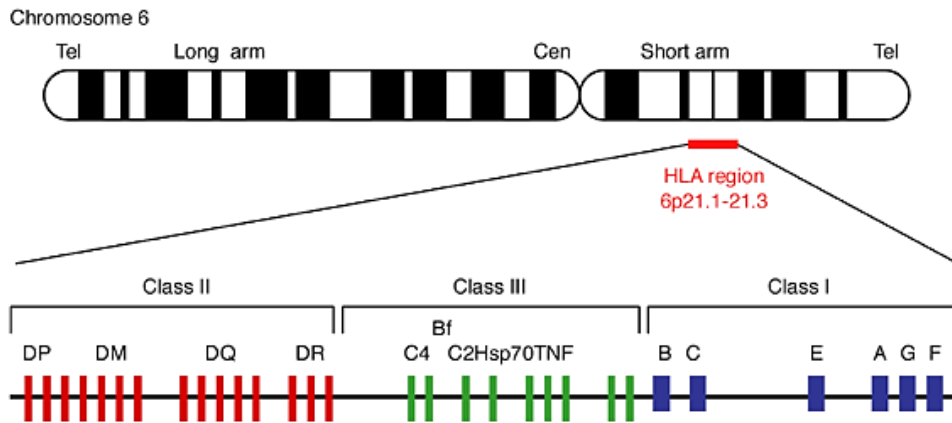
Por más de veinte años la funcionalidad de estos genes supuso una interrogante, pues siendo el transplante un evento extraño, permanecía incierta la función natural de estos genes evolutivamente conservados entre muchas especies. Alrededor de 1974, Zinkernagel y Doherty identificaron proteínas similares presentes en los leucocitos humanos y que provenían de genes agrupados de forma similar a los ya descritos. Fue así como se definieron las proteínas del Antígeno Leucocitario Humano (HLA, en inglés) y se observó que una clase de estas moléculas estaban directamente involucradas en la habilidad de las células T CD8<sup>+</sup> para reconocer y matar células infectadas por virus, un fenómeno que fue denominado restricción por HLA<sup>35</sup>.

Las proteínas HLA se encuentran codificadas en una región de 3.6 Mpb de longitud en el brazo corto del cromosoma 6. Los genes se agrupan en tres clases distintas, las dos primeras asociadas a la respuesta de linfocitos T y la tercera codificante para proteínas del complemento y diversas citosinas asociadas a respuesta inmune innata (Figura 18).

Los genes que producen las proteínas de Clase II se denominan DP, DM, DQ, DR y DO. Éstos se expresan únicamente en células especializadas presentadoras de antígenos como Células Dendríticas (CD), Macrófagos (M $\phi$ ) y Linfocitos B (LB). Las proteínas HLA-II son intermediarias en la presentación de antígenos extracelulares a los linfocitos T cooperadores (CD4<sup>+</sup>).

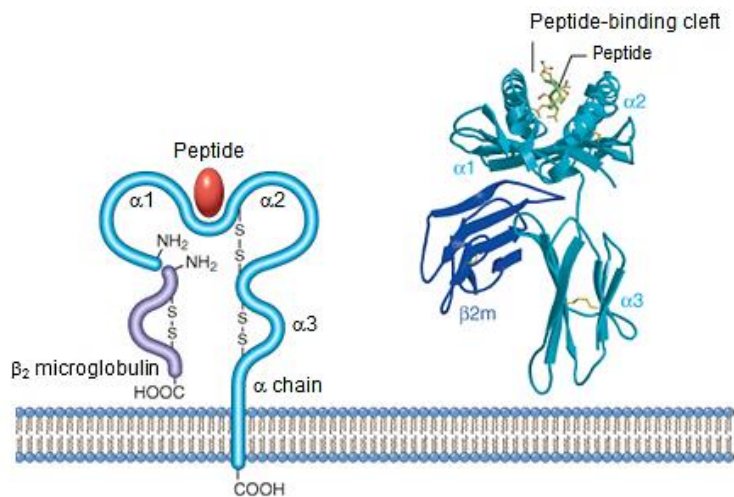
Por su parte, la clase I de HLA está conformada por los genes A, B, C, E, F y G que codifican para seis proteínas divididas en dos grupos, moléculas HLA-I “Clásicas” –A, B y C- y “No clásicas” –E, F y G-. Las proteínas no clásicas se han involucrado en la unión a receptores de células NK e impedimento del rechazo del feto; mientras que el HLA-A, B y C participan en la presentación de péptidos intracelulares, como virus, a Linfocitos T CD8<sup>+</sup>. Estas últimas se expresan en prácticamente todas las células nucleadas, excepto neuronas y células del trofoblasto. Cada célula expresa de forma co-dominante un conjunto de seis moléculas derivadas de los dos alelos de los genes A, B y C que se han heredado de los progenitores.

Al profundizar en el estudio de los genes del MHC-I, se determinó que es la región más polimórfica del genoma humano. Actualmente (agosto, 2015) se cuenta con el registro de 3192 alelos distintos para el gen A, 3977 para B y 2740 para C<sup>36</sup>.



**Figura 18 Mapa cromosómico del Antígeno Leucocitario Humano.** Las proteínas del antígeno leucocitario humano (HLA, inglés) están codificadas en el brazo corto del cromosoma 6 y se dividen en tres grupos. HLA Clase I media la presentación antigénica a linfocitos CD8<sup>+</sup> y HLA-II permite la presentación a linfocitos T CD4<sup>+</sup>. En la región de clase III se encuentran codificadas citosinas y proteínas del complemento. Modificada de <sup>37</sup>

En 1987, Bjorkman describió la estructura 3D de la molécula HLA-I. Cada molécula está conformada por dos proteínas unidas de forma no covalente: la cadena  $\alpha$  y la  $\beta$ 2-microglobulina. La cadena  $\alpha$  (44-47KD) está dividida en tres dominios, dos de los cuales ( $\alpha$ 1 y  $\alpha$ 2) se exponen al medio extracelular y forman un surco que permite la unión de péptidos de 8 a 11 aminoácidos. La enorme variabilidad de las moléculas HLA-I se distribuye entre los 57 aminoácidos que conforman surco. Los polimorfismos presentes pueden afectar en distinta medida la unión del péptido al HLA y la interacción del HLA con los receptores de los linfocitos CD8<sup>+</sup>. El segmento  $\alpha$ 3, cuya secuencia primaria está conservada en todas las moléculas de HLA, contiene un bucle de unión a CD8. En el segmento carboxilo-terminal posee una extensión de aproximadamente 55 aminoácidos que atraviesa la membrana celular favoreciendo su anclaje (Figura 19)<sup>34</sup>. Al dominio  $\alpha$ 3 se une la  $\beta$ 2 microglobulina (12 KD), la cual está codificada fuera de la región MHC. Esta proteína no varía entre las moléculas de clase I. La molécula HLA-I totalmente ensamblada es un heterotrímero que consta de la cadena  $\alpha$ , una  $\beta$ 2-microglobulina y un péptido antigénico unido al surco. Únicamente se expresan en la superficie celular las moléculas cargadas con péptidos (Figura 19).



**Figura 19 Estructura de una molécula HLA de clase I.** Representación esquemática (izquierda) de las distintas regiones de una molécula de HLA de clase I. Está compuesta por una cadena  $\alpha$ , dividida en tres dominios, unida de forma no covalente a la  $\beta_2$  microglobulina. A la derecha se muestra un diagrama de cintas de una molécula de HLA-I unido a un péptido. Tomada de [http://intranet.tdmu.edu.ua/data/kafedra/interna/patologanatom/classes\\_stud/en/med/lik/ptn/P\\_athomorphology/3/04\\_Pathomorph\\_immune\\_syst.htm](http://intranet.tdmu.edu.ua/data/kafedra/interna/patologanatom/classes_stud/en/med/lik/ptn/P_athomorphology/3/04_Pathomorph_immune_syst.htm).

### III.2 PRESENTACIÓN ANTIGÉNICA MEDIADA POR HLA-I

En la respuesta inmune, los linfocitos T CD4<sup>+</sup> activan otras células como macrófagos y linfocitos B para fagocitar patógenos y producir anticuerpos, respectivamente. Por su parte, los linfocitos T CD8<sup>+</sup> eliminan células infectadas por patógenos intracelulares como virus o células tumorales. La inducción de estas respuestas está desencadenada por el reconocimiento específico de antígenos, que en el caso de los linfocitos T, son fragmentos de péptidos que derivan de antígenos proteínicos y que están unidos a moléculas de HLA expresadas en la membrana de células presentadoras de antígeno (HLA-II) o en la superficie de células nucleadas (HLA-I).

El reconocimiento antigénico únicamente se da cuando los péptidos son presentados en las moléculas propias del individuo; es decir, los linfocitos T tienen la capacidad de reconocer HLA propios y extraños. A este reconocimiento específico se le conoce como **restricción por HLA** y permite a las células CD4<sup>+</sup> reconocer y reaccionar a moléculas de clase II del HLA y a los linfocitos CD8<sup>+</sup> reconocer HLA-I. Además de delinear la respuesta inmune, la restricción por HLA es la base del rechazo/aceptación de trasplantes.

Dado que la presentación antigénica en moléculas de HLA-I puede ser llevada a cabo por cualquier célula nucleada, los CD8<sup>+</sup> deben ser capaces de reconocer los péptidos propios que de forma natural se encuentran unidos al HLA de cada célula. El reconocimiento de

un péptido que no corresponda a los propios, por ejemplo proteínas virales o proteínas mutadas, desencadenan la respuesta lítica contra la célula que actúa como presentadora.

La presentación de antígenos en el contexto de moléculas de HLA de clase I a los linfocitos CD8+ se puede dividir en cuatro procesos:

### **1. DEGRADACIÓN PROTEOLÍTICA DE LAS PROTEÍNAS CITOSÓLICAS**

Los antígenos extraños del citosol pueden ser producto de una infección viral o de algún otro microorganismo intracelular capaz de romper fagosomas; o de un proceso tumoral que favorezca la producción de proteínas mutadas.

El principal mecanismo de generación de péptidos es la proteólisis de los antígenos citosólicos por el proteasoma. El proteasoma es un complejo proteico presente en la mayoría de las células y cuya función básica es la degradación de proteínas celulares que están mal plegadas, defectuosas o ya son innecesarias para la célula. En el caso de antígenos derivados de infecciones, el proteasoma degrada a las proteínas virales en pequeños fragmentos que en su extremo C terminal son óptimos para unirse al HLA-I pero en el extremo N-terminal poseen una extensión de algunos aminoácido que serán retirados posteriormente<sup>34</sup> (Figura 20).

### **2. TRANSPORTE DE PÉPTIDOS AL RETÍCULO ENDOPLÁSMICO**

Las moléculas HLA-I son sintetizadas en el retículo endoplásmico (RE) e inestables si están libres, por lo que los péptidos procesados deben ser transportados del citosol a este organelo. Para ello, existe un heterodímero denominado transportador asociado al procesamiento antigénico (TAP). TAP se une de forma efectiva a proteínas de 6 a 30 aminoácidos y que contienen extremos carboxílicos básicos, por lo que los péptidos originados por el proteasoma, resultan ideales para interactuar con él.

TAP permite el paso del péptido del citosol hacia la luz del RE donde se alojan las moléculas HLA-I recién sintetizadas. Al interior, las moléculas de HLA-I se encuentran unidas a un complejo conocido como complejo de carga del péptido (PLC por sus siglas en inglés). El PLC se compone por un dímero TAP que se une al HLA-I por medio de tapasina, así como ERp57, calnexina y calreticulina que estabilizan el complejo “vacío”

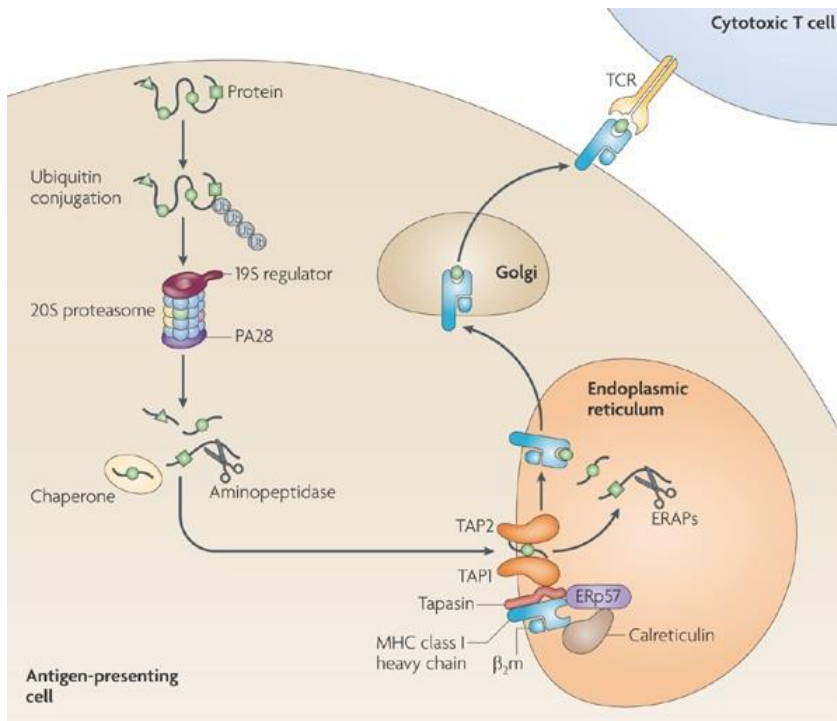
(Figura 20). En el RE también se encuentran proteínas denominadas ERAAP que cortan la extensión en el dominio N-terminal de los péptidos ingresados; de este modo se generan cadenas de 8 a 11 aminoácidos optimizadas para unirse al surco de HLA-I<sup>38</sup>.

### 3. ENSAMBLAJE DE LOS COMPLEJOS PÉPTIDO-HLA-I

Una vez dentro, el péptido se enlaza al surco del HLA-I a través de los aminoácidos 2 y 9 que interactúan con las cadenas  $\alpha 1$  y  $\alpha 2$ . Dado que el surco de esta clase de moléculas es cerrado, únicamente péptidos con el tamaño adecuado pueden ser cargados. La unión de HLA y péptido favorece la liberación de tapasina. Una vez liberados de TAP, los complejos pueden ser transportados a la superficie celular. Cuando no hay ningún péptido unido, propio o extraño, los complejos son degradados en el mismo RE.

### 4. EXPRESIÓN DE COMPLEJOS HLA-PÉPTIDO

Los complejos formados se mueven a través del aparato de Golgi hacia la superficie mediante vesículas. Las vesículas exocíticas pueden fusionarse con la membrana celular y favorecer la expresión de los complejos HLA-péptido para su reconocimiento (Figura 20).



**Figura 20 Procesamiento y presentación antigénica restringida por moléculas HLA de Clase I.** Las moléculas de HLA-I presentan péptidos citosólicos derivados de proteínas procesadas en el proteasoma. Los fragmentos resultantes pueden unirse a proteínas chaperonas que lo proteja de degradación hasta llegar al retículo endoplásmico (RE). Cada péptido ingresa al RE a través del heterodímero TAP. Dentro del RE las moléculas HLA recién sintetizadas se encuentran unidas a TAP a través de tapasina, a ERp57, calnexina y calreticulina para conformar un complejo de cargado de péptidos. Las proteínas ERAAP cortan el extremo N-terminal de los péptidos para optimizarlos a 8-11 aminoácidos. El péptido se une al surco del HLA-I y es exportado a la superficie a través del sistema vesicular trans Golgi. Tomada de [http://www.nature.com/nri/journal/v10/n1/fig\\_tab/nri2687\\_F1.html](http://www.nature.com/nri/journal/v10/n1/fig_tab/nri2687_F1.html)

### III.3 MADURACIÓN Y ACTIVACIÓN DE LOS LINFOCITOS CD8<sup>+</sup>

Los linfocitos proceden de precursores originados en el hígado fetal y la médula ósea adulta y que pueblan el timo a partir de la semana 8 de gestación.

En el estadio más temprano en el timo, a estas células se les denomina timocitos. Inicialmente, los timocitos migran hacia la corteza del timo y carecen de la expresión de las moléculas CD4 y CD8 en su superficie, por lo que son llamados timocitos dobles negativos o linfocitos pro-T.

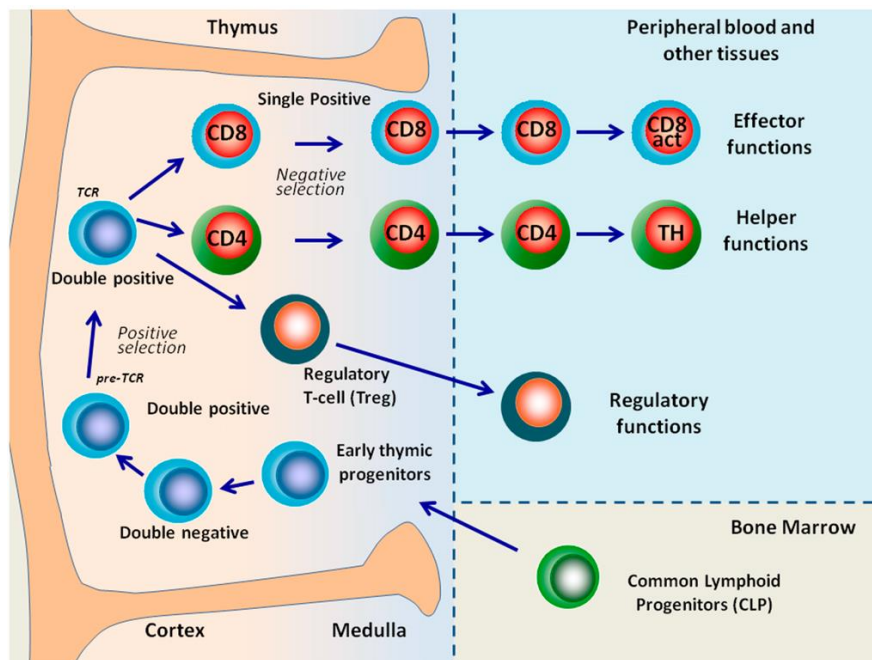
Los linfocitos pro-T contienen genes codificantes de las cadenas  $\alpha$  y  $\beta$  que conformarán el receptor del linfocito T o TCR (T-cell receptor), por sus siglas inglesas. La expresión de dichos genes, requiere de un complejo proceso de re-arreglo y recombinación génica orquestado por las proteínas Rag-1 y Rag-2. Este mecanismo asegura la producción de un repertorio amplio de clones de linfocitos con receptores distintos y únicos entre sí para el reconocimiento de prácticamente cualquier antígeno.

En el timocito se da únicamente la recombinación de los genes codificantes de la cadena  $\beta$ . Si el resultado de este re-arreglo génico produce una proteína completa y funcional, ésta se une a una cadena preliminar invariable  $\alpha$  y las proteínas CD3 y  $\zeta$  para conformar el receptor de linfocitos pre-T. Las señales procedentes de este nuevo pre-receptor favorecen la recombinación en la cadena  $\alpha$  y dirigen la transición de la fase doble negativa a la doble positiva, es decir, favorece la expresión de CD4 y CD8 en el mismo linfocito.

Como dobles positivos, los timocitos expresan la quimiocina CCR7 que los dirige a la médula tímica. Migrarán únicamente los linfocitos con TCR completos y funcionales, es decir, los que expresan las cadenas  $\alpha$  y  $\beta$  definitivas estabilizadas por la unión adecuada a CD3 y  $\zeta$ .

El repertorio completo que llega a la médula, incluye células cuyos receptores pueden reconocer cualquier antígeno peptídico (propio o extraño) presentado en cualquier molécula HLA. Cuando los linfocitos inmaduros expresan por primera vez el TCR  $\alpha\beta$ , entran en contacto con péptidos propios presentados en moléculas de HLA propio.

Únicamente aquellos que se unen con poca avidéz (débilmente) al complejo HLA-péptido son estimulados para proliferar. De este modo, se eliminan linfocitos que no reconocen las moléculas de HLA propio. Además, se establece la restricción para que los linfocitos CD8+ reconozcan HLA-I y los CD4+ reconozcan HLA-II. A este mecanismo se le conoce como selección positiva. Para terminar de depurar linfocitos potencialmente peligrosos, se procede a la selección negativa que induce la apoptosis en aquellas células que reconocen con demasiada avidéz al HLA. Esto garantiza la eliminación de clones que pudieran provocar autoinmunidad (Figura 21).



**Figura 21 Maduración de los linfocitos T.** En la médula ósea se alojan progenitores multi potenciales (CLP) que migran al timo para diferenciarse en linfocitos. Migran de la médula a la corteza en un estado inicial de dobles negativos. En la corteza ocurre la selección positiva permitiendo la aparición de linfocitos dobles positivos (CD4+/CD8+). La diferenciación final puede dirigirse a linfocito T CD8, T CD4 o Treg. Los linfocitos T CD8 ejercen su acción efectora como células citotóxicas y los T CD4+ actúan como cooperadoras en diversas respuestas inmunes. Modificada de <sup>39</sup>

El resultado neto de estos procesos es un repertorio de linfocitos T maduros restringidos por el HLA propio y con tolerancia autoantigénica que pueden salir del timo para ser activados.

Del mismo precursor y bajo un proceso similar al descrito, se pueden desarrollar poblaciones celulares que regulan la respuesta inmune (Treg y Th17). La diferenciación de estas poblaciones descritas recientemente no es objeto del presente trabajo.

Los linfocitos deben ser activados, una vez que han madurado. La activación de los linfocitos T CD8<sup>+</sup> vírgenes requiere del reconocimiento del antígeno presentado por HLA-I y segundas señales que potencian la respuesta.

La activación de células T-CD8<sup>+</sup> se puede dar por la presentación cruzada de células dendríticas o por el reconocimiento antigénico de péptidos presentados por células infectadas. En el primer caso, una célula dendrítica fagocita a células infectadas con patógenos intracelulares (p. ej. virus), procesa antígenos y los carga en moléculas HLA-I para ser reconocidas por el linfocito CD8<sup>+</sup>. Esta interacción se favorece con la participación de proteínas co-estimuladoras; la vía más estudiada involucra a la molécula CD28 expresada en los linfocitos y CD80-CD86 expresada en las células dendríticas que potencia la supervivencia del CD8, la producción de IL-2 para proliferación y la diferenciación.

En el caso de la presentación antigénica directa entre una célula infectada y un linfocito CD8<sup>+</sup>, diversas citosinas actúan como potenciador de la activación. Las moléculas que cumplen esta función y han sido caracterizadas de manera más detallada en infecciones virales pertenecen al grupo de los interferones (IFN). Específicamente el IFN $\alpha$  e IFN $\beta$  son producidos por las células infectadas para aumentar su susceptibilidad a la muerte mediada por linfocitos CD8<sup>+</sup> y actúan en células vecinas no infectadas para establecer un “estado antivírico” que bloquee la replicación viral.

La activación de los linfocitos citotóxicos se da por ambos mecanismos cuando hay infección por VIH. Cuando las células dendríticas fagocitan al virus en el sitio de infección y migran a los nódulos linfáticos, inducen la diferenciación de CD8<sup>+</sup> específicos contra péptidos virales por presentación cruzada. La activación puede ocurrir también si la célula dendrítica fagocita a un linfocito CD4<sup>+</sup> infectado, o bien, si es infectada directamente por el VIH (fenómeno que ocurre con poca frecuencia)<sup>40</sup>. Además, las células dendríticas pueden actuar como “caballos de Troya” al transportar partículas virales entre los pseudópodos de su membrana y acercarlos a los linfocitos CD4<sup>+</sup> residentes de los ganglios linfáticos. Por su parte, los linfocitos CD4<sup>+</sup> al ser el principal blanco del VIH y expresar en su membrana proteínas HLA-I, son capaces de activar de forma independiente a los linfocitos CD8<sup>+</sup>.



Para la activación de la respuesta citotóxica es necesaria una cascada de señales que se da a través del receptor del linfocito CD8<sup>+</sup> y que se desencadena por la interacción con el HLA. Este reconocimiento conforma la sinapsis inmunológica.

En los linfocitos T maduros y activados, el TCR, en su totalidad, está formado las cadenas  $\alpha$  y  $\beta$  que interactúan directamente con el HLA y tres dímeros ( $\gamma\epsilon$ ,  $\epsilon\delta$  y  $\zeta\zeta$ ) provenientes del CD3 y que sirven como transductores de señales. Las proteínas de CD3 poseen en su estructura dominios conocidos como ITAM susceptibles de fosforilación.

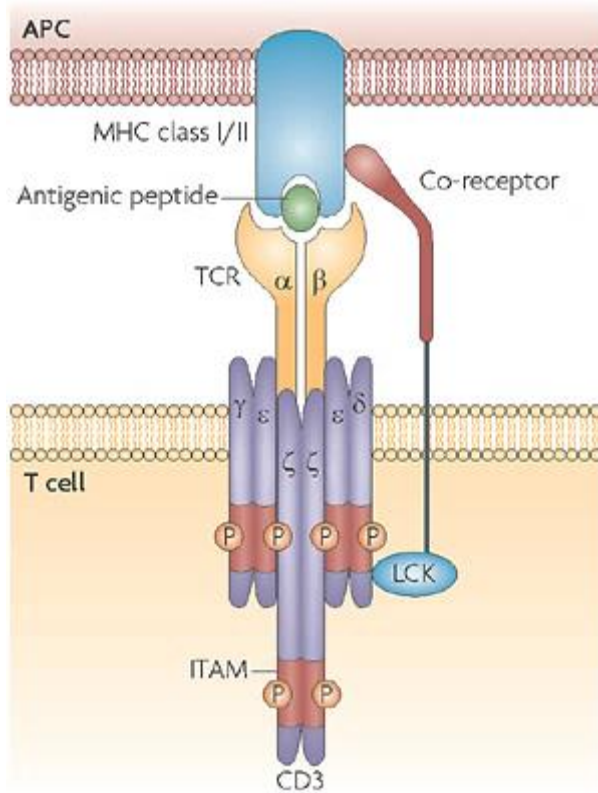
La cascada de señalización en el linfocito inicia con la fosforilación de los dominios ITAM que permite el reclutamiento de una cinasa denominada ZAP70. Una vez que ZAP70 se une a CD3, se favorece la unión del co-receptor CD8 que naturalmente se encuentra asociado a las cinasas Lck. La interacción de CD8 con CD3 estabiliza al complejo TCR-HLA (Figura 22) y desencadena señales que deriven en la proliferación y diferenciación del linfocito activado.

La activación celular permite la expansión clonal de linfocitos específicos a los antígenos presentados, de tal forma que existan suficientes elementos capaces de responder ante la infección. Este fenómeno se realiza de la mano de la diferenciación celular. Los linfocitos T activados pueden diferenciarse en efectores o de memoria que pueden permanecer en los órganos linfáticos o migrar.

La diferenciación a un fenotipo efector supone la aparición de orgánulos citoplásmicos unidos a la membrana que contienen proteínas como perforina y granzima cuyas funciones se describen más adelante. Además, los CD8<sup>+</sup> adquieren la capacidad de secretar citocinas como IFN $\gamma$ , linfoxina y TNF (Factor de necrosis tumoral) que actúan activando fagocitos e induciendo inflamación.

Los linfocitos que se diferencian a células de memoria antígeno-específicas, pueden persistir por muchos años. Estos linfocitos son responsables de respuestas rápidas y amplificadas ante un segundo contacto con el mismo antígeno. Este tipo de células diferenciadas se subdivide en linfocitos de memoria central y de memoria efectora. Los primeros, residen en los ganglios linfáticos y proliferan de forma acelerada ante el

reconocimiento del antígeno y pueden producir abundantes linfocitos efectoros. Los de memoria efectora se alojan en ejidos periféricos como mucosas y después de la estimulación antigénica producen  $IFN\gamma$  para una respuesta inicial rápida.



**Figura 22 Receptor de linfocito T.** El receptor está conformado por dos complejos proteicos llamados TCR y CD3. Las unidades  $\alpha$  y  $\beta$  que forman el TCR interactúan directamente con la molécula de HLA de la célula presentadora de antígeno o de una célula infectada. Los dímeros  $\gamma\epsilon$ ,  $\epsilon\delta$  y  $\zeta\zeta$  conforman el CD3 y poseen dominios ITAM que al ser fosforilados actúan como transductores de señales. Cuando el CD3 se fosforila, se recluta a la tirosina ZAP70 (no mostrada en la imagen) que a su vez interactúa con las proteínas Lck. Las tirosinas Lck se mantienen unidas a los co-receptores (CD8 o CD4). Al interactuar el co-receptor con CD3 se estabiliza la unión TCR-HLA y se promueve una cascada de señalización para activar la respuesta linfocitaria. En el caso de la respuesta citotóxica, participan moléculas HLA de clase I y el co-receptor CD8. Tomada de <sup>41</sup>.

### III.4 RESPUESTA CITOTÓXICA MEDIADA POR LINFOCITOS T CD8+

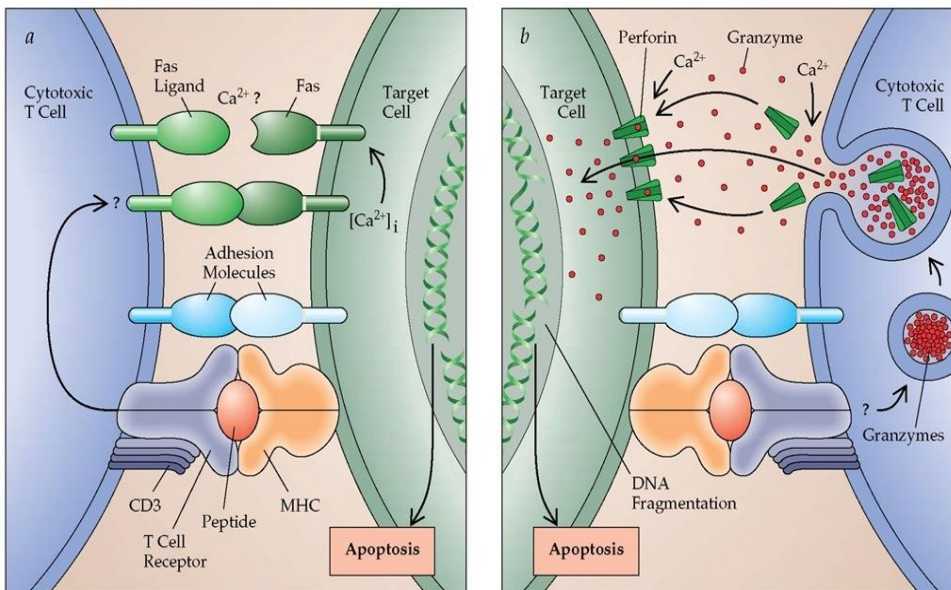
Aquellos linfocitos que se diferencian en efectoros, se denominan linfocitos citotóxicos o CTLs (Cytotoxic T Lymphocytes) y tienen como función primordial la destrucción de células infectadas. Los CTLs destruyen blancos celulares que expresan el mismo antígeno asociado a HLA-I que desencadenó la proliferación y diferenciación del linfocito del que derivan. Esta especificidad garantiza que las células no infectadas no sean destruidas por acción de los CTLs.

Para desencadenar la respuesta citotóxica, es necesario que el TCR y la proteína LFA-1 (Lymphocyte function-associated antigen 1) del CTL, reconozcan el complejo HLA-péptido e ICAM-1 (molécula de adhesión intercelular-1) de la célula blanco, respectivamente. La

respuesta lítica es altamente dirigida, ya que es mediada por proteínas secretadas exclusivamente hacia el punto de sinapsis inmunológica<sup>34</sup>.

El principal mecanismo de destrucción mediado por CTLs es la liberación de proteínas citotóxicas almacenadas en gránulos conocidos como lisosomas secretores. Los lisosomas pueden contener a las **granzimas** A, B y C que actúan como serin-proteasas; a la **perforina** que perturba la integridad de la membrana de la célula blanco; a la **serglicina** que ensambla un complejo de granzima y perforina; o a la **granulisina** que altera la permeabilidad de la membrana. Una vez dentro de la célula blanco, las granzimas escinden sustratos que desencadenan la respuesta apoptótica (p. ej. Caspasa-3)<sup>34</sup>. La apoptosis de células infectadas interrumpe la producción de nuevos virus y además favorece la liberación del patógeno para ser ingerido por fagocitos (Figura 23).

Un mecanismo secundario de lisis celular es mediado por la interacción de otra proteína de membrana del CTL denominada ligando de Fas (FasL), que se une al receptor de muerte celular, Fas, expresado en diversos tipos celulares<sup>34</sup>. La cascada de señalización producida, deriva en la activación de vías apoptóticas (Figura 23). La muerte de la célula blanco, por cualquier mecanismo, se efectúa de 2 a 6 horas posteriores a la activación del CTL y avanza incluso si las células se separan.



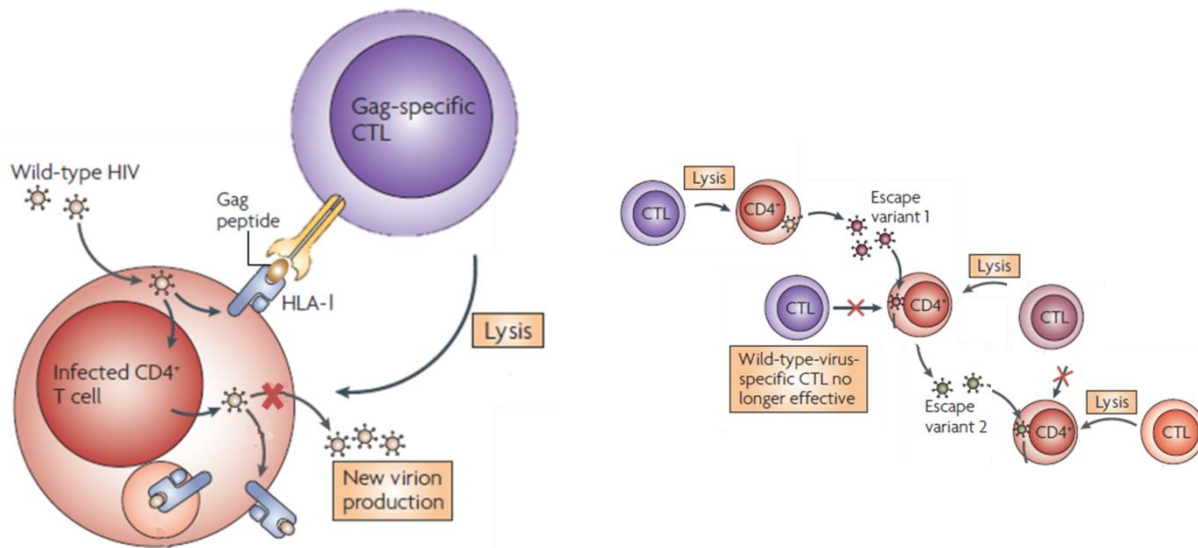
**Figura 23 Mecanismos efectoros de la respuesta lítica mediada por linfocitos T CD8+.** Los linfocitos T citotóxicos ejercen su acción efectora mediante la unión de Fas-FasL que induce apoptosis (izquierda) o por la secreción dirigida de perforina y granzima cuyos receptores se encuentran en la superficie del linfocito blanco (derecha). Tomada de <http://what-when-how.com/acp-medicine/cell-cell-interactions-cytokines-and-chemokines-in-immune-response-mechanisms-part-3/>.

### **III.5 INTERACCIÓN ENTRE EL VIH Y EL HOSPEDERO: EVOLUCIÓN MEDIADA POR HLA**

La relación temporal existente entre la aparición de la respuesta citotóxica y la disminución de la viremia remarca la importancia de este mecanismo inmune para el control de la infección. Numerosos estudios han establecido una fuerte relación entre la eficiencia de la respuesta citotóxica y el haplotipo HLA (conjunto específico de alelos) a través de observaciones como: i) la asociación epidemiológica entre alelos específicos de HLA y el grado de progresión de la infección<sup>42</sup>, ii) la observación de una progresión más rápida en individuos homocigotos de algunos alelos HLA-B<sup>43</sup> y iii) la asociación entre la aparición de ciertas mutaciones y una pérdida temporal del control de la infección. Estas evidencias sugieren que la calidad de la respuesta citotóxica está fuertemente determinada por los alelos presentes en un individuo y por lo tanto, de las características de los epítomos presentados.

La enorme variabilidad presente en los alelos HLA reside principalmente en los aminoácidos que conforman el surco de unión a péptidos. La respuesta mediada por CTLs, al estar basada en el reconocimiento de secuencias lineales de aminoácidos, actúa como una potente fuerza de selección de mutaciones dentro de los puntos de anclaje a las moléculas presentadoras. Dichas mutaciones pueden representar una ventaja evolutiva para el virus al favorecer el escape inmune (Figura 24) o aumentar la inmunogenicidad del péptido y por lo tanto fortalecer la respuesta citotóxica. Así mismo, se sabe que existen mutaciones que favorecen el escape inmune pero pueden afectar la capacidad del virus para replicarse eficientemente (fitness viral). Este fenómeno puede derivar en la aparición de otras mutaciones que compensen el fitness deficiente o bien, que alteren el procesamiento antigénico, estabilizando las mutaciones de escape seleccionadas inicialmente<sup>44</sup>.

La selección de variantes de escape por la presencia de un haplotipo determinado fue documentada inicialmente en 1996<sup>45</sup>. Pero en la última década, se han observado mutaciones que son seleccionadas de forma sistémica por alelos HLA específicos. Esto sugiere que la evolución del VIH sigue patrones predecibles cuando una fuerza inmune se aplica sobre él.



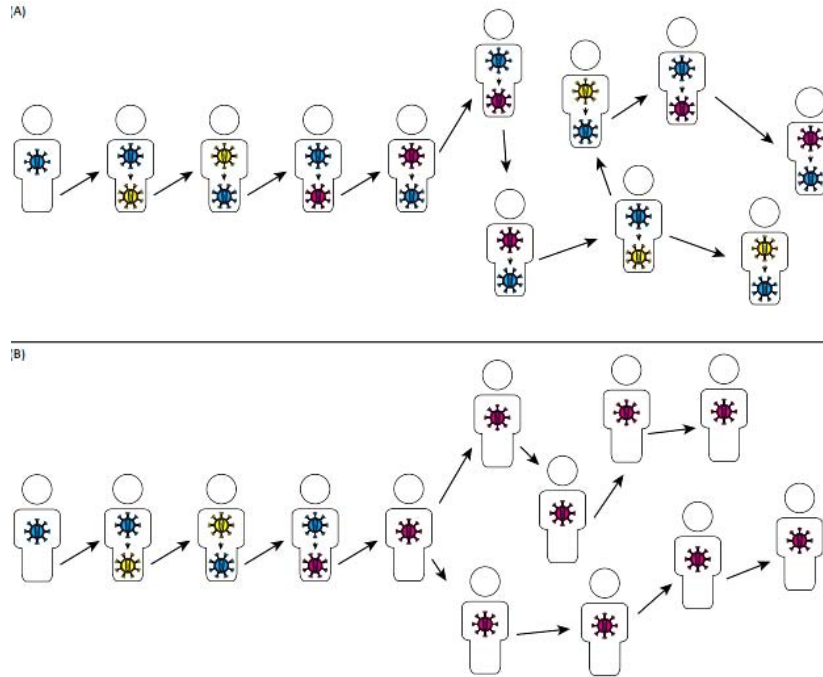
**Figura 24 Presión selectiva mediada por HLA.** A la izquierda; la respuesta citotóxica está restringida por el reconocimiento de péptidos virales presentados en moléculas de HLA-I. Dicho reconocimiento induce la lisis de la célula infectada, inhibiendo la producción de nuevos viriones. A la derecha; la restricción del HLA-I actúa como potente fuerza selectora de mutaciones que favorecen la aparición de variantes virales de escape. El aumento en la variabilidad vuelve inefectiva la respuesta citotóxica inicial. Tomada de <sup>44</sup>.

El hecho de que los sitios y las vías de escape sean predecibles indica que a pesar de la enorme diversidad que existe en el VIH, existen límites estrictos que dirigen su evolución. Además, si el escape viral dentro de un individuo es previsible basado en sus alelos HLA, entonces la frecuencia y distribución de los alelos HLA en los humanos dirige la evolución del virus a nivel poblacional<sup>44</sup>.

Estudios sucesivos al respecto, determinaron que las mutaciones seleccionadas pueden o no conservarse. Es decir, en algunas ocasiones, las mutaciones asociadas a un alelo HLA específico pueden persistir si ese alelo es prevalente; o por el contrario, revertir cuando el virus mutado es transmitido a un individuo carente del alelo selector. Esta dinámica establece, por lo tanto, que la evolución viral mediada por alelos HLA depende del balance entre el beneficio de las mutaciones de escape, el costo en fitness asociado y la prevalencia de alelos selectores (Figura 25).

Desde sus inicios, uno de los casos más ampliamente estudiados corresponde al fitness mediado por la proteína Gag y restringida por alelos HLA-B. Las observaciones al respecto han concluido que esta relación es determinante en el establecimiento del set point viral y la rapidez de progresión de la enfermedad <sup>46-49</sup>. Lo cual puede deberse a que

los péptidos derivados de estas proteínas estructurales son altamente inmunogénicos, lo que inicialmente permite el establecimiento de una respuesta inmune enérgica pero también dirige la selección de mutaciones de escape. No obstante, dichas mutaciones pueden derivar en la formación de viriones deficientes y por lo tanto disminuir la efectividad de la replicación.



**Figura 25 Reversión y persistencia de mutaciones de escape.** Algunas mutaciones de escape (rosa y amarillo) pueden revertir después de la transmisión hacia su forma “wild-type” (azul). A) Si la reversión se da de forma consistente entre pacientes con distintos haplotipos, la frecuencia de las cepas wild-type y mutadas se mantendrá medianamente constante. B) Si una mutación de escape persiste de forma consistente a pesar de la transmisión entre pacientes inmunogenéticamente diferentes, probablemente por un costo mínimo en fitness o por una compensación efectiva del daño, la frecuencia de la forma mutada aumentará gradualmente; aumentando también el progreso de la epidemia. Tomada de <sup>50</sup>.

El primer estudio de asociación entre HLA y progresión de la infección, reveló que los alelos B\*27 y B\*57 impactan de forma severa en el retraso de la progresión<sup>45</sup>. Específicamente, en las personas portadoras del alelo HLA B\*57 se favorece la mutación T242N que corresponde a la posición 3 del epítipo TW10 de p24 (proteína de matriz) y que impide la unión del péptido a la molécula de HLA. Esta mutación se expresa de forma temprana (primeras semanas de la infección) y favorece el control de la viremia<sup>42</sup>. Sin embargo, disminuye de forma significativa la capacidad del virus para replicarse. Esto se ha demostrado por una rápida reversión cuando el virus se transmite a un individuo no B\*57<sup>42</sup>.

Por su parte, el alelo B\*27 favorece la sustitución de una L a M en la posición 6 del epítipo KK10 de Gag y años después de ésta, aparece la mutación R a K en la posición 2 del mismo epítipo<sup>51</sup> que actúa como mutación compensatoria.

Existen más de 40 publicaciones que han identificado alelos con un rol “protector” en la enfermedad, y alrededor de 25 artículos que han demostrado la existencia de alelos asociados a “riesgo” de progresión. La mayoría de las asociaciones involucran alelos HLA-A y HLA-B<sup>52</sup>. Los alelos HLA-C que han sido asociados con progresión de la infección, generalmente se encuentran en desequilibrio de ligamiento con HLA-B. Aunque en los últimos años se han encontrado algunos HLA-C con un efecto propio sobre la progresión pero los mecanismos por los que afecta la progresión de la infección parecen ser distintos.

En cualquier caso, los datos existentes se han obtenido de estudios realizados en cohortes mayoritariamente caucásicas, africanas y más recientemente, asiáticas. Entre las que se han observado diferencias en el impacto del HLA sobre el virus debido a la variabilidad de prevalencias alélicas de cada población (Figura 26). Poca información se ha generado para poblaciones con orígenes distintos, como la latinoamericana; y por lo tanto, se desconoce si entre éstas existan mecanismos de control de la replicación viral o evasión inmune distintos a los reportados.

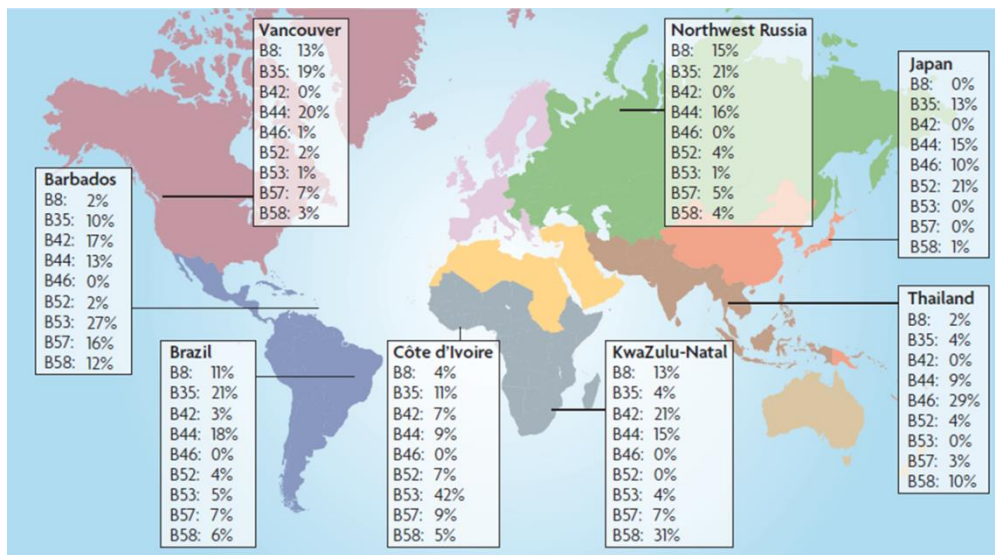


Figura 26 Ejemplo de distribución poblacional de algunos alelos HLA-B en el mundo. Se muestran las prevalencias (expresadas en %) de algunos alelos en distintas cohortes alrededor del mundo. Tomada de<sup>52</sup>

Es por ello, que caracterizar y tratar de establecer mecanismos particulares de interacción entre la respuesta citotóxica y la replicación del virus en una población claramente distinta a las estudiadas, representa un avance importante en el conocimiento de la patogénesis del virus y de su evolución poblacional en un contexto particular.

Por otro lado, los avances en el campo de vacunas parecen dirigirse al fortalecimiento de la respuesta citotóxica. Los intentos que hasta el momento se han realizado no han sido capaces de producir una respuesta sostenida ni suficientemente fuerte contra diversos epítomos, ya que el efecto parece estar dirigido hacia porciones virales que toleran alta variabilidad, por lo que la respuesta termina por ser inservible.

Esto sugiere que el éxito de una vacuna basada en CTLs podría estar cerca, si se inducen respuestas potentes contra epítomos múltiples y poco variables, por ejemplo de Gag, que permitan un efecto persistente, similares a las que ocurren de forma natural en algunos individuos.

El conocimiento de la interacción entre HLA y VIH permitiría, además, la caracterización de aquellas mutaciones que tienden a conservarse en determinada población. De tal forma, que el efecto de una vacuna pudiera ser independiente del HLA al menos, dentro del mismo contexto evolutivo del virus.



## **CAPITULO 2. OBJETIVOS**

### **OBJETIVO GENERAL**

Evaluar el efecto de los alelos HLA en la capacidad replicativa del VIH circulante e identificar las mutaciones seleccionadas por dichos alelos con potencial efecto en la progresión o control de la infección por VIH-1 en el contexto inmunogenético de la población guatemalteca y en comparación con observaciones en otras poblaciones.

### **OBJETIVOS PARTICULARES**

- ⌘ Realizar la tipificación de HLA en alta resolución de una cohorte de pacientes provenientes de Guatemala con infección por VIH y sin exposición a Tratamiento Antirretroviral.
- ⌘ Realizar la secuenciación del gen *gag* de virus de una cohorte de individuos guatemaltecos y observar su variabilidad.
- ⌘ Evaluar la capacidad de replicación viral derivada del gen *gag*, mediante un ensayo *in vitro* basado en la generación de virus recombinantes.
- ⌘ Evaluar la asociación entre la capacidad replicativa viral y los indicadores clínicos de progresión en la infección por VIH (carga Viral y CD4+).
- ⌘ Analizar el impacto de los alelos HLA encontrados sobre la capacidad replicativa del virus.
- ⌘ Investigar el impacto de mutaciones encontradas en posiciones bajo presión selectiva del gen *gag* del VIH, sobre la capacidad replicativa viral.

### **CAPITULO 3. ANTECEDENTES**

Debido a la coincidencia temporal entre la aparición de la respuesta citotóxica y la disminución significativa de la replicación viral durante la etapa aguda de la infección por VIH-1, el control del virus ha sido atribuido a los linfocitos T CD8+ citotóxicos (CTL, inglés), cuya respuesta está dirigida y restringida por los alelos HLA-I.

El control del VIH-I es consecuencia de la unión de péptidos virales, particularmente derivados del polipéptido Gag, a moléculas de HLA que pueden ser reconocidas por CTLs, con la consecuente eliminación de las células infectadas. Sin embargo, el VIH es capaz de evadir rápida y eficazmente esta presión inmune a través de la adquisición de mutaciones que alteran la unión al HLA o el reconocimiento de péptidos virales por parte del CTL. Adicionalmente, se ha observado que algunas de las mutaciones de escape inmune pueden afectar la capacidad replicativa viral (CRV), por lo que tienden a ser compensadas por otras mutaciones o a revertir en ausencia del alelo selector <sup>53-55</sup>.

Las asociaciones estudiadas inicialmente dilucidaron uno de los mecanismos por los que pacientes conocidos como controladores de élite son capaces de mantener la replicación viral indetectable sin necesidad de tratamiento antirretroviral. En estos pacientes, hay un enriquecimiento de alelos B\*27 y B\*57 que seleccionan mutaciones en epítopes de Gag con alto costo en la CRV <sup>42,53,54</sup>. Caso contrario, es la asociación del alelo B\*35:01 a una rápida progresión a SIDA <sup>52</sup>; aunque el mecanismo exacto se continúa estudiando, las hipótesis apuntan a la selección de una o varias mutaciones de escape inmune que no afectan CRV.

De este modo, se ha observado que el escape viral es predecible con base en los alelos HLA presentes en un individuo. Y por lo tanto, lo es también en una población con una prevalencia alélica particular. Sin embargo, la distribución tan divergente de los alelos HLA-I alrededor del mundo, sugiere la existencia de mecanismos de control y progresión específicos para cada población, es decir, existe un grado de adaptación distinto entre el virus y el sistema inmune que es dependiente de los alelos HLA-I más prevalentes en una región.

Se han realizado numerosos estudios del impacto de los alelos HLA en la progresión de la infección, mayormente en poblaciones caucásicas, africanas y algunas cohortes asiáticas<sup>52</sup>. Ello ha permitido establecer asociaciones estadísticas entre la presencia/ausencia de determinado alelo y una variación en la carga viral o conteo de T CD4+. La aparición consistente de dichas asociaciones ha permitido la identificación de alelos de riesgo y alelos protectores.

Para el estudio profundo de los mecanismos de progresión o control de la infección, dependientes de HLA, se han abordado distintas estrategias. Una de las más utilizadas, es la determinación *in vitro* de la capacidad replicativa viral que se basa en un sistema de virus recombinantes (NL4-3 y un gen de estudio) y una línea celular linfocitaria, desarrollada por el grupo de Marck Brockman y Cols<sup>56</sup>. Esta estrategia facilita la evaluación de la capacidad de replicación de los virus circulantes en un paciente, que es producto de la adaptación al sistema inmune.

Ambos enfoques, el estadístico y el experimental, han resultado en la descripción del contexto inmunogenético de cada población estudiada, el impacto de los alelos más prevalentes en la progresión de la enfermedad y la propuesta de mecanismos de evasión inmune y de control viral.

Como se mencionó anteriormente, la interacción HLA-VIH ha sido mayormente estudiada en poblaciones no latinas. Por ello se desconoce si existen alelos particulares que impacten en la progresión de la infección, distintos a los ya estudiados.

El conocimiento de la adaptación viral al sistema inmune, no solo nos permite conocer más profundamente la interacción entre el virus y el hospedero, o el virus y una población; sino que podría tener una aplicación en el campo de la vacunación, ya que el diseño actual de estrategias profilácticas y terapéuticas busca inducir la respuesta citotóxica hacia fracciones del virus que no toleren variabilidad o cuya mutación repercuta en la capacidad replicativa. De este modo, se podría superar la ineficacia de las vacunas universales que se han probado hasta el momento y mediante la identificación de mecanismos de control o progresión asociados a HLA-I, favorecer el diseño de vacunas eficaces específicas para una población o región geográfica.

## **CAPITULO 4. METODOLOGÍA**

### **4.1 SELECCIÓN DE PACIENTES**

El grupo de pacientes estudiado deriva de la cohorte mesoamericana del Departamento de Investigación en Enfermedades Infecciosas, CIENI, del Instituto Nacional de Enfermedades Respiratorias, INER. Esta cohorte fue conformada con la colaboración del Hospital Roosevelt de la ciudad de Guatemala.

En el Hospital Roosevelt fueron reclutados, desde el año 2010, individuos guatemaltecos con infección documentada por VIH sin previa exposición a Tratamiento Antirretroviral (TAR). A cada uno se le realizó una toma de sangre única que fue enviada al CIENI para su procesamiento. Todos los pacientes participantes otorgaron consentimiento escrito para la donación de muestras sanguíneas y el protocolo de investigación fue evaluado y aprobado por el Comité de Ciencia y Bioética del INER.

### **4.2 PROCESAMIENTO DE MUESTRAS**

Dentro del CIENI, el Laboratorio de Diagnóstico Viroológico realizó la determinación de carga viral plasmática, mediante PCR en tiempo real y conteo de linfocitos T CD4+ por citometría de flujo, para cada paciente. Adicionalmente, a partir de cada muestra sanguínea se obtuvieron leucocitos y plasma que fueron almacenados para posteriores estudios.

De la cohorte mesoamericana se seleccionaron de manera aleatoria 160 muestras que conformaron la cohorte de estudio del proyecto detallado en este reporte. Para cada una, se realizó la extracción de ácidos nucleicos como se describe a continuación.

#### **4.2.1 EXTRACCIÓN DE ADN**

Se realizó la purificación de ADN genómico a partir de la capa de leucocitos recuperada de la centrifugación de las muestras sanguíneas. Se utilizó el estuche de reactivos QIAmp DNA Blood Mini Kit (QIAGEN, Valencia, CA) de acuerdo a las especificaciones del fabricante. De manera breve, se tomaron 200 uL de muestra de leucocitos que fueron lisados mediante la adición de una proteinasa y su buffer correspondiente, provistos en el estuche de reactivos. Después de una incubación de 10 minutos a 56°C, se agregó etanol

absoluto y se colocó la mezcla en una columna que posee una membrana absorbente de sílice que permite la unión selectiva del ADN. Se centrifugó la muestra favoreciendo la unión del ácido nucleico a la membrana y se realizaron dos lavados sucesivos con distintas soluciones para remover restos celulares, proteínas u otros contaminantes. Finalmente, se eluyó el ácido nucleico en 100 uL de buffer y se almacenó a -20°C hasta su uso.

#### **4.2.2 EXTRACCIÓN DE ARN VIRAL**

Utilizando las alícuotas de plasma de cada paciente, se extrajo ARN viral con los reactivos de QIAamp Viral RNA Mini kit (QIAGEN, Valencia CA), de acuerdo a las especificaciones del fabricante. Brevemente, para esta técnica se centrifugaron alícuotas de 1 mL de plasma para concentrar al virus. Mediante el uso de una solución de lisis altamente desnaturante de tiocianato de guanidina se realizó la lisis viral, promoviendo la liberación del ácido nucleico y posteriormente se colocó la muestra en una columna de sílice permitiendo la unión del ARN viral a la membrana de la columna. Ésta se lavó en dos ocasiones con distintas soluciones amortiguadoras para eliminar contaminantes y finalmente, se eluyó el ARN puro en 50 uL de una solución de baja fuerza iónica.

El procesamiento posterior de los ácidos nucleicos contempló:

- I. Secuenciación
- II. Ensayos de Capacidad Replicativa
- III. Tipificación HLA

### **4.3 SECUENCIACIÓN**

#### **4.3.1 AMPLIFICACIÓN**

Se utilizó la técnica de Reacción en Cadena de la Polimerasa (PCR) anidada para amplificar un fragmento aproximado de 1600 pb de la región génica *gag-pro* del VIH a partir del RNA extraído.

Esta técnica consta de dos fases. La primera de ellas, amplificación externa, requirió el uso del “primer sentido” **623Fi** 5’ AAATCTCTAGCAGTGGCGCCCGAACAG 3’ y el “primer antisentido” **2cRx** 5’ TAACCCTGCGGGATGTGGTATTCC 3’. Para la amplificación interna o segunda ronda, los primers utilizados fueron **G1** 5’ GCAGGACTCGGCTTGCTGAA 3’ y **G10** 5’ TATCATCTGCTCCTGTATC 3’ como sentido y antisentido, respectivamente.

La amplificación externa requirió inicialmente 30 min. a 55 °C, 2 min. a 94°C. Posteriormente, 35 ciclos de 15 seg. a 94°C, 30 seg. a 55°C y 2 min. a 68°C. Finalmente, fueron necesarios 5 min. a 68°C. Por otro lado, el protocolo para realizar la amplificación interna constó de 3 min. a 94°C seguidos de 35 ciclos de 30 seg. a 94°C, 30 seg. a 56°C y 2 min. a 72°C. La elongación final consistió en 5 min. a 72°C.

Los reactivos y proporciones para cada ronda de amplificación se indican a continuación:

♦ Amplificación externa (Primera ronda)

♦ Amplificación interna (Segunda ronda)

Reactivo	1X (uL)	Reactivo	1X (uL)
H <sub>2</sub> O	8.5	H <sub>2</sub> O	13.75
2X Buffer (contiene dNTPs)	12.5	10X Buffer	2.5
Primer 623 Fi (20pM)	0.5	dNTP (1.25 mM)	4.0
Primer 2cRx (20pM)	0.5	MgCl <sub>2</sub> (50 mM)	1.0
Superscript III enzyme	0.5	Primer G1	1.0
ARN viral	5	Primer G10	1.0
<b>Volumen total</b>	<b>25</b>	Taq polymerase	0.25
		Producto de 1ª ronda	2.5
		<b>Volumen total</b>	<b>25</b>

Tras ambas rondas de amplificación, se realizó un corrimiento electroforético (80 mV / 40 min) en gel de Agarosa al 1% y se buscó una banda de 1600 pb para comprobar el éxito del procedimiento. Para aquellas muestras que resultó positivo, se realizó la purificación del ADN mediante el Kit QIAquick PCR Purification Kit (QIAGEN, Valencia CA), siguiendo el procedimiento especificado por el fabricante. Para ello, el kit contiene columnas con una membrana de sílice que permite la adsorción del ADN en presencia de una solución saturada de sales y un pH cercano a 7.5. Una vez lavada la membrana para eliminar

contaminantes de la PCR, se eluyó el ADN en condiciones de poca saturación de sales y un pH alcalino (8.5) con una solución de Tris.

#### 4.3.2 SECUENCIACIÓN AUTOMÁTICA

El ADN obtenido fue secuenciado mediante la técnica Sanger con el secuenciador capilar automático 3730xl Genetic Analyzer instrument (Life Technologies, Thermo Fisher Scientific, USA). Los primers “sentido” usados para este procedimiento fueron G2 (734-750 de HXB2), GAS3F (1401-1420), GAS3F (1401-1420), GAS5F (1696-1716), HPR1977F (1957 – 1977) y los antisentido GS1R (2313-2294), GAS4R (1481-1461), GAS6R (1790-1771) y GA2274R (2296 – 2274). El uso de 8 primers redundantes favoreció la obtención de al menos un duplicado en la secuencia de cada fragmento de la región de interés. La reacción se realizó empleando la química BigDye terminator v3.1 (Applied Biosystems, Carlsbad CA) en las siguientes proporciones:

<b>Reactivo</b>	<b>1X (uL)</b>
Big Dye v3.1	0.5
Oil Buffer 5X	3.5
H <sub>2</sub> O	2
Primer (2uM)	2
ADN (10-40 ng/uL)	2

El protocolo para la reacción de secuenciación precisó 1 min. a 96°C, 25 ciclos de 10 seg. a 96°C, 5 seg. a 50°C y finalmente 4 min. a 60°C.

La precipitación del ADN previa a su secuenciación, requirió la adición de 1 uL de acetato de sodio 3 M y 50 uL de etanol absoluto. La placa fue centrifugada a 2000g por 30 min para disolver impurezas de la reacción. El exceso de etanol fue retirado mediante una centrifugación de la placa invertida a 500 g por 1 min. Posteriormente, se adicionaron 150 uL de etanol al 70% y se llevó a cabo una centrifugación a 1650g por 15 min. Una vez terminado este ciclo, se invirtió la placa nuevamente para retirar el etanol, centrifugándose en tres ocasiones a 500g/1 min para favorecer el secado de la misma. Finalmente se resuspendió el ADN en 10 uL de formamida,

Los fragmentos secuenciados obtenidos se ensamblaron usando el software Geneious 8.0.3 y la secuencia del virus de referencia HXB2 como molde. De este modo, se obtuvo una secuencia consenso por cada muestra.

El mismo procedimiento fue utilizado para la obtención de una secuencia control posterior al ensayo de capacidad replicativa. La única variación fue la fuente del ARN, siendo plasma para la primera secuencia y cultivo viral para la segunda.

### **4.3.3 ALINEAMIENTO Y ANÁLISIS FILOGENÉTICO**

Con la ayuda del programa ClustalW del software Mega 6.0<sup>57</sup> se realizó el alineamiento de las secuencias virales previamente obtenidas, usando como referencia la secuencia consenso de subtipo B de VIH-1 obtenida de la Base de Datos de Los Alamos (<http://www.hiv.lanl.gov>).

Una vez alineadas las secuencias, se hizo el análisis manual de variación respecto a la secuencia consenso y se registraron todas aquellas posiciones con polimorfismos, inserciones o deleciones.

Por otro lado, se construyó un árbol filogenético para el análisis evolutivo usando el Software Mega 6.0. Para ello, se utilizó el método Maximum Likelihood y el modelo General Time Reversible para calcular la distancia evolutiva. La significancia fue determinada mediante 1000 réplicas de bootstrap. La edición del árbol filogenético se realizó con el software FigTree v.1.4.0.

De forma independiente y para corroborar la identidad de las secuencias antes y después del ensayo de capacidad replicativa, se utilizó el mismo método para la construcción de un árbol filogenético que permitiera la observación de pares de secuencias. Este análisis se realizó únicamente con un subgrupo de 25 pares de secuencias.

### **4.3.4 ANÁLISIS DE PRESIÓN SELECTIVA POR MODELO DE CODONES.**

Usando el servidor Data Monkey ([www.datamonkey.org](http://www.datamonkey.org)), el alineamiento obtenido se sometió a un análisis de presión selectiva por determinación de cociente de sustituciones sinónimas (dS) y no sinónimas (dN) utilizando el método SLAC (Single-Likelihood



Ancestor Counting). Este método se basa en la construcción de secuencias ancestrales sitio por sitio, utilizando métodos de máxima verosimilitud (maximum likelihood, ML), hasta obtener la de mayor congruencia y estimando valores de variación asociados a cada posición.

El cociente dN-dS es graficado a partir de cero y hacia valores positivos si en el codón analizado se favorece la variación de aminoácidos, y de cero hacia valores negativos si en determinado codón es preferencial la aparición del aminoácido silvestre. Data Monkey también evalúa la significancia de dicha variación y tabula las posiciones con cambios significativos,  $p < 0.01$ .

#### **4.4 ENSAYO DE CAPACIDAD REPLICATIVA (ECR)**

La técnica de medición de capacidad replicativa viral se basa en la construcción de virus quiméricos o recombinantes con una única fracción génica variable correspondiente a la región *gag-pro* de virus circulantes en cada paciente. El fragmento se inserta en un vector viral de la cepa de referencia de VIH-1, NL4-3 y los virus recombinantes se replican en una línea celular linfoide que contiene un gen reportero. El procedimiento se detalla a continuación:

##### **4.4.1 AMPLIFICACIÓN DEL FRAGMENTO GAG-PRO**

El producto de amplificación de la primera ronda de PCR detallada anteriormente, fue usado como sustrato para una segunda amplificación específica para el ensayo. Esta PCR requirió el uso de un par de oligonucleótidos largos que añaden al amplicón una región de recombinación con el vector NL4-3. **GagPro chimeric recomb\_F\_Toshi**  
GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGA  
GGGGCGGCGACTGGTGAGTACGCCAAAATTTTACTAGCGGAGGCTAGAAGGAGA  
GAGATGGG y **GagPro chimeric recomb\_R\_new**  
GGCCCAATTTTTGAAATTTTCTTCCTTTCCATTTCTGTACAAATTTCTACTAATGCT  
TTTATTTTTTCTTCTGTCAATGGCCATTGTTAACTTTTG fueron los primers sentido y antisentido utilizados.

La preparación de esta ronda de amplificación se realizó con las siguientes concentraciones de cada reactivo:

Reactivo	X 1 (uL)
H2O	38.5
Buffer 10X	5
MgCl <sub>2</sub> (50mM)	2
dNTP (Mix 10mM)	1
Primer recombinante F (20mM [20pmol/ul])	0.5
Primer recombinante R (20mM solution)	0.5
Enzima (High Fidelity)	0.2
Amplicón de primera ronda	2.5
<b>Volumen total</b>	<b>50</b>

El protocolo utilizado inició con 2 min. a 94°C, posteriormente 40 ciclos de 30 seg. a 94°C, 30 seg. a 60°C, 2 min. a 68°C y la elongación final fue de 7 min. a 68°C.

Se verificó la amplificación mediante una electroforesis en gel de Agarosa al 1%, siendo positiva al encontrar una banda alrededor de 1800 pb. Los productos de amplificación que cumplieron con la especificación fueron purificados mediante el sistema de columnas de sílice del estuche Qiagen detallado en la sección 5.3.1 de este documento.

#### 4.4.2 PREPARACIÓN DEL VECTOR

Se utilizó un vector plasmídico que contiene todos los genes del virus de referencia NL4-3 excepto la región correspondiente a *gag-pro* que ha sido “deletada” mediante la inserción de sitios de restricción para la enzima BstE II.

Previo a su uso, el vector fue linealizado mediante una digestión enzimática con BstE II (New Englan Biolabs). Para ello, se incubó un vial por 2 hrs en un baño a 60°C con los siguientes reactivos:

Reactivo	Cantidad (50 co-transfecciones)
Plásmido NL4-3 (~1000 µg/mL)	~500 µl (= 500 µg ADN)
10x NEB #3 Buffer	65 µl
100x BSA	6 µl
Enzima BstE II (10 U/µl)	100 µl (= 1000 Unidades, o ~2 U/µg ADN)

El sitio de reconocimiento de la enzima es  $5' \dots \overset{\nabla}{\text{GGT}}\text{NACC} \dots 3'$  y la digestión se verificó mediante un corrimiento electroforético en gel de agarosa al 1% que presentó el patrón de digestión en correspondencia con la información del fabricante.

#### 4.4.3 CULTIVO CELULAR

La línea celular usada es llamada GXR. Ésta deriva de linfocitos que expresan el receptor CD4, así como los dos co-receptores más ampliamente usados por el VIH para internalizarse en la célula, CCR5 y CXCR4. Estas células, además, poseen en su genoma el gen de la proteína verde fluorescente (GFP) acoplado a una región promotora retroviral LTR. La unión del factor de transactivación viral Tat a este promotor, permite la expresión del gen GFP que actúa como reportero. De este modo, la infección de una célula por VIH culminará en la expresión de GFP dentro de la misma, permitiendo la cuantificación del porcentaje de células infectadas por citometría de flujo.

Las células GXR se cultivaron en medio líquido a una densidad de 500,000 cel/mL en medio R10; el cual contiene como base RPMI 1640 (Lonza BioWhittaker, Suiza) adicionado con 10% de Suero Fetal Bovino (Lonza BioWhittaker, Suiza), 1% de L-Glutamina y 1% de Penicilina-Estreptomicina. El cultivo fue incubado a 37°C, 0.5 % CO<sub>2</sub> y 85 % de humedad.

#### 4.4.4 CO-TRANSFECCIÓN

Para permitir la conformación y propagación de los virus quiméricos, se co-transfectó el producto *gag-pro* amplificado y el vector digerido en la línea celular linfoide GXR, mediante electroporación. Para este procedimiento se utilizaron los reactivos y cantidades siguientes:

Reactivo	Cantidad por muestra
Plásmido pNL4-3ΔGag-Pro linealizado (~750 µg/mL)	13.5 µl (= 10 µg)
Amplicón de Gag-pro (producto de PCR)	45 µl (~2.5 µg)
Cultivo celular GXR	2x10 <sup>6</sup> células

Se aplicó un pulso eléctrico a cada celda con el equipo BioRad Gene Pulser Xcell (BioRad, CA USA) utilizando las siguientes condiciones:

- 300 Volts - Voltaje
- 500 uF - Capacitancia
- $\infty$  - Resistencia
- 4mm – Grosor de celda

Una vez transfectadas, se permitió el reposo de las células por 45 minutos y posteriormente, fueron colocadas en botellas de cultivo de 20 mL, junto con 5 mL de medio R10 previamente calentado a 37°C. Se incubaron en las mismas condiciones que las células GXR, mencionadas con anterioridad.

#### **4.4.5 PROPAGACIÓN DE VIRUS QUIMÉRICOS**

Tras dos días de incubación, cada cultivo fue adicionado con 10 mL de medio R10 para favorecer la división celular y la propagación del virus en caso de infección. Durante los cinco días subsecuentes se permitió la infección de las células y consecuente producción de nuevas partículas virales. A partir del día 8 se remplazaron 5 mL de cultivo con una cantidad equivalente de medio R10, cada dos días. Los cultivos virales se incubaron por aproximadamente 21 días más. Después de este periodo se evaluó la cantidad de virus producidos en cada botella mediante citometría de flujo. Para ello, se tomó de cada cultivo una alícuota de 2 mL que fue centrifugada a 1500 rpm por 10 min; se retiró el sobrenadante y las células se lavaron con 1 mL Buffer PBS + EDTA [2uM] en las condiciones de centrifugación antes mencionadas. Después de dos lavados, se añadieron 500 uL de una solución de Formalina al 1% (99% PBS + 1% formaldehído) para fijar las células. La cantidad extraída del cultivo fue reemplazada con medio R10.

Se midió la producción de proteína verde fluorescente (GFP) con el citómetro Attune NxT (Life technologies, Thermo Fisher Scientific, USA). El % de células positivas a GFP se tomó como una medida indirecta de la replicación viral. La producción de virus quiméricos fue monitoreada cada dos días hasta que cada cultivo alcanzara un mínimo de 20% de células GFP<sup>+</sup>.

Posteriormente, se realizó la cosecha de virus que supuso la centrifugación de todo el cultivo por 10 minutos a 1500 rpm. El sobrenadante fue almacenado en alícuotas de 1.5 mL a -80°C para posteriores evaluaciones y las células remanentes de cultivo fueron desechadas.

La estrategia de análisis que fue seguida para el monitoreo de la producción de virus quiméricos, así como para la titulación y ensayo de capacidad replicativa se detallan más adelante. Inicialmente, se seleccionaron células únicas o singlets (Figura 27-A). A partir de esta población se seleccionaron células vivas (Figura 27-B) y dentro de éstas se analizaron únicamente aquellas positivas para fluorescencia de GFP (Figura 27 C y D).

#### 4.4.6 TITULACIÓN

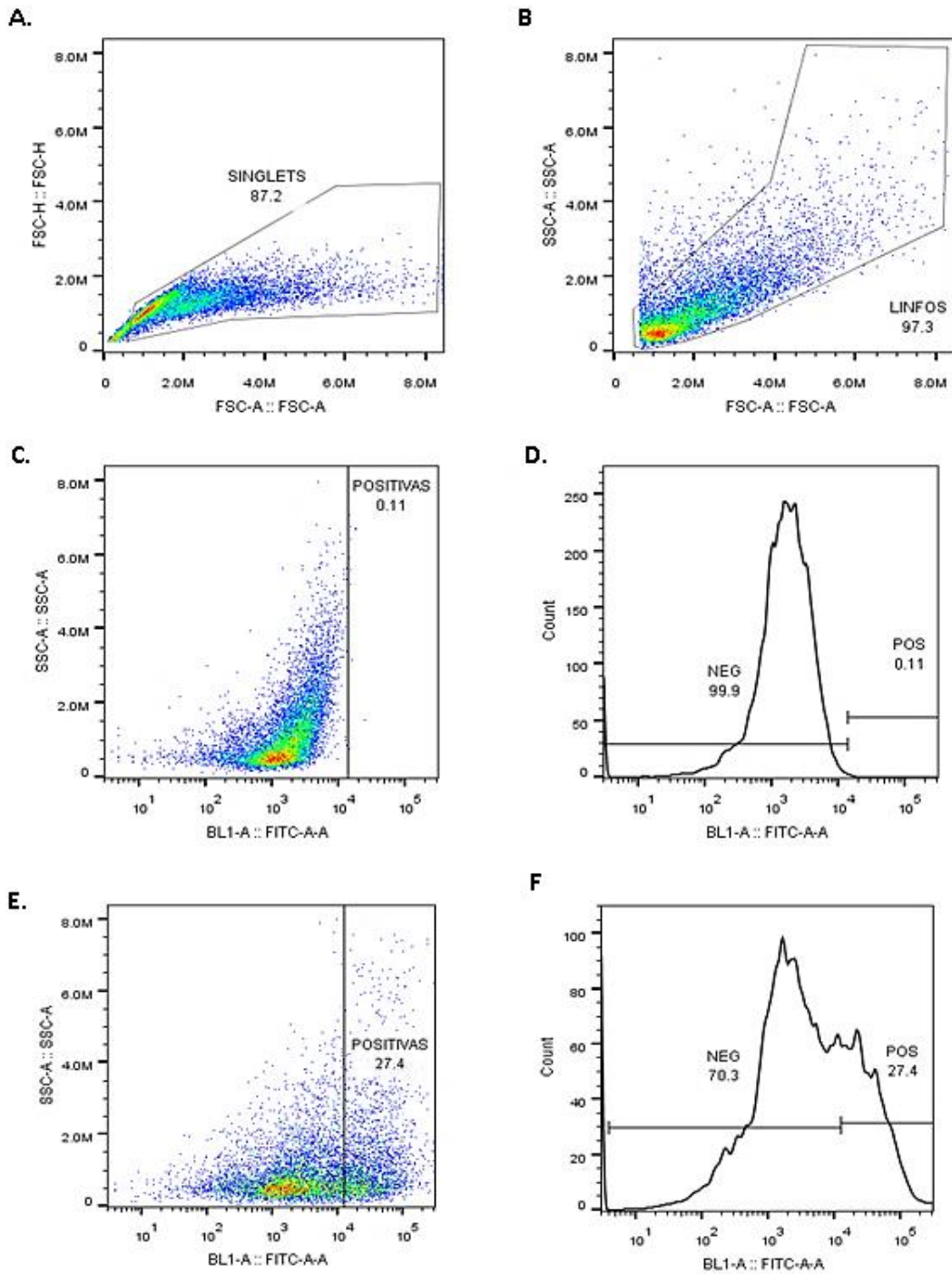
Para estandarizar el inóculo de cada virus utilizado en el ensayo de capacidad replicativa, se procedió a la titulación de una alícuota de cada cultivo viral, considerando un valor de Multiplicidad de Infección (MOI), es decir, la cantidad de virus presentes en cultivo por cada célula, de 0.3.

Para la titulación, se colocaron en una placa de 24 pozos de fondo plano, 1 millón de células GXR y 400 uL de alícuota de cada virus quimérico. Se utilizó como control positivo al virus NL4-3 completo y como control negativo 500 uL de medio R10. Tanto las muestras como el control positivo se cultivaron por duplicado.

El día posterior a la colocación del cultivo, se añadió a cada pozo 1 mL de medio R10 y se dejó en incubación por 24 horas más. Transcurrido este periodo, se tomaron 500 uL de cada cultivo y se colocaron en una placa de 96 pozos de fondo redondo. Se realizó la medición del % de células GFP<sup>+</sup> y se analizaron los datos con el software Flow Jo como se detalló en la sección anterior.

El porcentaje de células positivas obtenido de cada réplica fue promediado para obtener un valor único por muestra. Considerado este valor, se calculó la cantidad de alícuota viral necesaria para obtener un 0.3% de células positivas al día 2 de infección del ECR, mediante la siguiente fórmula:

$$\text{Cantidad de virus requerido } (\mu\text{L}) = 0.3\% \left( \frac{400 \mu\text{L}}{\%GFP} \right)$$



**Figura 27** Estrategia de análisis de citometría de flujo usada durante el ensayo de capacidad replicativa viral. Se muestra el análisis de cultivo usado como control negativo. Inicialmente, se seleccionaron células únicas o singlets (A). Dentro de esta población se analizaron linfocitos vivos (B) y posteriormente se identificaron aquellos GFP positivos (C y D). Los diagramas E y F corresponden a distintas visualizaciones de un ejemplo representativo del cultivo de un virus quimérico con 27.4% de células GFP positivas.

#### 4.4.7 CÁLCULO DE CAPACIDAD REPLICATIVA VIRAL (CRV)

Para la determinación de la Capacidad Replicativa Viral (CRV) se colocaron nuevos cultivos celulares en placas de 24 pozos. Al igual que en la titulación, se expuso  $1 \times 10^6$  células GXR a infección por virus recombinantes. Se realizó cada ensayo por triplicado, agregándose la cantidad de virus determinada con anterioridad diluida en 500  $\mu$ L de medio R10. Se incluyó un control positivo por duplicado y un único control negativo. Este día se identificó como Día 0.

En el día 1 se añadió 1 mL de medio R10 a cada pozo y el día 2 se inició el monitoreo de la CRV mediante citometría de flujo. La medición de células positivas se realizó en placa de 96 pozos en el equipo Attune NxT (Life technologies, Thermo Fisher Scientific, USA). Durante 8 días se registró diariamente el % de positividad a GFP en cada cultivo. Con la finalidad de descartar la autofluorescencia celular, al % GFP<sup>+</sup> obtenido en cada cultivo/día, le fue sustraído el % GFP del control negativo.

Una vez transcurridos los 8 días de medición, fue posible construir una curva de crecimiento de cada virus quimérico, para observar el progreso de la infección a lo largo del tiempo. El valor que refleja la CRV correspondió a la pendiente de cada curva, calculada mediante una regresión semilogarítmica en base a los siguientes criterios:

- a) Si al día 5 el valor de GFP<sup>+</sup> fue < 15%, se calculó la pendiente usando los valores de los días 3 al 6.
- b) Si al día 5 el valor de GFP<sup>+</sup> fue > 15%, se calculó la pendiente usando los valores de los días 3 al 5

El valor obtenido de cada virus recombinante fue dividido entre el promedio de CRV del control positivo, de este modo, se obtuvo un valor de capacidad replicativa viral relativo al virus de referencia usado en cada ensayo.

Finalmente, para corroborar la identidad genética de cada virus antes y después del ensayo *in vitro*, se extrajo RNA a partir de los virus cosechados restantes, se amplificó y secuenció la región *gag-pro* de estos virus y se comparó con la secuencia inicial. El

procedimiento para la obtención de la secuencia post-ensayo (S2) fue el mismo detallado para la secuenciación preliminar (Sección 5.3).

#### 4.5 TIPIFICACIÓN HLA

Se realizó la tipificación a alta resolución de los genes HLA de clase I A, B y C mediante la secuenciación automatizada de cada locus (SBT; sequence-based typing). La tipificación comprendió los exones 2 y 3 que codifican para los dominios extracelulares de las cadenas  $\alpha 1$  y  $\alpha 2$ .

Para ello, se extrajo ADN genómico a partir de aproximadamente  $5 \times 10^6$  PBMCs de cada paciente, el cual se utilizó en dos procedimientos independientes. El primero de ellos para la obtención de las secuencias de los exones 2,3 y 4 del loci A y B con el estuche de reactivos SeCore Locus Sequencing Kit (Invitrogen, USA). Y el segundo para la identificación de los exones 2,3 y 4 locus C con los reactivos del estuche AlleleSEQR® HLA PCR/Sequencing (Celera, USA). La realización de dos procedimientos obedeció a la disponibilidad exclusiva de primers adicionales para la amplificación de los exones 5,6 y 7 de HLA-C de marca Celera; dichos exones fueron requeridos cuando el software utilizado no fue capaz de determinar los alelos de un paciente con la información inicial.

Primeramente, la amplificación de los genes A y B requirió 19.8 uL de la mezcla de amplificación contenida en el estuche de reactivos, 0.2 uL de enzima (FastStart Taq) y 5 uL de ADN (15-30 ng/uL). Para el locus B esta mezcla se realizó por duplicado para cada paciente con la finalidad de amplificar cada alelo de forma independiente. El perfil de amplificación constó de:

Ciclos	Temperatura	Tiempo
1	95°C	4 min
35	95°C	20 seg
	63°C	20 seg
	72°C	40 seg
1	72°C	5 min
1	4°C	$\infty$



El éxito de la reacción de amplificación fue comprobada mediante una electroforesis en gel de Agarosa al 1.5%. Las muestras se consideraron positivas cuando se observaron dos bandas de 1100 y 900 pb para A; y dos de 1250 y 720 para cada prueba de locus B.

Posterior al corrimiento electroforético, se realizó la purificación de los productos de PCR mediante el uso de ExoSAP-IT. Este procedimiento requirió la adición de 4 uL de reactivo a cada muestra, para posteriormente colocar los tubos en un termociclador por 20 min. a 37°C, seguidos de 20 min. a 80°C para completar la purificación.

Para la reacción de secuenciación, se utilizaron 8 uL de mezcla locus-especifica de secuenciación y 2 uL de producto purificado. Para este procedimiento se utilizó un primer sentido y otro antisentido por cada exón para cada locus. Posteriormente, se realizó la precipitación de los productos de la reacción de secuencia. Este procedimiento requirió inicialmente la adición de 2 uL de buffer PPT, 40 uL de etanol al 100% y agitación vigorosa para la eliminación de contaminantes. A continuación, se centrifugaron las placas a 2000 g por 30 min. El exceso de etanol se retiró mediante la inversión de la placa sobre una toalla absorbente y la centrifugación a 500 g por 1 min. Para la precipitación final se utilizaron 100 uL de etanol al 70% y una centrifugación a 2000 g por 5 min. El sobrenadante se retiró mediante la centrifugación de la placa invertida como se mencionó anteriormente. Finalmente, se resuspendió el ADN en 15 uL de Formamida y se llevó a cabo su desnaturalización a 95°C por 2 min.

Por otro lado, la obtención de las secuencias de HLA-C requirió una reacción de PCR entre 16 uL de pre-Mix, 0.3 uL enzima (AmpliTaq Gold) y 4 uL de ADN (20 ng/uL) con el siguiente protocolo de amplificación:

<b>Ciclos</b>	<b>Temperatura</b>	<b>Tiempo.</b>
1	95 °C	10 min.
	96 °C	20 seg.
36	60 °C	30 seg.
	72 °C	3 min.
1	4°C	∞

La amplificación se corroboró en un gel de agarosa al 1% y las muestras positivas se purificaron mediante la adición de 3 uL de ExoSAP-IT y el seguimiento del siguiente perfil de termociclado: 37°C por 30 minutos y 80°C por 15 minutos. Una vez purificados los

productos, se llevó a cabo la reacción de secuenciación con los reactivos y cantidades que a continuación se especifican:

**Mezcla de reacción**

8 uL de mezcla de secuencia (incluida en el estuche de reactivos)	+	2 uL ADN (20 ng/uL)
<b>Protocolo de amplificación</b>		
Ciclos	Temperatura	Tiempo.
25	96 °C	20 seg.
	50 °C	30 seg.
	60 °C	2 min.
1	4°C	∞

La purificación de la placa se llevó a cabo con una centrifugación a 2000 g por 30 minutos posterior a la adición de 2 uL de acetato de sodio/EDTA y 25 uL de etanol al 100%. Se retiró el exceso de sobrenadante mediante una centrifugación con la placa invertida y enseguida se precipitó el ADN con 50 uL de etanol al 80% y una centrifugación de 2000 g por 5 minutos. El exceso de etanol se retiró de la manera antes mencionada y finalmente se diluyó el ácido nucleico en 15 uL de formamida.

Para ambos procedimientos se utilizó el Secuenciador Automático 3730xl Genetic Analyzer instrument (Life Technologies, Thermo Fisher Scientific, USA). Una vez obtenidas las secuencias derivadas de cada primer, éstas fueron ensambladas y analizadas con el programa Utype Dx Sequencing Analysis Software (Life Technologies, Thermo Fisher Scientific, USA). El cual, reconoce y acopla las secuencias correspondientes a los exones 2, 3 y 4 de cada locus obteniendo una secuencia continua, que es comparada contra una completa base de datos de alelos HLA. El programa arroja como resultado todos aquellos alelos con los que se encontró similitud y la cantidad de diferencias (missmatches) entre cada uno. Se realizó la revisión manual de cada nucleótido para evitar ambigüedades en la tipificación y se registraron únicamente los alelos asignados con 100% de identidad (cero missmatches).

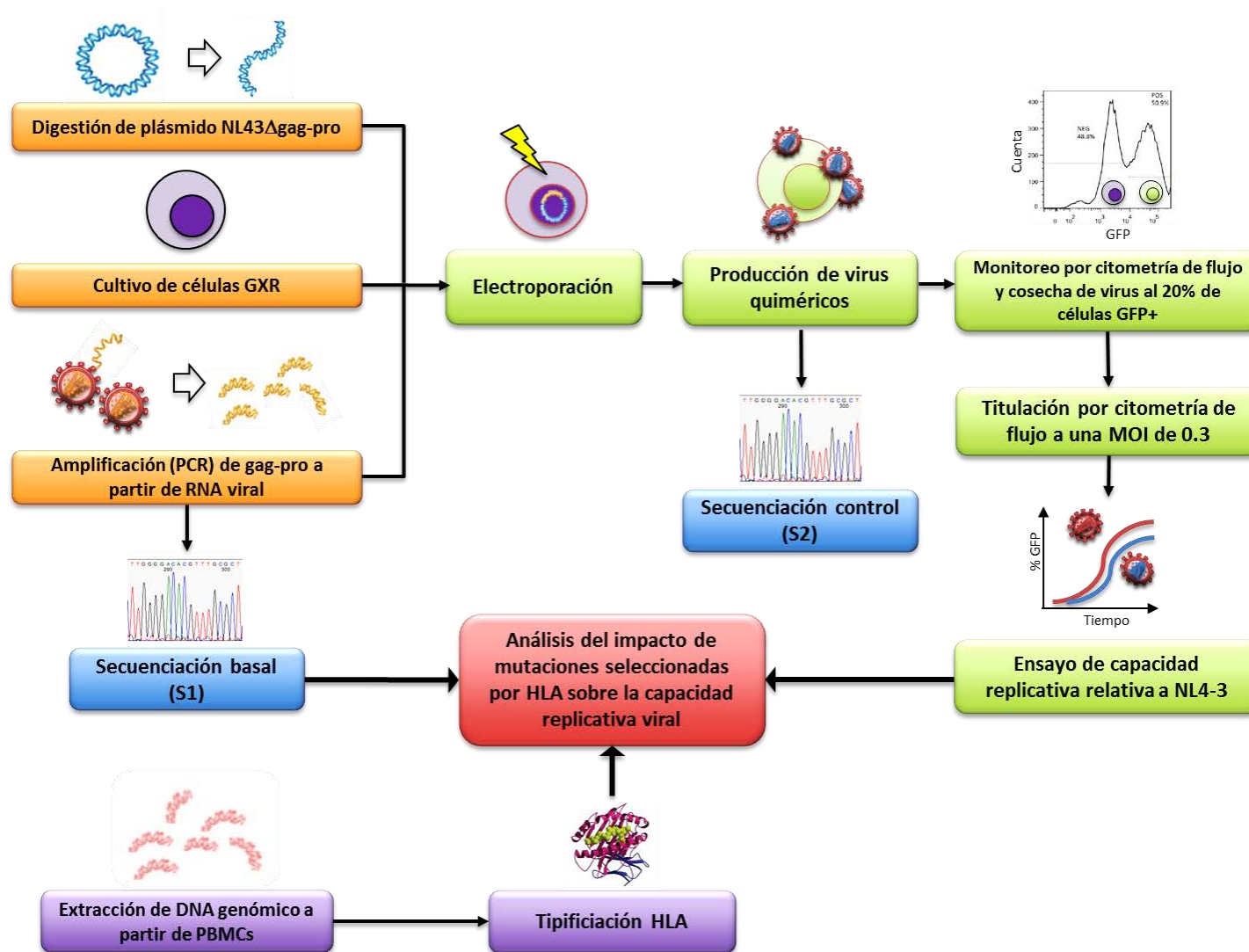
## 4.6 ANÁLISIS ESTADÍSTICOS

Para el análisis de prevalencia de HLA fue necesario calcular las frecuencias alélicas de cada gen. Para ello, se utilizó la herramienta “HLA Freq Analysis”/ HLA Graphing” encontrada en la Base de datos de Los Alamos ([http://www.hiv.lanl.gov/content/immunology/hla/hla\\_graph.html](http://www.hiv.lanl.gov/content/immunology/hla/hla_graph.html)).

Para estudiar la existencia de correlaciones entre distintos parámetros (carga viral vs capacidad replicativa y cuenta de CD4<sup>+</sup> vs capacidad replicativa) se utilizó la regresión lineal de Spearman. En el caso de comparación entre grupos independientes, ésta se llevó a cabo mediante la prueba de U de Mann-Whitney; considerándose significativos, valores de  $p < 0.05$ . Los análisis antes descritos, así como la construcción y edición de las gráficas correspondientes, se realizaron con el programa GraphPad Prism 6.0.

La comparación de prevalencia alélica en grupos con determinadas mutaciones, fue realizada a través de exactas de Fisher con una  $p < 0.05$  y mediante el uso de la herramienta HAPLOID (HLA Associated PoLymOrphism IDentifier), desarrollada por el grupo de la Dra. Zabrina Brumme y el Dr. Mark Brockman de la Universidad Simon Fraiser en British Columbia, Canadá, disponible en <http://oakdrum.com/HAPLOID/html/HAPLOID.html>.

La metodología completa se resume y esquematiza en la Figura 28.



**Figura 28 Diagrama de ensayo de capacidad replicativa viral.** Se construyeron virus recombinantes del fragmento de Gag amplificado de cada virus y el esqueleto del virus NL43. Se transfectaron por electroporación en una línea celular linfóide con un mecanismo reportero acoplado a proteína verde fluorescente (GFP). Se realizó el seguimiento por citometría de flujo. Por otro lado, se realizó la tipificación a alta resolución de los alelos HLA-I locus A, B y C. Se realizó la secuenciación de Gag antes y después del ensayo de capacidad replicativa.

## CAPITULO 5. RESULTADOS

### 5.1 DESCRIPCIÓN DE LA COHORTE DE ESTUDIO

El estudio se realizó en un subgrupo de 160 pacientes guatemaltecos de la cohorte Mesoamericana del CIENI. Este grupo se conformó de pacientes sin previa exposición a tratamiento antirretroviral, mayormente de género masculino (54%) en etapa crónica de la infección y con las características clínicas detalladas en la Tabla 1.

Parámetro	Mediana (Intervalo)
Edad (años)	36 (18-70)
Carga Viral (copias/mL)	52,030 ( 39 – 4'784,976)
Cuenta de CD4+ (células/uL)	186 (13 – 1,027)

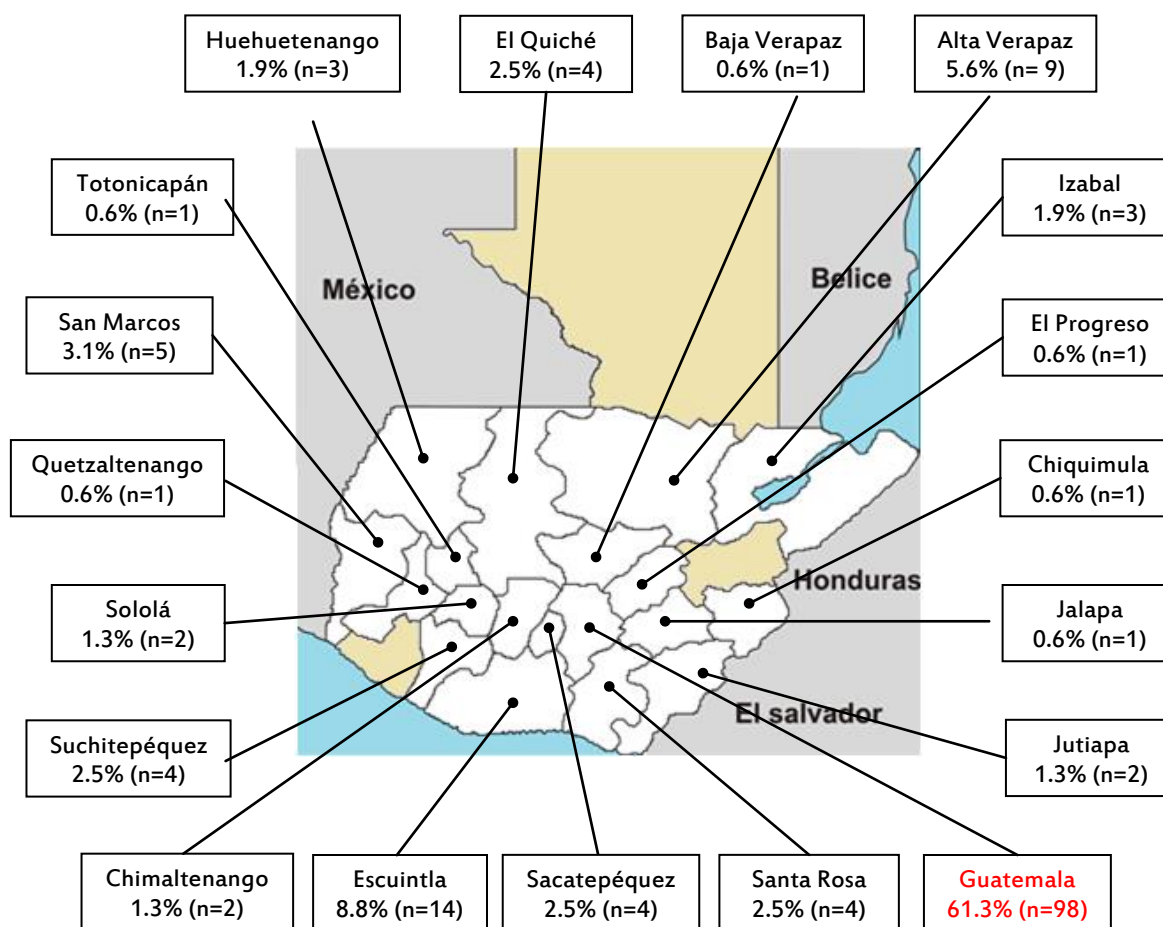
**Tabla 1 Características clínicas de la cohorte de estudio.** La cohorte de estudio se conformó por 160 pacientes guatemaltecos sin TAR mayormente del género masculino y en etapa crónica-tardía de la infección, caracterizada por las cargas virales y cuentas de linfocitos T CD4+ descritos.

En el grupo de estudio se contaron con muestras de 19 de los 22 departamentos que componen el país. El mayor porcentaje correspondió a pacientes provenientes de la ciudad capital, Guatemala (61.3%, n=98) que coincide con la región de mayor incidencia de VIH en aquel país. Por otro lado, los departamentos que no se hallaron representados en este grupo fueron El Petén, Retalhuleu y Zacapa (Figura 29).

### 5.2 FRECUENCIAS ALÉLICAS HLA-I

Para establecer el contexto inmunogenético de la cohorte de estudio, se realizó la tipificación a alta resolución de los alelos HLA-I locus A, B y C; se obtuvieron 160 tipificaciones del locus A, 157 del locus B y 158 del locus C. El 98.7% de los sujetos de estudio contaron con la tipificación completa de los 3 loci.

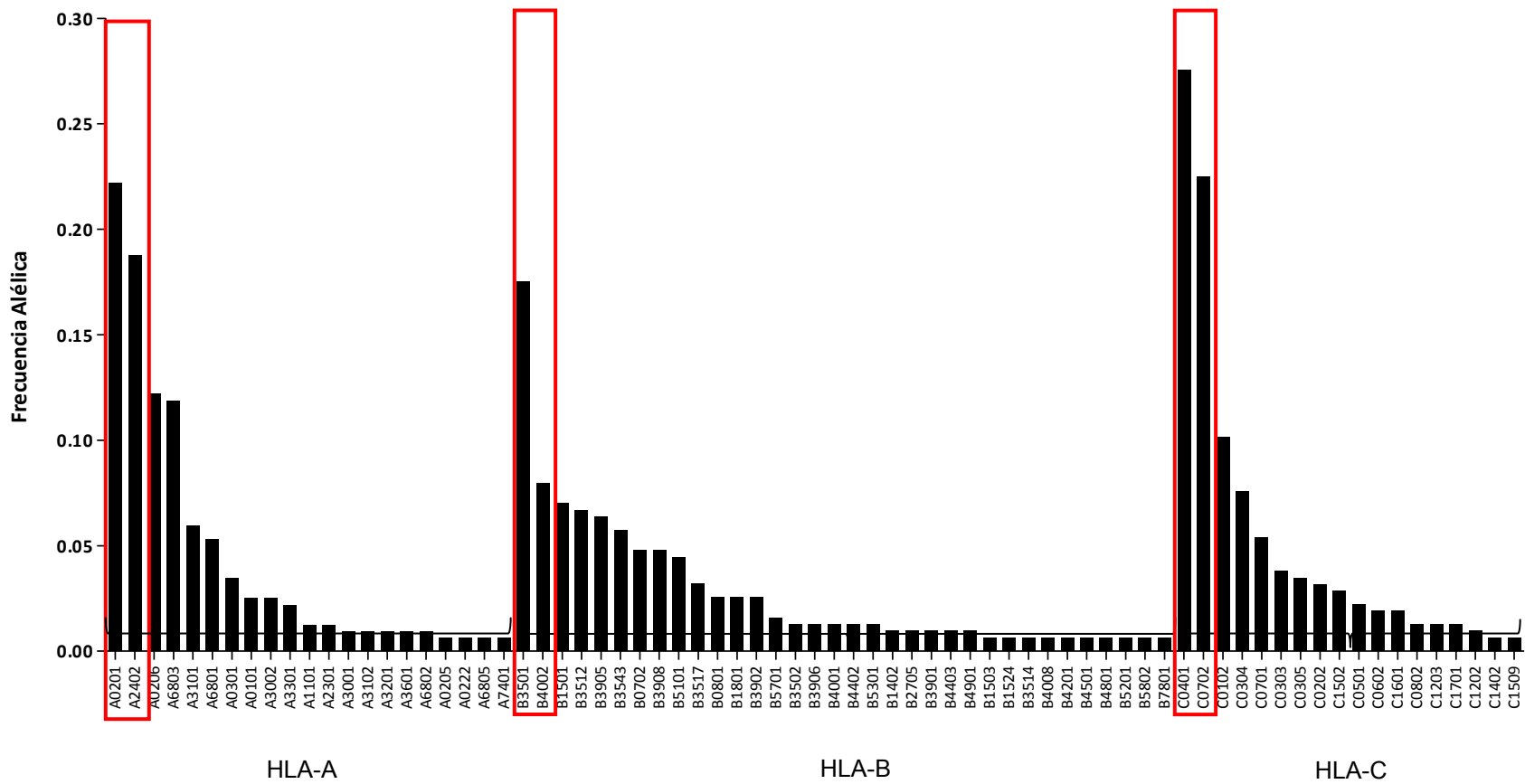
El análisis de los datos de tipificación arrojó la presencia de 32 alelos distintos del locus A, 49 del locus B y 23 del locus C en el grupo de estudio. Aquellos alelos presentes en al menos 2 pacientes fueron graficados y se muestran en la Figura 30.



**Figura 29 Distribución geográfica de la cohorte de estudio.** Se muestra la distribución de los 160 pacientes estudiados. La mayoría provinieron de la ciudad capital Guatemala (61.3%, n=98) y tres estados no estuvieron representados en esta cohorte y se muestran en amarillo: El Petén (al norte), Retalhuleu (occidente) y Zacapa (oriente).

Los alelos con los valores de frecuencia alélica más alta para cada locus fueron: A\*02:01, A\*24:02, B\*35:01, B\*40:02, C\*04:01 y C\*07:02 (Figura 30). Todos ellos, de origen amerindio.

Un estudio descriptivo previo de nuestro laboratorio mostró que en una cohorte de 415 pacientes, los alelos más frecuentes en esta población son A\*02:01 (frecuencia alélica 0.2206), A\*24:02 (0.1933), B\*35:01 (0.1401), B\*39:05 (0.079), B\*35:12 (0.0739), B\*40:02 (0.0662), C\*07:02 (0.2624) y C\*04:01 (0.2289) (Valenzuela P., comunicación personal). Ello sugiere que los datos obtenidos en el presente estudio pueden ser una representación de lo acontecido a nivel poblacional en dicha región.

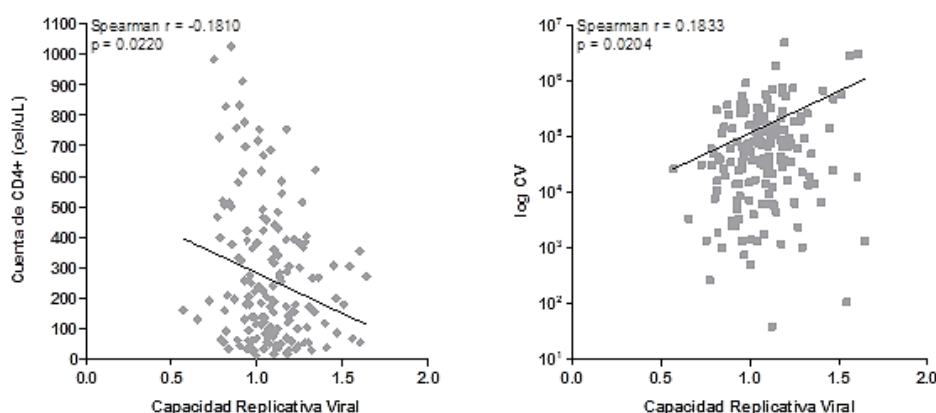


**Figura 30 Frecuencias alélicas de HLA-I en el grupo de estudio.** Se realizó la tipificación HLA a alta resolución de 160 pacientes guatemaltecos, obteniéndose 31 alelos HLA-A, 48 HLA-B y 22 HLA-C distintos. La frecuencia alélica de cada alelo fue graficada. Los dos alelos más frecuentes por locus se resaltan con un recuadro rojo. Para efectos de claridad, solo se graficaron los alelos presentes en al menos dos pacientes, las frecuencias alélicas totales se enlistan en el Apéndice I.

### 5.3 ANÁLISIS DE CAPACIDAD REPLICATIVA VIRAL (CRV)

El análisis de la capacidad replicativa viral (CRV) de los virus provenientes de los 160 pacientes participantes mostró una mediana de 1.0679 (min. 0.5689 – max. 1.6443).

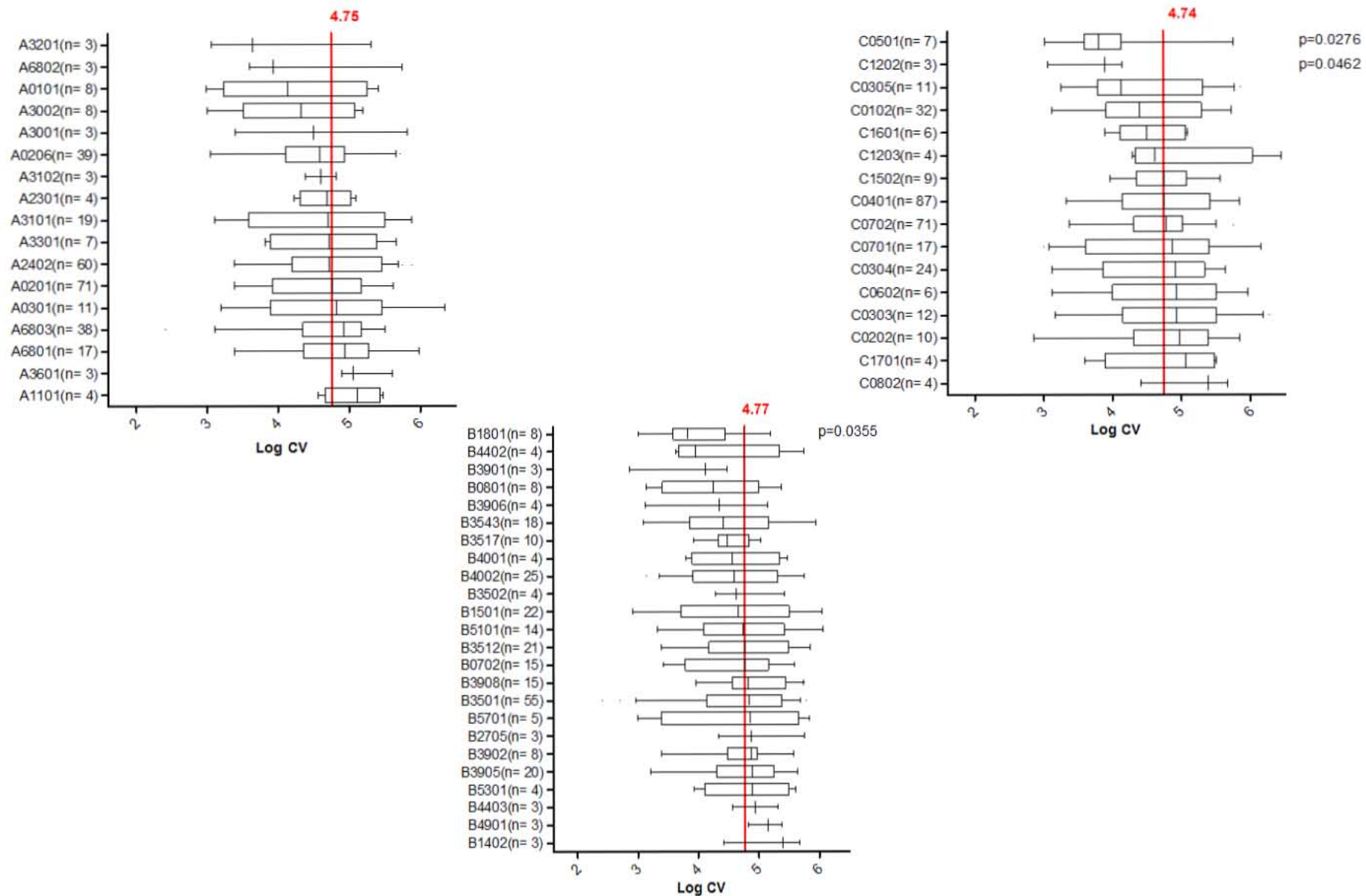
El análisis de correlación entre la capacidad replicativa y los parámetros de progresión clínica (cuenta de CD4<sup>+</sup> y carga viral) arrojó una correlación negativa ( $r=-0.1810$ ) entre CRV y cuenta de CD4<sup>+</sup> cuya significancia fue comprobada mediante una prueba de Spearman ( $p=0.0220$ ). Por otro lado, se observó una correlación positiva entre CRV y Carga Viral (CV) ( $r=0.1833$ ,  $p=0.0204$ ) (Figura 31).



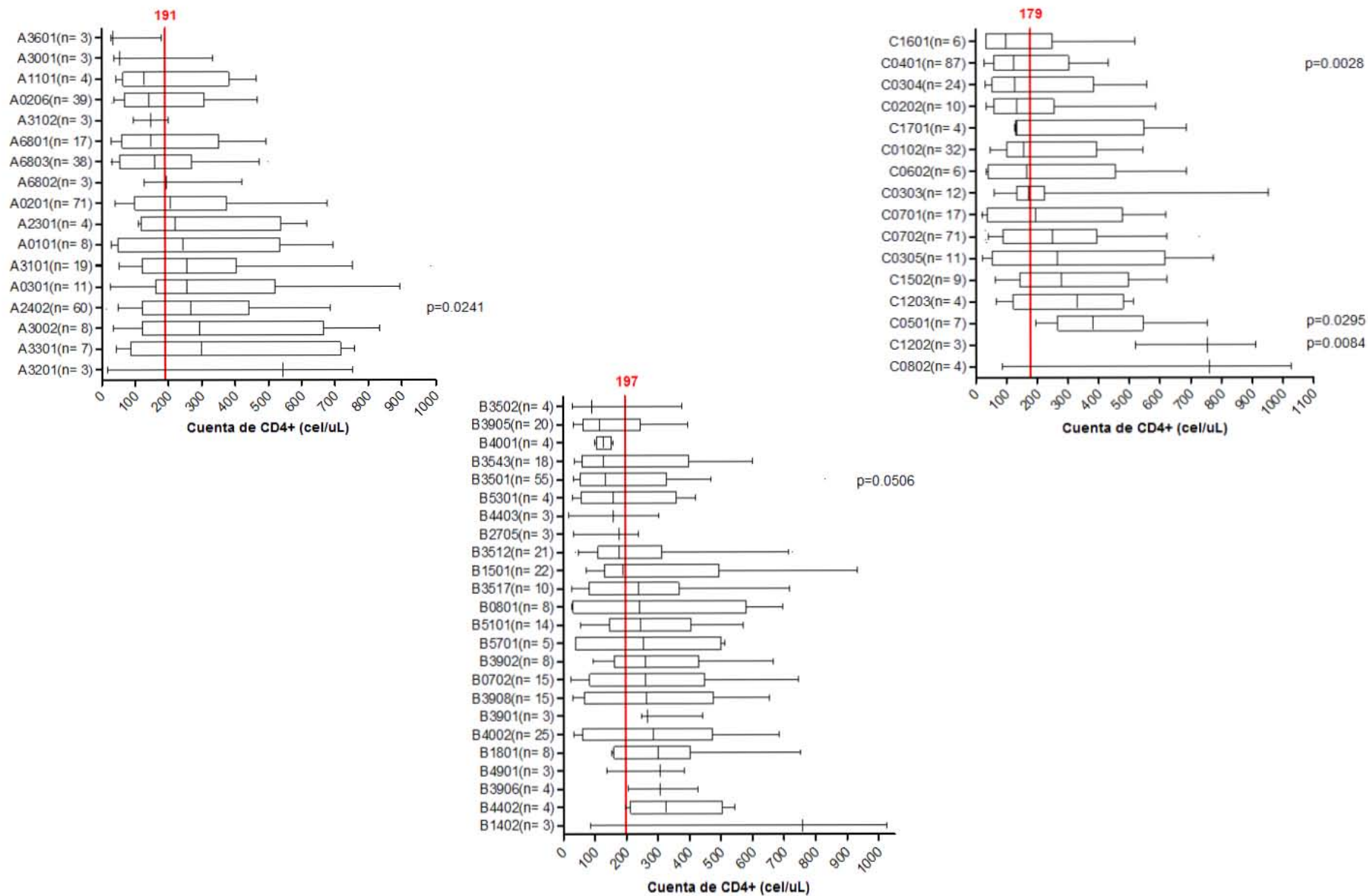
**Figura 31 Correlación de parámetros clínicos y capacidad replicativa viral.** A la izquierda se muestra la correlación negativa entre la cuenta de células T CD4<sup>+</sup> y la capacidad replicativa viral. A la derecha la gráfica representa la correlación positiva entre la carga viral plasmática expresada en Log y la capacidad replicativa viral. Ambas asociaciones se evaluaron por una  $r$  de Spearman y para fines de visualización se añadió la línea de tendencia en negro.

Para explorar la posible asociación entre la presencia de un alelo HLA-I específico y la CV, cuenta de CD4<sup>+</sup> o CRV, se utilizó la prueba de U de Mann-Whitney y se diseñaron los diagramas de cajas (boxplots) que se describen a continuación. Respecto a la primera asociación, los alelos B\*18:01 ( $p= 0.0355$ ), C\*05:01 ( $p=0.0276$ ) y C\*12:02 ( $p=0.0462$ ) mostraron relación a altos valores de carga viral (Figura 32). Por otro lado, los alelos A\*24:02 ( $p=0.0241$ ), C0501 ( $p=0.0295$ ) y C\*12:02 ( $p=0.0084$ ) fueron asociados a altas cuentas de células CD4<sup>+</sup>; mientras que B\*35:01 ( $p= 0.0506$ ) y C\*04:01 ( $p=0.0028$ ) se asociaron a bajas cuentas (Figura 33). Finalmente, el análisis de asociación con capacidad replicativa viral arrojó que los alelos A\*31:01 ( $p=0.0507$ ), A\*68:01 ( $p=0.0164$ ), B\*49:01 (0.0291), C\*04:01 ( $p=0.016$ ) y C\*05:01 ( $p=0.0285$ ) están relacionados con altos valores de CRV (Figura 34).

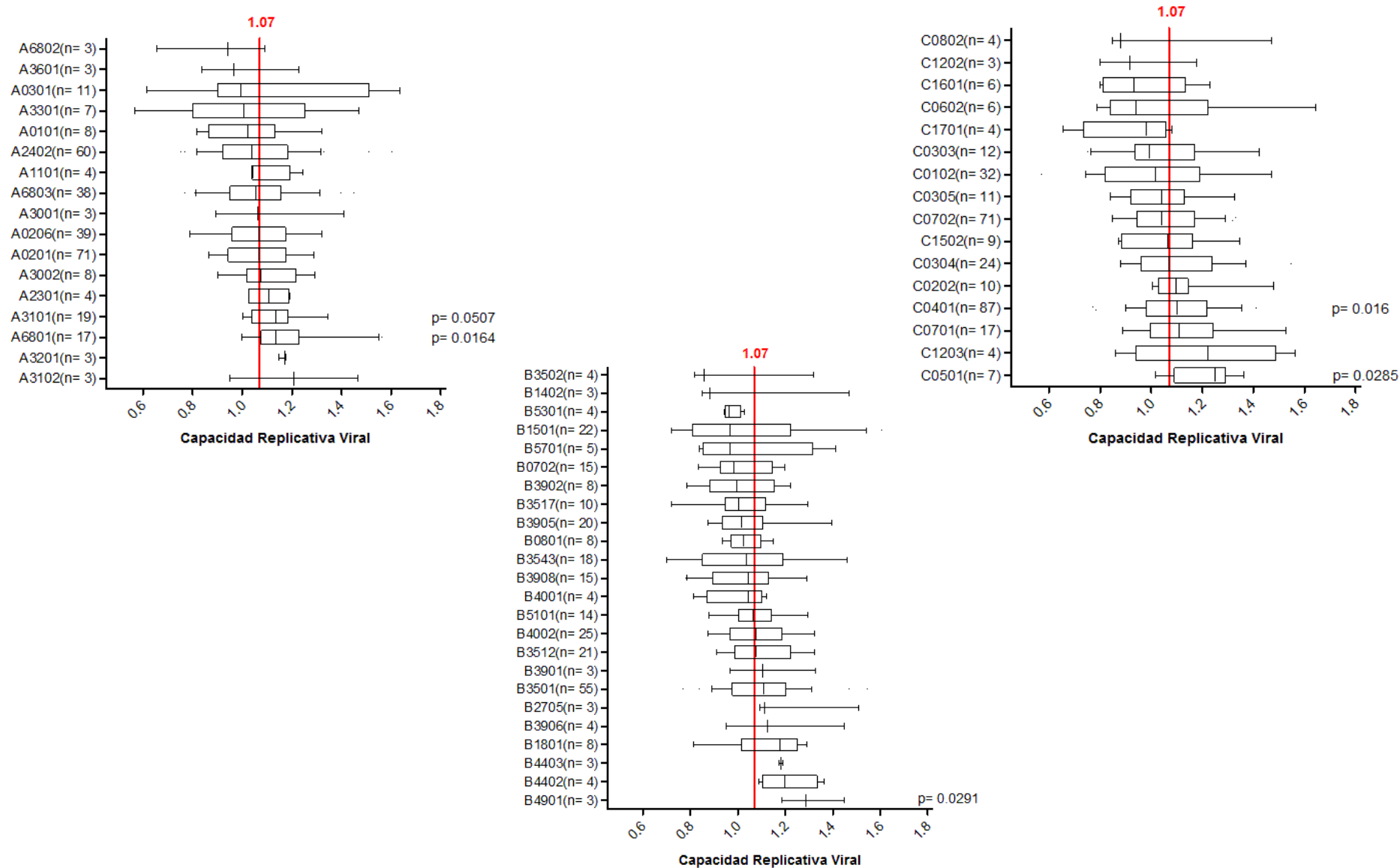




**Figura 32 Asociación entre Carga Viral y alelos HLA-I.** Se muestran gráficas de asociación entre carga viral (LogCV) y HLA-A (arriba/izq), HLA-B (abajo) y HLA-C (arriba/der). La mediana de Log CV de toda la cohorte se indica con una línea roja. Se consideró significativa una  $p \leq 0.05$ .



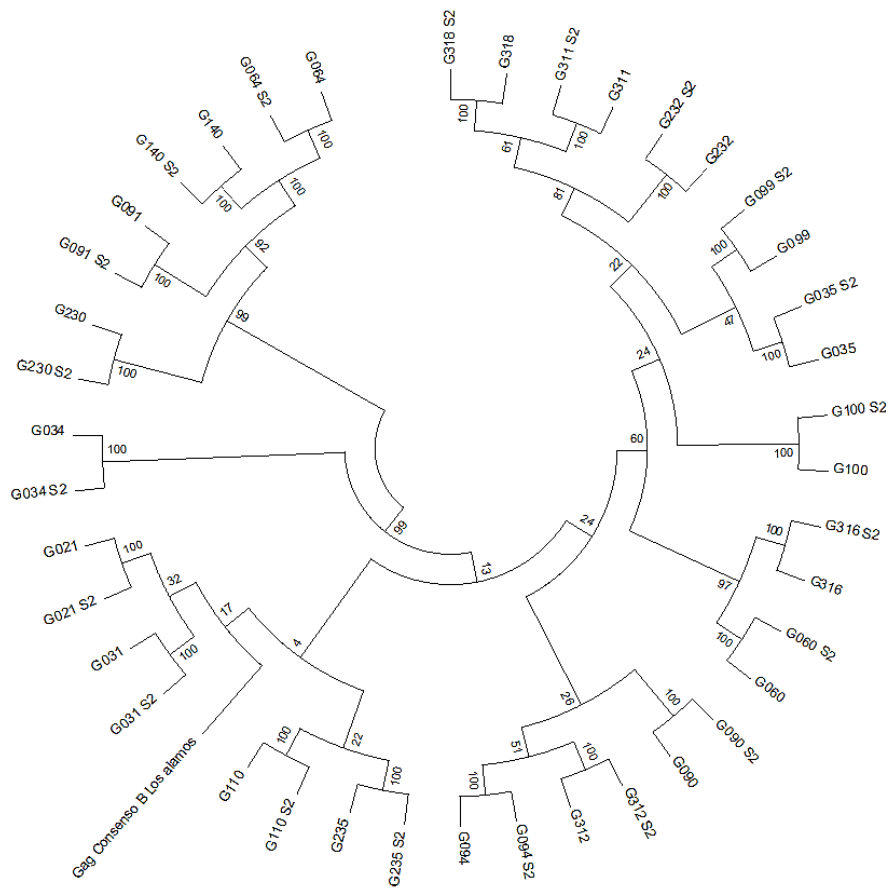
**Figura 33 Asociación entre Cuenta de Linfocitos T-CD4+ y HLA-I.** Se muestran gráficas de asociación entre cuenta de linfocitos T-CD4+ (cel/uL) y HLA-A (arriba/izq), HLA-B (abajo) y HLA-C (arriba/der). La mediana de cel/uL de toda la cohorte se indica con una línea roja. Se consideró significativa una  $p \leq 0.05$ .



**Figura 34 Asociación entre Capacidad Replicativa Viral (CRV) y alelos HLA-I.** Se muestran gráficas de asociación entre capacidad replicativa viral (CRV) y HLA-A (arriba/izq), HLA-B (abajo) y HLA-C (arriba/der). La mediana de CRV de toda la cohorte se indica con una línea roja. Se consideró significativa una  $p \leq 0.05$

## 5.4 ANÁLISIS MOLECULAR DE GAG-VIH-1: VARIABILIDAD Y PRESIÓN SELECTIVA

Se realizó el análisis de las secuencias del fragmento gag-pro obtenidas a partir del grupo de estudio. Para validar que las observaciones realizadas a partir del ensayo *in vitro* de capacidad replicativa fueran reflejo de la selección inmune de mutaciones en *Gag-pro in vivo* y no de mutaciones producidas por el ensayo, se construyó un árbol filogenético con 20 pares de secuencias obtenidas antes y después del ensayo y seleccionadas de forma aleatoria. Como resultado, se observó identidad en todas las secuencias (Figura 35).

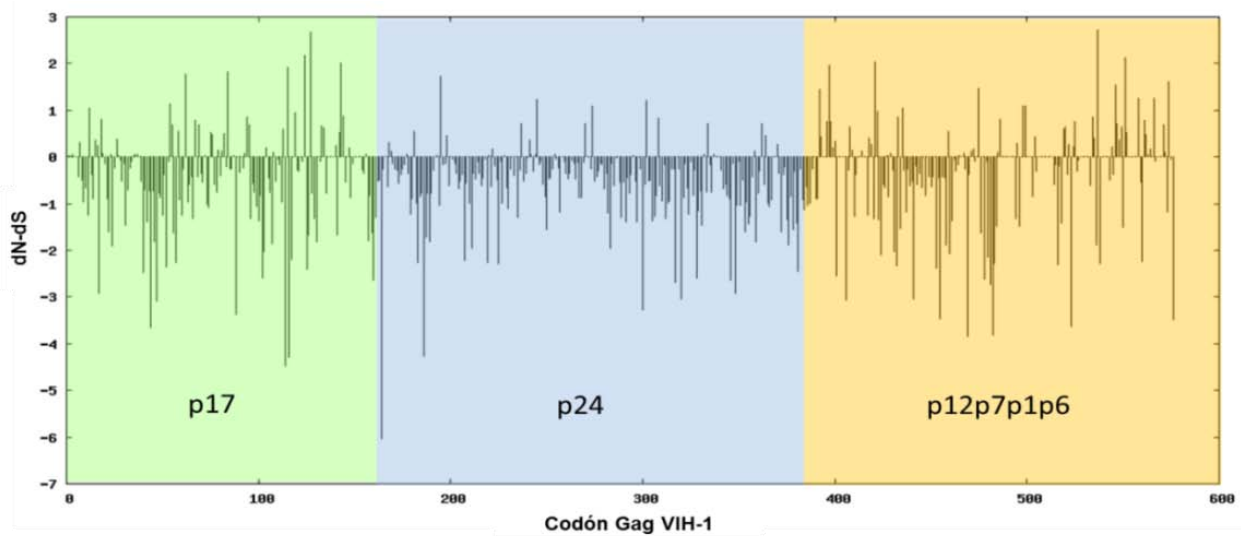


**Figura 35 Comparación filogenética de secuencias de Gag antes y después del ensayo de capacidad replicativa.** Se construyó un árbol filogenético con el método Maximum Likelihood y el modelo General Time Reversible para calcular la distancia evolutiva. La significancia fue determinada mediante 1000 réplicas de bootstrap. Se utilizaron 20 pares de secuencias elegidas al azar y obtenidas antes y después del ensayo de capacidad replicativa.

Por otro lado, se estimó la variabilidad protéica lo largo del fragmento *gag-pro* en las secuencias basales (antes del ensayo de CRV). De los 160 pacientes, se obtuvieron 149

secuencias completas de Gag, 8 secuencias incompletas pero analizables y 3 no pudieron obtenerse.

Para obtener la cantidad de mutaciones no sinónimas (dN) y sinónimas (dS) y el cociente de ambos (dN/dS) en cada codón del fragmento dado, se usó la herramienta en línea Data Monkey. El cociente dN-dS se graficó hacia valores positivos en los codones variables y hacia valores negativos en los codones conservados. Visualmente, se identificó que el dominio p24 fue el más conservado, mientras que la región terminal (p2p7p1p6) mostró alta variabilidad (Figura 36).



**Figura 36 Gráfica de variabilidad en la secuencia de Gag.** Se usó la herramienta en línea Data Monkey para estimar y graficar la razón de mutaciones no sinónimas (dN) y sinónimas (dS). Se muestra tres regiones de Gag, los codones variables (líneas hacia arriba) y las posiciones conservadas (líneas hacia abajo).

Mediante el uso de Data Monkey, también se obtuvo una lista de posiciones a lo largo de Gag donde la cantidad de sustituciones no sinónimas fue significativamente mayor a la de sustituciones sinónimas ( $dN > dS$ ) ( $p < 0.01$ ), es decir, codones sobre los que se ejerce presión de selección positiva que favorece su variación (Tabla 2). Se identificaron 6 codones bajo presión selectiva en p17 (CA, verde), 3 en p24 (MA, azul) y 9 en la región terminal de Gag (p2p7p1p6, anaranjado).

Codon	dN-dS	dN-dS normalizado	Valor p
12	6.11383	1.04602	0.000615
18	4.78268	0.818274	0.0083182
54	6.62826	1.13404	0.0007665
62	10.3354	1.76829	2.10E-05
84	10.6133	1.81584	0.0001857
115	5.58331	0.955255	0.0021027
173	10.1581	1.73797	9.48E-08
223	7.19736	1.2314	0.0016791
280	7.01983	1.20103	0.0088907
370	8.46817	1.44883	0.000626
373	11.4012	1.95064	6.71E-05
389	11.9215	2.03967	1.36E-05
390	5.6193	0.961412	0.0037275
441	8.58296	1.46847	0.0004221
451	4.69962	0.804063	0.0075357
478	8.9732	1.53523	0.0014499
490	7.36865	1.26071	0.0009191
498	9.35048	1.59978	0.0014134

**Tabla 2 Posiciones bajo presión selectiva en Gag.** Se analizó el alineamiento de secuencias de Gag obtenidas a partir de virus de 160 pacientes guatemaltecos. Se encontraron 18 posiciones bajo presión selectiva ( $p < 0.01$ ), de las cuales 6 corresponden a p17 (verde), 3 a p24 (azul) y 9 a la porción terminal, p2p7p1p6 (anaranjado).

Posteriormente, se identificaron las mutaciones presentes en cada una de las posiciones bajo presión selectiva. Para ello, se comparó cada secuencia viral con la secuencia consenso del grupo de estudio y se calculó la prevalencia de cada variación (%). Para este análisis se tomaron en cuenta incluso las secuencias incompletas y se compararon las porciones con las que se contaba.

Dentro de las posiciones variables de p17, los codones E12 y G62 presentaron la mayor cantidad de variantes (Tabla 3). Sin embargo, en ambas posiciones se identificaron 4 mutaciones presentes solamente en 1 paciente (0.7%) cada una. Dentro de las tres posiciones bajo presión selectiva de p24 (Tabla 3), las variantes más frecuentes fueron S173T (41.5%), I223V (32.7%) y T280V (34%). En la porción terminal se concentró el mayor número de codones bajo presión selectiva (Tabla 4). La variabilidad de cada posición fue amplia, observándose desde 1 variante, por ejemplo en la posición V390, hasta codones con 10 variantes como I389, Y441 y P478. Notablemente, en esta porción se observaron deleciones en los codones V370, P478, R490 y S498; la más frecuente fue esta última, observada en el 30.6% de las secuencias analizadas.

Gag - p17							Gag - p24		
Codón/Posición	12	18	54	62	84	115	173	223	280
Referencia	E	K	S	G	T	A	S	I	T
Prevalencia (%)	WT - 64.6	WT - 82.3	WT - 81.6	WT - 61.2	WT - 63.9	WT - 77.6	WT - 51.0	WT - 59.2	WT - 55.1
VARIANTES - PREVALENCIA (%)	K - 22.4	R - 14.3	A - 9.5	A - 8.8	V - 34.0	T - 8.2	T - 41.5	V - 32.7	V - 34.0
	Q - 6.1	G - 0.7	G - 2.7	E - 8.8	A - 0.7	V - 5.4	A - 1.4	N - 2.0	S - 4.8
	D - 2.7	S - 0.7	P - 2.0	I - 5.4		E - 3.4	V - 1.4	A - 1.4	I - 2.7
	A - 0.7		T - 2.0	R - 4.1		K - 2.0	L - 0.7	L - 1.4	P - 0.7
	G - 0.7		E - 0.7	Q - 3.4		D - 0.7		T - 0.7	T - 0.7
	N - 0.7			T - 2.0					
	R - 0.7			K - 0.7					
				M - 0.7					
				N - 0.7					
				V - 0.7					

**Tabla 3 Mutaciones seleccionadas en Gag-p17 y Gag-p24.** Se identificó la prevalencia de todas las variantes encontradas en cada posición bajo presión selectiva identificada en p17 (izquierda) y p24 (derecha). La prevalencia se muestra en porcentaje y se resaltan en amarillo aquellas posiciones presentes en al menos 3 pacientes de la cohorte (2%). WT; Wild-type.

### Gag - p2p7p1p6

Codón Referencia	370 V	373 S	389 I	390 V	441 Y	451 S	478 P	490 R	498 S
Prevalencia (%)	WT - 63.3	WT - 49.7	WT - 34.7	WT - 66.0	WT - 6.8	WT - 72.8	WT - 54.4	WT - 30.6	WT - 51.0
VARIANTES – PREVALENCIA (%)	L - 12.9	P - 24.5	T - 29.3	I - 32.0	H - 63.3	N - 26.5	Q - 11.6	K - 61.9	Del - 30.6
	A - 11.6	A - 6.8	P - 18.4		N - 15.6	G - 0.7	T - 9.5	Del - 2.7	L - 14.3
	M - 3.4	Q - 5.4	S - 5.4		S - 4.1		L - 4.8	T - 1.4	T - 0.7
	I - 2.7	N - 4.8	A - 1.4		Q - 3.4		S - 4.8	E - 0.7	Y - 0.7
	Del - 2.7	T - 4.1	M - 1.4		C - 1.4		Del - 4.8	S - 0.7	
		I - 0.7	N - 1.4		L - 1.4		K - 4.1		
			Q - 1.4		R - 1.4		A - 1.4		
			V - 1.4		K - 0.7		G - 1.4		
			H - 0.7		P - 0.7		D - 0.7		
			K - 0.7		T - 0.7		V - 0.7		

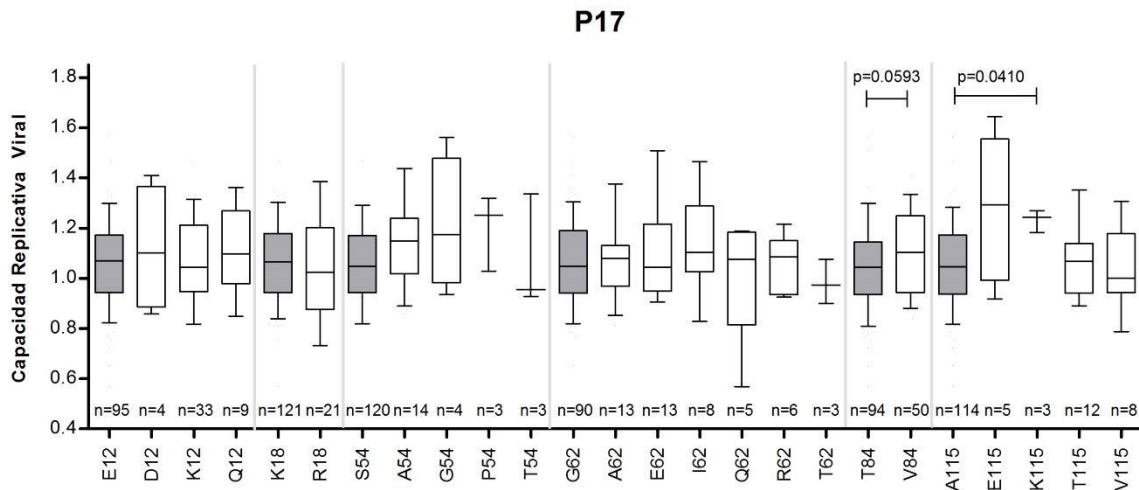
**Tabla 4 Mutaciones seleccionadas en la región terminal de Gag.** Se identificó la prevalencia de todas las variantes encontradas en cada posición bajo presión selectiva identificada en la porción terminal. La prevalencia se muestra en porcentaje y se resaltan en amarillo aquellas posiciones presentes en al menos 3 pacientes de la cohorte (2%). WT; Wild-type. Del; deleción.



## 5.5 IMPACTO DE MUTACIONES EN GAG SOBRE LA CAPACIDAD REPLICATIVA DEL VIH

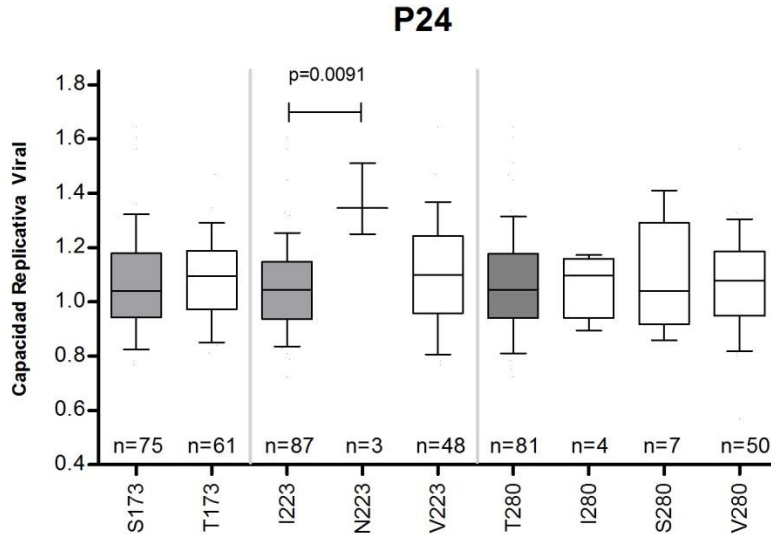
Se realizó el escrutinio de asociaciones entre la presencia/ausencia de una mutación y variación en la capacidad replicativa viral. Para ello se comparó el valor de capacidad replicativa de los virus portadores de cada mutación, presente en al menos 2% de la cohorte, en una posición determinada y el valor de CRV de los virus portadores del aminoácido silvestre (WT) en la misma posición. La significancia de esta comparación se evaluó mediante la prueba estadística U-Mann Whitney y un valor de  $p \leq 0.05$ .

Se analizaron 6 posiciones de la región p17 y se determinó que la presencia de las mutaciones T84V ( $p=0.0593$ ) y A115K ( $p=0.0410$ ) se asocia con un aumento de la capacidad replicativa viral (Figura 37).

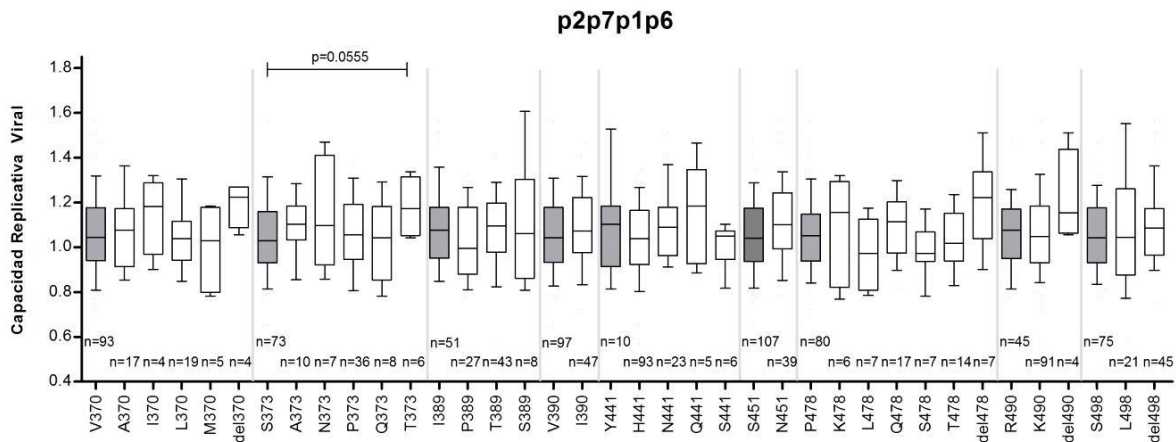


**Figura 37 Efecto de la presencia de mutaciones en p17 sobre la capacidad replicativa viral.** Se graficó la capacidad replicativa viral de cada virus con el aminoácido WT o una mutante determinada. Se compararon las medianas de cada grupo de virus portadores de mutación contra la mediana de CRV de los virus WT mediante U-Mann Whitney. Se resalta en gris el aminoácido WT para cada posición. Se consideró significativo un valor de  $p \leq 0.05$ .

Por otro lado, el análisis de los codones variables de p24 arrojó una asociación de la mutación I223N con valores mayores de CRV ( $p=0.0091$ ), así como la mutación S373T ( $p=0.055$ ) de la porción terminal (p1p2p7p6). Estas diferencias se muestran en las figuras 38 y 39, respectivamente.



**Figura 38 Efecto de la presencia de mutaciones en p24 sobre la capacidad replicativa viral.** Se graficó la capacidad replicativa viral de cada virus con el aminoácido WT o una mutante determinada. Se compararon las medianas de cada grupo de virus portadores de mutación contra la mediana de CRV de los virus WT mediante U-Mann Whitney. Se resalta en gris el aminoácido WT para cada posición. Se consideró significativo un valor de  $p \leq 0.05$ .

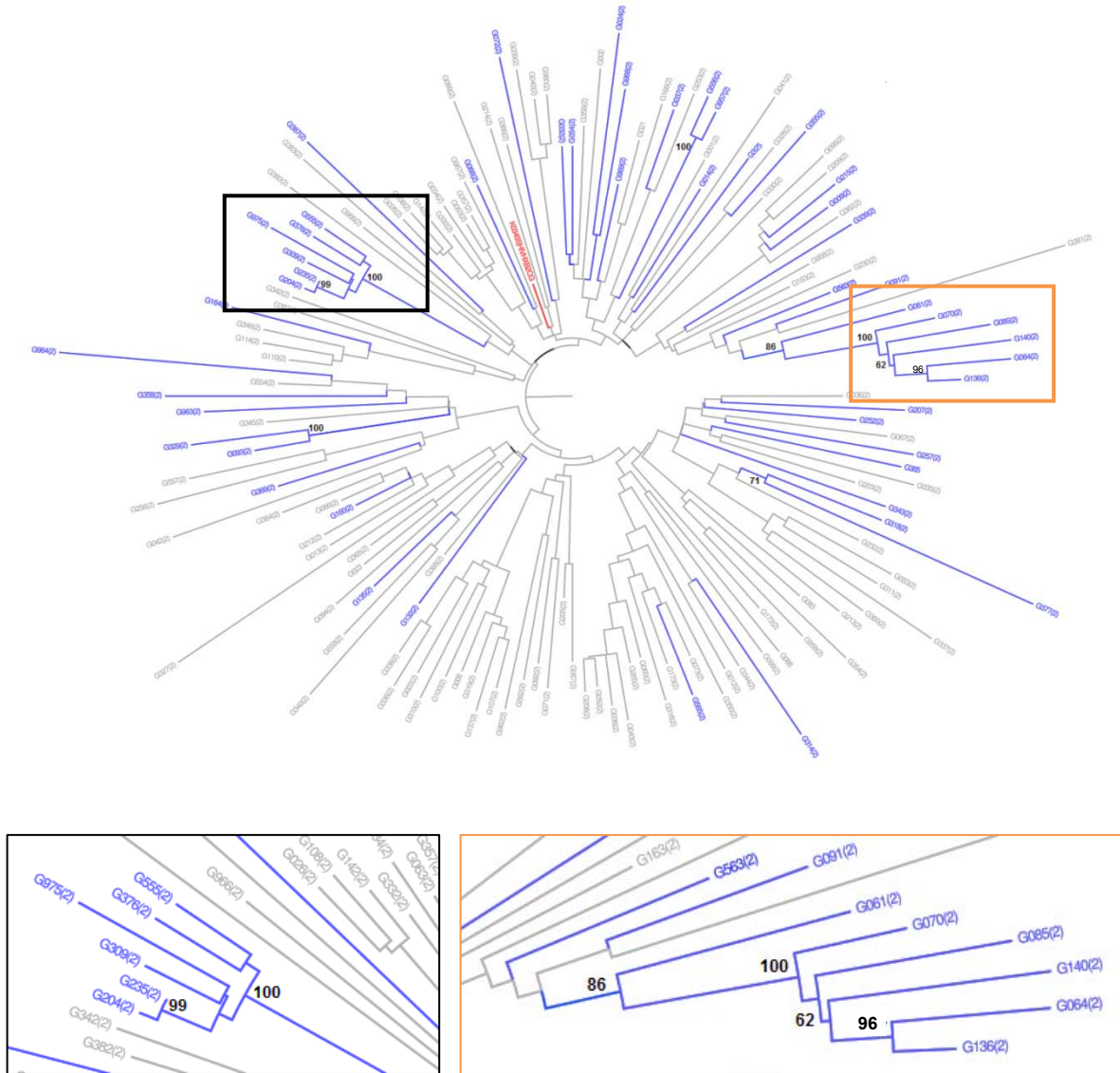


**Figura 39 Efecto de la presencia de mutaciones en p2p7p1p6 sobre la capacidad replicativa viral.** Se graficó la capacidad replicativa viral de cada virus con el aminoácido WT o una mutante determinada. Se compararon las medianas de cada grupo de virus portadores de mutación contra la mediana de CRV de los virus WT mediante U-Mann Whitney. Se resalta en gris el aminoácido WT para cada posición. Se consideró significativo un valor de  $p \leq 0.05$ .

Las mutaciones que fueron relacionadas con variación de la CRV, fueron analizadas filogenéticamente para buscar posibles patrones de transmisión en la cohorte estudiada. Para ello, se construyeron árboles filogenéticos con el modelo Maximum-Likelihood y 1000 réplicas de bootstrap.

De las mutaciones evaluadas, la sustitución T84V fue la más prevalente (34%). Se pudieron observar numerosos pares y grupos de transmisión con altos valores de bootstrap, siendo relevantes dos de ellos resaltados en la Figura 40, por estar

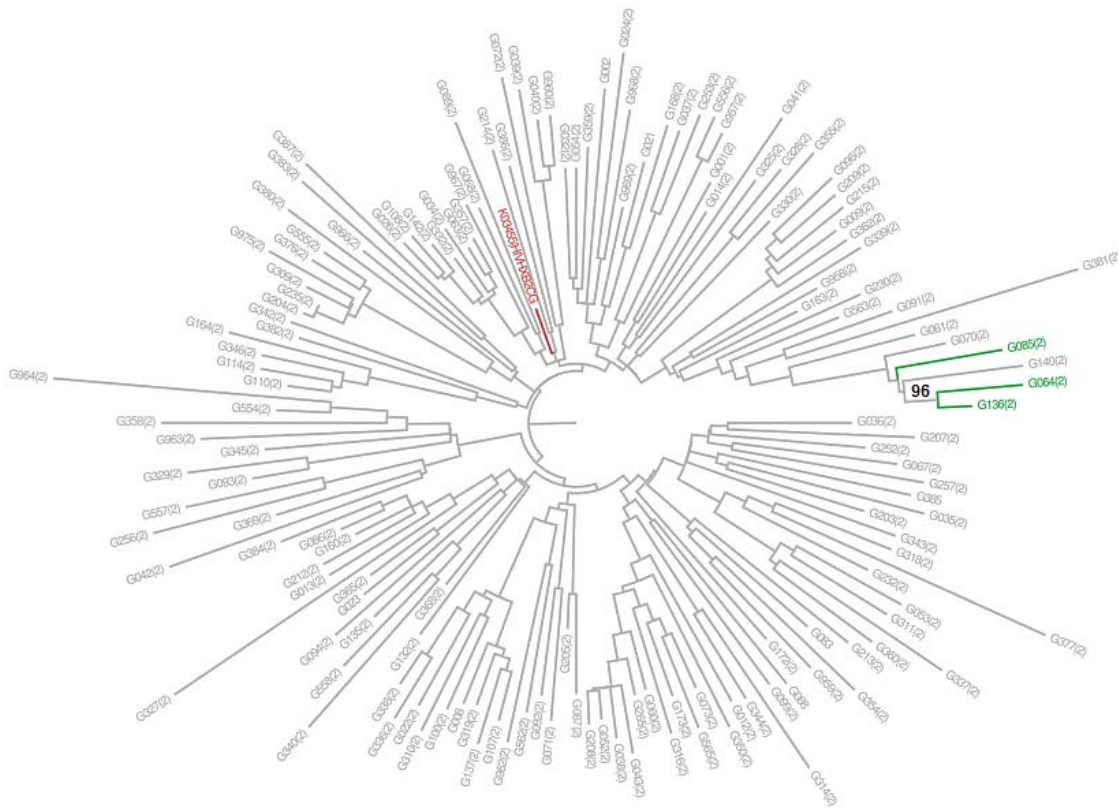
compuestos de al menos 5 secuencias cada uno. Se revisaron las características inmunogenéticas de ambos grupos de transmisión y se observó que el 64.3% de los individuos involucrados poseen al menos una copia del alelo C0401 y algún alelo del grupo B35 (B\*35:01, B\*35:12, B\*35:17, B\*35:43).



**Figura 40 Análisis filogenético de la mutación T84V.** Se construyó un árbol filogenético con el método Maximum Likelihood y el modelo General Time Reversible para calcular la distancia evolutiva. La significancia fue determinada mediante 1000 réplicas de bootstrap. Se resaltan en azul las secuencias portadoras de la mutación T84V (n=50). Se observaron dos importantes clusters de transmisión (abajo) con valores de bootstrap elevados. La secuencia de referencia del virus HXB2 se resalta en rojo.

La mutación A115K fue identificada únicamente en 3 secuencias, mismas que también mostraron la mutación T84V. Los tres individuos de donde provinieron las secuencias antes mencionadas, poseen el haplotipo HLA B\*35:01-C\*04:01.

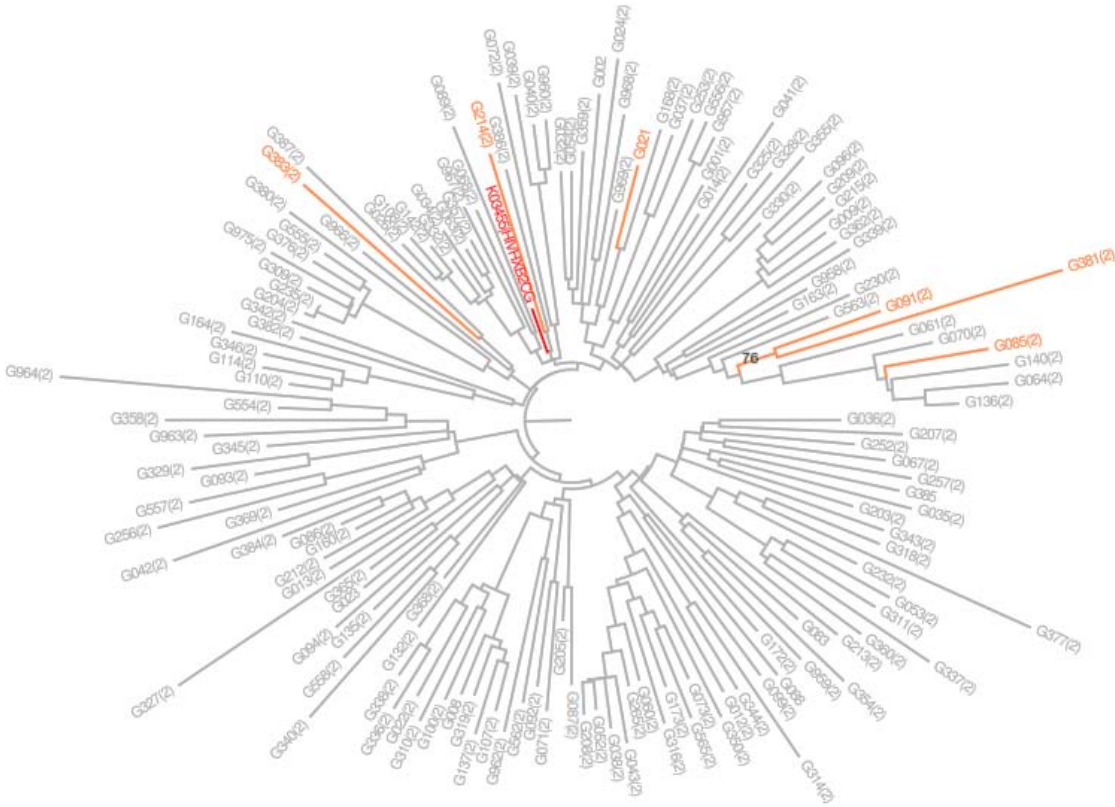
Dos de los virus portadores de dicho cambio mostraron relación filogenética con un bootstrap de 96 (Figura 41).



**Figura 41 Análisis filogenético de la mutación A115K.** Se construyó un árbol filogenético con el método Maximum Likelihood y el modelo General Time Reversible para calcular la distancia evolutiva. La significancia fue determinada mediante 1000 réplicas de bootstrap. Únicamente se observaron 3 secuencias portadoras de la mutación A115K (verde), dos de las cuales formaron un par de transmisión con un valor de bootstrap de 96. La secuencia de referencia del virus HXB2 se resalta en rojo.

El análisis filogenético de la mutación I223N no arrojó ningún par o grupo de secuencias asociadas filogenéticamente.

Finalmente, al construir el árbol filogenético de la mutación S373T se pudo observar que las secuencias G091 y G038 comparten un ancestro común y podrían constituir un par de transmisión de la mutación antes mencionada (Figura 42). En la secuencia G091 también se observó la mutación T84V. No hubo coincidencia en el haplotipo HLA de los individuos involucrados en este par de transmisión.



**Figura 42 Análisis filogenético de la mutación S373T.** Se construyó un árbol filogenético con el método Maximum Likelihood y el modelo General Time Reversible para calcular la distancia evolutiva. La significancia fue determinada mediante 1000 réplicas de bootstrap. Seis secuencias resaltadas en anaranjado fueron portadoras de la mutación S373T. Únicamente se observó un par de transmisión con valor de bootstrap de 76. La secuencia de referencia del virus HXB2 se resalta en rojo.

## 5.6 EFECTO DEL HLA EN LA SELECCIÓN DE MUTACIONES Y ALTERACIÓN DE LA CAPACIDAD REPLICATIVA VIRAL

Una vez que las mutaciones con impacto en la CRV fueron identificadas, se realizó la búsqueda de los epítomos en Gag que las incluyeran y fueran i) reconocidos *in vitro* por alelos asociados a cambios en CRV, CV o cuenta de CD4+, ii) predichos a partir de los motivos teóricos de reconocimiento de alelos que impactan la CRV, CV o cuenta de CD4+, iii) reconocidos, teórica o experimentalmente, por alelos significativamente frecuentes en la población guatemalteca sin un efecto directo en los parámetros de progresión de la infección.

La identificación de los epítomos probados experimentalmente y reconocidos por alelos HLA importantes se realizó mediante el uso de la base de datos inmunológica de Los

Alamos (<http://www.hiv.lanl.gov>). La predicción de epítomos reconocidos por algunos alelos HLA de interés, se realizó mediante la herramienta MotifScan (<http://www.hiv.lanl.gov>). Esta herramienta, ubica a lo largo de una secuencia proteica los aminoácidos de anclaje a una molécula HLA dada y predice el epítomo que sería reconocido por dicha molécula.

Los resultados de esta búsqueda fueron representados en un mapa de reconocimiento de epítomos de Gag-VIH-1 (Figura 43). Para la construcción de dicho mapa se obtuvo una secuencia consenso a partir de las 157 secuencias virales de Gag-pro con que se contaba, usando la herramienta "Consensus Maker" (<http://www.hiv.lanl.gov>). Se representaron en mayúsculas aquellos aminoácidos presentes en, al menos, el 90% de las secuencias analizadas (aminoácidos unánimes) y en minúsculas, los aminoácidos presentes en menos del 90% de las secuencias analizadas, pero mas del 60% (aminoácidos mayoritarios). En la posición 473 ningún aminoácido de los encontrados entre las 157 secuencias rebasó el umbral de mayoría, por lo que se representó con un signo de interrogación (?). Aquellos aminoácidos, mayoritarios o unánimes, que fueran totalmente distintos a la secuencia de referencia del virus HXB2 fueron subrayados en rojo.

Sobre la secuencia consenso, se señalaron las 18 posiciones bajo presión selectiva con un sombreado gris, a excepción de los 4 aminoácidos cuya mutación se vio asociada a elevación de CRV, sombreadas en azul. El cambio específico que se relacionó a CRV alta se muestra en rojo, debajo del aminoácido WT correspondiente.

A lo largo de la secuencia proteica de Gag, se señalaron los epítomos reconocidos por alelos de interés con una flecha y el alelo correspondiente en la parte superior. Solo se fueron representados aquellos epítomos que incluyeran una posición bajo presión selectiva. Se utilizó un color por alelo para facilitar su ubicación.

El mapa (Figura 43) comprende epítomos teóricamente reconocidos (predichos por Motif Scan) por: A\*31:01, A\*68:01 y C\*04:01 que impactan la CRV; por B\*18:01, asociado a disminución de la CV y por B\*35:01 y C\*04:01, asociados a disminución de la cuenta de células T-CD4+. Adicionalmente, aquellos epítomos que ya han sido probados experimentalmente y que fueron reportados en la base de datos de Los Alamos están

señalados en el mapa con un asterisco. Por otro lado, se mapearon los epítomos reconocidos por los alelos más frecuentes de la cohorte: A\*02:01, B\*35:01, B\*40:02, C\*04:01 y C\*07:02.

Se identificó que el alelo A\*68:01 reconoce el epítomo AK9 (Gag 83-91) que en la posición 2 tiene la mutante asociada a alta CRV, T84V (anaranjado); es decir, reconoce la secuencia AVLYCVHQQ, donde la primera Valina ha sustituido al aminoácido silvestre Treonina. El alelo A\*31:01 reconoce el epítomo AVLYCVHQR (café) que posee una modificación en el último aminoácido en donde la R ha sustituido a la K silvestre. El resto de los epítomos representados en el mapa corresponden a la secuencia consenso que señalan.

Un número importante de epítomos se observó reconocido por los alelos B\*35:01 (verde) y C\*04:01 (rosa). El alelo frecuente C\*07:02 fue asociado al reconocimiento de 9 epítomos distintos. Este alelo aparece constantemente en desequilibrio de ligamiento con B\*39:05, por lo que el efecto sobre el VIH podría ser realmente por el HLA-B. Sin embargo, no se hallaron motivos de reconocimiento para B\*39:05, pues la herramienta utilizada no contiene referencias suficientes de alelos típicos de poblaciones latinas – mestizas.

En gris se muestran los epítomos reconocidos por excelencia por los alelos canónicos de protección B\*27 y B\*57. Ninguno de dichos epítomos incluyó alguna posición bajo presión de selección. Es importante remarcar que la frecuencia de ambos alelos, en la cohorte de estudio, es muy baja <1%.

Posteriormente, se realizó un análisis exploratorio para definir si el aminoácido mutante aparece sobre-representado en un grupo de pacientes enriquecido con el HLA selector, lo cual sugeriría que un alelo HLA específico selecciona significativamente una mutación determinada. Dicho análisis se realizó con la herramienta HAPLOID, desarrollada por el grupo del Dr. Mark Broockman y la Dra. Zabrinna Brumme de la universidad Simon Fraser en British Columbia, Canadá. La cual evalúa, mediante exactas de Fisher con  $p \leq 0.05$ , los siguientes parámetros:

- a) La presencia o ausencia de un alelo HLA determinado

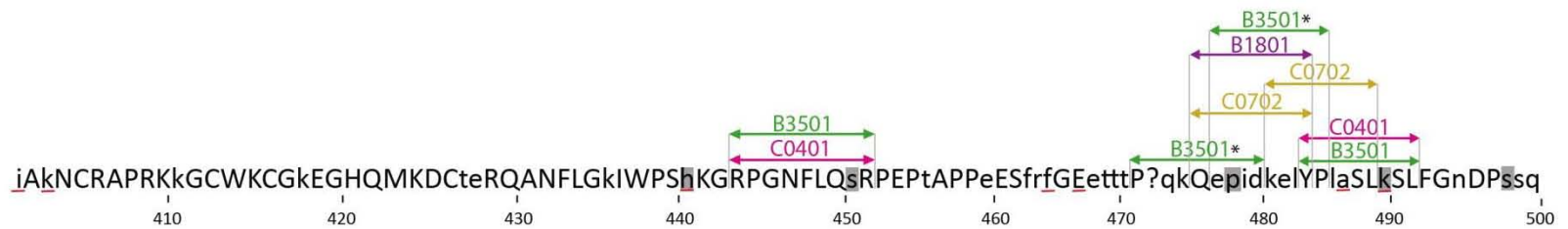
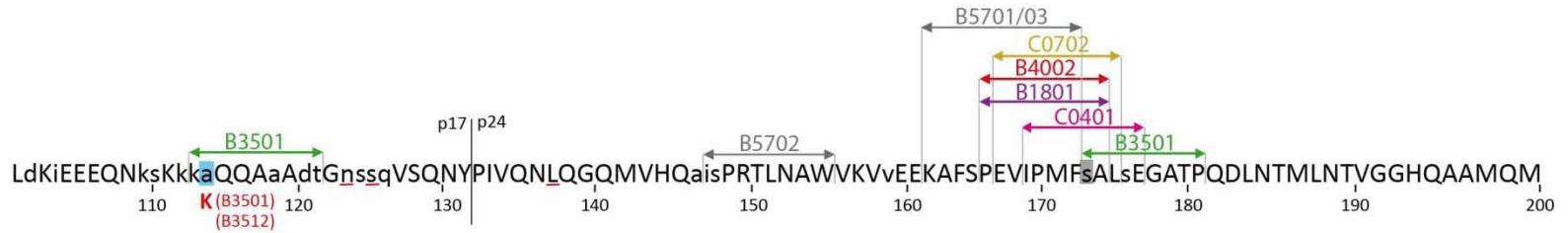
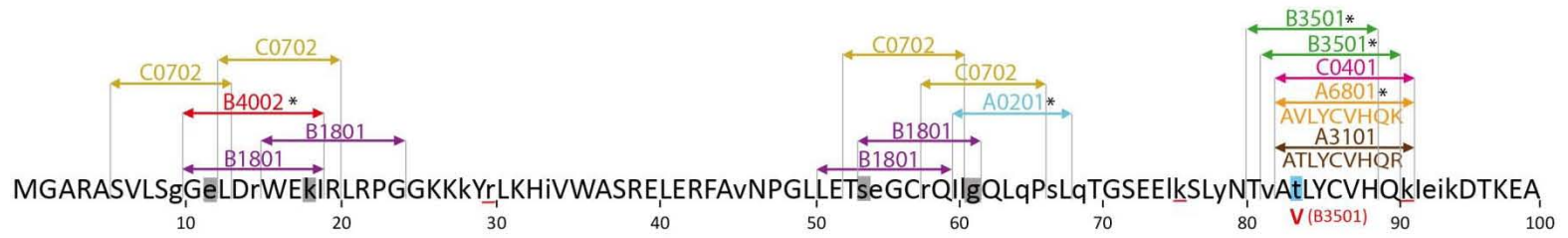
- b) La presencia o ausencia de un cambio, respecto a la referencia, en cada posición del fragmento proteico de interés
- c) La presencia de ambos parámetros simultáneamente, es decir, un alelo HLA específico y una mutación determinada.
- d) La ausencia de ambos parámetros.

Los resultados sugieren que la mutación T84V se ve favorecida significativamente cuando está presente el alelo B\*35:01 ( $p=0.034$ ), al igual que la mutación A115K ( $p=0.022$ ). En la posición 115, la mutación antes mencionada también se ve favorecida significativamente por el alelo B\*35:12 ( $p=0.042$ ). Este alelo ha sido asociado consistentemente a mayor progresión de la infección en diversos estudios de nuestro laboratorio sobre la población guatemalteca [Valenzuela P., comunicación personal], pero no se conocen aún los motivos de reconocimiento del alelo ni los probables epítomos que se pueden unir a él.

Los resultados antes descritos se indican entre paréntesis rojos en la Figura 43. A pesar de que los valores de  $p$  fueron significativos en las asociaciones mencionadas, los datos obtenidos con la herramienta Haploid no fueron corregidos por múltiples comparaciones, dado el número de secuencias disponibles y el carácter preliminar de este análisis.

**Figura 43 Mapa de presión selectiva de HLA sobre Gag de VIH-1.** Se muestra la secuencia proteica consenso de Gag circulante en la población guatemalteca. Los aminoácidos en mayúsculas son unánimes (presentes en >90% de las secuencias) y los aminoácidos en minúsculas representan posiciones mayoritarias (<90%, >60%). Los sitios que fueron completamente distintos a la referencia de Gag HXB2 se subrayan en rojo. Las posiciones bajo presión selectiva se indican con una sombra gris, a excepción de las cuatro posiciones en donde se reportaron mutaciones asociadas a aumento de CRV, sombreadas en azul. Los epítomos restringidos por un alelo determinado se limitan con flechas y en la parte superior se indica el alelo selector. Los epítomos que ya han sido probados experimentalmente se señalan con un asterisco. Las mutaciones asociadas a cambios en la CRV se muestran en rojo, debajo del aminoácido silvestre correspondiente. Los alelos que se asocian a la aparición de dicha mutación se señalan entre paréntesis rojos.





S

## **CAPITULO 6. DISCUSIÓN**

Desde su inicio, la epidemia del Virus de Inmunodeficiencia Humana ha sido ampliamente estudiada por su rápida diseminación en todo el mundo y la alta tasa de mortalidad que presentó en un principio. Actualmente, se contabilizan alrededor de 35 millones de personas infectadas con el virus en todo el mundo y a pesar de que el tratamiento antirretroviral ha incrementado la expectativa de vida de quienes lo reciben, las nuevas infecciones no han podido ser detenidas.

Una de las dificultades más significativas para el diseño de estrategias de erradicación y vacunación contra el VIH, es la alta tasa de mutación que posee y la elevada tolerancia de variabilidad en su genoma. Se ha demostrado, que el virus es capaz de adaptarse con rapidez a cambios en el ambiente donde se replica, especialmente, a la presión selectiva ejercida naturalmente por la respuesta inmune del hospedero, o a la presión artificial del tratamiento antirretroviral.

A lo largo de la infección se establece una dinámica entre el hospedero y el virus, porque la respuesta inmune es adaptable y el virus variable a lo largo del tiempo dentro del mismo individuo. Uno de los ámbitos en los que esta dinámica es relevante es la respuesta inmune mediada por los linfocitos T CD8+. Estas células eliminan linfocitos infectados o cualquier otra célula que presenten en su membrana péptidos virales. La respuesta inducida es específica y está restringida por las moléculas HLA de clase I que están codificadas en la región más polimórfica del genoma humano.

La coincidencia temporal de la respuesta citotóxica y la disminución de la viremia durante la etapa aguda de la infección, sugieren un papel trascendental de los linfocitos T CD8+ en la progresión de la enfermedad. Específicamente, el haplotipo HLA de clase I ha sido repetidamente predictivo del curso de la infección en diversos estudios.

La caracterización de la respuesta citotóxica ha ido en aumento desde principios de la década de los 90s, por ser de gran interés entender los mecanismos responsables del control natural de la infección en algunos pacientes. La investigación en esta área, permitió la identificación de los alelos HLA B\*27 y B\*57, que permiten a ciertos individuos controlar la replicación del virus sin tratamiento y tener una progresión notablemente lenta

de la enfermedad<sup>42,48</sup>. Otro de los hallazgos de la respuesta mediada por HLA fue que la acción citotóxica es más efectiva cuando se dirige a péptidos derivados de Gag. Estudios posteriores permitieron la identificación de mutaciones de escape a CTLs que se seleccionan de forma reproducible en un contexto específico de restricción de HLA. Esto quiere decir que “diferentes moléculas HLA-I seleccionan distintos epítopes de VIH para estimular la respuesta citotóxica. Por lo tanto, el tipo de moléculas HLA puede tener un efecto en el escape viral”<sup>58</sup>.

Si bien es cierto que la enorme variabilidad de las moléculas HLA impacta en la incidencia y la supervivencia de una gran variedad de enfermedades infecciosas, condiciones inflamatorias y problemas autoinmunes, dicha diversidad aumenta la complejidad de los análisis de interacción hospedero-virus.

A pesar de ello, se ha logrado estudiar la presión inmune que se ejerce sobre la población viral de un individuo. Se sabe que después de la transmisión hay un cuello de botella que explica que el 80% de las transmisiones heterosexuales se den por una única sepa, a partir de la cual se generan múltiples y variadas cuasiespecies, todas relacionadas entre sí y que son sujetas a la presión selectiva del sistema inmune de forma variable a lo largo de la infección.

Dentro de los primeros 21 días de infección, la respuesta inmune selecciona mutaciones de escape rápidamente. Al parecer, esas mutaciones solo afectan discretamente la capacidad de reconocimiento del péptido por parte del HLA. Posteriormente, la respuesta de CTLs selecciona una segunda oleada de mutantes de escape, generalmente en los residuos de anclaje al HLA, que son más efectivas para evadir el sistema inmune. Con el paso del tiempo, la secuencia viral consenso dentro de un individuo, se moldea en función de la selección inmune. Brumme y cols. determinó que cerca del 30% de los cambios en el consenso de *gag-pol* y 60% en *nef*, dentro del mismo individuo, son atribuibles a la presión del HLA<sup>54</sup>.

Si la presencia de ciertos alelos puede determinar la velocidad de progresión a SIDA en un individuo, es lógico pensar que ese fenómeno puede ser extrapolado a una población en donde existen alelos determinados con altas prevalencias. En este sentido, existe la hipótesis del efecto “huella digital del HLA”, que indica que la secuencia del VIH circulante

actualmente refleja la adaptación del virus a los alelos más comúnmente expresados en una población determinada a lo largo del tiempo<sup>59</sup>.

El costo en la capacidad de replicación, la presencia de mutaciones compensatorias y un complejo arreglo de otras presiones selectivas del hospedero y el virus, determinan las mutaciones que alcanzaran un nivel suficiente para ser apreciadas a nivel poblacional y que pueden seleccionarse, transmitirse o revertir en función de los alelos HLA presentes. Es decir, la especificidad y reproducibilidad del escape de CTLs significa que la frecuencia de un alelo HLA específico en una población, puede determinar la prevalencia de esa mutación.

A través del estudio de la interacción a nivel individual y poblacional entre el HLA y el VIH se pueden identificar mecanismos, vías y dinámicas de escape o control en una población determinada. Tal como se pretendió hacer en el presente estudio, donde se buscó el efecto de los alelos más prevalentes en una población hasta el momento no descrita sobre la infección por VIH.

Inicialmente, la determinación de los alelos más prevalentes entre pacientes infectados con VIH-1 provenientes de Guatemala, permitió establecer el contexto inmunogenético en que los virus, posteriormente estudiados, se encontraban. Los alelos más frecuentes en la cohorte de estudio fueron de origen amerindio, tal como se esperaba por la naturaleza de la población. Sin embargo, se encontraron diversos alelos sobre los que no se tiene suficiente información y que podrían representar una característica notable de la población estudiada. Por ejemplo, los alelos A\*68:03, B\*39:08, B\*35:12 y B\*35:43 no se encuentran representados en poblaciones caucásicas, africanas o japonesas previamente estudiadas. Además, se observó una prevalencia notablemente alta del alelo B\*35:01. En conjunto, los alelos del grupo B\*35 se han relacionado con una mayor progresión a la infección, por lo que el contexto inmunogenético guatemalteco podría estar asociado a una mayor predisposición de seleccionar mutaciones efectivas de escape y por lo tanto de control deficiente de la infección.

Una de las consecuencias más importantes que se ha observado al analizar la huella del HLA, es la selección diferencial de polimorfismos. Éstos pueden ser no adaptados, si el

a.a. mutado esta sub-representado en un grupo de pacientes donde determinado alelo esta enriquecido, lo cual significaría que la forma susceptible de ser reconocida por el sistema inmune es la más abundante. O adaptados, que son a.a. sobre-representados en un grupo de pacientes con determinado HLA, que implicaría que la forma de escape es más frecuente.

El impacto de la presión por HLA en el curso de la infección recae de forma importante, más no exclusiva, sobre la capacidad replicativa viral al seleccionar una mutación determinada. Por ello, se verificó que existiera correlación de CRV con los parámetros típicos de progresión de la infección en el grupo de estudio, pues un virus que replica eficazmente podría elevar el título viral en determinado momento. Nuestros resultados confirman las observaciones hechas previamente en otras cohortes, de que la CRV puede ser asociada con la progresión de la infección, aunque la correlación no sea perfecta, pero si significativa. Se observó que cuando un virus replica con mayor eficacia, definido por valores superiores a la media de CRV calculados en el grupo de estudio, la carga viral tiende a aumentar y las células CD4+ a disminuir. Si un virus replica con mayor rapidez, la aparición del mismo aumentará en plasma y al haber mayor cantidad de virus circulantes, la probabilidad de infección de los linfocitos también aumenta.

Como se mencionó anteriormente, existe evidencia que asocia alelos específicos de HLA-I a un aumento o disminución de la CRV y por lo tanto a un impacto en la rapidez de progresión. Sin embargo, dichas asociaciones deben poseer un impacto notable o ser muy frecuentes en una población para poder ser detectadas. A través de un modelo matemático, Jonathan Carlson de Microsoft Research Group, estimó que la tasa de falsos negativos (asociaciones HLA-CRV existentes no detectadas) es del 80% en una muestra con  $n=100$  y en el caso de los polimorfismos seleccionados por un HLA específico, se puede detectar solo el 50% de las asociaciones en un grupo de  $n=700$ .

Mediante el análisis de 160 pacientes, únicamente se identificaron alelos HLA-I asociados a potencial riesgo de progresión. Se observó que los alelos A\*31:01 y A\*68:01 correlacionan de forma consistente con una mayor capacidad de replicación. Ninguno de estos alelos ha sido catalogado como HLA de riesgo en otras poblaciones, sin embargo,

al ser alelo de origen amerindio no se descarta que en un futuro cercano se cuenten con datos adicionales de estos alelos que aparecen enriquecidos en poblaciones latinas.

Además, se observó el efecto de B\*49:01 sobre la CRV. Sin embargo, no se cuenta con mayor información sobre este alelo; a pesar de ser amerindio, no es un alelo predominante en poblaciones latinoamericanas.

Se sabe que entre los alelos HLA-I, los pertenecientes al locus B ejercen la fuerza selectora más grande sobre el virus. Esto implica que algunas de las asociaciones observadas con HLA-C se puedan dar por efecto indirecto al encontrarse en desequilibrio de ligamiento con alelos HLA-B fuertemente selectores. En este estudio, los alelos C\*04:01 y C\*05:01 correlacionaron con el aumento de CRV. Sin embargo, C\*04:01 presenta fuerte desequilibrio de ligamiento con B\*35:01, mientras que C\*05:01 se encuentra asociado a B\*18:01 y B\*44:02. El desequilibrio de ligamiento se evaluó en una cohorte de 400 pacientes guatemaltecos tipificados a alta resolución de HLA A, B y C (Valenzuela P., datos no publicados) que incluye a los 160 en quienes se evaluó también la capacidad replicativa viral.

Se obtuvieron resultados controversiales para B\*18:01, ya que este alelo se asoció significativamente a baja CV mientras que C\*05:01 se asoció a alta CV, alta CRV y bajas cuenta de T CD4+. El efecto de C\*05:01 en algunos casos podría estar explicado por el desequilibrio de ligamiento con el alelo HLA B\*44:02 que, aunque no alcanzó la significancia estadística, se ubica por encima de la mediana de CRV. Sin embargo, en aquellos pacientes en los que B\*18:01 se encuentra presente, es necesaria una exploración más profunda del mecanismo mediante el cual este alelo impacta en la progresión de la infección.

En estudios de poblaciones africanas infectadas con VIH-1 clado C, B\*18:01 ha sido caracterizado como alelo de riesgo. La divergencia del efecto de dicho alelo puede deberse en gran medida al subtipo viral y a la población. La observación de efectos aparentemente contrarios de un mismo alelo, ya ha sido previamente documentada por Kiepiela<sup>48</sup> y Leslie<sup>60</sup>. Ellos observaron que en una población caucásica donde B\*58:02, que es un alelo de riesgo está ausente, pero B\*57:01 aparece ligado a C\*06:02, C\*06:02 se define como alelo protector (por efecto de B57). Pero en poblaciones africanas, donde

B\*57:01 está ausente y B\*58:02 aparece en desequilibrio de ligamiento con C\*06:02, C\*06:02 es considera un alelo de riesgo.

Ya se han descrito alelos HLA-B y C que se encuentran en desequilibrio de ligamiento. Por ello, numerosos grupos de investigación han atribuido la protección o susceptibilidad de progresión asociada a un alelo HLA-C, a un efecto indirecto de los alelos HLA-B a los que se encuentra “ligado”. Sin embargo, recientemente se han identificado potenciales mecanismos que explican la protección o riesgo asociados a HLA-C independientes del desequilibrio de ligamiento con HLA-B. Se cree que la restricción tan fuerte de la respuesta por HLA esta influenciada por la acción de la proteína viral Nef. Esta proteína disminuye la expresión de HLA-A de forma efectiva evitando que la célula infectada sea detectada por el sistema inmune<sup>61</sup>, pero no tiene la misma eficiencia al inhibir a HLA-B, lo que podría explicar que sea una fuerza selectora mayor. Sin embargo, Nef no inhibe la expresión de HLA-C en la célula infectada, probablemente para disminuir la susceptibilidad a las células NK que responden a la ausencia de HLA en la superficie de una célula y así, mantener la replicación viral. Diversos estudios en la población caucásica han apuntado a que HLA-C impacta en la progresión de la infección por la presencia de un SNP 35 kb río arriba del locus C, que regula la expresión de la molécula en superficie de forma dependiente de un microRNA. Es decir, HLA-C puede afectar la CRV porque al expresarse más o menos en la superficie del linfocito infectado, modula la fuerza de la respuesta citotóxica. Mientras que HLA-B impacta en la progresión a través de la selección de mutaciones en una secuencia viral determinada.

Por otro lado, C\*04:01, asociado a alta CRV, está en desequilibrio de ligamiento con un alelo de riesgo significativamente más frecuente en Guatemala que en otras poblaciones, B\*35:01. Este alelo ha sido ampliamente estudiado por asociarse de forma consistente con una mayor progresión de la infección. Se han propuesto algunos mecanismos que explican el aumento de CV y CRV en pacientes portadores de este alelo. Entre ellos, se sugiere que la “educación tímica” de las CTLs después de la estimulación con HLA B\*35:01 puede jugar un papel importante en la efectividad de las funciones efectoras de las células. Este mecanismo ha sido estudiado para B\*27 y B\*57, en donde la función efectora de los CTLs estimulados con dichos alelos resulta más eficiente que la función de las células estimuladas con otro alelo<sup>62-65</sup>. Otra hipótesis, apunta a que ciertos

clonotipos de TCR tienen una capacidad disminuida de reconocer complejos péptido-B\*35:01 y por lo tanto la eliminación de células infectadas se ve disminuida, favoreciendo la replicación viral. Este mecanismo ha sido estudiado con anterioridad por Chen y cols.<sup>64</sup> para otros alelos HLA-B. Nosotros encontramos que en la secuencia viral consenso de una cohorte guatemalteca, existen múltiples epítomos reconocidos por ambos alelos, C\*04:01 y B\*35:01. Esto sugiere un aumento en la presión evolutiva ejercida sobre Gag, que pudiera seleccionar mutaciones de escape en un contexto inmunogenético donde los alelos selectores son significativamente frecuentes. Esta hipótesis se respalda en:

- La identificación de mutaciones asociadas a alta CRV dentro de los epítomos reconocidos por B\*35:01 y C\*04:01 en la cohorte estudiada.
- Las observaciones de Gao<sup>66</sup>, Lazaryan<sup>67</sup> y Pereyra<sup>68</sup> *et al.* quienes identificaron que B\*35:01 selecciona mutaciones de escape en el epítomo NPPIPVGDIY (Gag 253-262), favoreciendo la versión NPPIPVG~~I~~Y que no es inmunogénica y disminuye la estabilidad de unión al HLA, evadiendo la respuesta citotóxica.
- Reportes previos de Carrington y cols.<sup>42</sup> en donde la alta prevalencia de mutaciones de escape circulantes, restringidas por un HLA frecuente, explica por qué los alelos más comunes en una población no son protectores.

Los epítomos y mutaciones identificadas, seleccionadas por B\*35:01 y C\*04:01, se discuten más adelante.

Adicionalmente, se analizó la interacción de los alelos más frecuentes en la población guatemalteca con la proteína Gag. A pesar de no mostrar un impacto directo en la CRV, los alelos más prevalentes de una población moldean la secuencia viral circulante. Dicha interacción se da por fenómenos de variación y adaptación del virus y del sistema inmune, lo que puede provocar la diseminación de una cepa de VIH pre-adaptado conforme la pandemia progresa. Esa cepa puede tener mutaciones no deletéreas que sean huellas de adaptación, mutaciones que hagan a la cepa más susceptible a la respuesta inmune en un contexto específico o mutaciones de escape que no afecten la CRV o estén bien compensadas para permitir su replicación. Por ejemplo, hay evidencia de que B\*51:01 fue un alelo protector durante el inicio de la epidemia de VIH-1 subtipo B, pero la rápida acumulación de mutaciones de escape a nivel poblacional dentro del epítomo TAFTIPSI restringido por B\*51:01, redujo su efecto protector<sup>69</sup>.



Desde el inicio de la caracterización de la interacción HLA-VIH, se ha observado la selección de mutaciones en ambos sentidos, las que son huella de adaptación y aquellas que constituyen vías de escape. El hecho de que la mayor fuerza selectora (HLA) se ejerza sobre Gag, al parecer obedece a una ventaja cinética del procesamiento de péptidos. Esta proteína, junto con Pol, entran a la célula rápidamente después de la infección y son fácilmente degradadas por la vía de procesamiento de HLA-I; contrario a la proteína Env, que al quedarse en la membrana del linfocito infectado tiene que ser procesada por la vía endocítica.

Es por eso que se analizó la secuencia de Gag para definir las vías involucradas en el aumento de CRV en la población guatemalteca. Se observaron 18 posiciones bajo presión selectiva y en cuatro de ellas se identificaron mutaciones que impactan directamente la CRV (T84V, A115K, I223N y S373T). El aminoácido 84 se encuentra dentro del epítipo ATLYCVHQQR que es reconocido por A\*31:01, la versión mutada de este epítipo, AVLYCVHQK, está restringida por A\*68:01. El reconocimiento de la forma mutada, debería implicar una respuesta citotóxica efectiva contra el virus portador del cambio, sin embargo, A\*68:01 fue asociado significativamente a altos valores de CRV. Esto sugiere que pueden existir mutaciones compensatorias seleccionadas por el mismo alelo en otra región de Gag o incluso, en otro gen del VIH. Este fenómeno, conocido como co-variación de aminoácidos es complicado de evaluar por métodos estadísticos convencionales.

Se predijeron epítipos restringidos por A\*31:01 que comprenden las mutaciones T84V, I223N y S373T. Así mismo, las cuatro posiciones en donde se identificaron mutaciones asociadas directamente a CRV elevada, están ubicadas dentro de epítipos (predichos) restringidos por B\*35:01 y C\*04:01. Numerosos epítipos restringidos por B\*35:01 ya han sido probados experimentalmente, y se ha demostrado que desencadenan respuestas inmunodominantes; sin embargo, el efecto neto del reconocimiento del epítipo y el HLA, parece favorecer al virus. Para evaluar la correlación directa entre polimorfismos en Gag y alelos importantes, se buscaron los alelos sobre-representados en el grupo de pacientes portadores de cada mutación. Esto se realizó con la herramienta HAPLOID y se determinó que las mutaciones T84V y A115K están asociadas a la presencia de B\*35:01. Por otra parte, en el análisis evolutivo de las secuencias portadoras de estos dos cambios, se

observó que el 64% de los pacientes eran portadores C\*04:01, alelo ligado a B\*35:01, con impacto directo en la CRV y que restringe los mismo epítomos que B\*35:01. Este resultado nos permitió hipotetizar que B\*35:01 y C\*04:01 ejercen una amplia y conjunta presión evolutiva a lo largo de Gag, seleccionando mutaciones que se transmiten y que por lo tanto, le confieren una ventaja de replicación al VIH. Es decir, estos dos alelos, podrían constituir un importante mecanismo de riesgo predominante en esta población.

El efecto de A115K parece ser muy fuerte, dado que fue observado en un muy reducido grupo de pacientes (n=3). El análisis de correlación HLA-polimorfismo (HAPLOID) arrojó que esta mutación se asocia a la presencia de B\*35:01 y B\*35:12. Este último alelo, ha sido encontrado de forma significativa y consistente, asociado a riesgo de progresión de la enfermedad en otros estudios del laboratorio. Esta caracterización no se ha reportado con anterioridad, porque B\*35:12 no apareció en otras poblaciones estudiadas hasta el momento. Por lo que podría representar un alelo de progresión particular de Guatemala o de la población centroamericana.

Es importante resaltar, que todas las secuencias portadoras de A115K también fueron portadoras de T84V. Ambas mutaciones aparecen seleccionadas por alelos de riesgo (B\*35:01 y B\*35:12) y podrían constituir un evento de co-variación o de compensación. Las estrategias estadísticas usadas en este estudio, no permiten la evaluación estricta de estos eventos, pues considera a cada mutación y cada HLA como eventos independientes. Para poder analizar la variación simultanea de dos o más aminoácidos y la presión conjunta de dos o más HLAs existen modelos que corrigen los resultados por filogenia y múltiples comparaciones. Jonathan Carlson de Microsoft Research Group desarrolló un algoritmo denominado "Red de Dependencia Filogenética" (PDN) capaz de identificar la selección conjunta de mutaciones, la variación en direcciones opuestas de dos a.a. o el efecto de más de un alelo HLA sobre una secuencia dada. Esto se hace a través de la construcción de la secuencia ancestral de cada posición de la proteína, la evaluación de la selección por el alelo HLA asociado a esa secuencia y el efecto de variación de ese aminoácido respecto a cada uno de los restantes. Es por eso, que se requieren mínimo 500 secuencias para obtener asociaciones confiables.

A pesar de la limitación del análisis estadístico, la observación de asociaciones que desde distintos enfoques (filogenético, CRV, presión positiva, etc...), son consistentes permite, por primera vez, la caracterización de la relación entre los marcadores de escape y la progresión de la enfermedad en una población, hasta ahora desconocida.

Finalmente, al evaluar la restricción sobre Gag de otros alelos frecuentes en Guatemala, se encontraron epítomos que abarcan todas las posiciones bajo presión selectiva localizadas en dicha secuencia. Por eso, es importante continuar la caracterización de estas mutaciones restringidas por alelos predominantes en la población, a pesar de que en este estudio no mostraron un impacto significativo en la CRV.

El estudio de otros alelos como A\*68:03, B\*35:12, B\*35:43, B\*39:05, B\*39:08 y C\*07:02 que son significativamente más frecuentes en Guatemala que en otras poblaciones caracterizadas al momento, nos permitirá entender mejor la adaptación del virus desde el inicio de la epidemia hasta la actualidad en esa región y determinar otras vías de progresión o control de la infección por VIH.

El mapa preliminar de selección inmune mediada por HLA producido a partir de este estudio, puede ser complementado con pruebas funcionales que permitan conocer más a fondo la respuesta citotóxica. Ya sea por la medición de la respuesta citotóxica mediante ELISPOT, la medición de la avidéz de unión de un alelo y el péptido, por la evaluación del papel que pueden jugar los clonotipos de TCR en el reconocimiento del complejo péptido-HLA o por la estimación del impacto de otros genes en la capacidad replicativa viral.

Con el conocimiento de la importancia de la respuesta citotóxica en la rapidez de progresión de la infección por VIH, las estrategias recientes de vacunación, buscan inducir una respuesta mediada por linfocitos T CD8+ efectiva. Debido a que esta respuesta inmune ejerce una notable fuerza selectora sobre el VIH y favorece su variación, se ha propuesto limitar el diseño de una vacuna a regiones conservadas altamente inmunogénicas, de tal forma que el escape viral solo pueda dirigirse a mutaciones con alto costo en replicación. O bien, diseñar inmunógenos que incorporen la forma wild-type y la forma de escape, siempre que ésta mantenga cierto grado de unión al HLA, para bloquear las rutas preferentes de escape viral.

El mosaico que se pretende en una vacuna, debe asegurar que aun los epitopos con variantes de escape sean inmunogénicos y seguros, por lo que deben estudiarse las consecuencias inmunológicas de las variantes de escape, su transmisión y diseminación. Así como las consecuencias en la replicación y la virulencia en un contexto inmunogenético determinado. Si bien, el desarrollo de una vacuna personalizada es teóricamente posible, la ejecución de este proyecto estaría ampliamente limitado por el reparto inequitativo de recursos de investigación y desarrollo entre países. Sin embargo, la creación de una vacuna regional o poblacional, basada en el haplotipo HLA más prevalente, sí es probable.

Por ello, a pesar de que existen muchos factores que influyen la cinética del escape viral, haciendo sus consecuencias clínicas difíciles de cuantificar, complementar la imagen de los sitios, las vías y la cinética del escape viral no solo ayudara a entender en qué medida la inmunidad del hospedero favorece la evolución del VIH, sino que dirigirá el diseño racional de futuros inmunógenos vacunales.

## CAPITULO 7. CONCLUSIONES

- ✂ Se describió por primera vez el contexto inmunogenético de los individuos infectados con VIH-1 de Guatemala en el que los alelos más frecuentes fueron A\*02:01, A\*24:02, B\*35:01, B\*40:02, C\*04:01 y C\*07:02.
- ✂ Se identificaron alelos particularmente frecuentes en Guatemala como A\*68:03, B\*35:12, B\*35:43, B\*39:05 y B\*39:08 que aún no han sido bien caracterizados.
- ✂ Se corroboró la asociación positiva entre la capacidad replicativa viral y la carga viral y la correlación negativa de la capacidad replicativa y la cuenta de T CD4+.
- ✂ Se identificó que los alelos A\*31:01, A\*68:01, B\*18:01, B\*35:01, B\*49:01, C\*04:01 y C\*05:01 tienen impacto en la capacidad replicativa viral.
- ✂ Se identificó un papel controversial y aparentemente opuesto de los alelos en el desequilibrio de ligamiento B\*18:01 y C\*05:01, por lo que un mecanismo independiente de impacto en la capacidad replicativa viral de C\*05:01 no se descarta.
- ✂ Se mapeó un epítipo mutado restringido por A\*68:01, el alelo fue asociado a alta CRV, lo que sugiere la existencia de mutaciones compensatorias dentro o fuera de Gag.
- ✂ Los resultados sugieren una posible vía de progresión a la infección basada en la selección simultánea de mutaciones de escape por parte de los alelos B\*35:01 y C\*04:01.
- ✂ Se observó que las mutaciones de escape T84V y A115K, que son seleccionadas por B\*35:01 y C\*04:01, podrían representar un evento de co-variación o compensación.
- ✂ Se obtuvo evidencia de la interacción entre los alelos más frecuentes en Guatemala y la proteína Gag circulante, al mapearse numerosos epítipos que comprenden posición bajo presión selectiva y están restringidos por dichos alelos.
- ✂ Se corroboró que el efecto de alelos típicos de protección (B27 y B57) no es notable en la población estudiada.
- ✂ Se construyó el primer mapa de selección inmune mediada por HLA en la secuencia de Gag circulante en la población guatemalteca.

## LITERATURA CITADA

1. AB NM. Nobel Prize. 2014; [http://www.nobelprize.org/faq/questions\\_in\\_category.php?id=22](http://www.nobelprize.org/faq/questions_in_category.php?id=22)>. Accessed Enero 9, 2015.
2. (UNAIDS) JUNPoHA. *Global Report. UNAIDS report on the global AIDS epidemic 2013*. Switzerland: Joint United Nations Programme on HIV/AIDS (UNAIDS); November 2013 2013.
3. (UNAIDS) JUPoHA. *The GAP report*. Switzerland: Joint United Programme on HIV/AIDS (UNAIDS); July 2014 2014.
4. CENSIDA. La epidemia del VIH y el SIDA en México. In: Secretaría de Salud CNpIPyeCdVyeS, ed. México, D.F.2014.
5. CENSIDA. Vigilancia Epidemiológica de casos de VIH/SIDA en México. Registro Nacional de Casos de SIDA. Actualización al 30 de septiembre de 2014. In: Secretaría de Salud DGdE, ed. México: Secretaría de Salud; 2014.
6. Monterroso LE, Castañeda X., García J. Boletín No- 1-2014: Estadísticas de VIH y VIH avanzado. Guatemala, enero 1984-Septiembre 2014. In: Epidemiología. MdSPyASCNd, ed. Guatemala, Guatemala: MSPAS, CNE; 2014.
7. *Mortalidad por VIH en Guatemala 2005-2013*. Guatemala, Guatemala2014.
8. *La sala situacional. Indicadores de vigilancia, el financiamiento y de la respuesta al VIH en Guatemala*. Guatemala, Guatemala2014.
9. Hemelaar J. The origin and diversity of the HIV-1 pandemic. *Trends in molecular medicine*. Mar 2012;18(3):182-192.
10. Cochrane. A. Human Immunodeficiency Virus. In: Wilkins WKHLW, ed. *Fields virology*. 5th ed. USA2007.
11. Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet*. Aug 5 2006;368(9534):489-504.
12. Hunt R. Virology: Human Immunodeficiency Virus. In: <http://www.microbiologybook.org/mhunt/flu.htm>., ed. *Microbiology and Immunology On-line*2013.
13. Foley B. LT, Apetrie C., Hahn B., Mizrachi I, Mullins J., Rambaut A., Wolinsky S., Korber E. *HIV Sequence Compendium 2013*. Vol LA-UR 13-26007: Theoretical Biology and Biophysics Group; 2013.
14. Freed EO MM. *HIVs and their replication*. Vol II. Quinta ed. Philadelphia: Krieger DM; 2007.
15. Richard H. Virology. In: *Medicine UoSCSo*, ed. USA2007: <http://pathmicro.med.sc.edu/book/virol-sta.htm>.
16. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *The Journal of experimental medicine*. Feb 17 1997;185(4):621-628.
17. Córdoba Villalobos J. A. PdLRS, Valdespino J. L. *25 años de SIDA en México. Logros, desaciertos y retos*. Segunda ed. Cuernavaca, Morelos,México: Instituto Nacional de Salud Pública; 2009.
18. Turner BG, Summers MF. Structural biology of HIV. *Journal of molecular biology*. Jan 8 1999;285(1):1-32.
19. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature reviews. Microbiology*. Apr 2012;10(4):279-290.

20. Simon V, Ho DD. HIV-1 dynamics in vivo: implications for therapy. *Nature reviews. Microbiology*. Dec 2003;1(3):181-190.
21. Bank RPD. The Structural Biology of HIV. 2011.
22. Database LA. HIV-1 Landmarks. <http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html>, 2015.
23. Melikyan GB. Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology*. 2008;5:111.
24. Wilen CB, Tilton JC, Doms RW. HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine*. 2012;2(8).
25. Arhel NJ, Souquere-Besse S, Munier S, et al. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J*. Jun 20 2007;26(12):3025-3037.
26. Bowerman B, Brown PO, Bishop JM, Varmus HE. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes & development*. Apr 1989;3(4):469-478.
27. Delelis O, Carayon K, Saib A, Deprez E, Mouscadet JF. Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology*. 2008;5:114.
28. Engelman A, Cherepanov P. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS pathogens*. Mar 2008;4(3):e1000046.
29. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews. Immunology*. Jan 2010;10(1):11-23.
30. Weber J. The pathogenesis of HIV-1 infection. *British medical bulletin*. 2001;58:61-72.
31. Kahn JO, Walker BD. Acute human immunodeficiency virus type 1 infection. *The New England journal of medicine*. Jul 2 1998;339(1):33-39.
32. Schacker TW, Hughes JP, Shea T, Coombs RW, Corey L. Biological and virologic characteristics of primary HIV infection. *Annals of internal medicine*. Apr 15 1998;128(8):613-620.
33. Mogensen TH, Melchjorsen J, Larsen CS, Paludan SR. Innate immune recognition and activation during HIV infection. *Retrovirology*. 2010;7:54.
34. Abbas AK, Lichtman AH, Pillai S. *Cellular and molecular immunology*. 6th ed. Philadelphia: Saunders/Elsevier; 2010.
35. Blackwell JM, Jamieson SE, Burgner D. HLA and infectious diseases. *Clinical microbiology reviews*. Apr 2009;22(2):370-385, Table of Contents.
36. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. *Nucleic acids research*. Jan 2015;43(Database issue):D423-431.
37. Mehra NK, Kaur G. MHC-based vaccination approaches: progress and perspectives. *Expert reviews in molecular medicine*. Feb 2003;5(7):1-17.
38. Raghavan M, Del Cid N, Rizvi SM, Peters LR. MHC class I assembly: out and about. *Trends in immunology*. Sep 2008;29(9):436-443.
39. De Tullio G, De Fazio V, Sgherza N, et al. Challenges and opportunities of microRNAs in lymphomas. *Molecules*. 2014;19(9):14723-14781.
40. Ahlers JD, Belyakov IM. New paradigms for generating effective CD8+ T cell responses against HIV-1/AIDS. *Discovery medicine*. Jun 2010;9(49):528-537.

41. Gascoigne NR. Do T cells need endogenous peptides for activation? *Nature reviews. Immunology*. Nov 2008;8(11):895-900.
42. Carrington M, O'Brien SJ. The influence of HLA genotype on AIDS. *Annual review of medicine*. 2003;54:535-551.
43. Tang J, Costello C, Keet IP, et al. HLA class I homozygosity accelerates disease progression in human immunodeficiency virus type 1 infection. *AIDS research and human retroviruses*. Mar 1 1999;15(4):317-324.
44. Carlson JM, Brumme ZL. HIV evolution in response to HLA-restricted CTL selection pressures: a population-based perspective. *Microbes Infect*. Apr 2008;10(5):455-461.
45. Kaslow RA, Carrington M, Apple R, et al. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nature medicine*. Apr 1996;2(4):405-411.
46. Campbell TB, Schneider K, Wrin T, Petropoulos CJ, Connick E. Relationship between in vitro human immunodeficiency virus type 1 replication rate and virus load in plasma. *J Virol*. Nov 2003;77(22):12105-12112.
47. Quinones-Mateu ME. Is HIV-1 evolving to a less virulent (pathogenic) virus? *Aids*. Oct 14 2005;19(15):1689-1690.
48. Kiepiela P, Leslie AJ, Honeyborne I, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature*. Dec 9 2004;432(7018):769-775.
49. Kiepiela P, Ngumbela K, Thobakgale C, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med*. Jan 2007;13(1):46-53.
50. Carlson JM, Le AQ, Shahid A, Brumme ZL. HIV-1 adaptation to HLA: a window into virus-host immune interactions. *Trends in microbiology*. Apr 2015;23(4):212-224.
51. Goulder PJ, Watkins DI. HIV and SIV CTL escape: implications for vaccine design. *Nature reviews. Immunology*. Aug 2004;4(8):630-640.
52. Goulder PJ, Walker BD. HIV and HLA class I: an evolving relationship. *Immunity*. Sep 21 2012;37(3):426-440.
53. Leslie AJ, Pfafferott KJ, Chetty P, et al. HIV evolution: CTL escape mutation and reversion after transmission. *Nature medicine*. Mar 2004;10(3):282-289.
54. Brumme ZL, Brumme CJ, Carlson J, et al. Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection. *Journal of virology*. Sep 2008;82(18):9216-9227.
55. Li B, Gladden AD, Altfeld M, et al. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *Journal of virology*. Jan 2007;81(1):193-201.
56. Brockman MA, Tanzi GO, Walker BD, Allen TM. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. *J Virol Methods*. Feb 2006;131(2):134-142.
57. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. Aug 2007;24(8):1596-1599.
58. Phillips RE, Rowland-Jones S, Nixon DF, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature*. Dec 12 1991;354(6353):453-459.



59. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science*. May 24 2002;296(5572):1439-1443.
60. Leslie A, Matthews PC, Listgarten J, et al. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *Journal of virology*. Oct 2010;84(19):9879-9888.
61. Cohen GB, Ghandi, R.T., Davis, D.M., Mandelboim, O., B.K., Strominger, J.L., and Baltimore, D. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity*. 1999;10:661-671.
62. Kosmrlj A, Read EL, Qi Y, et al. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature*. May 20 2010;465(7296):350-354.
63. Almeida JR, Price DA, Papagno L, et al. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *The Journal of experimental medicine*. Oct 1 2007;204(10):2473-2485.
64. Chen H, Ndhlovu ZM, Liu D, et al. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nature immunology*. Jul 2012;13(7):691-700.
65. Yu XG, Addo MM, Perkins BA, et al. Differences in the expressed HLA class I alleles effect the differential clustering of HIV type 1-specific T cell responses in infected Chinese and caucasians. *AIDS research and human retroviruses*. May 2004;20(5):557-564.
66. Gao X, Nelson GW, Karacki P, et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *The New England journal of medicine*. May 31 2001;344(22):1668-1675.
67. Lazaryan A, Song W, Lobashevsky E, et al. The influence of human leukocyte antigen class I alleles and their population frequencies on human immunodeficiency virus type 1 control among African Americans. *Human immunology*. Apr 2011;72(4):312-318.
68. International HIVCS, Pereyra F, Jia X, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science*. Dec 10 2010;330(6010):1551-1557.
69. Kawashima Y, Kuse N, Gatanaga H, et al. Long-term control of HIV-1 in hemophiliacs carrying slow-progressing allele HLA-B\*5101. *Journal of virology*. Jul 2010;84(14):7151-7160.

## APENDICE I.

Frecuencias alélicas de 160 individuos Guatemaltecos, VIH positivos y sin tratamiento antirretroviral incluidos en el presente estudio.

Alelo HLA	Frecuencia alélica	Frecuencia poblacional
A0201	0.2219	0.4062
A2402	0.1875	0.3438
A0206	0.1219	0.2375
A6803	0.1187	0.2313
A3101	0.0594	0.1187
A6801	0.0531	0.1
A0301	0.0344	0.0688
A0101	0.025	0.05
A3002	0.025	0.05
A3301	0.0219	0.0437
A1101	0.0125	0.025
A2301	0.0125	0.025
A3001	0.0094	0.0187
A3102	0.0094	0.0125
A3201	0.0094	0.0187
A3601	0.0094	0.0187
A6802	0.0094	0.0187
A0205	0.0063	0.0125
A0222	0.0063	0.0063
A6805	0.0063	0.0125
A7401	0.0063	0.0125
A0202	0.0031	0.0063
A2401	0.0031	0.0063
A2501	0.0031	0.0063
A2601	0.0031	0.0063
A2901	0.0031	0.0063
A2902	0.0031	0.0063
A3004	0.0031	0.0063
A3402	0.0031	0.0063
A6601	0.0031	0.0063
A6901	0.0031	0.0063
A8001	0.0031	0.0063

Alelo HLA	Frecuencia alélica	Frecuencia poblacional
B3501	0.1752	0.293
B4002	0.0796	0.1529
B1501	0.0701	0.1019
B3512	0.0669	0.1338
B3905	0.0637	0.1019
B3543	0.0573	0.0955
B0702	0.0478	0.0892
B3908	0.0478	0.0892
B5101	0.0446	0.0764
B3517	0.0318	0.0573
B0801	0.0255	0.0382
B1801	0.0255	0.0446
B3902	0.0255	0.051
B5701	0.0159	0.0255
B3502	0.0127	0.0191
B3906	0.0127	0.0191
B4001	0.0127	0.0255
B4402	0.0127	0.0255
B5301	0.0127	0.0255
B1402	0.0096	0.0191
B2705	0.0096	0.0191
B3901	0.0096	0.0191
B4403	0.0096	0.0127
B4901	0.0096	0.0191
B1503	0.0064	0.0127
B1524	0.0064	0.0127
B3514	0.0064	0.0064
B4008	0.0064	0.0127
B4201	0.0064	0.0127
B4501	0.0064	0.0127
B4801	0.0064	0.0127
B5201	0.0064	0.0127
B5802	0.0064	0.0127
B7801	0.0064	0.0127
B1302	0.0032	0.0064
B1401	0.0032	0.0064
B2702	0.0032	0.0064
B3503	0.0032	0.0064
B3504	0.0032	0.0064
B3520	0.0032	0.0064
B3545	0.0032	0.0064
B3801	0.0032	0.0064
B3903	0.0032	0.0064
B3913	0.0032	0.0064
B4502	0.0032	0.0064
B5001	0.0032	0.0064
B5305	0.0032	0.0064
B5601	0.0032	0.0064
B5703	0.0032	0.0064

Alelo HLA	Frecuencia alélica	Frecuencia poblacional
<b>C0401</b>	0.2753	0.4177
<b>C0702</b>	0.2247	0.3608
<b>C0102</b>	0.1013	0.1772
<b>C0304</b>	0.0759	0.1392
<b>C0701</b>	0.0538	0.0886
<b>C0303</b>	0.038	0.0696
<b>C0305</b>	0.0348	0.0696
<b>C0202</b>	0.0316	0.0633
<b>C1502</b>	0.0285	0.038
<b>C0501</b>	0.0222	0.0443
<b>C0602</b>	0.019	0.038
<b>C1601</b>	0.019	0.038
<b>C0802</b>	0.0127	0.019
<b>C1203</b>	0.0127	0.0253
<b>C1701</b>	0.0127	0.0253
<b>C1202</b>	0.0095	0.019
<b>C1402</b>	0.0063	0.0127
<b>C1509</b>	0.0063	0.0127
<b>C0302</b>	0.0032	0.0063
<b>C0509</b>	0.0032	0.0063
<b>C0604</b>	0.0032	0.0063
<b>C0704</b>	0.0032	0.0063
<b>C1801</b>	0.0032	0.0063