



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE CIENCIAS

“Epidemiología Molecular del Virus de la  
Inmunodeficiencia Humana 1 (VIH-1) en pacientes  
mexicanos: proyecto piloto”

**T E S I S**  
QUE PARA OBTENER EL TÍTULO DE:  
**BIÓLOGA**

PRESENTA:

Daniela Garrido Rodríguez



Tutor: Dr. Gustavo Reyes Terán



Universidad Nacional  
Autónoma de México



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

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

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

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

FACULTAD DE CIENCIAS



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

División de Estudios Profesionales

**ACT. MAURICIO AGUILAR GONZÁLEZ**  
**Jefe de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Por este medio hacemos de su conocimiento que hemos revisado el trabajo escrito titulado:  
"Epidemiología Molecular del Virus de la Inmunodeficiencia Humana 1  
(VIH-1) en pacientes mexicanos: proyecto piloto"

realizado por Daniela Garrido Rodríguez

con número de cuenta 09911379-5 , quien cubrió los créditos de la licenciatura en  
Biología

Dicho trabajo cuenta con nuestro voto aprobatorio.

Tutor (a)  
Propietario Dr. Gustavo Reyes Terán

Propietario Dr. Hector Enrique Espinosa Arciniega

Propietario Dr. Joel Armando Vazquez Pérez

Suplente M. en C. Santiago Avila Rios

Suplente Dr. Ricardo Stanley Vega Barrientos

**Atentamente**  
**"POR MI RAZA HABLARÁ EL ESPÍRITU"**  
Ciudad Universitaria, D.F., a 23 de Octubre  
**CONSEJO DEPARTAMENTAL DE Biología**

del 2006

  
Dr. Zenón Cano Santana

FACULTAD DE CIENCIAS



UNIDAD DE ENSEÑANZA  
DE BIOLOGÍA

A mamá y papá,  
El tiempo y el esfuerzo para culminar  
una carrera no se compara con el tiempo y esfuerzo  
que implica educar a un hijo, pero esto es una pequeña  
muestra del bueno trabajo que hicieron al educarme.  
Gracias por sus palabras, cariño y apoyo en mi camino.

A Gaby,  
Por siempre estar a mi lado, apoyándome y dándome consejos.  
Muchas personas pasan toda su vida buscando un amigo,  
pero yo encontré la amistad desde pequeña contigo  
Siempre serás mi mejor amiga.

A Efren,  
Por esos días en los que cerraba los ojos creyendo que el  
mundo se derrumbaba y tú estabas a mi lado ayudándome  
a ver las cosas con otra perspectiva.  
Por enseñarme que el amor y la amistad van de la mano.  
Te amo.

## AGRADECIMIENTOS

Al Dr. Gustavo Reyes Terán, por la paciencia y confianza que me brindó a lo largo de este trabajo. Gracias por ser un amigo y un tutor excelente.

Al Dr. Enrique Espinosa, por todos los consejos que me brindó para concluir este trabajo. Gracias por ser mi guía en el ámbito científico.

Al Dr. Joel Vazquez, por su apoyo y paciencia dentro del laboratorio. Gracias por todos los consejos que me diste.

Al M. En C. Santiago Avila, por todas las charlas en las que me ayudaste a entender un poco más de la ciencia. Gracias por ser un excelente amigo.

A mi familia en Cuba y Chile, por todo el apoyo que me han dado a lo largo de mi vida y que a pesar de su lejanía siempre han estado cerca de mi corazón.

A mi familia en México, por convertirse en una familia más para mí y compartir conmigo todos los momentos significativos de la vida.

A mis amigos; Rebeca, Martín, Jessica, German, Christian, Mariana, Carlos y Angela, por todas las experiencias que pasamos juntos, por los consejos que me dieron y por su amistad.

A todos los integrantes del CIENI, que directa o indirectamente me ayudaron en todo este proceso, así como por todos los momentos que compartimos juntos en el laboratorio.

A todos mis profesores que sin su ayuda no hubiera llegado a culminar una carrera. Su esfuerzo y paciencia en mis años de estudiante me han hecho una persona más preparada.

## ÍNDICE

Introducción .....	1
1. Virología	
1.1 Estructura y Genes.....	2
1.2 Ciclo de Replicación.....	5
1.3 Origen y Evolución.....	10
2. Antecedentes.....	13
2.1 Tratamiento Antirretroviral.....	23
2.2 Mutaciones de Resistencia a Fármacos.....	24
3. Método	
3.1 Justificación.....	25
3.2 Hipótesis.....	26
3.3 Objetivos	
3.3.1 Objetivo General.....	26
3.3.2 Objetivos Particulares.....	26
3.4 Metodología	
3.4.1 Población de Estudio.....	27
3.4.2 Envío y Almacenamiento de Muestras.....	28
3.4.3 Genotipificación.....	28
3.4.4 Análisis Estadístico.....	38
3.4.5 Tipificación del VIH-1.....	38
3.4.6 Análisis Filogenético.....	38
3.4.6.1 ClustalX.....	39
3.4.6.2 Neighbor-Joining y Bootstrapping.....	39
3.4.6.3 Tree-View.....	40
3.4.6.4 Tree-Puzzle.....	40
4. Resultados	
4.1 Población de Estudio.....	42
4.2 Mutaciones de Resistencia a Fármacos.....	44
4.3 Comparación de Frecuencias entre Poblaciones.....	48
4.4 Tipificación del VIH-1.....	50
4.5 Análisis Filogenético.....	50
5. Discusión.....	52
6. Conclusión.....	56
7. Anexo.....	57
Referencias.....	59

## **Introducción**

La ciencia de la epidemiología se originó y evolucionó en respuesta a las grandes epidemias como el cólera, la fiebre tifoidea, la viruela y la fiebre amarilla. Por definición, la epidemiología es la ciencia que estudia la aparición, determinantes, distribución y control de la salud y la enfermedad en una población humana definida<sup>1</sup>.

La pandemia de VIH/SIDA, considerada como una de las más destructivas en la historia, ha sido estudiada desde sus inicios a partir de métodos epidemiológicos. Dentro de algunos de los métodos usados en el estudio de esta enfermedad se han utilizado las secuencias genómicas de virus. De esta forma se ha podido establecer un enfoque epidemiológico diferente: el de la epidemiología molecular. Este método ha permitido establecer la distribución geográfica de las diferentes formas del virus y establecer el transmisor de infección entre personas relacionadas, así como entre poblaciones cercanas.

Los objetivos planteados para este estudio consisten en establecer la epidemiología del VIH/SIDA en México, las relaciones filogenéticas del virus circulante en nuestra población con virus circulantes en otros países y la prevalencia de mutaciones primarias del virus en la población mexicana.

En el primer capítulo se da una breve revisión de la virología, el origen, la evolución y epidemiología actual del VIH con el fin de establecer los parámetros biológicos del virus, su interacción con el hospedero y la distribución de la enfermedad en las diferentes poblaciones.

El segundo capítulo es una revisión bibliográfica de la distribución geográfica actual de las diferentes formas del VIH.

En el tercer capítulo se explica: el planteamiento y justificación del problema, la hipótesis, los objetivos del estudio, el diseño de estudio, los sujetos, la secuencia de las pruebas de laboratorio y el tipo de análisis realizado.

En el cuarto capítulo se presentan los resultados, comenzando por la prevalencia de mutaciones encontradas en la población de estudio y seguido por el análisis filogenético con la comparación de los diferentes países.

Y finalmente, en el quinto capítulo se discuten los resultados, relacionándolos con las investigaciones previas en la población mexicana y en diferentes poblaciones, asociándolo a problemas sociales y clínicos de la actualidad.



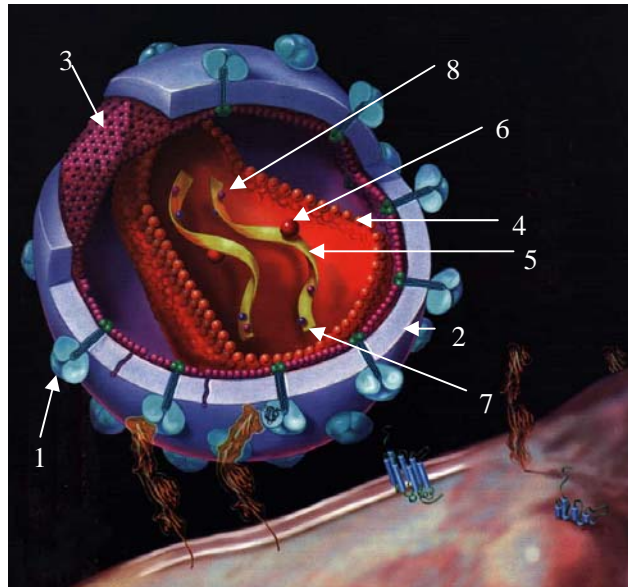
## 1. Virología

### 1.1 Estructura y genes

El VIH es miembro de la familia Retroviridae perteneciente al género Lentivirus<sup>2</sup>. Los retrovirus presentan características distintivas del grupo. Dichas características son: a) un genoma compuesto de ARN, b) una estructura viral organizada por tres genes poliprotéicos llamados grupo antígeno específico (*gag*), polimerasa (*pol*) y envoltura (*env*), c) un ciclo de replicación en donde se incorpora la inserción del genoma viral dentro del material genético del hospedero mediante la transcripción inversa, y d) la habilidad de alterar su genoma mediante las mutaciones en respuesta a las diferentes presiones de selección. Los lentivirus además de los genes estructurales de los retrovirus se caracterizan por presentar una combinación compleja de genes accesorios o genes no estructurales y de tener largos periodos de latencia clínica durante la infección.

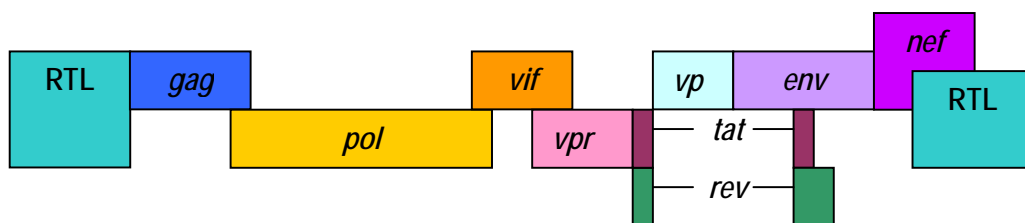
El VIH es un virus envuelto, la partícula es de forma icosaédrica de 110 nm de diámetro. Posee cerca de 72 espículas en su superficie que constan de dos glicoproteínas llamadas gp120 y gp41, derivadas de un solo precursor protéico (gp160) codificado por el gen *env*. La molécula gp41 es una glicoproteína transmembranal que interactúa no-covalentemente con la molécula gp120 formando así una estructura oligomérica de carácter trimérico. La forma y la integridad de los viriones está dada por los productos del gen *gag*. La matriz está compuesta por la proteína p17 de *gag* la cual está sujeta a la cara interna de la membrana lipídica. Una segunda proteína *gag* llamada p24 genera la forma cónica característica de la cápside. Dos proteínas más de *gag* conocidas como p7 y p6 generan la nucleocápside en donde se encuentra el genoma viral, el cual está compuesto de dos cadenas simples/positivas de ARN.

La nucleocápside contiene toda la información genética y bioquímica necesaria para la replicación. Esta información incluye dos copias idénticas de ARN genómico, un ARN de transferencia reclutado durante el proceso de liberación del virión y tres enzimas llamadas transcriptasa inversa (RT), integrasa (IN) y proteasa (PR) (Figura 1).



**Figura 1.** Virus de la Inmunodeficiencia Humana (VIH). (1) glicoproteínas gp120 y gp41, (2) membrana lipídica. (3) matriz. (4) nucleocápside, (5) material genético, dos cadenas de ARN positivas. (6) enzima RT.(7) enzima PR. (8) enzima IN.

El genoma ARN del VIH tiene una longitud aproximada de 9.2 kb y una organización básica de las secuencias del ácido nucleico que es característica de todos los retrovirus conocidos (Figura 2).



**Figura 2.** Genoma del VIH

Las repeticiones terminales largas (LTR, por sus siglas en inglés) en los dos extremos del genoma regulan la integración del virus en el genoma del hospedero, la expresión de los genes y la replicación del virus. Las secuencias *gag* codifican proteínas estructurales nucleares, mientras que *env* codifica las glicoproteínas gp120 y gp41 de la envoltura, necesarias para la infección de las células. Las secuencias *pol* codifican la transcriptasa inversa, la integrasa y la proteasa del virus, todas ellas esenciales para su replicación. Además de estos genes típicos de los retrovirus, el VIH posee asimismo otros seis genes reguladores: *tat*, *rev*, *vif*, *nef*, *vpr* y *vpu*, cuyos productos regulan la reproducción vírica de varias formas. En el cuadro 1 se resumen las funciones de los genes<sup>3</sup>.

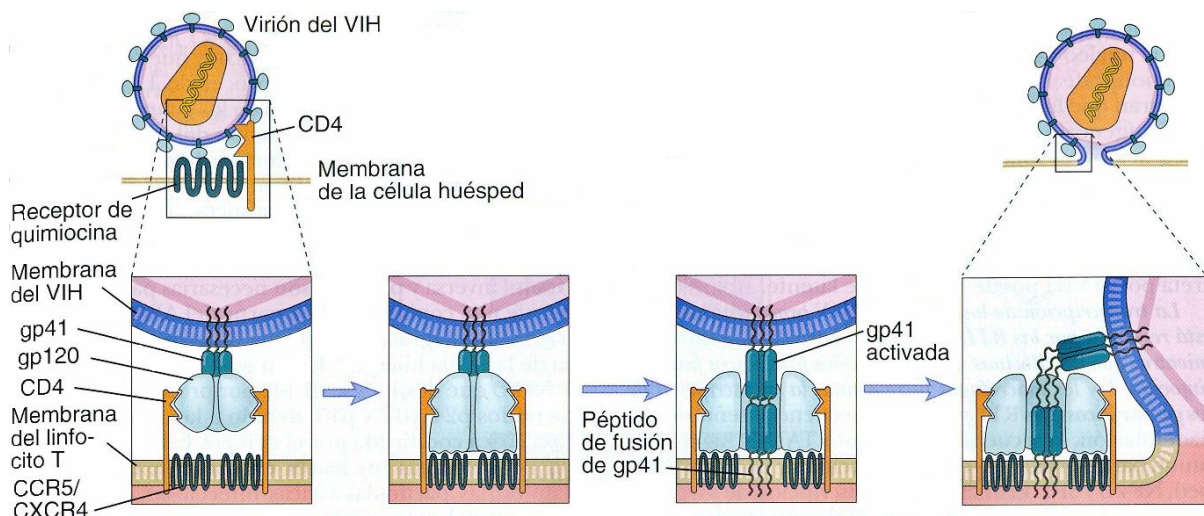
**Cuadro 1.** Función de los genes del VIH

<b>Gen</b>	<b>Función</b>
<i>LTR</i>	Integración del ADN viral en el genoma de la célula; lugar de unión de los factores de transcripción.
<i>gag</i>	Proteínas de matriz (MA, p17), proteínas de cápside (CA, p24), proteínas de nucleocápside (NC, p7) y p6
<i>pol</i>	Codifica una gran variedad de enzimas incluyendo la proteasa (p10), la transcriptasa inversa y la RNasa H (p66/51) y la integrasa (p32).
<i>env</i>	Proteínas de la cubierta vírica (gp120 y gp41) que intervienen en la unión con CD4, el receptor de quimiocina y la fusión de la membrana.
<i>vif</i>	Regula la capacidad de infección de las partículas víricas.
<i>vpr</i>	Proteína viral R (p15). Favorece la infección en macrófagos y la importación de ADN vírico al núcleo, arresta el ciclo celular en G <sub>2</sub> .
<i>tat</i>	Activador transcripcional (p14), necesario para el alargamiento de los productos de transcripción.
<i>rev</i>	Regulador de la expresión de genes virales (p19), favorece la salida del ARN vírico no procesado o con procesamiento sencillo desde el núcleo.
<i>vpu</i>	Proteína viral U, reduce la expresión de CD4 en la célula y potencia la liberación de virus a partir de las células.
<i>nef</i>	Factor negativo (p24), reduce la expresión de CD4 en la célula y potencia la liberación del virus a partir de las células; reduce la expresión del Complejo Mayor de Histocompatibilidad de clase I de la célula.

## 1.2 Ciclo de replicación

Al igual que otros virus, el VIH requiere infectar una célula para llevar a cabo su ciclo de replicación. El mecanismo de entrada a la célula, en este caso un linfocito T, comienza con el reconocimiento del receptor de membrana CD4<sup>+</sup> por medio de la glicoproteína gp120. Esta unión provoca un cambio conformacional en ambas membranas exponiendo la glicoproteína gp41 del virión y el correceptor de quimiocina CCR5 o CXCR4 de la célula.

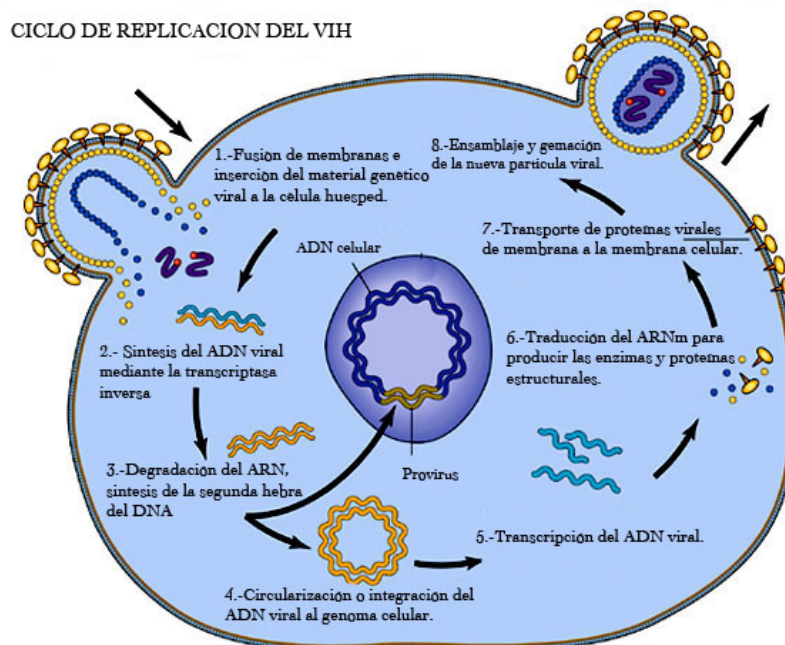
Posteriormente, la glicoproteína gp120 reconoce al correceptor CCR5 o CXCR4, mientras que la glicoproteína gp41 se fusiona a la membrana provocando otro cambio conformacional entre ambas membranas dando pie a la fusión de las mismas (Figura 3).



**Figura 3.** Unión de receptores. Reconocimiento del receptor CD4<sup>+</sup> por la glicoproteína gp120 y posterior reconocimiento del correceptor CCR5 o CXCR4 y la glicoproteína gp41

Una vez fusionadas las membranas, el core se rompe vaciando su contenido, las enzimas y el genoma viral son liberados al citoplasma celular. La liberación exitosa de las enzimas da lugar al complejo de retro-transcripción, lo que involucra el genoma viral de ARN, el primer tARN<sup>Lys</sup>, transcriptasa inversa (RT), integrasa (IN), matriz (MA), la

nucleocápside, la proteína viral R (Vpr) y varias proteínas celulares. Este complejo se ancla a los microfilamentos de actina para que se pueda llevar a cabo de forma eficiente la síntesis del ADN viral. Muchas veces este complejo se desestabiliza por efecto de algunas proteínas celulares; la proteína viral Vif ayuda a mantener la estabilidad del complejo. Una vez completa la síntesis del ADN viral por medio de la RT se forma el complejo de preintegración que está compuesto por: una doble cadena de cADN, IN, MA, Vpr y RT. Este complejo viaja al núcleo con ayuda de los microtúbulos y se sitúa cerca de un poro nuclear. Una vez anclado al poro nuclear se lleva a cabo la importación al interior del núcleo. Ya dentro del núcleo, la adecuada integración del ADN viral al ADN celular está mediada por la IN, la cual primero remueve nucleótidos de los extremos del ADN viral y después cataliza su unión con el ADN celular. Al ADN viral integrado en el ADN celular se le conoce como provirus<sup>3</sup>(Figura 4).



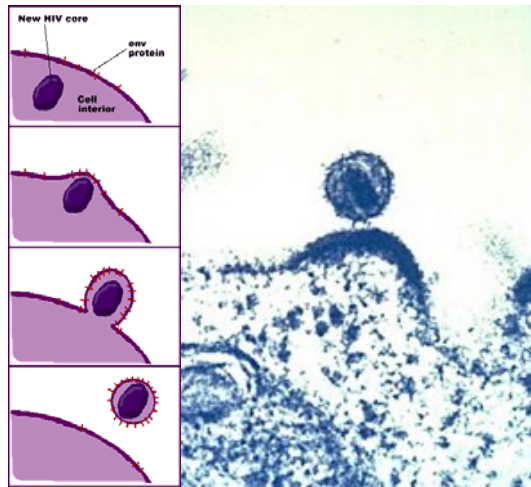
**Figura 4.** Ciclo de replicación del VIH

La transcripción de los genes del ADN proviral está regulado por las LTR de los genes estructurales del virus. Las LTR contienen secuencias señal de poliadenilación, la secuencia promotora TATA y lugares de unión para los diversos factores de transcripción de la célula, como NF- $\kappa$ B, SP1 y NFAT entre otros. El inicio de la transcripción de los genes del VIH en los linfocitos T está relacionada con su activación fisiológica por antígenos o citocinas. Para que se expresen los genes del VIH se necesita la presencia de la proteína Tat, que actúa uniéndose a una secuencia corta de RNA al extremo 5' del transcrito de RNA conocido como TAR (por sus siglas en inglés, *trans-acting response element*) aumentando así, la producción de transcritos completos de ARNm viral. La síntesis de partículas víricas comienza una vez completada la transcripción del ARN vírico y cuando los genes víricos se expresan en forma de las proteínas correspondientes. Los ARNm víricos que codifican las diversas proteínas del virus proceden de una sola transcripción completa del genoma la cual es sometida a varios procesos de corte y empalme ("splicing"). La expresión de los genes del VIH puede dividirse en una fase inicial, durante la cual se expresan los genes reguladores, y una fase avanzada, en la que se expresan los genes estructurales y se empaquetan los genomas víricos completos. En la fase inicial se generan las proteínas *rev*, *tat* y *nef* que son producto del multiprocesamiento del ARNm viral, los cuales son exportados desde el núcleo y se transcritos a proteínas en el citoplasma inmediatamente después de la infección de la célula. En la fase avanzada se generan los genes tardíos *env*, *gag*, y *pol*, que codifican componentes estructurales del virus y que se transcriben a partir de ARNm viral sin procesar. La proteína Rev inicia el cambio de la expresión de los genes iniciales a los tardíos, estimulando la exportación del ARNm sin procesar de éstos genes tardíos hacia el exterior del núcleo.

El producto del gen *pol* es una proteína precursora que se divide de forma sucesiva para formar las enzimas transcriptasa inversa, proteasa, ribonucleasa e integrasa. El gen *gag* codifica una proteína de 55 kD que después es cortada por proteólisis en los polipéptidos p24, p17 y p15, debido a la acción de la proteasa vírica codificada por el gen *pol*. Estos polipéptidos son las proteínas nucleares esenciales para el ensamblaje de las partículas víricas infecciosas.

El producto primario del gen *env* es una glicoproteína de 160 kD (gp160), a la que las proteasas celulares del retículo endoplásmico dividen en las proteínas gp120 y gp41, imprescindibles para la unión del VIH a las células.

Una vez transcritos varios genes, las proteínas del virus se sintetizan en el citoplasma. Comienza entonces el ensamblaje de las partículas víricas infecciosas mediante el empaquetamiento del ARN completamente transcrito del genoma provírico en el interior de un complejo nucleoprotéico formado por las proteínas nucleares *gag* y las enzimas codificadas por *pol* necesarias para el siguiente ciclo de integración. Este complejo nucleoprotéico se rodea de membrana y es liberado de la célula mediante gemación a partir de la membrana plasmática (Figura 5). Una vez que la partícula viral ha sido liberada al espacio extracelular la nueva partícula vírica comienza su maduración en donde la enzima proteasa acaba de cortar las cadenas de proteína del VIH en proteínas individuales que se combinan para formar un nuevo virus funcional e infeccioso<sup>4</sup>.



**Figura 5.** Gemación de una nueva partícula viral

### 1.3 Origen y Evolución

El Virus de la Inmunodeficiencia Humana (VIH) es el agente causal del Síndrome de Inmunodeficiencia Adquirida (SIDA). La epidemia de VIH/SIDA ha continuado su dispersión a lo largo del mundo desde su descubrimiento en 1981. Ahora, dos décadas después, 40 millones de personas viven con VIH/SIDA y cerca de 5 millones de personas adquirieron la enfermedad en el 2005<sup>5</sup>.

Para el mejor entendimiento de la pandemia se han desarrollado nuevas líneas de investigación y se han usado otras ya existentes. Dentro de estas últimas la evolución, el origen y el desarrollo siempre han sido de gran ayuda para el manejo adecuado del patógeno, así como para una mejor perspectiva de su posible futuro. El origen del VIH y su transmisión al ser humano ha sido una de las preguntas más difíciles de contestar. Ahora con base en los nuevos análisis de reconstrucción filogenética se sabe que el virus causante de esta pandemia tuvo su origen a partir del Virus de la Inmunodeficiencia del Simio (VIS), mediante transmisiones zoonóticas hacia el humano.



Actualmente se sabe que existen dos tipos de VIH: el VIH-1 y el VIH-2. El VIH-1 y el VIH-2 son filogenéticamente diferentes entre sí, cada uno es resultado de una transmisión zoonótica diferente; el VIH-1 a partir del SIVcpz de los chimpancés (*Pan troglodytes troglodytes* y *Pan troglodytes schweinfurthii*) ubicados en la zona ecuatorial del oeste y centro de África<sup>6,7</sup> y el VIH-2 a partir del SIVsm de los sooty mangabey (*Cercocebus atys*) localizados en la zona oeste de África<sup>8</sup>. Debido a la evolución que han presentado ambos tipos de VIH, éstos a su vez han divergido y han dado origen a diferentes subtipos, para esto se cree que existieron al menos de dos a tres saltos zoonóticos separados del chimpancé al humano para permitir la dispersión de los grupos M, N y O del VIH-1<sup>9</sup>. Mientras que se necesitaron aproximadamente siete linajes de virus del sooty mangabey para dar origen a los subtipos A - G del VIH-2<sup>10</sup>.

Existen diversos factores involucrados en la rápida evolución del VIH. Dentro de los más significativos se encuentran: a) la alta tasa de errores de la enzima transcriptasa inversa en la que comete un error por cada 1,000 a 10,000 bases incorporadas en la transcripción de ARN a ADN en las primeras etapas del ciclo de replicación viral<sup>11</sup>, b) la alta tasa de replicación del VIH; aproximadamente se producen de  $10^9$  virus cada día<sup>12</sup>; c) la recombinación (ver más adelante) del virus dentro de las células infectadas, ya que juega un papel importante en la variabilidad de éste siendo una de las formas más comunes por las que el virus puede adquirir genes que le confieran resistencia o mutaciones favorables<sup>13</sup>; y d) la presión de selección a la que está sometido el virus, ejercida de forma artificial por los fármacos antirretrovirales (ART) o natural por la respuesta inmune del hospedero.

A grandes rasgos la recombinación es el proceso mediante el cual la información genética se redistribuye. En los retrovirus, la recombinación ha sido una estrategia para intercambiar información genética entre dos moléculas de RNA heterogéneo.

Este proceso se lleva a cabo durante la transcripción inversa del RNA. Los retrovirus comúnmente empaquetan dos moléculas idénticas, o casi idénticas, de RNA. Las consecuencias genéticas de usar una sola molécula de RNA versus porciones de las dos moléculas como templates para la síntesis de DNA es usualmente lo mismo. Sin embargo, bajo ciertas condiciones, dos moléculas de RNA genéticamente diferentes pueden ser empaquetadas en un mismo virion y porciones de cada molécula pueden ser usadas para la síntesis de DNA produciendo de esta forma un recombinante. La recombinación ocurre con una alta frecuencia en aquellos viriones producidos en una célula co-infectada por dos tipos diferentes de virus, demostrando que la recombinación requiere del empaquetamiento de dos genomas virales en la misma partícula viral.

La pandemia de VIH/SIDA es provocada principalmente por el VIH-1, mientras que el VIH-2 está restringido a algunas regiones del este y centro de África. Dentro del VIH-1, el grupo M es el causante del 90% de las infecciones a nivel mundial. Este grupo está representado por 27 formas genéticas, distribuidas como: 11 subtipos y sub-subtipos (A1, A2, B, C, D, F1, F2, G, H, J y K), 16 formas recombinantes circulantes (CRFs) y formas únicas recombinantes (URFs)<sup>14,15</sup>. Para realizar la clasificación de las diferentes formas del virus se han usado las diferencias en las secuencias nucleotídicas de algunos genes, cabe mencionar que las formas recombinantes son virus cuyo genoma está compuesto por más de dos subtipos y se clasifican basándose en la secuenciación completa del genoma viral y que se presenten en al menos 3 pacientes sin relaciones epidemiológicas. Por otro lado las formas únicas recombinantes son aquellos virus igual que los recombinantes pero su estructura genómica solo se presenta en pacientes con relaciones epidemiológicas.

El gen de envoltura (*env*), ha sido el más utilizado para establecer estas diferencias, mostrando una diferencia del 30-50% entre el grupo M y O mientras que la variación intersubtipo es del 20-30% e intrasubtipo del 10-15%<sup>16</sup>. En contraste, la variación en el gen proteasa (*pol*) es de dos a tres veces menos variable mostrando una diferencia de 6-10% intersubtipo y de 3-6% intrasubtipo<sup>17</sup>. En el resto de los genes virales las variaciones son aún menores debido a la alta conservación en sus secuencias por motivos funcionales ya que mutaciones en estos sitios pueden disminuir la viabilidad del virus. El estudio de la variación en el gen *pol* ha resultado de gran importancia ya que ha permitido establecer filogenéticamente y geográficamente la distribución de los diferentes subtipos del VIH-1<sup>18</sup>. Asimismo, las terapias antivirales actuales se centran en las enzimas codificadas por este gen permitiendo de esta forma estudiar la prevalencia de mutaciones de resistencia primaria (inicial) y secundaria a los antirretrovirales y la transmisión de los virus resistentes en diferentes poblaciones<sup>19,20</sup>.

El estudio de las fuentes de variación genética y la evolución del VIH a nivel global es necesario, no solo para conocer su origen y entender la epidemiología del VIH-1, sino también para entender y tal vez predecir el surgimiento de nuevos subtipos, formas recombinantes circulantes e intra-subtipos con diferentes características de transmisibilidad y virulencia. Estos estudios tienen un gran valor en el desarrollo de posibles vacunas permitiendo seleccionar antígenos relevantes para virus circulantes en diferentes poblaciones.

## 2. ANTECEDENTES

Los estudios de epidemiología molecular se basan en: establecer la prevalencia de formas genéticas en una región, los flujos de transmisión entre personas relacionadas así como entre poblaciones cercanas y la distribución geográfica de los diferentes subtipos o formas recombinantes. Para esto existen diversos métodos que permiten entender la heterogeneidad del VIH. El método de referencia es la secuenciación y el análisis filogenético, aunque existen métodos más accesibles y económicos como el ensayo de movilidad de heteroduplex y los fragmentos polimórficos de restricción<sup>21,22</sup>.

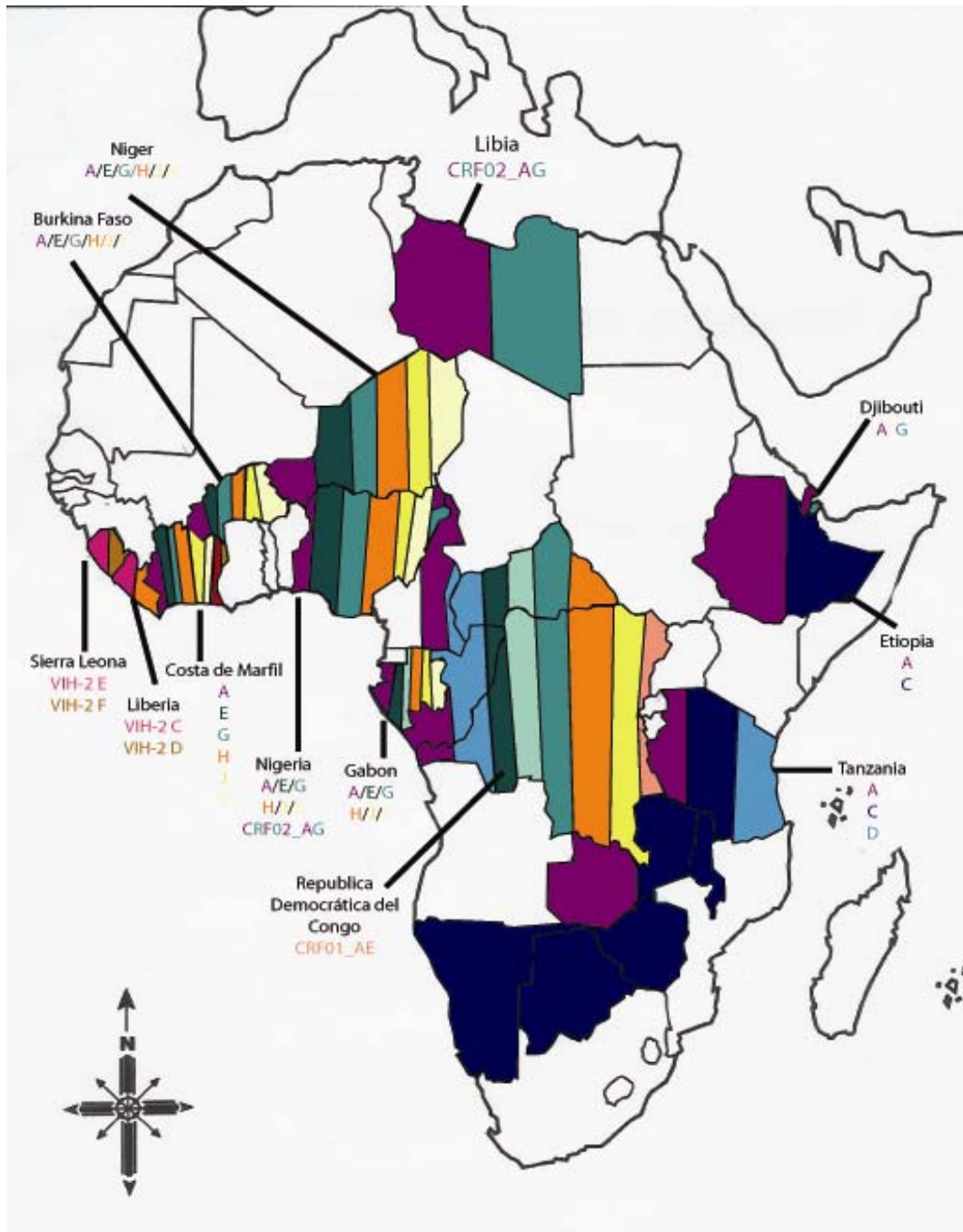
Con base en estos estudios se sabe que en la actualidad el VIH-1 es el responsable de la mayoría de las infecciones a nivel mundial. Los genotipos encontrados con mayor frecuencia en la población mundial son: el subtipo C (47%), A (27.2%), B (12.3%), D (5.3%) y CRF01\_AE (3.2%)<sup>23</sup>. Cerca del 70% de las infecciones mundiales por VIH-1 se encuentran en África sub-sahariana. África central es la única región en donde se han identificado todos los grupos del VIH-1 y todos los subtipos del grupo M.

Esta diversidad es consistente con la hipótesis de que África central es el epicentro y origen de la pandemia actual de VIH. La distribución mundial de las diferentes formas genéticas del VIH es la siguiente:

### *África*

En África el subtipo A se encuentra predominantemente distribuido en el centro y oeste. Este subtipo es responsable del 80% de las infecciones en el oeste y del 30% en el centro. En las ciudades de África central se encuentran prevalencias del subtipo A del 69% en Bwamanda, 53% en Kinshasa y 46% en Mbuji-Maya. En esta misma región, los subtipos C, D, G y H presentan una prevalencia del 7 al 9%, mientras que los subtipos F, J, K y CRF01\_AE presentan un 2 a 4%<sup>24</sup>.

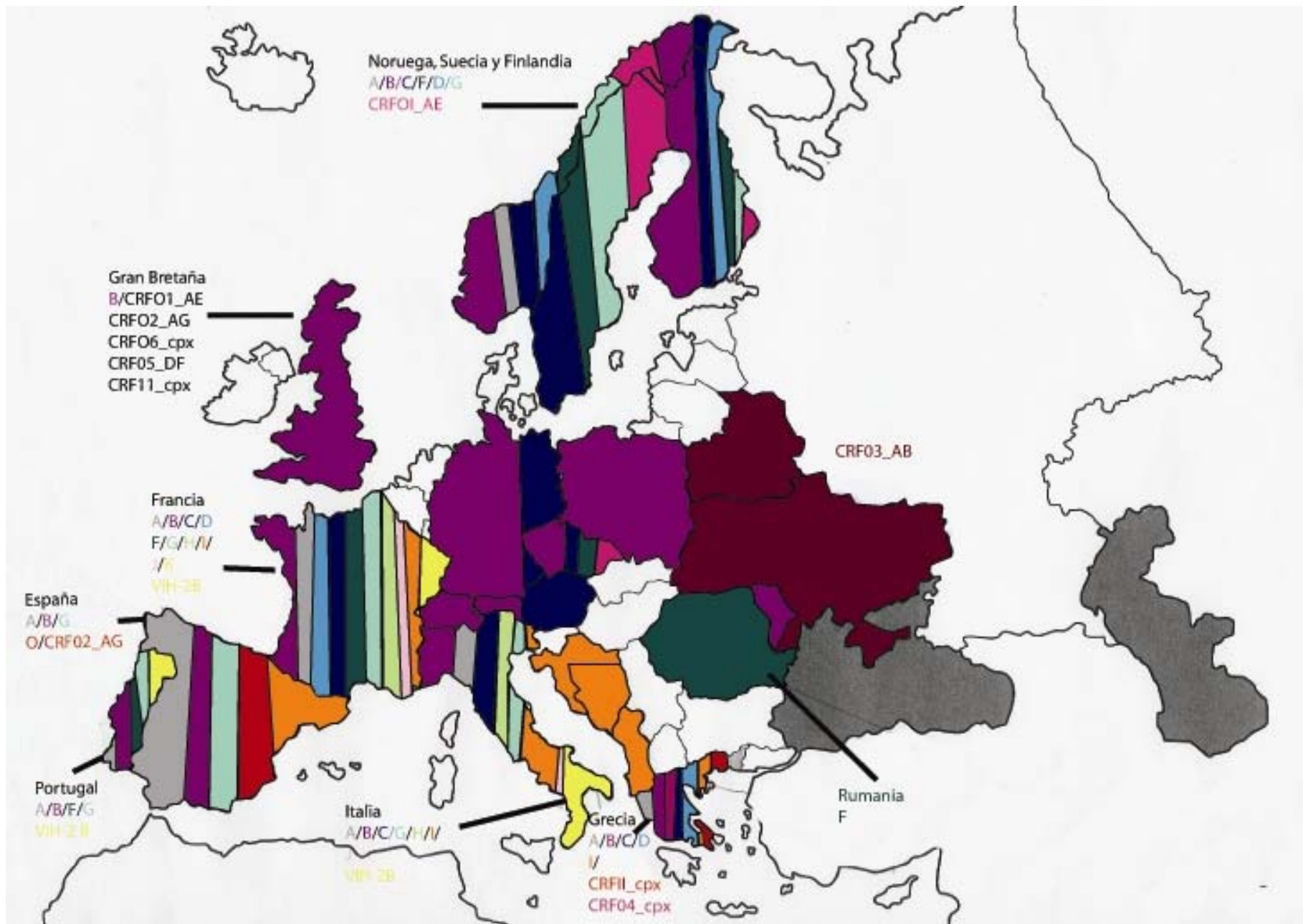
El subtipo C también es predominante en este continente y es el responsable del 60% de las infecciones mundiales, se cree que la distribución de este subtipo comenzó en el sur de África siendo Botswana, Zimbabwe, Malawi, Zambia, Tanzania y Namibia los países con mayor prevalencia de este subtipo<sup>25,26</sup>. El subtipo D es responsable del 5 al 40% de las infecciones en países del este y centro en donde co-circula con el subtipo A. En países como Burkina Faso, Malí, Nigeria, Costa de Marfil y Gabón, los subtipos A, E, G, H, J y K son predominantes<sup>27</sup>. Una secuencia mosaico compleja en donde se alternan el subtipo A y G reconocido como CRF02\_AG tuvo su origen en Ibadan, Nigeria, esta forma recombinante se ha distribuido a varios países del oeste y centro en donde representa un 50 a 70% de las formas recombinantes circulantes<sup>27-29</sup>. Otras formas recombinantes presentes son CRF11\_cpx, CRF10\_CD y varias formas únicas recombinantes como AD, AC, 02DO<sup>29,30</sup>. Con lo que respecta al VIH-2, su distribución es más limitada, el grupo A se encuentra en los países del oeste<sup>31,32</sup>. El grupo B y H están limitados en Costa de Marfil, mientras que los grupos E y F están limitados a Sierra Leona, y los grupos C y D están en Liberia. (Figura 6)



**Figura 6.** Distribución geográfica de las formas genéticas del VIH en el África.

## *Europa*

En Europa el subtipo B es la principal forma genética que conforma el 60% de la epidemia, mientras que el subtipo A el 11%, el subtipo C el 5% y otros subtipos conforman el 11%<sup>33,34</sup>. Entre los usuarios de drogas intravenosas (UDI) circula la forma CRF03\_AB principalmente en Stalingrado, Ucrania y Bielorusia, países en los que los subtipos A y B co-circulan. En Portugal el subtipo B es predominante en un 50.2% seguido del subtipo G (21.7%), A (17.5%) y F (5.5%), mientras que los subtipos C, D, H y J se encuentran en un 5.1%<sup>35</sup>. El subtipo F se encuentra en el este de Europa principalmente en Rumania<sup>36</sup>. El subtipo I se identificó por primera vez en Chipre y Grecia y se ha dispersado por todo el mediterráneo<sup>37</sup>. En Francia e Italia, a pesar de la migración africana, mantiene una prevalencia mínima de subtipos no-B como son el subtipo A, C, D, F, G, H, I, J y K estando éstos entre un 4 y 7% en la población<sup>18,38</sup>. En el Reino Unido los inmigrantes han sido responsables de la dispersión de una gran cantidad de formas recombinantes tales como: CRF01\_AE, CRF14\_BG, CRF05\_DF, CRF06\_cpx, CRF11\_cpx y CRF02\_AG así como diferentes formas únicas recombinantes (URF)<sup>39</sup>. El grupo B del VIH-2 se ha identificado en Portugal, España y Francia, esta dispersión se debe principalmente a inmigrantes africanos y a la proximidad de ambas poblaciones<sup>40-43</sup>.(Figura 7)

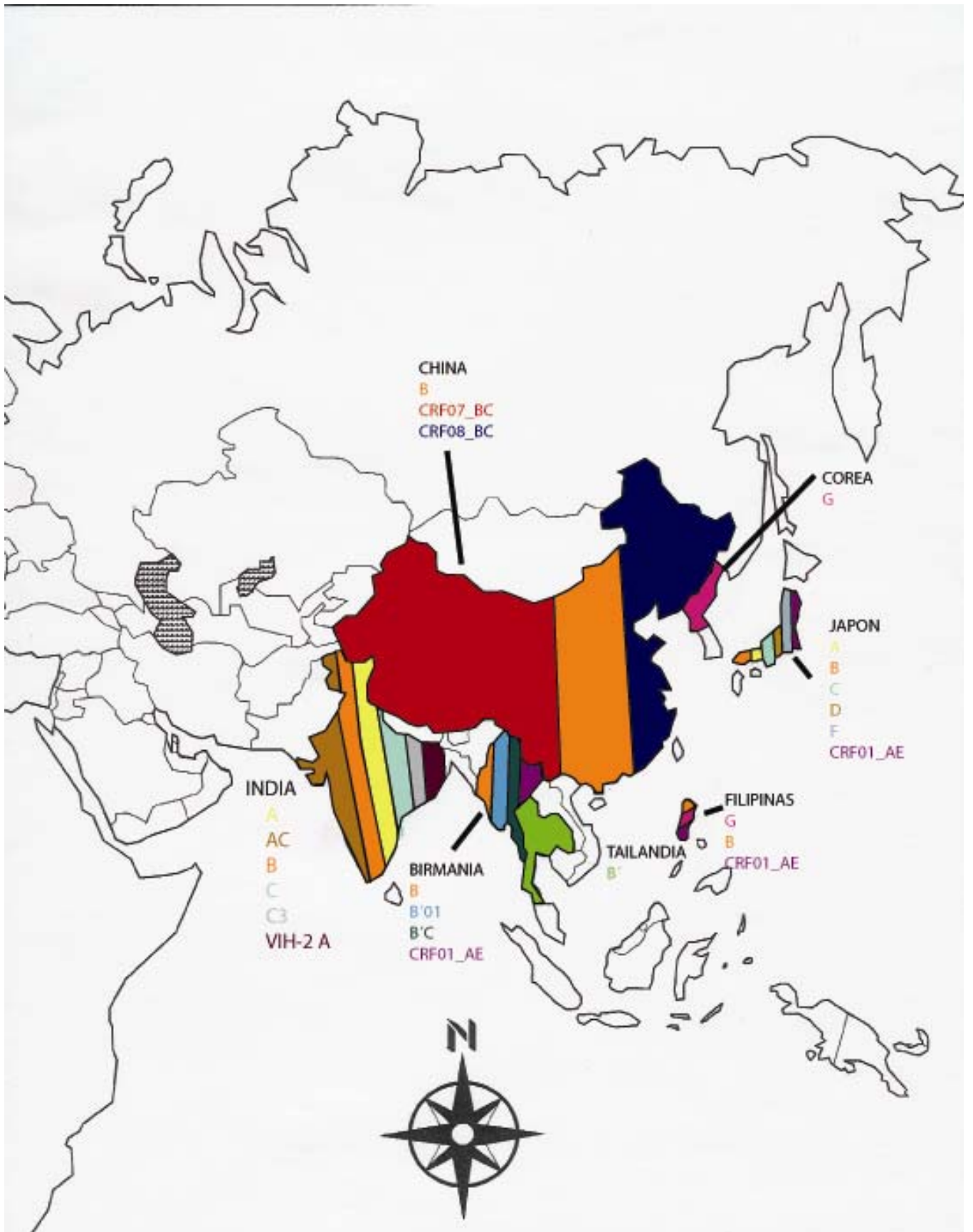


**Figura 7.** Distribución geográfica de las formas genéticas del VIH en el Europa.



## *Asia*

El subtipo C es predominante en India. En la región se ha observado una prevalencia del 78.4%, de los cuales el 60% corresponde a una variante conocida como C3<sup>44</sup>. Se han detectado, en menor proporción, el subtipo A, B y una variante recombinante AC<sup>45</sup>. Asimismo se ha detectado el subtipo A del VIH-2, igual al encontrado en África sin mostrar divergencia alguna<sup>46</sup>. En el sureste de Asia predomina el recombinante CRF01\_AE y una variante del subtipo B, reconocida como B' de Tailandia, esta variante usualmente predomina entre los usuarios de drogas intravenosas (UDI) y se ha dispersado principalmente entre este grupo de riesgo<sup>47-50</sup>. Existen dos recombinantes estrechamente relacionados, CRF07\_BC y CRF08\_BC, que se están dispersando entre los usuarios de drogas intravenosas en el noroeste (Provincia de Xinjiang) y sureste (Provincia de Guangxi) de China, respectivamente<sup>51,52</sup>. La distribución del VIH-1 en Japón es: subtipo B (74%), CRF01\_AE (20%), C(3.6%), A(2.0%), F(1.0%), D (0.5%)<sup>23</sup>. En Birmania se han encontrado formas únicas recombinantes como son B'01 y B'C, además del subtipo B y el recombinante CRF01\_AE<sup>45</sup>. El subtipo B tiene una menor predominancia en la mayor parte de Asia encontrándose en las Filipinas, Península Malaya, China y Japón. El subtipo H se ha encontrado en el centro de Asia y junto con el subtipo G, encontrado en Taiwán y Corea, son los menos predominantes<sup>23,53</sup>. (Figura 8)



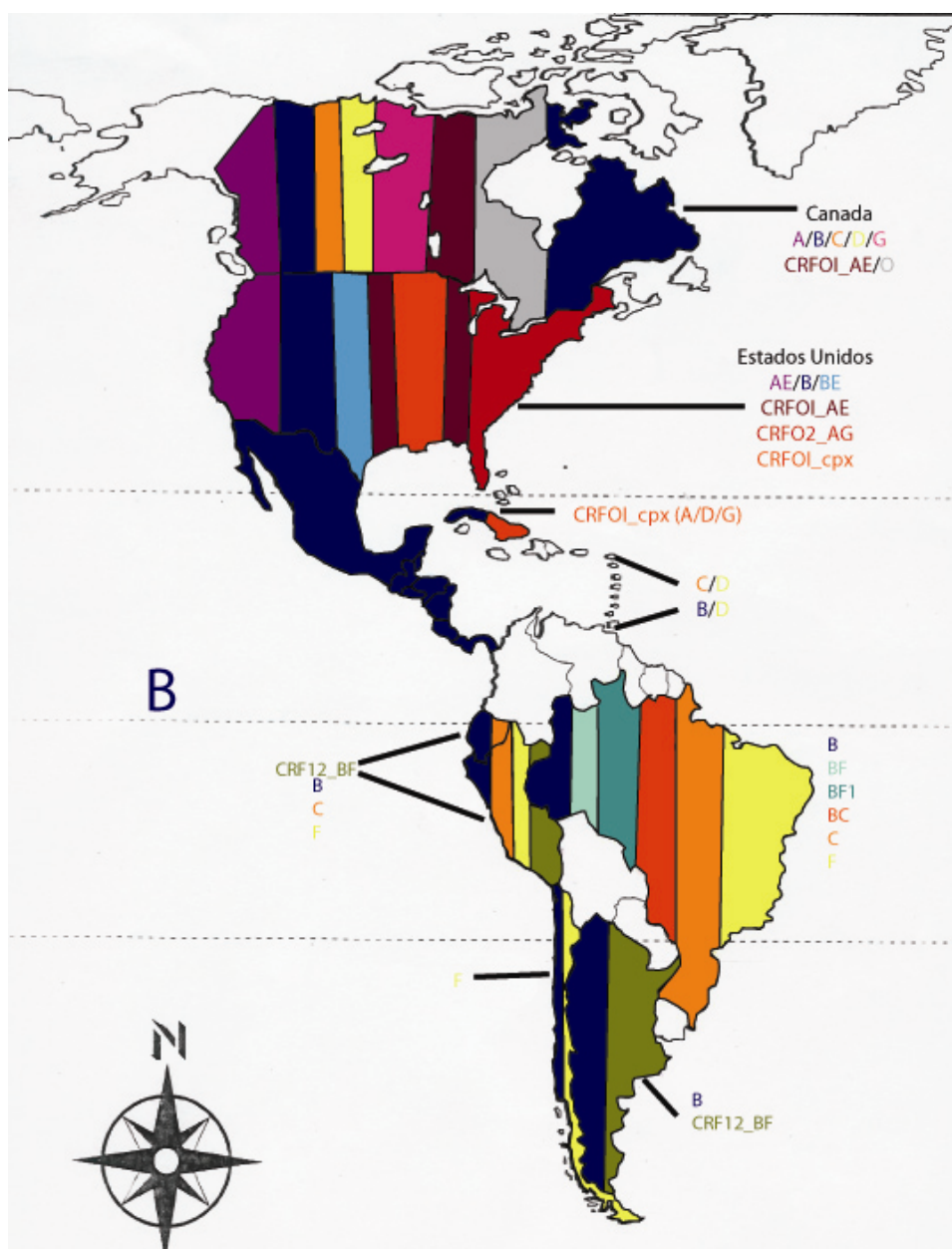
**Figura 8.** Distribución geográfica de las formas genéticas del VIH en el Asia.

### *América y el Caribe*

Al igual que en Europa, el subtipo B del VIH-1 es predominante en esta región del mundo, variando desde un 98% hasta un 48% en los diferentes países<sup>54</sup>. Aunque el subtipo B muestre un claro efecto fundador (poca variabilidad genética en una población debido a que su origen se dio con pocos individuos) en esta parte del mundo, los subtipos C, F, D y A también se han identificado aunque su prevalencia es menor. En Estados Unidos se han encontrado formas recombinantes tales como CRF01\_AE, CRF02\_AG, CRF09\_cpx , AE y BE en ciertos grupos de la población<sup>55,56</sup>. En Canadá se han identificado el grupo O, y los subtipos A (A1 y A2), C, D y G y el recombinante CRF01\_AE<sup>57,58</sup>. En México, por un solo estudio de epidemiología molecular, se ha identificado exclusivamente al subtipo B<sup>59</sup>. Los países de Centroamérica han sido poco estudiados, sin embargo se sabe que en Honduras predomina el subtipo B<sup>60</sup>. En Sudamérica, al igual que en el resto del continente, la epidemia es causada por el subtipo B, aunque se han identificado otros subtipos. En Brasil, los subtipos C y F han alcanzado una prevalencia del 8%, mientras que se han presentado de forma aislada los recombinantes BF, BF1 y BC<sup>61-63</sup>. En Argentina se han encontrado los recombinantes CRF12\_BF debido a la co-circulación de ambos subtipos, al parecer la prevalencia de este recombinante es mayor entre los usuarios de drogas intravenosas<sup>64</sup>. En Chile se muestra una mayor prevalencia del subtipo F, aunque el subtipo B sigue siendo más frecuente<sup>65</sup>. En Ecuador, Perú, Bolivia, Venezuela, Colombia y Uruguay se presenta el mismo patrón que en los países antes mencionados, el recombinante CRF12\_BF y los subtipos B, F y C co-circulan en esta zona dando pie al aumento de éstas formas genéticas pero sin rebasar el predominio del subtipo B en todas<sup>66-68</sup>.

En los países del caribe el subtipo B es igualmente el dominante y a pesar de los pocos estudios realizados en esta zona se han encontrado algunos casos de subtipos C y D.

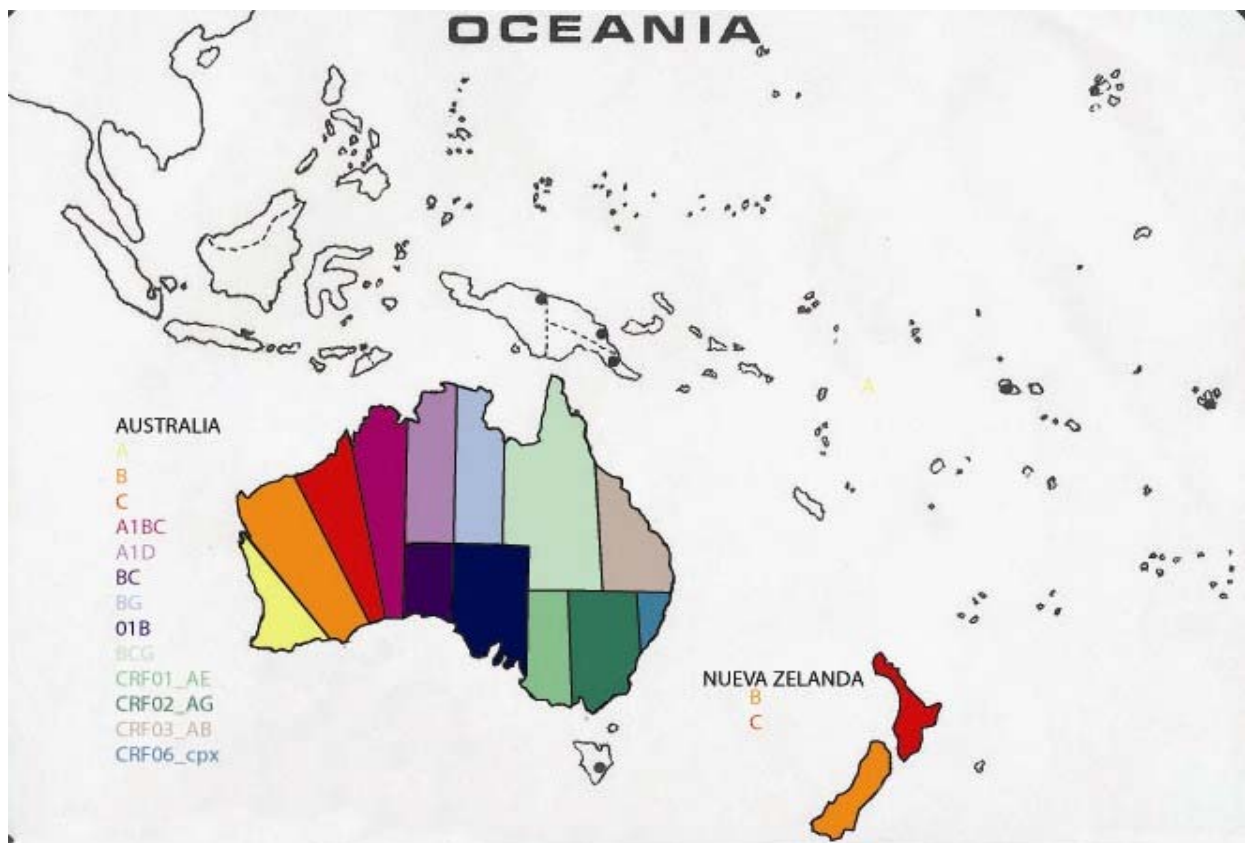
En Antigua y Trinidad y Tobago se han encontrado recombinantes BD<sup>69</sup>. En Cuba recientemente se ha encontrado un nuevo recombinante ADG clasificado como CRF19\_cpx<sup>70</sup>. Resulta difícil conocer la epidemiología molecular del VIH-1 en América Latina y el Caribe debido al alto costo de los estudios, de ahí que se tenga un conocimiento precario la distribución genética del VIH en esta zona. (Figura 9)



**Figura 9.** Distribución geográfica de las formas genéticas del VIH en América y el Caribe.

## Oceanía

En Australia, país poblado principalmente por inmigrantes, predomina el subtipo B. La población de vietnamitas usuarios de drogas intravenosas infectados por el virus presentan un 50% de infección por el recombinante CRF01\_AE<sup>71</sup>, mismo recombinante que predomina en el sureste de Asia. Otros subtipos, formas recombinantes y únicas recombinantes encontrados en Australia aunque con menor frecuencia son el C, A, BC, CRF06\_cpx, CRF02\_AG, CRF03\_AB, 01B, BG, A1D, A1BC y BCG<sup>54</sup>. En Nueva Zelanda solo circulan los subtipos B y C, y hasta el momento no se han encontrado formas recombinantes<sup>72</sup>. En las islas Fiji solo se encuentra el subtipo A, pero al igual que en los países de América Latina y el Caribe los estudios son escasos en esta región<sup>54</sup>. (Figura 10)



**Figura 10.** Distribución geográfica de las formas genéticas del VIH en Oceanía

## 2.1 Tratamiento Antirretroviral

Los fármacos antirretrovirales actúan principalmente en el ciclo de replicación del virus. Se dividen en tres familias de acuerdo a la enzima en la que actúan y son: Inhibidores de Proteasa (PI), Inhibidores de la Transcriptasa Reversa análogos de nucleósidos (NRTI) e Inhibidores de la Transcriptasa Reversa no análogos de nucleósidos (NNRTI).

El uso del tratamiento antirretroviral se comenzó a usar en Estados Unidos entre los años de 1995 y 1996, mientras que en México el uso de éstos comenzó entre los años de 1997 y 1998. Durante ésta época, el tratamiento antirretroviral era usado exclusivamente en las instituciones de seguridad social, y no fue sino hasta el año 2000 en que todas las personas infectadas por el virus tuvieron acceso a los fármacos (*Datos no publicados*). En los primeros años del uso del tratamiento antirretroviral, las instituciones de seguridad social emplearon los regimenes de forma semejante a Estados Unidos y Europa y las combinaciones usadas en ese entonces eran: 2 análogos de nucleósidos (generalmente AZT o D4T + 3TC) y 1 inhibidor de proteasa (Indinavir o Nelfinavir o Ritonavir o Saquinavir). En algunos centros la combinación más usada fue: 2 análogos de nucleósidos (AZT o D4T + 3TC) más 1 inhibidor de la transcriptasa reversa no análogo de nucleósidos (Efavirenz) en el 85% de los casos, mientras que en el 15% restante se usaba un inhibidor de proteasa en lugar del inhibidor de la transcriptasa reversa no análogo de nucleósidos. Sin embargo, en muchas regiones de la República se utilizaban únicamente 2 análogos de nucleósidos. Actualmente el régimen de primera intención consta de 2 inhibidores de la transcriptasa reversa análogo de nucleósidos más un inhibidor de la transcriptasa reversa no análogo de nucleósidos .

## **2.2 Mutaciones de Resistencia a Fármacos**

Las mutaciones de resistencia a fármacos se presentan en todos los patógenos atacados clínicamente por medicamentos. En el VIH-1 la aparición de mutaciones de resistencia se da principalmente por la presión de selección causada por éstos. Cabe mencionar que existen otros factores involucrados en la aparición de mutación como son: a) la alta producción de viriones, b) los errores de la enzima transcriptasa reversa durante el proceso de replicación y c) la presión de selección por el sistema inmune del hospedero.

Las mutaciones de resistencia a fármacos se han clasificado en dos tipos: las mutaciones de resistencia primaria y mutaciones de resistencia secundaria. Las mutaciones de resistencia primaria son aquellas adquiridas durante la infección y que por si solas confieren resistencia a alguno de los fármacos. Las mutaciones de resistencia secundaria son aquellas adquiridas durante el tratamiento antirretroviral, en donde se presenta una acumulación de mutaciones en el genoma y que como consecuencia se da la resistencia. Dentro de esta clasificación se han realizado diversas sub-clasificaciones para tratar de explicar todas las mutaciones adquiridas por el virus durante el ciclo de la enfermedad. Con motivo de un mejor entendimiento para este estudio se hablará de mutaciones de resistencia primaria a aquellas mutaciones que por si solas presenten resistencia a algún fármaco y en donde el paciente no haya tenido un tratamiento antirretroviral previo. Por otro lado, se mencionarán como: a) mutaciones menores a aquellas mutaciones que no le confieran resistencia a fármacos pero que son diferentes a la cepa silvestre, y b) mutaciones mayores a aquellas que han sido reportadas como mutaciones que confieran resistencia baja o intermedia a fármacos.

Actualmente se ha reportado un incremento de entre 14% al 7% en las mutaciones de resistencia primaria en países desarrollados como son el Reino Unido, Italia, Francia y Estados Unidos. Desafortunadamente en México no se conoce con exactitud la prevalencia de estas mutaciones en la población, se ha documentado que la frecuencia de mutaciones de resistencia primaria es del 1% al 6% en la población<sup>73</sup>.



### **3. MÉTODO**

#### **3.1 JUSTIFICACIÓN**

Cada día es de mayor importancia conocer la divergencia genética del VIH, así como su distribución geográfica. La entrada de nuevos subtipos o formas recombinantes a las poblaciones es cada vez mayor, por lo que el monitoreo constante es necesario para determinar la presencia de las nuevas formas en cada región.. Aunado a esto, las estrategias de una posible vacuna se han enfocado en utilizar virus regionales como inmunógenos debido a la gran divergencia del virus en las diferentes poblaciones, por lo que es necesario conocer la secuencia consenso de los virus predominantes de cada región geográfica.

En México los datos disponibles de la epidemia de VIH/SIDA son incompletos, no representativos y no actualizados. Debido a esto, no hay datos suficientes que nos indiquen el origen de la epidemia y las variaciones genéticas del virus en las diferentes regiones del país. Un estudio epidemiológico molecular bueno nos permitirá entender mejor la epidemia en nuestro país, no solo en el sentido de encontrar qué subtipos pueden estar circulando en nuestra población, sino en el aspecto de encontrar patrones de transmisión que permitan tomar acciones que ayuden en el control de la enfermedad.

Por otro lado, la prevalencia de mutaciones de resistencia primaria se ha incrementado en diversas poblaciones debido al uso de los fármacos antirretrovirales. En nuestro país no hay estudios que muestren de forma confiable la prevalencia de estas mutaciones en la población.

Por este motivo, es importante conocer cuál es la prevalencia de este tipo de mutaciones en nuestro país permitiéndonos de esta forma establecer mejores decisiones terapéuticas y políticas de manejo de la epidemia.

### **3.2 HIPÓTESIS**

En México, se encontrará como predominante el subtipo B del VIH-1 en los diferentes estados de la República, debido a la alta prevalencia de este subtipo en el Continente Americano.

En la población mexicana la prevalencia de mutaciones de resistencia primaria será menor que la encontrada en otros países, debido al uso relativamente reciente de los fármacos antirretrovirales.

### **3.3 OBJETIVOS**

#### **3.3.1 OBJETIVO GENERAL**

El objetivo de este estudio fue conocer la epidemiología molecular del VIH/SIDA en México, mediante la secuenciación de regiones del genoma del virus. Estableciendo las relaciones filogenéticas del virus en diferentes regiones del país así como con virus de diferentes regiones geográficas.

#### **3.3.2 OBJETIVOS PARTICULARES**

Determinar el subtipo de VIH-1 predominante en la población mexicana mediante la secuenciación del gen *pol* del virus.

Establecer las relaciones filogenéticas del virus en diferentes regiones del país así como con virus de diferentes regiones geográficas.

Establecer la prevalencia de mutaciones de resistencia primaria a fármacos en la población mexicana y compararlas con la prevalencia de éstas en otras poblaciones.

### **3.4 METODOLOGÍA**

#### **3.4.1 POBLACIÓN DE ESTUDIO**

El diseño del estudio es transversal. Los criterios de inclusión para el reclutamiento de pacientes al estudio fueron: a) que tuvieran infección por VIH, y b) que no hubieran estado expuestos previamente al tratamiento antirretroviral.

Para reclutar a los pacientes se invitó a médicos de diferentes estados de la República, pertenecientes a la red de prestadores de servicio de VIH/SIDA, para evaluar las condiciones clínicas de cada paciente candidato al estudio.

El reclutamiento de pacientes se llevó a cabo entre los meses de febrero del 2005 a marzo del 2006. Se reclutaron a 105 pacientes de 4 regiones del país, pertenecientes a diversos centros de salud de los Servicios Estatales de Salud (SESA) (Cuadro 2). Cabe mencionar que el tamaño de muestra para este estudio comprende 105 pacientes de 829 planteados para el estudio completo. Con motivos de obtener una muestra más representativa para un estudio de epidemiología molecular del VIH-1 en México, el tamaño de muestra para el estudio completo se obtuvo considerando el porcentaje de personas infectadas al año por entidad federativa.

A cada paciente se le dio a leer y firmar una carta de consentimiento informado para participar de forma voluntaria en el estudio. Se recolectaron los datos sociodemográficos de cada paciente mediante el uso de un cuestionario (Ver Anexo). De cada paciente seleccionado se obtuvo un muestra de 22 mL de sangre por venopunción utilizando el equipo Vacutainer y EDTA como anticoagulante.

### **3.4.2 ENVÍO Y ALMACENAMIENTO DE MUESTRAS**

El envío de muestras se llevo a cabo siguiendo los estándares nacionales e internacionales para el manejo de muestras biológico-infecciosas<sup>74</sup>. Las muestras sanguíneas fueron tomadas en tubos de plástico (contenedor primario) debidamente rotulados (clave y nombre del paciente) y sellados con papel parafilm para evitar derrames. El juego de tubos de cada paciente se envolvió con material absorbente y “papel burbuja” dentro de un contenedor resistente (contenedor secundario). El contenedor secundario se protegió con una red de unicel dentro del contenedor externo para evitar daños físicos o ambientales y permitir su manejo seguro. Se utilizó el servicio de mensajería con sistema de entrega de paquetería en 24 hrs.

Las muestras para conteo de subpoblaciones de linfocitos se procesaron inmediatamente. Para los procedimientos de carga viral y genotipo se separó el plasma por centrifugación a 2500 rpm por 20 minutos y se almacenó en tubos de plástico de 1 mL debidamente rotulados a -80° C, hasta su procesamiento.

### **3.4.3 GENOTIPIFICACIÓN**

El procedimiento de laboratorio empleado en este estudio se llevó a cabo utilizando el estuche comercial ViroSeq<sup>TM</sup> HIV-1 Genotyping System (Celera Diagnostics) para la extracción de RNA viral a partir de muestras de plasma, transcripción inversa, amplificación y la posterior secuenciación del gen *pol* del VIH. El procedimiento se divide en: i) extracción de ARN viral, ii) amplificación del gen *pol* y iii) secuenciación.

## **i. Extracción de ARN viral**

ViroSeq™ HIV-1 Genotyping System, está diseñado para el aislamiento de RNA viral del VIH-1 a partir de 1 mL de plasma humano en EDTA. La muestra de plasma debe centrifugarse a 17,000 rpm durante 60 minutos a 4°C para concentrar las partículas virales de VIH-1 en un “pellet”. Después de remover el sobrenadante, las partículas virales son lisadas para liberar el RNA viral usando el buffer de lisis viral el cual contiene como elementos principales detergentes como Tween 80 y SDS. El RNA viral es precipitado con isopropanol, centrifugado y posteriormente lavado con etanol al 70%. Finalmente, el RNA viral se debe dejar secar y resuspender en diluyente de RNA.

## **ii. Amplificación**

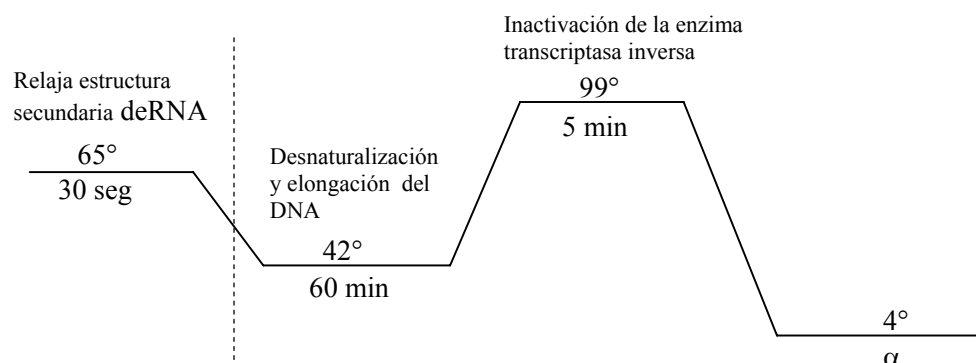
El proceso de amplificación mediante la reacción en cadena de la polimerasa (PCR: Polimerase Chain Reaction)<sup>75</sup> usado en este estudio consiste en dos tipos de amplificación. La primera consiste en llevar a cabo la transcripción inversa mediante una RT-PCR, es decir convertir el ARN viral en cADN. La segunda consiste en amplificar exponencialmente la muestra de cADN.

### 1. Transcripción inversa

Se amplifican una región de 1.3 Kb del gen *pol* del VIH-1, que incluye el gen completo de la proteasa y aproximadamente dos terceras partes del gen de la transcriptasa reversa. Después de la extracción de las muestras, se transfieren 10 µL del RNA resuspendido a un tubo MicroAmp<sup>R</sup> de 0.2 mL y se calienta a 65°C para relajar la estructura secundaria del RNA.

Posteriormente, se le agregan 10  $\mu\text{L}$  de la mezcla de reacción, la cual contiene la reversa transcriptasa de Virus de la Leucemia Murina (MuLV), un primer sencillo, ditiotreitól (DDT), inhibidor de RNasa y deoxinucleótidos trifosfatados (dNTPs). La reacción se coloca en el termociclador de Applied Biosystems (GeneAmp<sup>R</sup> PCR system 9700) y se deja incubar por una hora a 42° C, para generar hebras sencillas de DNA complementario (cADN).

La hebra sencilla de RNA del VIH-1 presente en la muestra es convertida a una doble hebra híbrida de RNA-cDNA durante este proceso. Las muestras son calentadas nuevamente a 99°C durante 5 minutos para inactivar la acción de la enzima transcriptasa inversa. Finalmente, las muestras son enfriadas a 4°C (Fig. 10).



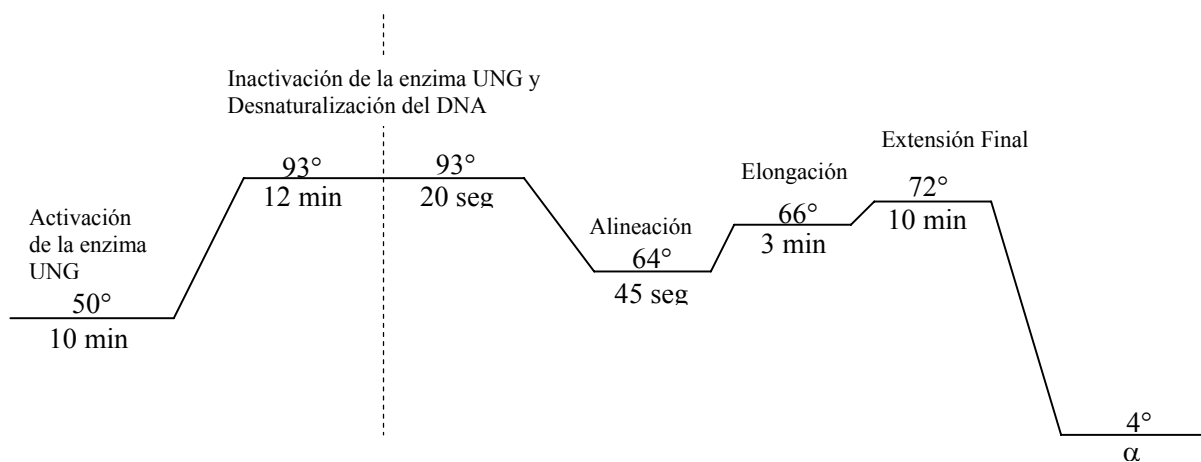
**Figura 10.** Condiciones de tiempo y temperatura para la reacción de transcripción inversa en PCR (RT-PCR)

## 2. PCR de las muestras

Seguido de la transcripción reversa del RNA blanco, se agregan 30  $\mu\text{L}$  a la muestra de la mezcla de reacción conteniendo AmpliTaq Gold<sup>R</sup> DNA polimerasa, iniciador sentido y antisentido, dNTPs y AmpErase<sup>R</sup> uracil-N-glicosilasa (UNG). Las muestras son colocadas en el termociclador (9700) y se deja correr dentro de un protocolo específico de variaciones en las condiciones de temperatura por 4.5 horas.

El primer paso en la reacción de PCR consiste en un ciclo sencillo a 50°C por 10 minutos. En este paso, la enzima UNG controla la contaminación destruyendo cualquier especie de DNA conteniendo deoxiuridina. La mezcla de PCR de ViroSeq contiene deoxiuridina trifosfatada (dUTP), la cual es incorporada al DNA amplificado durante el proceso de PCR. Cualquier resto de amplificación anterior conteniendo deoxiuridina es seleccionado y destruido por la UNG, mientras que la muestra de cDNA, que no contiene deoxiuridina, permanece intacta.

Después del control de contaminación, la muestra es calentada a 93°C por 12 minutos para inactivar la AmpErase<sup>R</sup> UNG, activar la AmpliTaq<sup>R</sup> DNA polimerasa y desnaturalizar el híbrido de RNA-cDNA. Posteriormente, se llevan a cabo 3 pasos que se repiten en 40 ciclos. Estos pasos son 93°C durante 20 segundos para desnaturalizar el DNA, 64°C durante 45 segundos para alinear los primers y 66°C durante 3 minutos para extender el primer y crear una doble hebra de DNA. Durante este proceso, un fragmento de 1.8 Kb del gen *pol* del VIH-1 es amplificado. Las moléculas de DNA amplificado, también llamados amplicones, son calentados a 72°C durante 10 minutos para permitir la extensión final, seguido de un descenso de temperatura a 4°C para preservar la muestra hasta el siguiente proceso (Figura 11).



**Figura 11.** Condiciones de tiempo y temperatura para la reacción de PCR

### iii. Secuenciación

La secuenciación se realizó mediante la técnica de Sanger modificada. Brevemente la secuenciación automatizada fluorescente, se basa en la incorporación de marcadores fluorescentes durante la extensión de DNA usando didesoxinucleótidos trifosfato (ddNTP) como terminadores. Con la marcación de terminadores, cada uno de los cuatro ddNTPs presenta una fluorescencia específica. La cadena en crecimiento es terminada y marcada simultáneamente cuando un ddNTP es adicionado. Los fragmentos generados son separados por tamaño mediante una electroforesis. Este proceso crea una secuencia de DNA líder de tal forma que cada producto de extensión es una base más grande que el anterior. Los fluorocromos son excitados mediante un láser y el sistema óptico en el secuenciador capta la fluorescencia emitida por cada terminador asignando una base a cada uno dependiendo de la longitud de onda de emisión. La fluorescencia de los terminadores es la siguiente: A = verde, C = azul, G = amarilla y T = roja.



El proceso del ciclo de secuenciación consta de 5 pasos:

- a. Purificación del producto de PCR
- b. Cuantificación del producto de PCR
- c. Reacción de Secuenciación
- d. Purificación del producto de la reacción de secuenciación
- e. Secuenciación Automatizada

**a. Purificación del producto PCR**

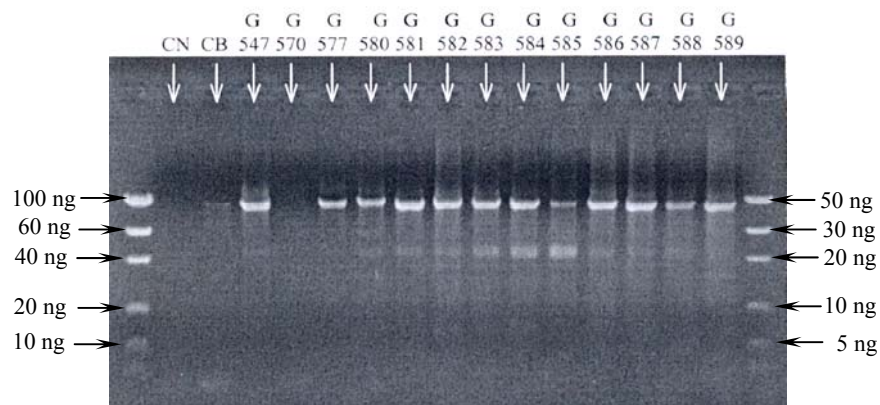
Después de la PCR, los productos obtenidos son purificados para remover cualquier primer o dNTP no incorporado. Para esto se utilizan membranas de ultrafiltración las cuales permiten el paso de los dNTPs y primers evitando el paso del fragmento de DNA. Para lograr esto, se coloca una columna dentro de los tubos de 1.5 mL para cada producto, se le agregan 300µL de agua bi-distilada junto con los 25µL del producto de PCR obtenido, se centrifuga a 3,000 rpm durante 15 minutos. Se le agregan 35µL de agua bi-distilada se invierte la columna en un tubo de 1.5 mL limpio y se vuelve a centrifugar a 3,00 rpm durante 5 minutos para obtener aproximadamente entre 30 – 40µL de producto de PCR purificado.

**b. Cuantificación del producto de PCR**

Los productos de PCR se cuantifican por comparación con un patrón de concentración conocida por corrimiento electroforético en gel de agarosa. Se toman 5 µL del producto de PCR purificado y se le agrega un volumen igual de buffer de carga para gel de agarosa, el buffer de carga es una mezcla de glicerol y un colorante para permitir la difusión del producto obtenido en el gel.

Las muestras se colocan en un gel de agarosa al 1% conteniendo bromuro de etidio (1µl de bromuro de etidio por cada 100 mL de gel). Dos cargas de marcador de peso molecular, de concentración conocida deben también correrse en el gel. El gel se coloca en la cámara de electroforesis con suficiente buffer TBE (Tris-Borato-EDTA) para permitir una difusión constante y pareja en el gel. Se deja correr aproximadamente por 1 hr. Posteriormente el gel se expone a luz ultravioleta (UV) para poder comparar las bandas obtenidas con las bandas del marcador de peso molecular.

A partir de esta comparación, se puede asignar una concentración de manera semi-cuantitativa al DNA obtenido (Fig. 12). El producto de PCR restante es diluido de forma óptima para las concentraciones necesarias para la reacción del ciclo de secuenciación.



**Figura 12.** Cuantificación del producto de PCR. Gel de agarosa al 1%, las bandas claras representan el producto de PCR, la intensidad de éstas se compara con las bandas de las orillas para determinar la cantidad de ADN obtenido en el proceso.

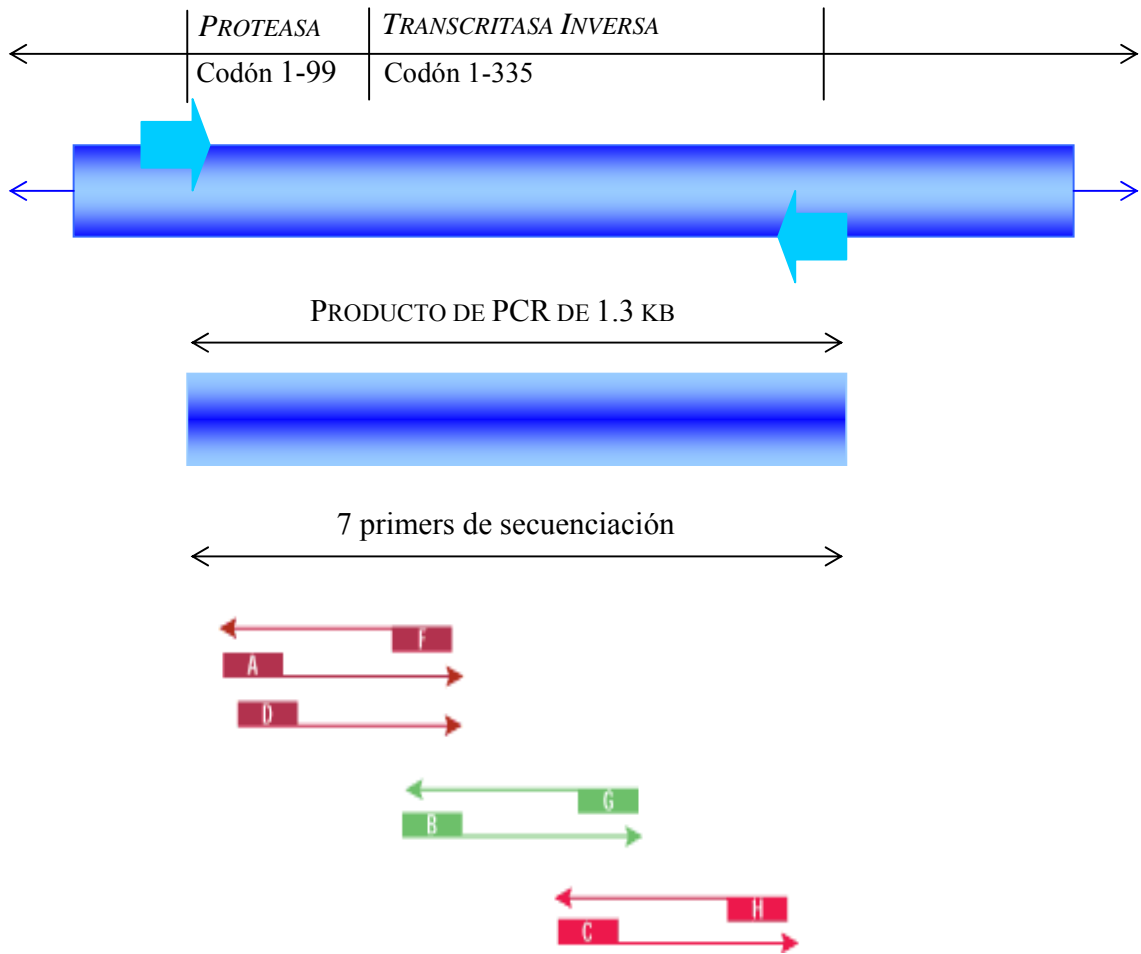
### c. **Reacción de Secuenciación**

El estuche empleado usa siete primers (cuatro sentido y tres antisentido) para secuenciar la región completa del gen proteasa (del codón 1 al 99) y dos terceras partes del gen TR (del codón 1 al 335) (Fig. 12).

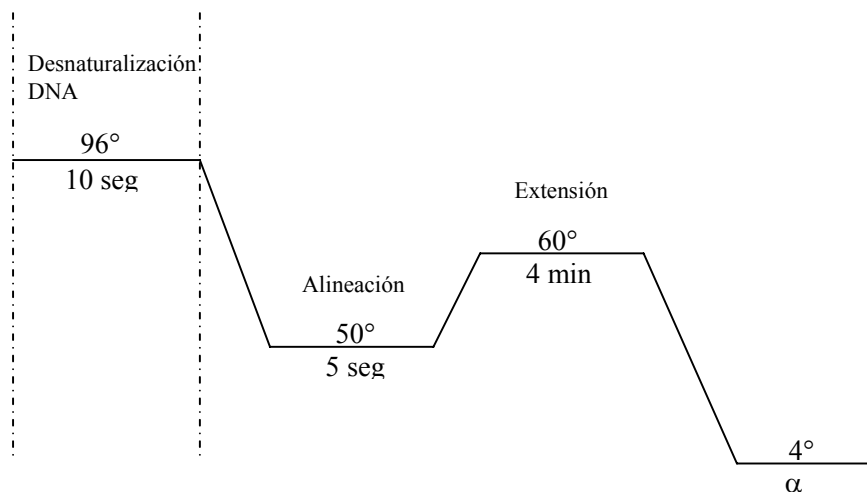
La química de BigDye<sup>R</sup> usada en el sistema ViroSeq permite una resolución de aproximadamente 600 bases en el modelo 3100-*Avant* Genetic Analyzers. Por lo tanto, se necesita de múltiples primers para poder completar el fragmento de 1.3 Kb de DNA. Con este arreglo de primers se asegura que la mayor parte del fragmento a analizar esté cubierto con por lo menos dos secuencias.

Se mezclan 6 µL de cada uno de los 7 primers de secuencia de ViroSeq, los cuales además contienen los demás reactivos necesarios para la PCR, con 4 µL de muestra diluida y purificada.

La mezcla de reacción se incuba en un termociclador durante aproximadamente 2.5 horas. Existen tres pasos repetidos para 25 ciclos lo cuales son: 96°C por 10 minutos para desnaturalizar el DNA, 50°C por 5 segundos para alinear los primers y 60°C por 4 minutos para extender los primers. Al finalizar, las muestras deben ser colocadas a 4°C hasta el próximo proceso (Fig. 13).



**Figura 12.** Arreglo de los primers antisentido y sentido para la secuenciación del gen *pol*.



**Figura 13.** Condiciones de tiempo y temperatura para la reacción de secuenciación.

#### **d. Purificación del producto de la reacción de secuenciación.**

Es necesario remover los terminadores BigDye<sup>R</sup> no incorporados de las muestras para evitar interferencia en la secuenciación y análisis. Se pueden realizar dos diferentes métodos para purificar los productos de PCR de las reacciones de secuencia:

- Precipitación por isopropanol
- Precipitación por acetato de sodio/etanol

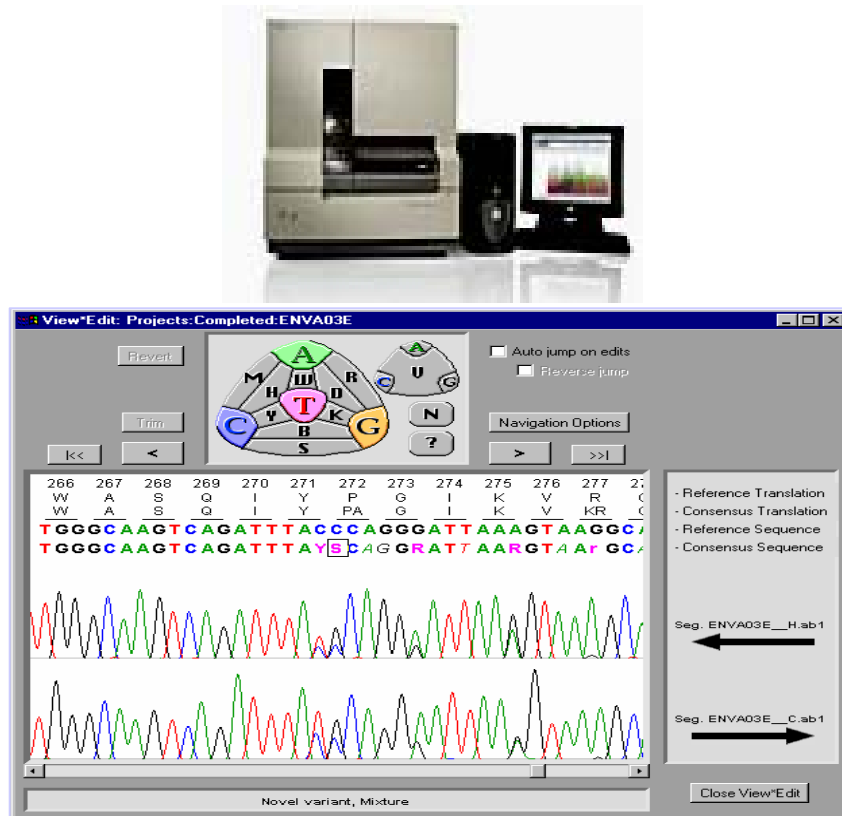
La precipitación por ambos métodos consiste en agregar isopropanol o acetato de sodio al producto de PCR, dejar incubar a temperatura ambiente durante 20 minutos y después centrifugar a 12,000 rpm durante 20 minutos. Se elimina el sobrenadante y se agrega etanol al 70%, se centrifuga a máxima velocidad y se elimina todo el etanol dejándolo secar por completo. El pellet se resuspende con formamida para poder ser leído en el secuenciador. La formamida evita la formación de estructuras secundarias y permite la separación de los fragmentos estrictamente por su tamaño.

#### **e. Secuenciación Automatizada**

La secuenciación se realizó en el 3100-*Avant* Genetic Analyzer (Applied Biosystems) de acuerdo a las instrucciones incluidos en el manual de operación.

Las secuencias obtenidas se analizaron en el ViroSeq HIV-1 Genotyping System Software, el cual toma los datos de secuencia de los siete primers de cada muestra y combina la información dentro de un solo proyecto. Durante el ensamblaje de las secuencias, una secuencia consenso es creada cubriendo el gen de la proteasa por completo y hasta el codón 335 del gen de la transcriptasa reversa.

Esta secuencia consenso es comparada con la secuencia HXB-2, secuencia de referencia del subtipo B, para marcar los puntos de variabilidad. Manualmente se confirman o modifican los puntos variables para asegurar líneas basales correctas (Fig. 14).



**Figura 14.** Secuenciador 3100-Avant Genetic Analyzer y Electroferograma.

### 3.4.4 Análisis Estadístico

Los datos de las variables sociodemográficas y biológicas de los pacientes se analizaron con estadísticas univariadas (medidas de tendencia central y de dispersión). En el caso de los cuadros de frecuencia generados para medir la relación entre las frecuencias de mutaciones de resistencia a fármacos en las diferentes poblaciones, así como la relación entre las etapas de la enfermedad y carga viral con las mutaciones a las diferentes familias de fármacos, se empleó el estadístico  $X^2$  (Chi cuadrada).

En los casos en que se generaba una tabla de 2x2 se usó la prueba exacta Fisehr. Todo el procedimiento de análisis estadístico se realizó con el paquete STATA versión 8.

### **3.4.5 Tipificación del VIH-1**

La tipificación del VIH-1 empleada en este estudio fue mediante el programa de tipificación HIVdb de la Universidad de Stanford (<http://hivdb.stanford.edu>), así como mediante el análisis filogenético con secuencias de referencia.

### **3.4.6 ANÁLISIS FILOGENÉTICO**

Para el análisis filogenético se utilizaron secuencias de referencia, así como secuencias aisladas en diferentes regiones geográficas. Las secuencias de referencia consideradas como consenso para los subtipos de VIH-1 se obtuvieron de la base de datos de VIH de Los Alamos National Laboratory (<http://hiv-web.lanl.gov>). Las secuencias de diferentes regiones geográficas se obtuvieron del Genbank siendo seleccionadas todas aquellas en donde el gen *pol* estuviera presente, las secuencias fueron: Argentina (AY037268, AY037269, AY563170), Brasil (AF005495, AY037282, AY727522), Ecuador (AY173959, AY173960), Colombia (AY561236, AY561237), Cuba (AY588971, AY900576), Estados Unidos (AY268493, AY314048), Uruguay (AF385934), Chile (AY536235), Holanda (AY423381), España (AF450097), Reino Unido (AJ271445), y Grecia (AY046058).

#### **3.4.6.1 CLUSTALX**

Se usó el programa CLUSTALX versión 1.81<sup>76</sup> para realizar las alineaciones de las secuencias obtenidas con las secuencias de referencia y las secuencias de diferentes regiones geográficas.

Este método permite alinear múltiples secuencias de tal forma que facilita el reconocimiento de regiones similares. De igual forma este programa permite el posterior análisis de las secuencias en los diferentes programas filogenéticos.

#### **3.4.6.2 NEIGHBOR-JOINING Y BOOTSTRAPPING**

Se realizaron los dendogramas y el bootstrap (100 replicas) mediante el método de Neighbor-Joining (PAUP 4.0b10) <sup>77</sup>. Las distancias filogenéticas entre las secuencias se calcularon usando el modelo GRT+I+G.

El método de Neighbor-Joining es un método de distancia y se basa en el principio de mínima evolución. Éste método emplea una serie de cálculos específicos para estimar un árbol. Los cálculos involucran la manipulación de una matriz de distancia generada en una alineación múltiple. A partir de ésta matriz base, el método primero calcula una red de divergencia del taxon (organismo o secuencia) con respecto al resto de los taxa (plural de taxon). Esta red se calcula como la suma individual de distancias del taxon. Posteriormente, utiliza la red de divergencia para calcular una matriz de distancia corregida. De esta forma, el método es capaz de encontrar el par de taxa con la menor distancia corregida y calcular la distancia de cada uno de esos taxa al nodo que los une (las distancias que existen desde los dos taxa al nodo no necesita ser idéntica).

Una nueva matriz es creada en donde el nuevo nodo es sustituido por los taxa encontrados. Este proceso se repite con todas los taxa que se desean agrupar.

A pesar de tener una cierta noción de la topología del árbol (el orden en que las diferentes secuencias divergen), no se tiene la noción de que tan confiable es la agrupación. Para obtener la confiabilidad de un árbol se emplea el método de bootstrapping.



Éste método utiliza una submuestra de la matriz de distancia obtenida en la múltiple alineación y crea diversos árboles basados en esa submuestra. Este proceso se repite múltiples veces (generalmente se usan 1,000 réplicas aunque 100 son consideradas como lo mínimo) y el resultado se recopila para permitir un estimado de la confiabilidad del agrupamiento.

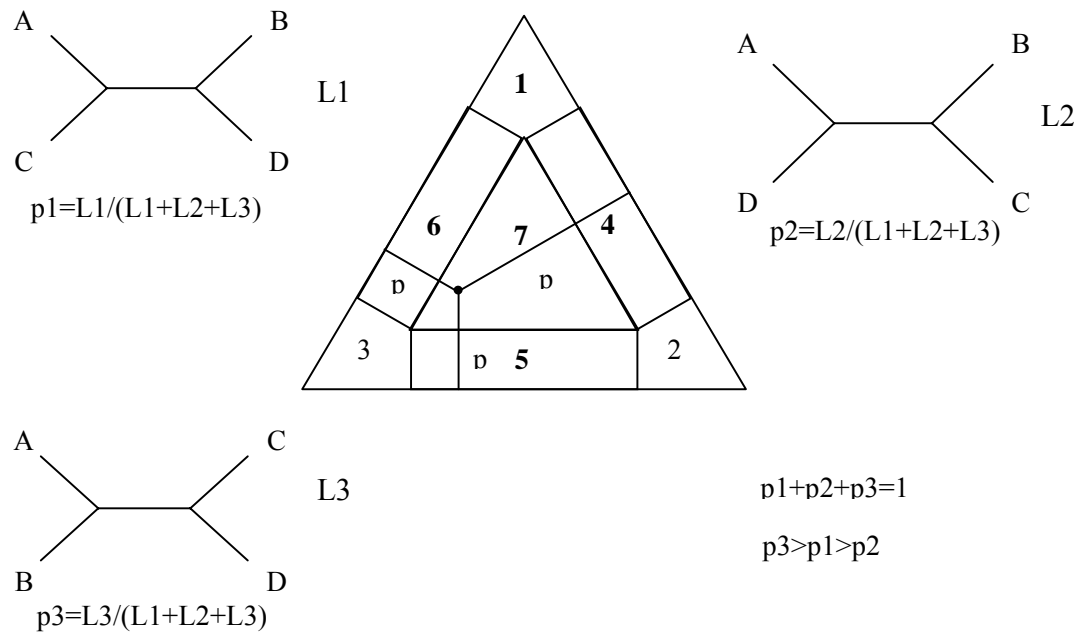
### **3.4.6.3 TREEVIEW**

Los diagramas de los árboles reconstruidos se trazaron en el programa TreeView versión 1.6. Éste programa permite exclusivamente dibujar los árboles filogenéticos a partir de los archivos obtenidos en los programas filogenéticos. De igual forma es posible modificar la apariencia del árbol de acuerdo a las necesidades y gusto de cada persona.

### **3.4.6.4 Tree Puzzle**

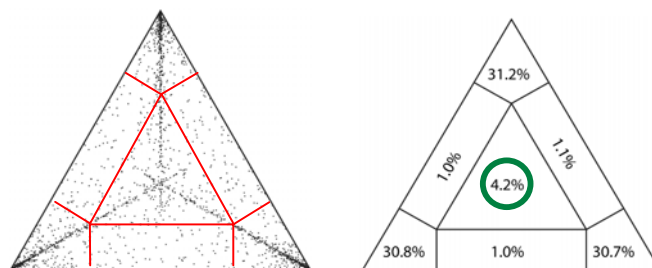
Para evaluar la cantidad de información filogenética en las secuencias se utilizó el método de mapeo de verosimilitud (likelihood mapping) mediante el programa Tree-Puzzle versión 5.2<sup>78</sup>. Éste método permite evaluar si la alineación de las secuencias contiene información suficiente para llevar a cabo un análisis filogenético. En resumen, para un grupo de 4 secuencias, existen 3 árboles posibles. Para cada uno de estos árboles se puede calcular el valor de verosimilitud (L1, L2 y L3), que es la probabilidad de observar este árbol bajo un modelo de evolución. Las probabilidades se estandarizan obteniendo los valores de p1, p2 y p3, los cuales se trazan en un triángulo equilátero. Cada esquina representa una topología del árbol. Si la intersección obtenida de los valores p queda en el área 1, 2 o 3, una de las tres topologías está claramente sostenida que las otras, en este caso el árbol está resuelto. Si la intersección es en las áreas 4, 5 o 6, es un árbol menos resuelto.

Pero si la intersección es en el área 7, los datos no contienen información suficiente que permita el análisis filogenético(Fig. 15).



**Figura 15.** Mapeo de verosimilitud. Valores de verosimilitud para un grupo de 4 secuencias.

Si se realiza éste mismo análisis, no con 4 secuencias sino con todas las 4 posibles combinaciones de un grupo de secuencias, se obtiene un triángulo similar al de la figura 16 en donde cada punto corresponde a un grupo de 4 secuencias. El porcentaje de puntos es entonces calculado. Se admite, desde el punto de vista biológico, que si el área central contiene más del 20% de los puntos, los datos no permiten la inferencia filogenética.



**Figura 16.** Mapeo de Verosimilitud. Porcentaje de puntos calculado por las 4 posibles combinaciones de un grupo de secuencias.

## 4. RESULTADOS

### 4.1 POBLACIÓN DE ESTUDIO

De los 105 pacientes reclutados, 77 (73%) fueron del sexo masculino y 28 (27%) del sexo femenino, el intervalo de edad fue de 17 a 60 años con un promedio de 33.81 años.

La distribución de los pacientes por entidad federativa fue de: 37% (n=39) provenientes del estado de Puebla, 29% (n=30) del estado de México, 18% (n=19) de Oaxaca, el 15% (n=16) del Distrito Federal, y el 1% (n=1) de Hidalgo. Los datos sociodemográficos de los pacientes se resumen en el cuadro 2.

**Cuadro 2.** Datos sociodemográficos

<b>Variables Sociodemográficas</b>	<b>Frecuencia</b>	<b>%</b>
<b>Sexo</b>		
Masculino	77	73.33
Femenino	28	26.66
<b>Total</b>	<b>105</b>	<b>100</b>
<b>Edad</b>		
< 20	2	1.90
20-29	37	35.23
30-39	40	38.09
40-49	21	20.0
50 +	5	4.76
<b>Total</b>	<b>105</b>	<b>100</b>
<b>Escolaridad</b>		
Analfabeta	8	7.61
Primaria	26	24.76
Secundaria	27	25.71
Preparatoria	12	11.42
Técnico	6	5.71
Licenciatura	21	20.0
Posgrado	5	4.76
<b>Total</b>	<b>105</b>	<b>100</b>

<b>Residencia Habitual</b>		
Distrito Federal	16	15.23
Estado de México	30	28.57
Oaxaca	19	18.1
Puebla	39	37.14
Hidalgo	1	0.95
<b>Total</b>	<b>105</b>	<b>100</b>
<b>Seguridad Social</b>		
IMSS	20	19.04
ISSSTE	2	1.90
SSA	83	79.04
<b>Total</b>	<b>105</b>	<b>100</b>
<b>Estado Civil</b>		
Soltero	65	61.90
Casado	14	13.33
Viudo	6	5.71
Divorciado	2	1.90
Unión Libre	18	17.14
<b>Total</b>	<b>105</b>	<b>100</b>

Los datos obtenidos de los marcadores clínicos muestran que el intervalo de carga viral va de un nivel indetectable (<50 copias/ml) a 1,130,000 copias/ml con un promedio de 266,038 copias/ml y una mediana de 238,574.8 copias/ml. Mientras que el intervalo de conteo de CD4<sup>+</sup> va de 1 célula/ml a 1,800 células/ml con un promedio de 267.02 células/ml y una mediana de 231.28 células/ml.

Ambos marcadores muestran una diferencia estadísticamente significativa ( $p < 0.001$ ), confirmando que a mayor carga viral menor cantidad de CD4<sup>+</sup> y por lo tanto un posible avance en la enfermedad.

Las etapas de la enfermedad se clasificaron de acuerdo al conteo de CD4<sup>+</sup> en: a) etapa temprana (CD4<sup>+</sup> > 499 células/ml); b) etapa intermedia (CD4<sup>+</sup> entre 200-499 células/ml), y c) etapa tardía (CD4<sup>+</sup> <200 células/ml). La distribución de la población de estudio en etapas de la enfermedad muestra que 10% de los pacientes se encontraban en la etapa temprana, 39% en la etapa intermedia y 51% en la etapa tardía (Cuadro3).

**Cuadro 3.** Etapa de la enfermedad según conteo de CD4+.

<b>ETAPA DE LA ENFERMEDAD*</b>	<b>FRECUENCIA</b>	<b>PORCENTAJE (%)</b>
Temprana (>499cel./ml)	10	10%
Intermedia (200-499 cel./ml)	41	39%
Tardía (<200cel./ml)	53	51%
<b>Total</b>	<b>104</b>	<b>100</b>

\* A un paciente no se le pudo realizar conteo de CD4+ por fragilidad capilar

#### **4.2 MUTACIONES DE RESISTENCIA A FÁRMACOS**

En la población de estudio, muy pocos pacientes (n=5) se encontraban libres de mutaciones, mientras que cerca del 50% presentaron de 2 a 3 mutaciones para cualquier familia de fármaco. Ver cuadro 4.

**Cuadro 4.** Número global de mutaciones de resistencia a fármacos

<b>Número de mutaciones</b>	<b>Frecuencia</b>	<b>%</b>
Ninguna	5	4.76
Una	15	14.28
Dos	22	20.95
Tres	31	29.52
Cuatro	17	16.19
Cinco	12	11.42
Seis	2	1.90
Siete	1	0.95
<b>Total</b>	<b>105</b>	<b>100.0</b>

Los datos genotípicos muestran las diferentes mutaciones presentes en el genoma viral para cada uno de los pacientes. Dichas mutaciones se dividen de acuerdo con la familia de fármacos utilizados durante el tratamiento antirretroviral frente a los que confieren resistencia. Las familias de fármacos se dividen en: a) inhibidores de proteasa (PI); b) inhibidores de la transcriptasa reversa análogos de nucleósidos (NRTI), y c) inhibidores de la transcriptasa reversa no análogos de nucleósidos (NNRTI).

De todos los pacientes que presentaron mutaciones, el 95% lo hicieron a los inhibidores de proteasa (PI), el 25% a los inhibidores de la transcriptasa reversa análogos de nucleósidos (NRTI), y el 16% a los inhibidores de la transcriptasa reversa no-análogos de nucleósidos (NNRTI). Ver cuadro 5.

**Cuadro 5.** Porcentaje de mutaciones por familia de fármaco.

<b>Mutaciones</b>	<b>Frecuencia</b>	<b>%</b>
PI	100	95.23
NRTI	26	24.76
NNRTI	17	16.19
Ninguna	5	4.76
<b>Total*</b>	<b>105</b>	<b>100</b>

\*El porcentaje marcado es con base en el total de pacientes y no en el total de mutaciones encontradas en la población de estudio, ya que hay pacientes que presentan mutaciones para cualquier familia de fármaco de forma conjunta.

Dentro de las mutaciones para inhibidores de proteasa (PI), 20% de los pacientes presentaron una mutación, mientras que el 50% de los pacientes presentaron de dos a tres mutaciones y cerca del 23% presentaron de cuatro a cinco mutaciones. Ver cuadro 6.

**Cuadro 6.** Número de mutaciones por paciente para inhibidores de proteasa encontradas en la población de estudio.

<b>Número de mutaciones</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Ninguna	5	4.76
1	21	20.00
2	28	26.67
3	25	23.81
4	12	11.43
5	12	11.43
6	2	1.90
<b>Total</b>	<b>105</b>	<b>100.00</b>

La mutación encontrada con mayor frecuencia fue la L63P que se presentó en el 86% (90/105) de la población, seguida de la mutación I93L que se presentó en el 53% (56/105). Ambas mutaciones son mutaciones menores. La mutación L33F, mutación mayor que confiere resistencia intermedia <sup>79</sup> se presentó en 3/105 pacientes. Ver cuadro 7.

**Cuadro7.** Mutaciones encontradas en la población de estudio para los Inhibidores de Proteasa (PI)

<b>PI</b>	<b>Frecuencia de mutaciones</b>	<b>%</b>
L10I/V	6	5.7
K20M/R	3	2.8
V32A	1	0.9
<b>L33F</b>	3	2.8
M36I	16	15.2
I54F	1	0.9
L63A/H/L/M/P/Q/S/T/V	90	85.7
A71V/T	24	22.8
V77I	48	45.71
I93L	56	53.3

Para la familia de los inhibidores de la transcriptasa reversa análogos de nucleósidos, el 4% de los pacientes presentaron 2 mutaciones conjuntas, mientras que el 21% presentaron una sola mutación (Cuadro 8).

**Cuadro 8.** Número de mutaciones por paciente para los Inhibidores de la Transcriptasa Reversa Análogos de Nucleósidos.

<b>Número de mutaciones</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Ninguna	79	75.24
1	22	20.95
2	4	3.81
<b>Total</b>	<b>105</b>	<b>100.00</b>

La mutación G333E/D, mutación mayor que favorece la resistencia a AZT, se presentó en un 11% (12/105) de la población, seguida de la V118I, mutación mayor que confiere resistencia intermedia, presente en 5 de los pacientes. Ver cuadro 9.

**Cuadro 9.** Frecuencia de mutaciones para análogos de nucleósidos de RT encontradas en la población.

<b>NRTI</b>	<b>Frecuencia</b>	<b>%</b>
M41R/L	2	1.9
E44D/K	2	1.9
T69N/S	3	2.8
D67N	1	0.9
V75L	1	0.9
Y115C/D/G/Y	1	0.9
V118I	5	4.7
G333E/D	12	11.4
L210F/L/S	4	3.8

Para la familia de los inhibidores de la transcriptasa reversa no-análogos de nucleósidos, el 13% de los pacientes presentaron una sola mutación y el 3% de los pacientes presentaron dos mutaciones (Cuadro 10).

**Cuadro 10.** Número de mutaciones por paciente para los Inhibidores de la Transcriptasa Reversa No-Análogos de Nucleósidos.

<b>Número de mutaciones</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Ninguna	88	83.81
1	14	13.33
2	3	2.86
<b>Total</b>	<b>105</b>	<b>100.00</b>



Las mutaciones mayores K103N/R y V179D/E/F se presentaron con igual frecuencia, encontrándose en 8/105 de los pacientes. La mutación G190A, mutación primaria, se encontró en un solo paciente (0.9%). Ver cuadro 11.

**Cuadro 11.** Frecuencia de mutaciones de resistencia para inhibidores de la transcriptasa reversa no-análogos de nucleósidos encontrados en la población de estudio.

<b>NNRTI</b>	<b>Frecuencia</b>	<b>%</b>
A98T	1	0.9
K103N/R	8	7.6
V108I	2	1.9
V179D/E/F	8	7.6
G190A	1	0.9

#### **4.3 COMPARACIÓN DE FRECUENCIAS ENTRE POBLACIONES**

Al hacer la comparación estadística entre las frecuencias de mutaciones primarias encontradas en la población del Reino Unido, la población de Estados Unidos y la población de estudio, se encontró que dentro de las mutaciones para los inhibidores de proteasa I54F, A71V, A71T, V77I e I93L se presentan en mayor frecuencia en la población de estudio que en las otras dos poblaciones. Por otro lado, las mutaciones M41L, K103N y G190A se encuentran en menor frecuencia en la población de estudio que en las dos poblaciones. Ver cuadro 12

**Cuadro 12.** Comparación de Frecuencias de Mutaciones entre las poblaciones del Reino Unido y Estados Unidos con México.

PI	México		UK		E.U.		p por Exacta de Fisher			
	N tot=105		N tot=335		N tot PI=2787		Mex vs. UK	Mex vs. USA		
	Mutación	Prevalencia	N con la Mut	Prevalencia	N con la Mut	Prevalencia	N con la Mut			
	L10I	4.76%	5	3.25%	11	8.70%	242			
	L10V	0.95%	1	.	.	2.60%	72			
	K20M	1.90%	2	1.63%	5	0.30%	8			
	K20R	0.95%	1	.	.	1.90%	53			
	V32A	0.95%	1	.	.	.	.			
	L33F	2.86%	3	.	.	0.30%	8		0.0066	
	M36I	15.24%	16	5.81%	19	14.00%	390	0.0077		
	I54F	0.95%	1	.	.	.	.	0.0001		
	L63P	85.71%	90	6.27%	21	.	.			
	A71V	9.52%	10	2.44%	8	2.60%	72	0.0047	0.001	
	A71T	13.33%	14	1.63%	5	5.40%	150	0.001	0.001	
	V77I	45.71%	48	4.18%	14	27.00%	752	0.0001		
	I93L	53.33%	56	.	.	25.00%	697		0.0001	
NRTI						N tot NRTI=839				
		M41R	0.95%	1	.	0.10%	1			
		M41L	0.95%	1	8.25%	28	0.60%	5	0.0097	
		E44D	0.95%	1	1.16%	4	0.30%	3		
		E44K	0.95%	1	.	.	.	.		
		T69N	0.95%	1	2.90%	10	0.70%	6		
		T69S	1.90%	2	.	.	0.60%	5		
		D67N	0.95%	1	4.07%	14	0.00%	0		
		V75L	0.95%	1	.	.	0.30%	3		
		Y115C/D/G	0.95%	1	.	.	.	.		
		V118I	4.76%	5	3.37%	11	2.80%	23		
		G333E	10.48%	11	2.09%	7	.	.		
		G333D	0.95%	1	.	.	.	.		
		L210S	0.95%	1	.	.	0.20%	2		
	L210F	2.86%	3	.	.	0.80%	7			
NNRTI						N tot NNRTI=2084				
		K103R	6.67%	7	.	2.70%	56		0.0343	
		K103N	0.95%	1	6.39%	21	0.60%	13	0.0363	
		V108I	1.90%	2	1.16%	4	0.60%	13		
		V179F	0.95%	1	.	.	.	.		
		V179E	1.90%	2	.	.	0.20%	4		0.0311
		V179D	5.71%	6	.	.	0.70%	15		0.0004
	G190A	0.95%	1	2.90%	10	20.00%	417		0.0001	

#### **4.4 TIPIFICACIÓN DEL VIH-1**

Se encontró que el subtipo circulante en la población mexicana es el B. El algoritmo empleado por el programa de tipificación HIVdb de la Universidad de Stanford muestra que las secuencias de la población mexicana tienen una semejanza entre 91% a 97% con la secuencia de referencia HXB2 del subtipo B del VIH-1. Este resultado se comprobó con el análisis filogenético (ver más adelante) en donde se muestra un agrupamiento de las secuencias obtenidas con la secuencia de referencia del consenso B.

Por otro lado la distribución de frecuencias de subtipos en la presente muestra es significativamente diferente de la distribución en la base multinacional de Stanford, mostrando un valor de  $X^2$  (chi cuadrada) de 72.812 con una  $p < 0.0001$ .

#### **4.5 ANÁLISIS FILOGENÉTICO**

Se realizó el análisis filogenético con las secuencias de referencia, las secuencias de las diferentes regiones geográficas y las secuencias obtenidas de los aislados mexicanos. Se realizaron dos árboles filogenéticos mediante el método de Neighbor-Joining.

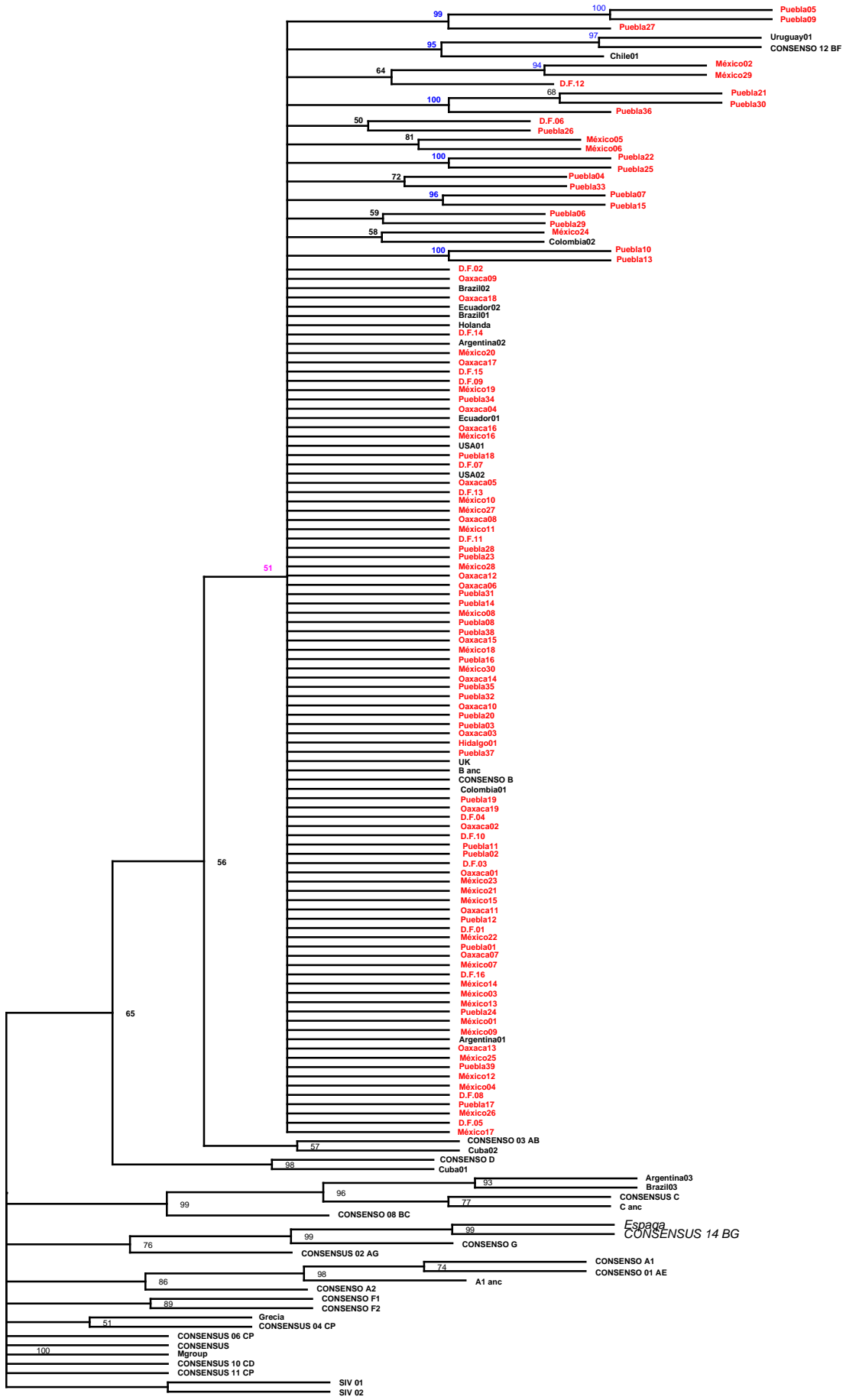
En el primer árbol (figura 17), se muestra la topología del árbol en donde de color rojo se representan todas las secuencias pertenecientes a los aislados mexicanos. Se puede observar en este árbol que las secuencias mexicanas se agrupan en un cluster junto con la secuencia consenso B, una secuencia del Reino Unido, secuencias de Sudamérica y secuencias de Estados Unidos. Todas éstas secuencias tienen en común el pertenecer al subtipo B, comprobando de esta forma que el subtipo circulante en México es el B. Para comprobar que tan confiable es el árbol obtenido se realizó un segundo árbol con los valores de bootstrap (100 réplicas). Dichos valores se muestran en la figura 18 localizándolos en los nodos de las ramas.

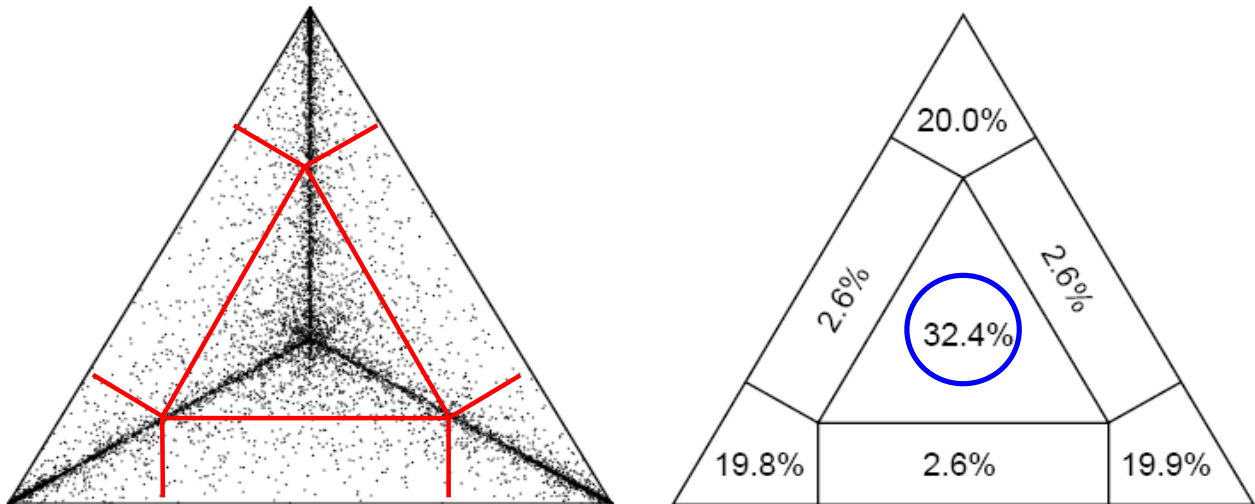
A pesar que la cantidad de réplicas empleadas para el bootstrap son consideradas como mínimas, los valores obtenidos son bajos mostrando que el árbol filogenético obtenido no tiene confiabilidad. Cabe mencionar que existen algunas ramas de los aislados mexicanos con valores altos de bootstrap (Fig. 18-a) lo que significa que esas ramas en particular son confiables, sin embargo el valor obtenido del nodo por el cual se desprenden esas ramas (Fig. 18-b) presenta un valor bajo. Debido a que los valores de bootstrap en el árbol son bajos, se decidió realizar un mapeo de verosimilitud (likelihood mapping) para determinar la cantidad de información filogenética presente en las secuencias. En el mapa de verosimilitud se encontró que en el área 7 del triángulo (centro) se presenta un 32.4% de los puntos o grupos de secuencias, estableciendo que la alineación obtenida de las secuencias de los aislados mexicanos contiene poca información filogenética (Figura 19). Este resultado explica el porqué de los valores bajos de bootstrap e impide realizar un análisis filogenético confiable.



0.1

SIV 02 SIV 01





**Fig. 19.** Mapeo de verosimilitud. El área central contiene 32.4% de los puntos por lo que los datos no son suficientes para hacer inferencias filogenéticas.

## 5. DISCUSIÓN

El objetivo principal de este estudio fue determinar la epidemiología molecular del VIH en la República Mexicana, determinando la forma genética del virus que circula de manera predominante en la población. A partir del análisis de la estructura genética del gen *pol* realizado con el programa de tipificación HIVdb de la Universidad de Stanford (<http://hivdb.stanford.edu>), así como mediante el análisis filogenético con secuencias de referencia, se pudo corroborar que el subtipo B del VIH-1 es el que continúa circulando de manera dominante en México. A pesar de las limitaciones de la muestra de casos de VIH seleccionados y del uso del gen *pol* para el análisis, se podría afirmar que la epidemia en México esta causada por un solo subtipo, ya que de encontrarse otro subtipo o forma recombinante en la población habría sido posible encontrarlo con el análisis elegido. Así mismo, aun cuando el resultado de este estudio es consistente con lo previamente encontrado por Rivera-Morales, habría que tomar ciertas precauciones a la hora de interpretar el resultado porque una de las limitaciones del estudio es la falta de representatividad de los casos de VIH que se están detectando en la región norte del país, así como en áreas de alta prevalencia de la enfermedad como es el Estado de Veracruz.

Muchas veces se ha dicho que la epidemia de VIH/SIDA en México es una continuación de la epidemia de Estados Unidos por la cercanía geográfica y por el constante flujo migratorio entre ambas poblaciones. Esta afirmación tiene una base biológica común: tanto en Estados Unidos como en México las infecciones del VIH-1 están dominadas por el subtipo B. Sin embargo, es necesario hacer notar que recientemente en Estados Unidos se han detectado, con muy bajas frecuencias, nuevas formas genéticas del VIH-1 como son los subtipos A, E y los recombinantes CRF01\_AE, CRF02\_AG y



CRF09\_cpx, las que hasta el momento no se han detectado en México y no se detectaron en este estudio.

Otro de los resultados obtenidos con el presente estudio es la frecuencia de mutaciones de resistencia primaria que se encontraron en los pacientes con VIH-1 en el país. Como se sabe, este tipo de resistencia tiene una importancia fundamental para guiar la selección y uso de los fármacos que deben emplearse en el tratamiento de los pacientes, los cuales contribuyen a mejorar la calidad de vida y la supervivencia de los mismos. Al respecto, se observó que en la población de estudio la prevalencia de mutaciones de resistencia primaria fue baja (1-3%), comparada con la que se presenta en países desarrollados, donde las cifras de mutaciones primarias varían de 11 a 50%<sup>79-81</sup>. Este resultado se esperaba debido al uso relativamente reciente de los fármacos antirretrovirales en el país. Las mutaciones encontradas en la población fueron: L33F (3%), D67N (1%), T69N (3%), K103N/R (8%), G333E (11%) y G190A (1%). Estos resultados son similares a los encontrados en estudios anteriores<sup>73</sup>. De acuerdo a la base de datos de la Universidad de Stanford de resistencia a fármacos, sólo las mutaciones K103N y G190A, encontradas en la población de estudio, son consideradas mutaciones primarias. Sin embargo, es importante considerar al resto de las mutaciones, ya que éstas favorecen una resistencia intermedia a los fármacos y pueden contribuir a probables fallas del tratamiento. Aunque muy pocos pacientes presentaron mutaciones primarias, la mayoría presentó mutaciones menores. Estas mutaciones por sí solas no confieren resistencia a los fármacos, sin embargo, deben tomarse en cuenta como posibles marcadores de resistencia al inicio del tratamiento antirretroviral en los pacientes. A pesar de la baja frecuencia en las mutaciones primarias, es importante mantener un monitoreo constante en la población, ya que un aumento en estas mutaciones puede traer consecuencias graves en la calidad de vida de las

personas recientemente infectadas con el virus. De igual forma, el desabastecimiento de los fármacos y la falta de adherencia de los pacientes a éstos pueden traducirse en un aumento de las frecuencias de mutaciones primarias, ya que éstos pacientes se convierten en posibles transmisores de estas nuevas formas de virus resistentes.

Muchas veces resulta difícil establecer comparaciones de las frecuencias de mutaciones de resistencia a fármacos entre las poblaciones debido a las diferencias metodológicas entre los estudios, así como a los diferentes criterios empleados para definir a las mutaciones como mutaciones de resistencia. En este estudio se realizó una comparación estadística para establecer las diferencias entre las frecuencias de mutaciones de resistencia a fármacos encontradas en la población de estudio con las publicadas en el Reino Unido y la base de datos de Resistencia a Fármacos de la Universidad de Stanford. En general, la frecuencia de mutaciones primarias encontradas en el estudio es mucho más baja que la encontrada en las dos poblaciones mencionadas, aunque cabe destacar que la frecuencia de mutaciones menores y mayores de los aislados mexicanos en este estudio son mayores que en las reportadas en las poblaciones del Reino Unido y Estados Unidos. Los datos estadísticos en estas últimas mutaciones nos obligan a prestar cierta atención en los esquemas de tratamiento que se les dan a los pacientes que inicien tratamiento, principalmente para aquellos que comiencen su tratamiento con inhibidores de proteasa. Debido a que los resultados obtenidos en este estudio muestran que la presencia de mutaciones menores o mayores en la enzima proteasa son mayores que en las poblaciones del Reino Unido y Estados Unidos, esto se puede traducir en nuestra población como fallas virológicas más tempranas en los pacientes durante el uso del tratamiento antirretroviral.

Para llevar a cabo el análisis filogenético en este estudio se utilizó el gen *pol*. La secuenciación de este gen es común para determinar la presencia de mutaciones de

resistencia a fármacos antirretrovirales en pacientes con falla virológica y en pacientes vírgenes a tratamiento. Actualmente la gran cantidad de secuencias disponibles de este gen ha permitido realizar reconstrucciones filogenéticas de eventos de transmisión. Sin embargo, en este estudio no fue posible generar un árbol filogenético confiable ya que los valores de bootstrap encontrados en el árbol fueron muy bajos mostrando poca consistencia en las ramas. Para comprobar si las secuencias obtenidas del gen *pol*, presentaban información suficiente para realizar un análisis filogenético se realizó un mapeo de verosimilitud en donde se muestra claramente (32.4% de puntos en el centro del triángulo) que las secuencias no contienen información suficiente para hacer filogenia. En estudios anteriores, se ha establecido que el gen *pol* es suficiente para establecer relaciones filogenéticas<sup>19</sup> aunque también se ha establecido que no es suficiente para realizar este tipo de estudios<sup>82</sup>. El argumento para el estudio en donde el gen *pol* no es suficiente para establecer relaciones filogenéticas se basa en que es un gen altamente conservado y por ende no muestra de forma confiable la evolución del virus. Por otro lado, el estudio que argumenta que el gen *pol* es suficiente para establecer relaciones filogenéticas, radica en que al ser un gen atacado por los fármacos antirretrovirales es posible detectar eventos de transmisión entre personas epidemiológicamente relacionados. Se ha propuesto, en varios estudios, el uso de más de un gen estructural para determinar relaciones filogenéticas, así como para incrementar la posibilidad de encontrar las diferentes formas genéticas del virus incluyendo los recombinantes y las formas únicas recombinantes.

## **6. Conclusión**

El subtipo B del VIH-1 es responsable de todas las infecciones en la muestra del presente estudio. Con estos resultados y con los resultados mostrados en estudios anteriores podemos concluir que hasta el momento la evolución del virus en México es diferente al del resto del mundo, ya que considerando el constante flujo migratorio y la entrada de nuevas formas genéticas del virus en las diferentes poblaciones, en México sigue de forma dominante el mismo subtipo. De igual forma la distribución de frecuencias de subtipos en la presente muestra es significativamente diferente de la distribución en la base multinacional de Stanford. Por otro lado, las mutaciones primarias K103N y G190A en la población mexicana se encuentran en frecuencias bajas, dándonos cierta ventaja en el control de la enfermedad ya que con estos nuevos resultados y con los resultados en otros estudios podemos mantener un monitoreo y detectar si alguna mutación primaria comienza a incrementar su frecuencia o bien si alguna mutación antes no detectada aparece en la población. De igual forma el haber detectado que la presencia de mutaciones mayores y menores es mayor en nuestra población que en otras poblaciones nos ayuda a tener una base para el monitoreo en los esquemas de tratamiento, evitando de esta forma tomar decisiones incorrectas que puedan afectar la calidad de vida de los pacientes recientemente infectados y que comienzan su esquema de tratamiento.

7. ANEXO



INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
CENTRO DE INVESTIGACIÓN EN ENFERMEDADES INFECCIOSAS



***Epidemiología Molecular del Virus de la Inmunodeficiencia Humana 1 (VIH-1)***

Fecha de la entrevista: \_\_\_\_\_

No. Folio Lab: \_\_\_\_\_

Datos del Paciente

- Nombre: \_\_\_\_\_  
Edad: \_\_\_\_\_
- Sexo: M ( ) F ( ) Edo. Civil: \_\_\_\_\_

Datos sociodemográficos

1. Lugar de nacimiento: \_\_\_\_\_  
—
2. Residencias anteriores (lugar y año en que vivió):
  - a. \_\_\_\_\_  
\_\_\_\_\_
  - b. \_\_\_\_\_  
\_\_\_\_\_
  - c. \_\_\_\_\_  
\_\_\_\_\_
  - d. \_\_\_\_\_  
\_\_\_\_\_
  - e. \_\_\_\_\_  
\_\_\_\_\_
3. Lugar de residencia actual: \_\_\_\_\_
4. Escolaridad:

Primaria	<input type="checkbox"/>
Secundaria	<input type="checkbox"/>
Técnico	<input type="checkbox"/>
Preparatoria	<input type="checkbox"/>
Licenciatura	<input type="checkbox"/>
Posgrado	<input type="checkbox"/>
Otro:	_____

5. ¿ Ha viajado a Sudamérica? SI NO  
 ¿Cuándo? \_\_\_\_\_  
 ¿A qué país? \_\_\_\_\_  
 ¿Por cuánto tiempo? \_\_\_\_\_
6. ¿ Ha viajado a Norteamérica? SI NO  
 ¿Cuándo? \_\_\_\_\_  
 ¿A qué país? \_\_\_\_\_  
 ¿Por cuánto tiempo? \_\_\_\_\_
7. ¿ Ha viajado a Europa? SI NO  
 ¿Cuándo? \_\_\_\_\_  
 ¿A qué país? \_\_\_\_\_  
 ¿Por cuánto tiempo? \_\_\_\_\_
8. ¿ Ha viajado a África? SI NO  
 ¿Cuándo? \_\_\_\_\_  
 ¿A qué país? \_\_\_\_\_  
 ¿Por cuánto tiempo? \_\_\_\_\_

Datos clínicos

- ¿Conoce su fecha aproximada de infección y/o diagnóstico? SI NO  
 ¿Cuándo? \_\_\_\_\_
- ¿Presenta síntomas clínicos? SI NO  
 ¿Cuáles? \_\_\_\_\_
- ¿A tomado fármacos antirretrovirales? SI NO
- ¿Está tomando medicamentos? SI NO  
 ¿Cuáles? \_\_\_\_\_

Células t CD4+/ $\mu$ L \_\_\_\_\_  
 VIH/mL \_\_\_\_\_

Copias RNA

Clasificación por etapa del paciente:

- |   |   |
|---|---|
| Etapa Temprana <input type="checkbox"/>   | Etapa Tardía <input type="checkbox"/>   |
| Etapa Intermedia <input type="checkbox"/> | Etapa Avanzada <input type="checkbox"/> |

Dr. \_\_\_\_\_  
 Fecha: \_\_\_\_\_

## Referencias

1. Kleinbaum, D., Kupper, L. & Morgenstern, H. *Epidemiologic Research.*, pp. 1-15 (Van Nostrand Reinhold Company, New York, 1982).
2. Fauquet. *Virus Taxonomy VIII, Eight report of the International Committee on Taxonomy of Viruses.*, pp. 433-436 (Elsevier Academic Press, 2005).
3. Green, W. & Matija, B. Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. *Nature Medicine* **8**, 673-680 (2002).
4. Abbas, A. & Lichtman, A. *Inmunología celular y molecular.*, pp. 453-476 (Elsevier, 2004).
5. UNAIDS/WHO Epidemic Update December 2005. WHO, 1-87. 2005. ONUSIDA /WHO.
6. Han, B., Shaw, G. & Sharp, P. AIDS as a zoonosis: scientific and public health. *Science* **287**, 607-614 (2000).
7. Gao, F., Robertson, D., Chen, Y., Rodenburg, C. & Michael, S. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436-441 (1999).
8. Gao, F., Yue, L. & White, A. Human infection by genetically diverse SIVsm-related HIV-2 in West Africa. *Nature* **358**, 495-499 (1996).
9. Sharp, P., Bailes, E. & Gao, F. Origins and evolution of AIDS viruses: estimating the time-scale. *Biochem Soc Trans* **28**, 275-282 (2000).
10. Gao, F., Yue, L. & Robertson, D. Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *Journal of Virology* **68**, 7433-7447 (1994).
11. Preston, B. & Dougherty, J. Mechanisms of retroviral mutation. *Trends of Microbiology* **4**, 16-21 (1996).
12. Ho, D. *et al.* Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123-126 (1995).
13. Robertson, D., Hahn, B. & Sharp, P. Recombination in AIDS viruses. *Journal Molecular Evolution* **40**, 249-259 (1995).
14. Robertson, D. *et al.* HIV-1 Nomenclature Proposal. Kuiken, CL, Foley, B, Hahn, BH, Rober, B, Cutchan, F, rx, PA, llors, JW, llins, JI, dorski, J, and linsky, S. <http://www.hiv.lanl.gov/>. 2005. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

15. Thomson,M., Perez-Alvarez,L. & Najera,R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infectious Disease* **2**, 461-471 (2002).
16. Rejinfo,B. & Essex,M. HIV-1 subtypes and recombinants. New York (2002).
17. Wainberg,M. HIV-1 subtype distribution and the problem of drug resistance. *AIDS* **18**, S63-S68 (2004).
18. Balotta,C., Facchi,G., Violin,M., VanDooren,S. & Cozzi-Lepri,A. Increasing prevalence of Non-clade B HIV-1 strains in heterosexual men and women, as monitored by analysis of reverse transcriptase and protease sequences. *Journal of AIDS* **27**, 499-505 (2001).
19. Hue,S., Clewley,J., Cane,P. & Pillay,D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* **18**, 719-728 (2004).
20. De Baar,M., Janssens,W., De Ronde,A., Fransen,K. & Colebunders,R. Natural residues versus antiretroviral drug-selected mutations in HIV type 1 group O reverse transcriptase and protease related to virological drug failure in vivo. *AIDS Research and Human Retroviruses* **16**, 1385-1394 (2000).
21. Delwart,E., Herring,B., Rodrigo,A. & Mullins,J. Genetic subtyping of human immunodeficiency virus using a heteroduplex mobility assay. *PCR Methods Appl.* **4**, S202-216 (1995).
22. Van Harmelen,J., Van der Ryst,E., Wood,R., Lyons,S. & Williamson,C. Restriction fragment length polymorphism analysis for rapid gag subtype determination of human immunodeficiency virus type 1 in South Africa. *Journal Virology Methods* **78**, 51-59 (1999).
23. Takebe,Y., Kusagawa,S. & Motomura,K. Molecular epidemiology of HIV: Tracking AIDS pandemic. *Pediatrics International* **24**, 236-244 (2004).
24. Vidal,N., Peeters,M., Mulanga-Kabeya,C. & Nzilambi,N. Unprecedented degree of Human Immunodeficiency Virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that HIV-1 pandemic originated in Central Africa. *Journal of Virology* **74**, 10498-10507 (2000).
25. Arroyo,M., Hoelscher,M. & Sanders-Buell,E. HIV type 1 subtypes among blood donors in the Mbeya region of southwest Tanzania. *AIDS Research and Human Retroviruses* **20**, 895-901 (2004).
26. Koch,N. *et al.* Genetic analysis in HIV type 1 strains in Bujumbura (Burundi): predominance of subtype C variant. *AIDS Research and Human Retroviruses* **17**, 269-273 (2001).



27. Requeji,H. Worldwide molecular epidemiology of HIV. *Rev Saúde Pública* **40**, 331-345 (2006).
28. Fischetti,L. *et al.* Molecular epidemiology of HIV in Ghana: dominance of CRF02\_AG. *Journal Medical Virology* **73**, 158-166 (2004).
29. Montavon,C. *et al.* CRF06\_cpx: a new circulating recombinant form of HIV-1 in West Africa involving subtypes A, G, K and J. *Journal of Acquired Immune Deficiency Syndrome Human Retrovirol* **29**, 522-530 (2002).
30. Pandrea,I. & Robertson,D. Analysis of partial pol and env sequences indicates a high prevalence of HIV type 1 recombinant strains circulating in Gabon. *AIDS Research and Human Retroviruses* **18**, 1103-1116 (2002).
31. Xiang,Z. *et al.* HIV type 2 pathogenecity is not related to subtype in rural Guinea Bissau. *AIDS Research and Human Retroviruses* **13**, 501-505 (1997).
32. Sarr,A. *et al.* Genetic analysis of HIV type 2 in monotypic and dual HIV infections. *AIDS Research and Human Retroviruses* **16**, 295-298 (2000).
33. Tatt,I., Barlow,K., Clewely,J., Gill,O. & Parry,J. Surveillance of HIV-1 subtypes among heterosexuals in England and Wales. *Journal of Acquired Immune Deficiency Syndrome* **36**, 1096-1099 (2004).
34. Stanojevic,M. *et al.* HIV-1 subtypes in Yugoslavia. *AIDS Research and Human Retroviruses* **18**, 519-522 (2002).
35. Esteves,A. Molecular epidemiology of HIV type 1 infection in Portugal: High prevalence of Non-B subtypes. *AIDS Research and Human Retroviruses* **18**, 325 (2002).
36. Op de Coul,E. *et al.* Genetic evidence of multiple transmission of HIV type 1 subtype F within Romania from adult blood donors to children. *AIDS Research and Human Retroviruses* **16**, 327-336 (2000).
37. Kostrikis,L. *et al.* Genetic analyses of human immunodeficiency virus type 1 strains from patients in Cyprus: identification of new subtype designated subtype I. *Journal of Virology* **69**, 6122-6130 (1995).
38. Yahi,N. *et al.* Use of drug resistance sequence data for the systematic detection of Non-B Human Immunodeficiency Virus Type 1 (HIV-1) subtypes: How to create a sentinel site for monitoring the genetic diversity of HIV-1 at a country scale. *Journal of Infectious Disease* **183**, 1311-1317 (2001).
39. Barlow,K., Tatt,I., Cane,P., Pillay,D. & Clewely,J. Recombinants strains of HIV type 1 in the United Kingdom. *AIDS Research and Human Retroviruses* **17**, 467-474 (2001).

40. Machuca,A. *et al.* Human immunodeficiency virus type 2 infection in Spain. The HIV-2 Spanish Study Group. *Intervirology* **42**, 37-42 (1999).
41. Soriano,V. *et al.* Human immunodeficiency virus type 2 (HIV-2) in Portugal: clinical spectrum, circulating sutypes, virus isolation and plasma viral load. *Journal Medical Virology* **61**, 111-116 (2000).
42. Damond,F. *et al.* Variability of human immunodeficiency virus type 2 (HIV-2) infecting patients living in France. *Virology* **280**, 19-30 (2001).
43. Heredia,A. *et al.* Genetic analysis of human immunodeficiency virus type 2 strains from Spain. *AIDS Research and Human Retroviruses* **14**, 91-94 (1998).
44. Lole,K., Bollinger,R. & Paranjape,R. Full-lenght human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology* **73**, 152-160 (1999).
45. Kandathil,A., Ramalingam,S., Kannangai,R., Shoba,D. & Sridharan,G. Molecular Epidemiology of HIV. *Indian Journal Medical Research* **121**, 333-344 (2005).
46. Grez,M. *et al.* Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses. *Journal of Virology* **68**, 2161-2168 (1994).
47. Motomura,K., Kusagawa,S. & Lwin,H. Different subtype distributions in two cities in Myanmar: evidence for independent clusters of HIV-1 transmission. *AIDS* **17**, 633-636 (2003).
48. Ou,C., Takebe,Y. & Luo,C. Wide distribution of two subtypes of HIV-1 in Thailand. *AIDS Research and Human Retroviruses* **8**, 1471-1472 (1992).
49. Ou,C., Takebe,Y. & Weinger,B. Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet* **341**, 1171-1174 (1993).
50. Weinger,B. & Brown,T. The march of AIDS through Asia. *N. England J. Med.* **335**, 343-345 (1996).
51. Piyasirisilp,S., McCutchan,F. & Carr,J. A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant from AE and a novel BC recombinant. *Journal of Virology* 11286-11295 (2000).
52. Su,L., Graf,M. & Zhang,Y. Characterization of a virtually full-lenght human immunodeficiency virus type 1 genome of a prevalent intersubtype (C/B') recombinant strain in China. *Journal of Virology* **74**, 11367-11376 (2000).

53. McCutchan,F. Understanding the genetic diversity of HIV-1. *AIDS* **14**, S31-S44 (2000).
54. Anonymous. Subtype distribution of collected HIV-1 sequences. Los Alamos National Laboratory, Los Alamos, New Mexico . 1-6-2006.
55. Brodine,S. *et al.* Diverse HIV-1 subtypes and clinical, laboratory and behavioral factors in a recently infected US military cohort. *AIDS* **17**, 2521-2527 (2003).
56. Delwart,E. Two percent of HIV-positive U.S. blood donors are infected with non-subtype B strains. *AIDS Research and Human Retroviruses* **19**, 1065-1070 (2003).
57. Estable,M., Merzovk,A. & Arella,M. Distinct clustering of HIV type 1 sequences derived from injection versus noninjection drug users in Vancouver, Canada. *AIDS Research and Human Retroviruses* **14**, 917-919 (1998).
58. Montpetit,M. *et al.* Molecular epidemiological analysis of human immunodeficiency virus type 1 in Newfoundland, Canada. *AIDS Research and Human Retroviruses* **14**, 1209 (1998).
59. Rivera-Morales,L. The molecular epidemiology of HIV type 1 of Men in Mexico. *AIDS Research and Human Retroviruses* **17**, 87-92 (2001).
60. Candal,D. *et al.* Genetic variability of HIV type 1 in Honduras. *AIDS Research and Human Retroviruses* **13**, 1349-1350 (1997).
61. Barreto,C. *et al.* Trends in antiretroviral drug resistance and clade distributions among HIV-1 infected blood donors in Sao Paulo, Brazil. *Journal of Acquired Immune Deficiency Syndrome* **41**, 338-341 (2006).
62. Bello,G., Guimaraes,M. & Morgado,M. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* **20**, 763-768 (2006).
63. Gadelha,S. & Shindo,N. Molecular epidemiology of Human Immunodeficiency Virus-1 in the state of Ceará, Northeast, Brazil. *Mem Inst Oswaldo Cruz, Rio de Janeiro* **98**, 461-464 (2003).
64. Espinosa,A. *et al.* Intersubtype BF recombinants of HIV-1 in a population of injectin drug users in Argentina. *Journal of Infectious Disease* **36**, 630-636 (2004).
65. Fernandez,J. *et al.* Molecular epidemiology of HIV-1 in Chile: diferencial geographic and transmission route distribution of B and F subtypes. *Journal of Infectious Disease* **21**, 835-840 (2005).
66. Grossman,Z. Genotypic variation of HIV-1 reverse transcriptase and protease: comparative analysis of clade C and clade B. *AIDS* **15**, 1453-1460 (2001).

67. Castro,E. Molecular epidemiology of HIV-1 in Venezuela: High prevalence of HIV-1 Subtype B and indentification of a B/F recombinant infection. *Journal of Acquired Immune Deficiency Syndrome* **32**, 338-344 (2003).
68. Hierholzer,J. Molecular epidemiology of HIV-1 in Ecuador, Peru, Bolivia, Uruguay, and Argentina. *AIDS Research and Human Retroviruses* **18**, 1339-1350 (2002).
69. Vaughan,H. Characterization of HIV type 1 clades in the Caribbean using pol gene sequences. *AIDS Research and Human Retroviruses* **19**, 929-932 (2003).
70. Casado,G., Michael,M., Thomson,M., Sierra,M. & Najera,R. Identification od a novel HIV-1 circulating ADG intersubtype recombinant form (CRF19\_cpx) in Cuba. *Journal of Acquired Immune Deficiency Syndrome* **40**, 532-537 (2005).
71. Ryan,C. & Elliot,J. The molecular epidemiology of HIV type 1 among Vietnamese Australian injecting drug users in Melbourne, Australia. *AIDS Research and Human Retroviruses* **20**, 1364-1367 (2004).
72. Dwyer,D. *et al.* Human immunodeficiency virus type 1 subtypes B and C detected in New Zealand. *AIDS Research and Human Retroviruses* **14**, 1105-1108 (1998).
73. Escoto-Delgadillo,M. *et al.* Drug-resistance mutations in antiretroviral-naive patients with established HIV-1 infection in Mexico. *HIV Medicine* **6**, 403-409 (2005).
74. WHO. Guidance on regulations for the transport of Infectious Substances. World Health Organization, Geneva (2005).
75. Yang,S. & Rothman,R. PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. *Lancet Infectious Diseases* **4**, 337-347 (2004).
76. Thompson,J., Gibson,T., Plewniak,F., Jeanmougin,F. & Higgins,D. The ClustalX windows interface: flexible strategies for multiple alugnment aided by quality analysis tools. *Nucleic Acids Research* **24**, 4876-4882 (1997).
77. Saitou,N. & Nei,M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology Evolution* **4**, 406-425 (1987).
78. Strimmer,K. & Von Haeseler,A. Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl. Acad. Sci.* **94**, 6815-6819 (1997).
79. HIV Drug Resistance Database. Stanford University . 2006. 22-9-2006.
80. Cane,P. *et al.* Time trends in primary resistance to HIV drugs in the United Kingdom: multicentre observational study. *British Medical Journal* (2005).

81. Wensing,A. *et al.* Prevalence of drug-resistance HIV-1 variants in untreated individual in Europe: implications for clinical management. *Journal of Infectious Disease* **192**, 958-966 (2005).
82. Stürmer,M., Preiser,W., Gute,P., Nisius,G. & Doerr,H. Phylogenetic analysis of HIV-1 transmission: pol gene sequences are insufficient to clarify true relationships between patient isolates. *AIDS* **18**, 2109-2113 (2004).