

**Universidad de Granada**



**Tesis doctoral**

**Las herramientas de análisis filogenético para el estudio de la propagación de la epidemia de VIH-1 en Andalucía Oriental.**

**Director de Tesis**

**Dr. Federico García García**

**Doctorando**

**Santiago Pérez Parra**

Editor: Universidad de Granada. Tesis Doctorales  
Autor: Santiago Pérez Parra  
ISBN: 978-84-9125-761-5  
URI: <http://hdl.handle.net/10481/43471>

## **AGRADECIMIENTOS**

Quiero agradecer en primer lugar a mi director de tesis Federico García García por la confianza depositada en mi trabajo.

A mis padres porque sin vosotros no habría llegado hasta aquí, gracias por quererme siempre y por vuestra confianza exagerada en mí. Gracias a mi hermana, por ser mi mayor confidente y consejera en todos estos años juntos. A mi rubia preferida, gracias por escucharme, aguantarme, quererme, y darme tantos momentos de amor. A mi abuelo, por ser tan grande y uno de mis referentes. A mis amigos, por brindarme tantos buenos momentos, por hacerme un hueco en vuestra agenda cada vez que os reclamo, por ser mi segunda familia. A mis compañeros de residencia, Jose Luis, Paz y Recio, por los buenos y no tan buenos momentos de residencia juntos. A todos los compañeros de trabajo con los que he aprendido microbiología. A Alejandro por enseñarme tanto desde su humildad. A Jose Antonio Pérez Lopez por compartir tan sabias y grandes palabras. A Vicente por ser mi primer adjunto y tan buena gente. A Juan Román y a Pepe Liebana por ser ante todo, adjuntos de buen corazón. A Natalia, Marta y a todo el personal de biología molecular, por su incansable y profesional trabajo y porque sin ellos esta tesis no habría sido posible. A los técnicos de laboratorio por ser esa gran familia, por hacer el trabajo más llevadero durante tantas guardias. A M<sup>a</sup> Carmen por ser la secretaria perfecta y a Jose Luis por su buen humor y esos regates mágicos en los partidos de fútbol sala.

A los que siguen en mi vida, y a los que no lo están, porque seguro que de todos ellos habré aprendido algo importante a lo largo de estos años. Esta tesis también es de vosotros.

## RESUMEN

Desde comienzos de los años 1982 la epidemia de VIH en España ha seguido una dinámica cambiante en la distribución de subtipos, prevalencia de las rutas de transmisión y en las características de los colectivos de riesgo de transmisión del virus. El análisis filogenético, a través del análisis de las relaciones filogenéticas entre las diferentes formas genéticas que infectan a los pacientes de una determinada región, aporta información valiosa sobre la existencia de ciertos colectivos de riesgo y ayuda a inferir una relación evolutiva que nos permita establecer un marco temporal de la historia evolutiva del virus. Dentro del análisis filogenético, los análisis filogeográficos a través del estudio de las relaciones epidemiológicas entre los eventos migratorios y la diseminación del VIH a nivel local, son herramientas especialmente importantes en escenarios donde existe una alta tasa de migración ya que ésta puede influir en la epidemiología local. Asimismo, la filodinámica aporta información acerca del tamaño efectivo poblacional de una determinada forma genética permitiendo analizar el crecimiento o decrecimiento del número efectivo de infecciones a lo largo del tiempo.

En los últimos años están apareciendo diferentes estudios en diferentes cohortes de pacientes naïve, que utilizan el análisis filogenético como herramienta para estudiar el comportamiento de la epidemia de VIH en una determinada área: Yebra et al demuestran en la comunidad de Madrid una mayor inclusión de homosexuales en cluster de transmisión y Ambrosioni et al concluyen que desde el periodo de 2008 a 2010 el 42% de los pacientes naïve de la cohorte suiza se agruparon en cluster de transmisión. Pernas et al describen un aumento de la prevalencia del subtipo F en Galicia debido a la existencia de un cluster de transmisión que se transmite en el norte de España y Delgado et al han identificado un cluster de transmisión de subtipo A entre pacientes homosexuales en diversas áreas de España, advirtiendo de la necesidad de una vigilancia epidemiológica continua. Recientemente, Patiño et al han alertado de la emergencia de la variante de VIH CRF19\_cpx entre homosexuales, una variante endémica de Cuba, relacionada con una mayor progresión de los pacientes a SIDA y asociada a mutaciones de resistencia a fármacos antiretrovirales de primera línea.

Nuestro centro sirve de referencia para el estudio de resistencia a fármacos antiretrovirales en todo el área de Andalucía Oriental. Disponemos de información de la secuencia del gen *pol*, información que se encuentra ligada a datos clínicos a través de un grupo colaborativo con las unidades de enfermedades infecciosas. Con estos antecedentes y dado que no existen hasta la fecha grandes estudios epidemiológicos realizados a través de análisis filogenético en Andalucía Oriental, creemos que este trabajo aporta información importante sobre la historia evolutiva del virus del VIH en los últimos años.

Para esta tesis doctoral se recopilieron todas las secuencias parciales del gen *pol* del VIH-1 (posiciones 4-99 de la proteasa y 38-247 de la transcriptasa reversa), obtenidas utilizando el ensayo Trugene HIV-1 genotyping Kit (Siemens), procedentes de pacientes naïve de diferentes hospitales de Granada, Almería y Jaén infectados en el periodo de 2005 a 2012. Cada una de estas secuencias fue vinculada a información demográfica (edad, sexo, fecha de diagnóstico y forma de infección por VIH), clínica (recuento de CD4) y virológica (carga viral) mediante la construcción de una base de datos anonimizada, encriptada y protegida. Los pacientes estudiados en nuestra cohorte fueron mayoritariamente hombres jóvenes (41 años), homosexuales (51.5%), nativos españoles (65.5%) y residentes principalmente en la zona de Granada (43,7%) o Almería (36.9%).

El estudio ha sido aprobado por el Comité Ético del Complejo Hospitalario Universitario de Granada.

Para caracterizar las formas genéticas de VIH-1 utilizamos en primer lugar herramientas de subtipado rápido (REGA 2.0 y REGA 3.0) y finalmente confirmamos estos subtipos puros y formas recombinantes circulantes mediante análisis filogenético.

El primer planteamiento que llevamos a cabo en esta tesis doctoral fue la realización de una primera caracterización filogenética que nos permitiera realizar un análisis global de las agrupaciones filogenéticas. Para evaluar este aspecto, realizamos un análisis filogenético “clásico” mediante el método de Neighbour Joining. En primer lugar estudiamos la presencia de pacientes incluidos en cluster de transmisión que presentaban mutaciones de resistencia a fármacos antiretrovirales de primera línea y posteriormente evaluamos las variables clínico-

epidemiológicas que estuvieran relacionadas con el agrupamiento de los pacientes en cluster de transmisión.

La mayoría de agrupaciones filogenéticas que encontramos se dieron entre parejas, fundamentalmente parejas de homosexuales. Asimismo, el agrupamiento en cluster de transmisión de gran tamaño se produjo fundamentalmente entre pacientes homosexuales, jóvenes y que presentaban un recuento elevado de CD4.

La presencia de resistencia a fármacos antiretrovirales de primera línea en pacientes que formaron cluster de transmisión fue poco frecuente y se dio fundamentalmente en pacientes que se agruparon en cluster de pequeño tamaño, datos que sugerían en gran medida que los virus con mutaciones de resistencia presentaron una desventaja biológica frente a cepas silvestres y que por tanto este hecho disminuyó la eficacia en la transmisión.

En segundo lugar, procedimos a analizar por separado los perfiles filogenéticos de la epidemia de VIH-1 subtipo B y subtipo no-B (dado que la filogenia molecular de una población no puede ser reducida a la reconstrucción simple de una única topología de un árbol filogenético) implementando para ello nuevos métodos de análisis filogenético. Esta vez se analizaron por separado los perfiles filogenéticos de las principales formas genéticas mediante el método de análisis filogenético “clásico” de Máxima Verosimilitud y confirmamos estas agrupaciones filogenéticas utilizando métodos avanzados de Inferencia Bayesiana que nos permitieron además establecer (para cada forma genética) una escala temporal, estudiar el tiempo del más reciente común antecesor para estos cluster de transmisión y su tasa evolutiva.

En esta tesis doctoral hemos demostrado la existencia de una epidemia de VIH compartimentada y condicionada fundamentalmente por las características sociales de estos pacientes. Existieron diferencias importantes entre los pacientes infectados con virus de VIH subtipo B en comparación con los infectados con otras formas genéticas. Dentro del subtipo B, encontramos los cluster de transmisión con mayor repercusión epidemiológica. Los primeros cluster de transmisión aparecieron en las zonas de costa durante la década de los 80, década en la que demostramos un crecimiento exponencial del VIH-1 subtipo B. La epidemia VIH subtipo B se estabilizó en la década de los 90, en gran medida como consecuencia de las

primeras campañas de prevención y la comercialización de los primeros fármacos antiretrovirales de primera línea en España. Sin embargo, los comportamientos de alto riesgo entre pacientes homosexuales en esta década causó la emergencia en Granada de la mayoría de cluster de transmisión de gran repercusión epidemiológica, formados fundamentalmente entre pacientes jóvenes (menores de 35 años), españoles y con un alto recuento de CD4.

Por otro lado, demostramos como la cercanía geográfica de Andalucía oriental con respecto al continente Africano y como consecuencia la alta tasa de migración en los últimos años en la zona de El Ejido, han desencadenado una alta prevalencia ( $\approx 23\%$ ) y diversidad genética de variantes de VIH-1 subtipo no-B en comparación con otras regiones de España. Asimismo estos pacientes fueron fundamentalmente hombres heterosexuales, mayores de 35 años, de nacionalidad Africana y que residían en la zona de El Ejido. Cabe destacar que la mayoría de cluster de transmisión para estas formas genéticas se originaron en la década de los 2000 (años en los que la migración en el sur de España aumentó considerablemente), formados fundamentalmente entre pacientes heterosexuales y con un recuento bajo de CD4.

En esta tesis doctoral cabe destacar, el hallazgo de colectivos de alto riesgo en la transmisión del virus. Como ejemplo de ello destacamos un cluster de transmisión formado por 58 pacientes homosexuales con un elevado recuento de CD4, datos clínicos que sugieren una alta proporción de pacientes con infección aguda reciente. Asimismo, demostramos como el subtipo A fue importado principalmente a través de mujeres trabajadoras del sexo de Europa del Este. En este subtipo hemos encontrado una gran proporción de pacientes relacionados epidemiológicamente que se agruparon en diferentes cluster de transmisión. Cluster que agruparon principalmente a mujeres de nacionalidad rusa y hombres españoles.

En conclusión, presentamos datos que demuestran que los análisis filogenéticos a través de el estudio de la filogeografía y la filodinámica de los diferentes subtipos de VIH-1 en Andalucía Oriental se presenta como una fuente de información de gran utilidad para la vigilancia y control epidemiológico de la propagación del VIH y que su utilización ayudaría a reforzar e implantar actuaciones eficaces para prevenir la diseminación del VIH.

## **PALABRAS CLAVE**

VIH-1, Epidemiología molecular, Subtipos de VIH-1, Análisis filogenético, Cluster de transmisión, TMRCA, tasa evolutiva.

## ÍNDICE DE ABREVIATURAS

<b>μl</b>	Microlitros
<b>aa</b>	Aminoácidos
<b>ADN</b>	Ácido desoxirribonucleico
<b>ARN</b>	Ácido ribonucleico
<b>ARV</b>	Antirretroviral
<b>BLAST</b>	Herramienta de búsqueda básica en alineamientos locales
<b>CRF</b>	Forma recombinante circulante
<b>CV</b>	Carga viral
<b>HLA</b>	Antígeno leucocitario humano
<b>HPD</b>	Intervalo de mayor densidad posterior
<b>HSH</b>	Homosexual
<b>HTX</b>	Heterosexual
<b>IC</b>	Intervalo de confianza
<b>IP</b>	Inhibidor de la proteasa
<b>ITIAN</b>	Inhibidores de la transcriptasa inversa análogos de nucleósidos
<b>ITINAN</b>	Inhibidores de la transcriptasa inversa no análogos de nucleósidos
<b>Log</b>	Logaritmo decimal
<b>LTR</b>	Repeticiones largas terminales
<b>MCC</b>	Máxima credibilidad de clados
<b>ML</b>	Máxima verosimilitud (Maximum Likelihood)
<b>MP</b>	Máxima parsimonia
<b>NJ</b>	Unión por vecindad (Neighbor Joining)
<b>nt</b>	Nucleótidos
<b>OMS</b>	Organización Mundial de la Salud

<b>pb</b>	Pares de bases
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>PR</b>	proteasa
<b>RT</b>	Transcriptasa reversa
<b>RIS</b>	Red de investigación de sida
<b>SIDA</b>	Síndrome de la inmunodeficiencia adquirida
<b>TAR</b>	Tratamiento antirretroviral
<b>TARGA</b>	Terapia antirretroviral de gran actividad
<b>TDR</b>	Mutación de resistencia transmitida
<b>TMRCA</b>	Tiempo del más Reciente común antecesor
<b>URF</b>	Forma recombinante única
<b>UDVP</b>	Usuarios de droga por vía parenteral
<b>VIH</b>	Virus de la inmunodeficiencia humana

### **Abreviaturas nucleótidos**

<b>A</b>	Adenina
<b>C</b>	Citosina
<b>G</b>	Guanina
<b>T (o U)</b>	Timina (o Uracilo)
<b>R</b>	A o G
<b>Y</b>	C o T
<b>S</b>	G o C
<b>W</b>	A o T
<b>K</b>	G o T
<b>M</b>	A o C
<b>B</b>	C o G o T
<b>D</b>	A o G o T
<b>H</b>	A o C o T
<b>V</b>	A o C o G
<b>N</b>	cualquier base

### Abreviaturas aminoácido

<b>A</b>	Ala	Alanina
<b>C</b>	Cys	Cisteína
<b>D</b>	Asp	Ác. Aspártico
<b>E</b>	Glu	Ác. Glutámico
<b>F</b>	Phe	Fenilalanina
<b>G</b>	Gly	Glicina
<b>H</b>	His	Histidina
<b>I</b>	Ile	Isoleucina
<b>K</b>	Lys	Lisina
<b>L</b>	Leu	Leucina
<b>M</b>	Met	Metionina
<b>N</b>	Asn	Asparagina
<b>P</b>	Pro	Prolina
<b>Q</b>	Gln	Glutamina
<b>R</b>	Arg	Arginina
<b>S</b>	Ser	Serina
<b>T</b>	Thr	Treonina
<b>V</b>	Val	Valina
<b>W</b>	Trp	Triptófano
<b>Y</b>	Tyr	Tirosina

# ÍNDICE

<b>Capítulo 1: INTRODUCCIÓN</b>	<b>15</b>
<b>1.1 Antecedentes generales</b>	<b>16</b>
<b>1.2 Características estructurales del VIH.</b>	<b>17</b>
<b>1.3 Ciclo biológico del VIH-1</b>	<b>20</b>
<b>1.4 Epidemiología del VIH-1</b>	<b>24</b>
<b>1.5 Variabilidad genética del VIH</b>	<b>27</b>
1.5.1 <i>Recombinantes inter-subtipos del VIH-1</i>	29
1.5.2 <i>Distribución mundial de los diferentes subtipos de VIH-1</i>	30
1.5.3 <i>Implicaciones clínicas y epidemiológicas de los subtipos del VIH.</i>	33
<b>1.6 Generalidades del tratamiento antiretroviral</b>	<b>34</b>
1.6.1 <i>Fármacos inhibidores de la RT.</i>	35
1.6.2 <i>Fármacos inhibidores de la proteasa</i>	36
<b>1.7 Interpretación de la resistencia del VIH-1 a los fármacos antiretrovirales.</b>	<b>37</b>
1.7.1 <i>Mutaciones de resistencia transmitidas</i>	40
<b>1.8 Las herramientas de análisis filogenético en el control de la epidemia de VIH.</b>	<b>43</b>
1.8.1 <i>Introducción a los métodos de análisis filogenético</i>	45
1.8.2 <i>Métodos de construcción de árboles filogenéticos</i>	49
-Métodos basados en distancias genéticas	50
-Métodos basados en caracteres	51
1.8.3 <i>Métodos de análisis filogenético utilizados en VIH-1.</i>	53
1.8.4 <i>Modelos de sustitución nucleotídica</i>	54
<b>Capítulo 2: OBJETIVOS</b>	<b>56</b>
<b>Capítulo 3: ANÁLISIS FILOGENÉTICO GLOBAL DE LA TRANSMISIÓN DE VIH EN ANDALUCÍA ORIENTAL.</b>	<b>58</b>
<b>3.1 Antecedentes</b>	<b>59</b>
<b>3.2 Pacientes y métodos</b>	<b>60</b>
<b>3.3 Resultados</b>	<b>67</b>
<b>3.4 Discusión</b>	<b>83</b>
<b>Capítulo 4: ESTUDIO DE LOS PERFILES FILOGENÉTICOS, FILODINÁMICOS Y FILOGEOGRÁFICOS DEL VIH-1 SUBTIPO B.</b>	<b>86</b>
<b>4.1 Antecedentes</b>	<b>87</b>
<b>4.2 Pacientes y métodos</b>	<b>88</b>
<b>4.3 Resultados</b>	<b>91</b>
<b>4.4 Discusión</b>	<b>104</b>
<b>Capítulo 5: ANÁLISIS DE LOS PERFILES FILOGENÉTICOS Y FILOGEOGRÁFICOS DE LAS FORMAS GENÉTICAS DE VIH-1 SUBTIPO NO-B.</b>	<b>107</b>
<b>5.1 Antecedentes</b>	<b>108</b>
<b>5.2 Pacientes y métodos</b>	<b>111</b>
<b>5.3 Resultados</b>	<b>114</b>
<b>5.4 Discusión</b>	<b>134</b>

<b>Capítulo 6: DISCUSIÓN FINAL</b>	<b>137</b>
<b>Capítulo 7: CONCLUSIONES</b>	<b>140</b>
<b>ANEXO I: RECURSOS DE INTERNET</b>	<b>143</b>
<b>ANEXO II: ESTRUCTURA DE LOS CRF INCLUIDOS EN ESTA TESIS</b>	<b>144</b>
<b>ANEXO III: PARTICIPACIÓN EN CONGRESOS</b>	<b>145</b>
<b>ANEXO IV: PUBLICACIONES SURGIDAS DE ESTA TESIS</b>	<b>152</b>
<b>BIBLIOGRAFÍA</b>	<b>153</b>

# Capítulo 1: INTRODUCCIÓN

## 1.1 Antecedentes generales

La enfermedad del virus de la inmunodeficiencia humana (VIH) fue descrita por primera vez en 1981 por dos grupos, uno en San Francisco y otro en Nueva York, asociada en ese momento con deficiencia inmune grave, debido a neumonía por *Pneumocystis carinii* y sarcoma de Kaposi agresivo (1). El virus VIH en sí no fue identificado hasta 2 años después (2); durante ese tiempo, varias otras causas fueron consideradas, incluyendo factores como estilo de vida, el abuso crónico de drogas y otros agentes infecciosos (3). Pronto se evidenció que presentaba el mismo perfil epidemiológico de hepatitis B en cuatro grupos bien diferenciados: homosexuales, hemofílicos, hemoperfundidos y heroínómanos. La similitud hizo pensar que la causa podría ser un agente infeccioso, probablemente viral con transmisión sexual y sanguínea (4). En 1986 el Comité Internacional de Taxonomía de Virus acepta el nombre definitivo de VIH para este nuevo agente.

Dos especies distintas del VIH (VIH-1 y VIH-2) han sido identificadas y cada una está compuesta de múltiples subtipos que tienden a causar una enfermedad similar, pero que difiere de su distribución mundial. El VIH-1 probablemente se originó de una o más transferencias entre especies de los chimpancés en el centro de África (5,6). El VIH-2 está estrechamente relacionado con los virus que infectan a los mangabeys en África occidental (7), el VIH-1 y VIH-2 son superficialmente similares, pero cada uno contiene genes únicos y su propio proceso de replicación característico.

La transmisión e infección de VIH-2 tiende a progresar más lentamente con el síndrome de inmunodeficiencia adquirida (SIDA). Las personas infectadas con el VIH-2 desarrollan una carga viral más baja que las personas con VIH-1 (8,9), una mayor carga viral está asociada con una progresión más rápida hacia el SIDA en infecciones VIH-1 (10,11).

## 1.2 Características estructurales del VIH.

El VIH es un virus envuelto, que forma partículas esféricas de 80-100 nm de diámetro. La cubierta vírica tiene una estructura de bicapa lipoproteica que incluye dos glucoproteínas víricas, la glucoproteína de superficie gp120, que se ancla a la envoltura mediante la proteína transmembrana gp41. La envoltura recubre la proteína de la matriz p17 y el core o núcleo del virus formado por la proteína de la cápside p24. La nucleocápside es de forma cónica o esférica y alberga el genoma del virus y una serie de proteínas víricas, esencialmente las enzimas requeridas para realizar el ciclo infectivo del virus como el complejo transcriptasa inversa, la proteasa vírica, la proteína de unión al ARN, proteínas de regulación y la integrasa vírica (Figura 1).

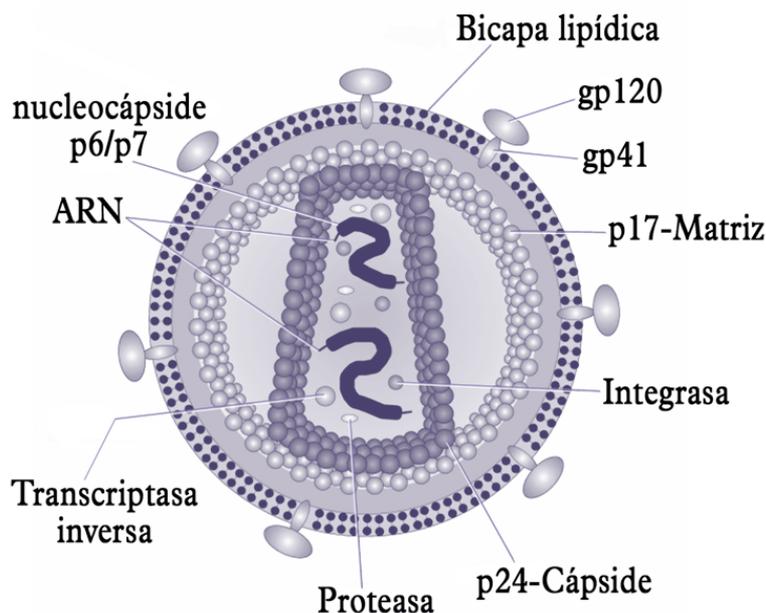


Figura 1. Representación esquemática de la morfología del VIH.

El material genético del virus VIH está formado por una doble hebra de ARN de polaridad positiva con un tamaño de 9,8 Kb y está constituido por tres genes estructurales (*gag*, *pol* y *env*) y seis genes reguladores (12) (*nef*, *tat*, *rev*, *vpr*, *vif* y *vpu*).

Los genes estructurales *gag*, *pol* y *env* codifican respectivamente, las proteínas del core, las del complejo enzimático proteasa vírica/ transcriptasa inversa y las de la envoltura:

- El gen *gag* codifica las proteínas p55 (proteína precursora), p6 (participa en la encapsidación), p7 (asociada al ácido nucleico), p24 (proteína de cápside) y p17 (proteína de la matriz).
- El gen *pol* codifica la transcriptasa inversa, la p66 y 51 (integrasa) y la p31 (proteasa).
- El gen *env* codifica para la gp160 (poliproteína precursora), gp120 (proteína de la envoltura que interacciona con el receptor CD4) y la gp41 (interviene en la fusión de membranas).

A su vez estos tres genes se encuentran flanqueados por dos secuencias repetitivas no codificantes situadas a ambos lados del genoma vírico en su forma de ADN completamente retrotranscrito (llamadas secuencias repetitivas largas o LTR) que constituyen el genoma mínimo de los retrovirus.

A la estructura genómica antes mencionada se le añade en el caso de los retrovirus complejos como es el VIH, una serie de genes reguladores del ciclo replicativo que reciben el nombre de *rev* (regulador del transporte y procesamiento de ARNm), *vif* (participa en el aumento de infectividad e interacción con factores celulares antivíricos) y *nef* (participa en la infectividad, aumento de retrotranscripción, inducción de apoptosis de CD8 citotóxicos y regulación negativa de CD4 y HLA-I). Otros genes que participan en la regulación del ciclo replicativo son el *vpr*, *tat* y *vpu* (Figura 2).

## HIV-1 Genome

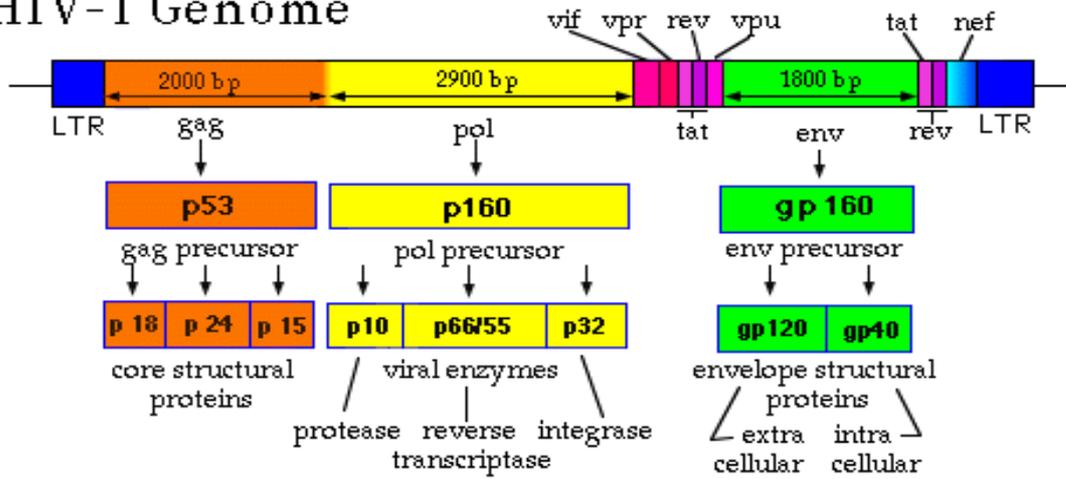


Figura 2. Representación esquemática del genoma del VIH.

### 1.3 Ciclo biológico del VIH-1

El mecanismo de transmisión del VIH se realiza por contacto directo entre el hospedador y el virus. Las formas más comunes de infección son la transmisión sexual, a través de la mucosa vaginal o colónica, por contacto con líquidos infectados (sangre y derivados mediante jeringuillas o transfusiones...), por transmisión de la madre al feto de forma congénita, perinatal o postnatal (leche materna).

El primer fenómeno que se produce consiste en la unión del virus a moléculas de membrana en determinados tipos celulares. Estas moléculas interactúan con la envuelta y “anclan” las partículas víricas a la superficie de la célula. Algunas de estas uniones son inespecíficas y tienen baja afinidad con el VIH. Otras, sin embargo tienen mayor importancia patogénica como las de las lectinas DC-SIGN y L-SIGN que unen a virus con envueltas glucosiladas. Estas moléculas se expresan de forma diferencial en células dendríticas y no pueden considerarse receptores en sentido estricto ya que su interacción con el virus no se produce a través de estructuras conformacionales. Tras la unión del VIH con las citadas lectinas, se produce endocitosis, presentación de antígenos en moléculas de HLA de clase II y adhesión intercelular que facilita el reconocimiento por el receptor T del linfocito T CD4, la entrada del virus en la célula se produce por la interacción de la gp120 del virus con los receptores CD4+ celulares. Tras la unión al CD4, la proteína vírica gp120 sufre un cambio conformacional que le permitirá la unión a los correceptores de quimiocinas (13) CCR5 y/o CXCR4.

Las quimiocinas constituyen una familia de proteínas de bajo peso molecular que incluye al menos 70 proteínas que se clasifican en 4 subfamilias: C, CXXXC, CXC y CC. Las quimiocinas se unen a receptores de tipo 7M o 7 dominios transmembrana que se clasifican también en función del tipo de quimiocina que unen en C, CXXXC, CXC y CC. En 1996 fue descubierto el primer correceptor del VIH, llamado inicialmente “fusin” y conocido en la actualidad como CXCR4. Posteriormente fue descubierto otro correceptor del VIH-1 y fue denominado CCR5. Los receptores de quimiocinas se expresan en múltiples tipos celulares. Sin embargo, como esquema general, podemos considerar que el receptor CXCR4 se expresa de forma constitutiva en un gran número de poblaciones, incluidos linfocitos y monocitos, mientras que los receptores de la familia CC, a la que pertenece el CCR5, se expresan forma

constitutiva en un porcentaje limitado de células y requieren activación celular para su expresión de novo.

Así pues, se facilita e incrementa la infección de los linfocitos mediante varios mecanismos: concentración de partículas en estas estructuras, activación de los linfocitos y propagación a través de células dendríticas infectadas. Los órganos linfoides y, en particular las células dendríticas, son el gran reservorio donde la infección se establece y se transmite a los linfocitos CD4.

La utilización de uno u otro correceptor va a depender de la secuencia de la envoltura vírica. En función del correceptor involucrado, clasificamos el VIH en tres variantes: R5 (usan exclusivamente correceptores CCR5), X4 (usan únicamente los CXCR4) y R5X4(usan indistintamente los dos para entrar en la célula).

Los ligandos naturales de CCR5 son las quimiocinas RANTES ó CCL-5 (Regulated on Activation Normal T-cell Expressed and Secreted), MIP-1 $\alpha$  ó CCL-3 (Macrophage Inflammatory Protein 1 $\alpha$ ) y MIP-1 $\beta$  ó CCL-4, y el único ligando natural conocido de CXCR4 es SDF-1 o CXCL12 (Stromal Cell Derived Factor 1). Las quimiocinas que se unen a CCR5 y CXCR4, muy especialmente RANTES, MIP y SDF, pueden interferir con el VIH en la unión con sus correceptores, lo que con toda probabilidad constituye un potente mecanismo de protección para la infección por el VIH in vivo. En concreto se ha demostrado que hay una alta producción de SDF1 por células epiteliales mucosas y células dendríticas que explicaría la baja tasa de transmisibilidad de las cepas cuyo correceptor es el CXCR4.

Este proceso es necesario para la inducir fusión del virus con la membrana de la célula hospedadora y en consecuencia, la entrada y desensamblaje de la cápsida en el interior citoplasmático.

Posteriormente el ARN vírico monocatenario es convertido en ADN bicatenario provírico. Este proceso lo lleva a cabo el complejo transcriptasa inversa (14) en el citoplasma de la célula hospedadora y requiere de una serie de factores celulares inducidos durante la activación linfocitaria. En linfocitos no activados, el genoma vírico es retrotranscrito de forma incompleta con lo cual no se produce la finalización de la retrotranscripción ni la integración hasta la activación del linfocito (15). Estas formas no integradas permanecen en

el citosol hasta una semana formando un reservorio de VIH. Una vez sintetizado, el ADN es transportado al núcleo y se integra en el genoma celular mediante la acción de la integrasa constituyendo lo que se conoce como forma “proviral” del VIH.

Al sintetizarse el ADN proviral, éste formará al unirse a factores celulares y víricos (Vpr, proteína de la matriz) el llamado complejo de preintegración que será transportado al núcleo e integrado en el genoma de la célula hospedadora mediante la acción de la integrasa vírica.

Al integrarse en el genoma del hospedador, el VIH puede tener un comportamiento variable: puede permanecer latente o experimentar una replicación masiva. Esto dependerá de factores celulares, entre los que destacan la familia de factores de transcripción Rel/NF-kB la cual no existe en su forma activa en los linfocitos CD4 en estado de reposo y es inducido únicamente en el proceso de activación inmunológica. La activación celular induce en el linfocito CD4 la síntesis de proteínas NF-kB necesarias para la reactivación del virus, el cual es muy rápido y agresivo, ya que se estima que, tras la activación linfocitaria, en 2 horas se produce la síntesis de todas las proteínas virales, detectándose viriones viables a las 4-6 horas de la reactivación(16).

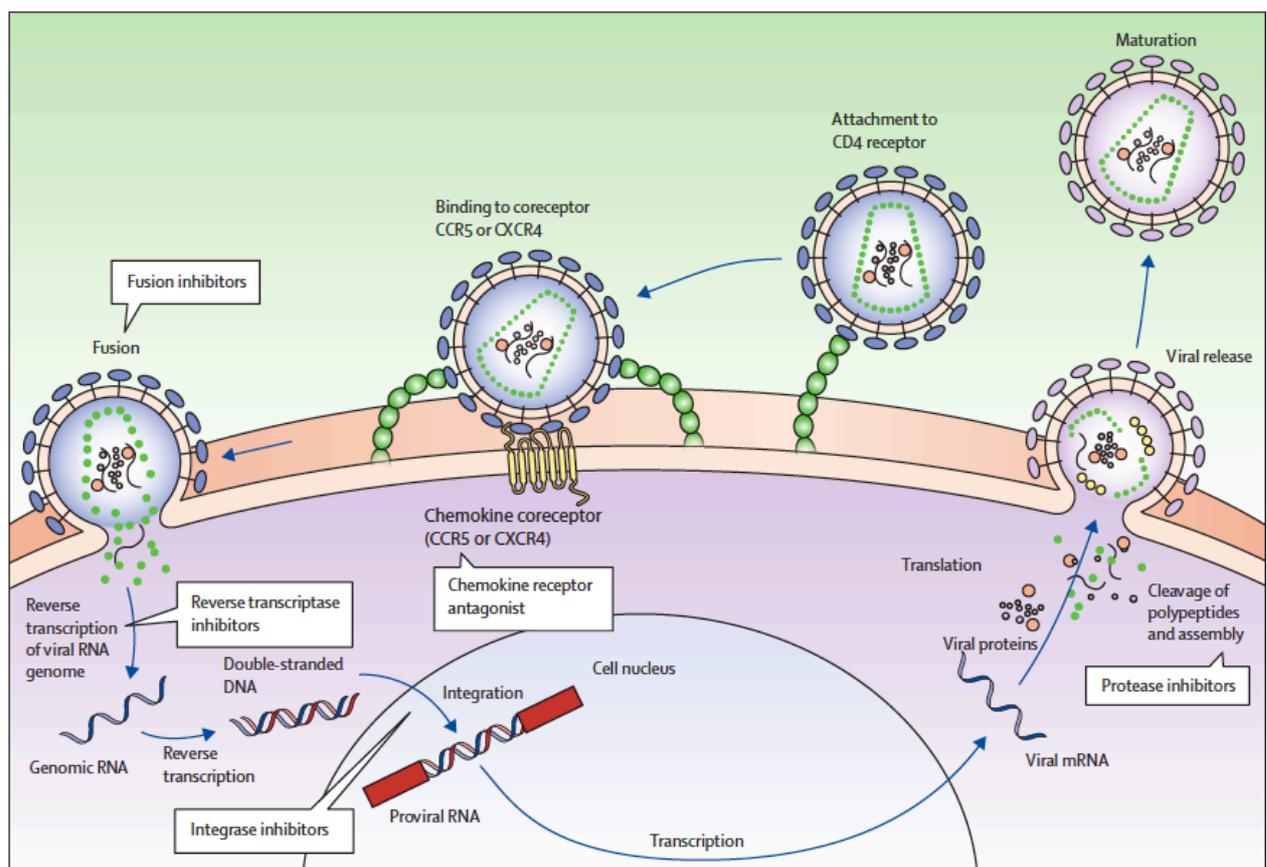
Al iniciarse la síntesis de ARN vírico, la proteína vírica tat aumenta del orden de  $10^2$ - $10^3$  la tasa de transcripción y, junto con otros factores, permite la síntesis de la totalidad del ARN del virus. Este transcrito es transportado al citoplasma mediante la acción de la proteína vírica rev.

Después de su síntesis, las proteínas víricas son procesadas mediante diferentes proteasas celulares y la proteasa vírica. Otras proteínas virales importantes en la fase final de la infección son vif y vpu. La proteína vpu aumenta la liberación de viriones y vif aumenta la infectividad entre 100 y 1000 veces. La proteína vif interacciona con una proteína celular con actividad citidina desaminasa llamada APOBEC3G y que representa un mecanismo de inmunidad innata antiretroviral, uniéndose al virión e interrumpiendo el proceso de retrotranscripción en células que serán infectadas en el siguiente ciclo de propagación.

La proteína vírica Vif contrarresta el efecto antivírico de la APOBEC3G impidiendo su incorporación al virión y promoviendo su degradación. La proteasa vírica procesa los

precursores proteicos gag y gag-pol en las proteínas de la nucleocápsida, la transcriptasa inversa y la proteasa.

El procesamiento de la gp160 en gp120 y gp41 se produce por una proteasa celular. La maduración final de los viriones y el ensamblaje correcto de las proteínas víricas se produce en el momento final del ciclo infeccioso, durante el proceso de gemación de los virus a través de la membrana celular, formándose una partícula vírica madura.



**Figura 3. Representación esquemática del proceso infeccioso del VIH**

## 1.4 Epidemiología del VIH-1

El VIH-1 se localizó inicialmente en países de África central y oriental. Debido a fenómenos de migración desde las áreas rurales en los cincuenta, el virus dejó de estar restringido a los núcleos rurales y se propagó a la población de las grandes ciudades. En la década de los sesenta, los fenómenos de despoblación en los cuales ciudadanos occidentales volvieron a sus países, las guerras civiles en las que participaron ejércitos extranjeros y el turismo ayudaron a propagar la epidemia. La diseminación se produjo en sucesivos pasos: una ruta epidemiológica lleva a las islas del Caribe y, desde allí, a las costas este y oeste de Estados Unidos. Otra ruta alcanza directamente Europa a partir de África central, y una tercera los países del sudeste asiático desde África central. Por último, cabe resaltar la ruta que siguieron los productos sanguíneos obtenidos por las multinacionales a partir de donaciones de individuos africanos de áreas afectadas.

A finales del 2003, la OMS estimó que entre 34 y 46 millones de personas estaban infectadas por VIH, aunque, probablemente la cifra real de infectados rondase los 200 millones. La incidencia fue especialmente elevada en los países del África subsahariana, en alguno de los cuales la incidencia alcanza el 10% de la población total y al 40% de las mujeres embarazadas. Se estimó que en Sudáfrica la vida media se reduciría desde 1980 en 10 años y que en el 2010 la esperanza de vida sería de 40 años.

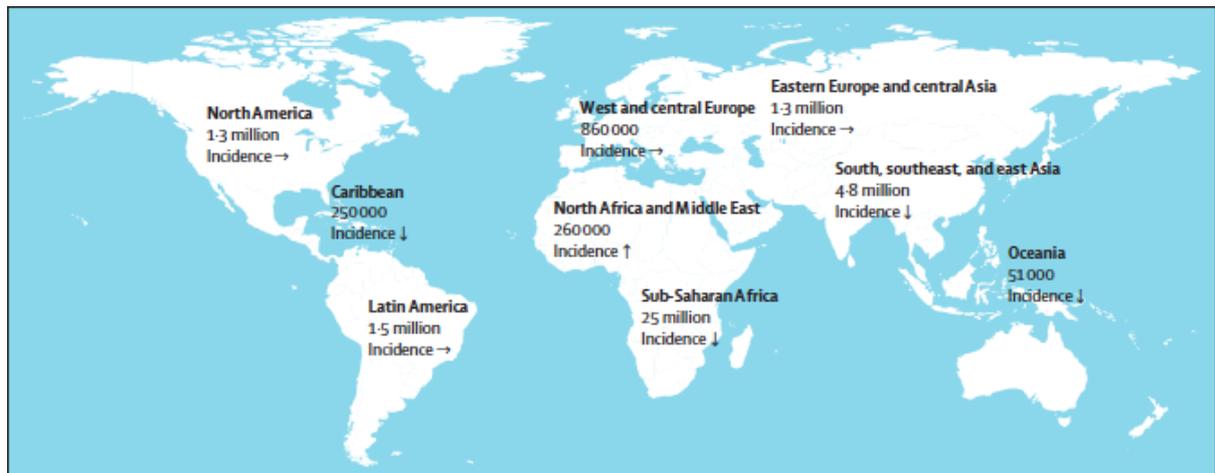
En el sudeste asiático, la situación no fue menos grave al vivir en esa zona la mitad de la población mundial. Asimismo, en los países latinoamericanos como Brasil y Argentina la situación es alarmante debido a la propagación de la enfermedad en la última década .

Esta situación contrasta con la de los países desarrollados que han podido mantener la epidemia en cifras estables en los últimos 5 años gracias a las medidas preventivas y a los tratamientos. En estos países, la tasa de mortalidad anual se sitúa en el 0,5-1,5% en contraposición al 7% de los países sin acceso a tratamiento.

El programa del sida en las Naciones Unidas (UNAIDS) estimó que en diciembre de 2007, el número total de personas en el mundo viviendo con el VIH/ sida era de 33 millones, destacando que 22 millones (67%) se encontraban en África subsahariana. A nivel mundial,

entre el 75-80% de las infecciones se produjeron por relaciones sexuales no protegidas, de las cuales más del 75% fueron heterosexuales. En la región Europea durante el 2007 se notificaron 48.892 nuevos diagnósticos de VIH representando una tasa de 76 por millón de habitantes. Una tercera parte (33%) fueron mujeres y con una edad comprendida entre 15 y 29 años. En global, las principales vías de transmisión fueron: HTX (36%), UDVP (32%) y HMX (20%). Asimismo, En 2012 se estimó que 35.3 millones de personas estaban infectadas por VIH y la prevalencia en África subsahariana se estimó entre el 70-80%.

Actualmente, la epidemiología global de la infección por el VIH ha cambiado como resultado del aumento de personas que tienen acceso a la terapia antiretroviral (TARGA); En 2012, 9.7 millones de personas con bajos ingresos e ingresos medios comenzaron la terapia antiretroviral. La prevalencia global ha incrementado de 31 millones de personas en 2002 a 35.3 millones en 2012, debido a la larga esperanza de vida de las personas que se encuentran con tratamiento antiretroviral. Sin embargo, la incidencia en la mayoría de los continentes tiende a decrecer (Figura 4), en este contexto la incidencia global que se estimó de 3.3 millones en 2002 ha decrecido hasta 2.3 millones en 2012(6). Esta reducción en la incidencia global del VIH se debe en gran medida a la reducción en la transmisión heterosexual. Por otra parte, campañas de prevención de la infección en personas adictas a drogas por vía parenteral (especialmente en países de Europa del este), han reducido significativamente la transmisión también en este colectivo. En regiones donde la principal ruta de transmisión del VIH se produce entre hombres que tienen sexo con hombres (Europa central y América), la incidencia se mantiene estable a pesar de la alta cobertura de la terapia antiretroviral. Los factores impulsores de la epidemia del VIH entre hombres que tienen sexo con hombres son complejos e incluyen el aumento de los comportamientos de riesgo desde la introducción de la terapia antiretroviral, el “optimismo terapéutico”, la existencia de un mayor riesgo de transmisión en el coito anal receptivo, las redes de transmisión sexual o clusters de transmisión, y el estigma social que disminuye el acceso a la terapia antiretroviral en algunos casos.



**Figura 4.** Número de personas que viven con VIH en 2012 y tendencia en la incidencia global de la infección por el virus de 2001 a 2012. Datos reportados a través de UNAIDS (actualizados a 2013)

Actualmente la tasa global de nuevos diagnósticos de VIH en España está en niveles similares a los de otros países de Europa occidental. Sin embargo, aunque la mejora respecto a décadas pasadas es indudable, la tasa es superior a la media de la Unión Europea. Las relaciones sexuales no protegidas entre hombres ocupan son el mecanismo más frecuente de transmisión en el conjunto global de datos. Estas transmisiones son mayoría entre las personas nacidas en España y entre los varones, sean españoles o extranjeros. Por ello, el colectivo de HSH es prioritario para los programas de prevención. Respecto a los nuevos casos de sida, la información aportada por el Registro Nacional indica que, tras más de una década de TARGA, el avance en la reducción de la incidencia de sida en España ha sido notable. Este descenso, inicialmente espectacular, se ha ido ralentizando en los últimos años.

## 1.5 Variabilidad genética del VIH

El VIH tiene varios mecanismos intrínsecos que aseguren una rápida evolución viral. La transcriptasa inversa carece de capacidad para confirmar que la transcripción del ADN es una copia exacta del código de ARN, y otorga una tasa de mutación de aproximadamente  $3.4 \times 10^{-5}$  mutaciones por par de bases por ciclo de replicación. Dado que el genoma del VIH se estima en  $10^4$  pares de bases de longitud y la tasa de referencia de la producción viral es de aproximadamente  $10^{10}$  viriones por día, millones de variantes virales se producen dentro de cualquier persona infectada en un solo día. Asimismo, la recombinación de VIH-1 puede conducir a una mayor diversidad viral producto de que una persona este coinfectada con dos cepas separadas de virus que se multipliquen en la misma célula (17).

El VIH se divide en tipos (VIH-1 y VIH-2), grupos, subtipos, sub-subtipos, formas recombinantes circulantes (CRF) y formas recombinantes únicas (URF).

En la clasificación que se utiliza actualmente propuesta por Robertson y Cols en 2000 se reconocen tres grupos filogenéticos dentro del VIH-1: M (principal, del inglés main), O (externo, del inglés outlier) y N (no M-noO, del inglés non-M non-O).

En la clasificación filogenética del VIH-1, existen subtipos diferentes subtipos dentro del grupo M: A-D, F-H, I y K. Los subtipos forman agrupaciones más o menos equidistantes unas de otras, en un árbol filogenético. En los subtipos A y F, se distinguen sub-subtipos designados como A1 y A2, y F1 y F2. Cada par de sub-subtipos está más estrechamente relacionado entre sí que con otros subtipos (Figura 5).

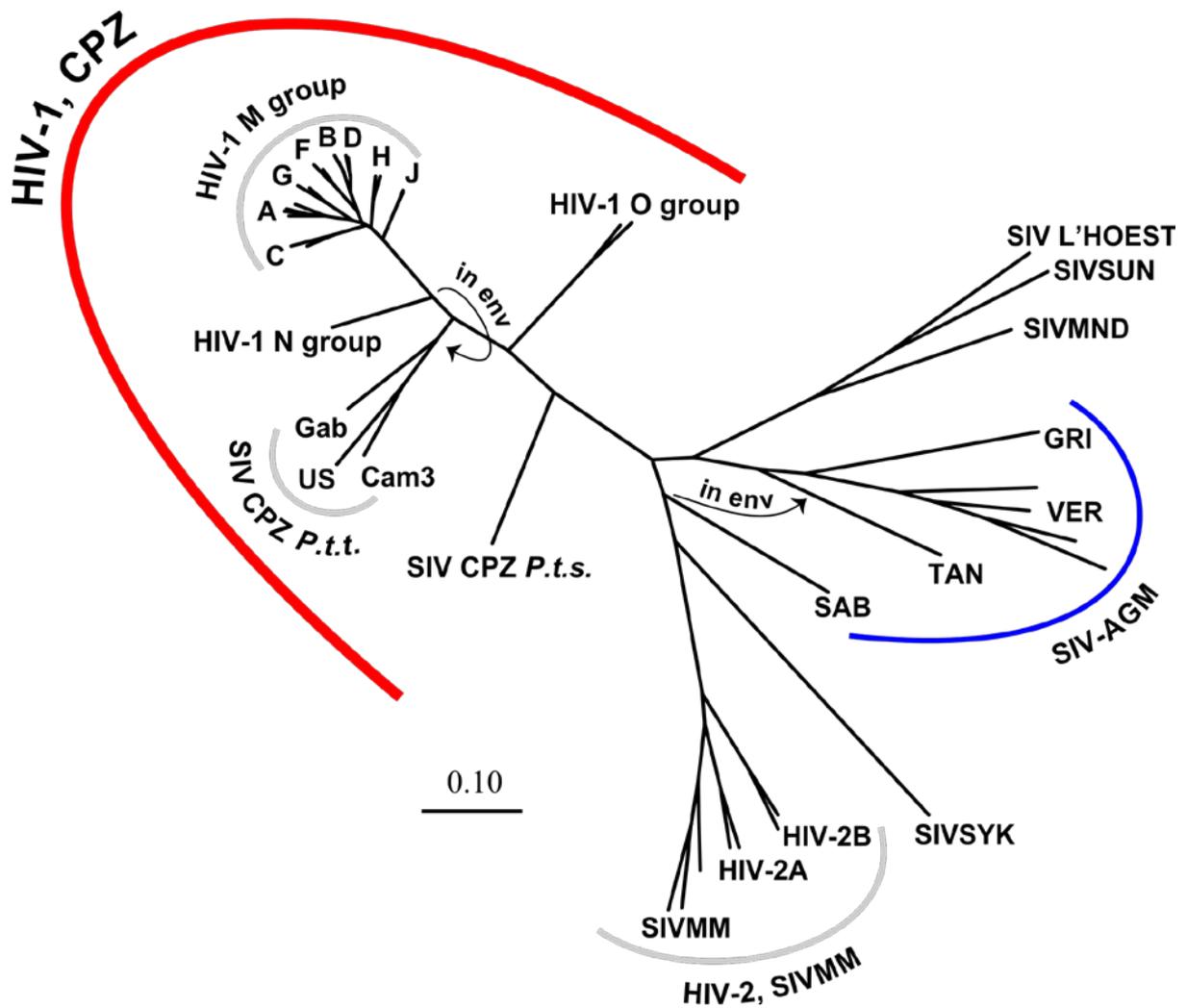


Figura 5. Arbol filogenético del virus de la inmunodeficiencia humana en humanos (VIH) y en simios (SIV).

### **1.5.1 Recombinantes inter-subtipos del VIH-1**

Se identifican dos tipos de formas recombinantes: las circulantes (CRF) y las únicas (URF). Cada CRF es designada por un número identificado con letras que indican los subtipos involucrados. Las letras son desplazadas por “cpx” (complejo, del inglés complex), que denota complejidad si más de dos subtipos están involucrados. Una gran mayoría de las CRF son genomas recombinantes inter-subtipos altamente complejos, algunos extremadamente estables y otros que tienen una apariencia de “retazos” debido a los numerosos puntos de entrecruzamiento.

Conjuntamente con las CRF se han reportado varias cepas de VIH-1 con estructuras de mosaico únicas. Estas formas se conocen como URF y no han mostrado ninguna evidencia de diseminación epidemiológica (formas presentes en menos de 3 individuos sin vínculo epidemiológico).

Los subtipos de la mayoría de las cepas VIH-1 pueden ser determinados por el análisis de la secuencia de cualquiera de las regiones genómicas. La variación genética dentro de un subtipo, puede ser de 15 a 20 %, mientras que la variación entre los subtipos suele ser del 25-35% (18). El subtipaje a partir de una única región del genoma, debe realizarse cuidadosamente porque pueden perderse cepas recombinantes. Para definir un nuevo subtipo, sub-subtipo o CRF, deben identificarse cepas representativas al menos en tres individuos que no tengan vínculo epidemiológico. La estructura de mosaico tiene que coincidir y la consistencia del agrupamiento filogenético debe caracterizarse. Al menos en dos de las cepas, el genoma debe secuenciarse casi en su totalidad.

### 1.5.2 Distribución mundial de los diferentes subtipos de VIH-1

La distribución del VIH-1, subtipos y recombinantes en cada región es notablemente diferente en todo el mundo (Figura 6). La mayor diversidad se encuentra en el centro de África, donde todos los subtipos y muchos CRF y URFs están representados. Los seis países de la región albergan una gran diversidad de subtipos, pero difieren en su presencia. En la República Democrática del Congo, se encuentran todos los subtipos, formas recombinantes CRF01\_AE, CRF02\_AG y muchos otros CRF y URFs, a excepción del subtipo B. En África occidental, se detectan todos los subtipos con la mayoría de variantes dominantes para CRF02\_AG y subtipo G. En África oriental, la mayoría de las infecciones se deben al subtipo A, mientras que el resto a los subtipos C, D y URFs. En el sur de África, Etiopía y la India, las epidemias son casi exclusivamente causadas por subtipo C. El subtipo B domina el norte de América, Caribe, América Latina, Europa central y occidental y Australia. En Europa occidental y central, se detectan todos los subtipos principales y muchos CRF y URFs. La epidemia en Europa oriental y Asia central está dominada por el subtipo A y B. En el sur y el sudeste de Asia, CRF01\_AE es responsable de la gran mayoría de las infecciones. En esta región, la proporción combinada de todas las infecciones recombinantes es de 86%, la más alta del mundo. En Asia, la epidemia está dominada por CRF07\_BC, CRF08\_BC, CRF01\_AE y subtipo B. El medio Oriente y el norte de África están afectados principalmente por subtipo B y CRFs diferentes. La mayoría de la distribución del subtipo C está presente en el sur de África y la India, en el este de África y Etiopía. El subtipo A se encuentra principalmente en el este de África, Europa oriental y Asia central y el resto en África occidental y central y en el sur y sudeste asiático. La epidemia por el subtipo B es más amplia y distribuida uniformemente que los otros subtipos. CRF02\_AG es a nivel mundial la cuarta en presencia y se concentra en el oeste de África, con números más pequeños en el centro de África, Oriente Medio y África del Norte. CRF01\_AE es el quinto subtipo, se encuentra en el sur y el sudeste de Asia, Asia oriental y un pequeño número en el centro de África. El subtipo G se concentra en África occidental y central. El subtipo D está presente principalmente en el este de África, con nuevas infecciones en África central y occidental. El subtipo F es distribuido uniformemente en todo el mundo, mientras que el subtipo H, J y K se encuentra en el centro, sur y oeste de África. CRF03\_AB no juega un papel importante a nivel mundial o regional. Otros CRF se distribuyen de manera diferencial en el oeste de África (principalmente CRF06\_cpx), Asia oriental (principalmente CRF07\_BC y CRF08\_BC en China), África central (CRF11\_cpx entre otros), América Latina (CRF12\_BF, CRF28\_BF,

CRF31\_BC, CRF38\_BF y otros), medio Oriente y el norte de África (principalmente CRF06\_cpx) y el sur y el sureste de Asia (principalmente CRF35\_AD y CRF07\_BC). Una amplia variedad de URFs se distribuyen en África subsahariana, América Latina (principalmente recombinantes únicas de subtipos B y F) y el sur y sureste de África (18) .

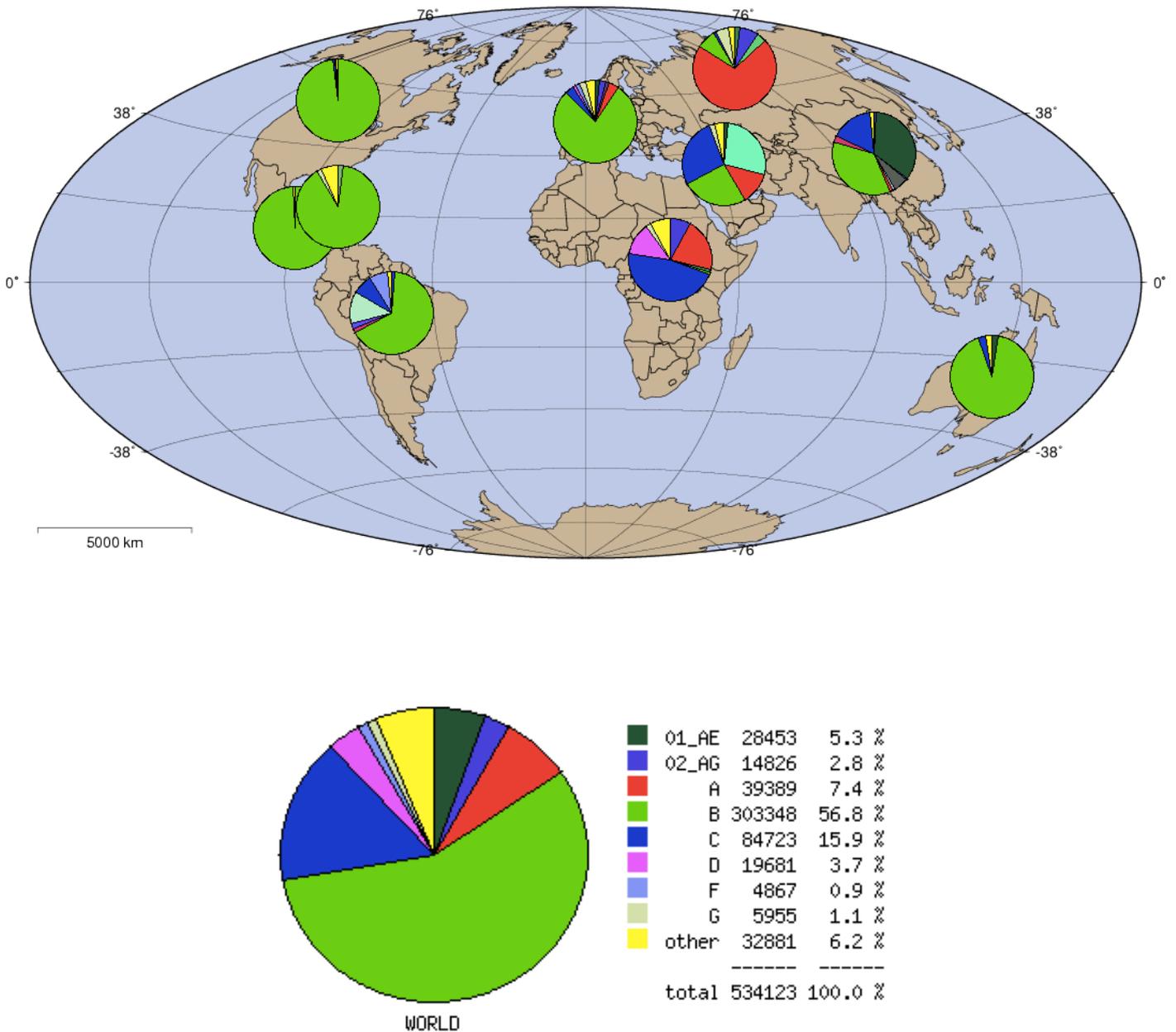
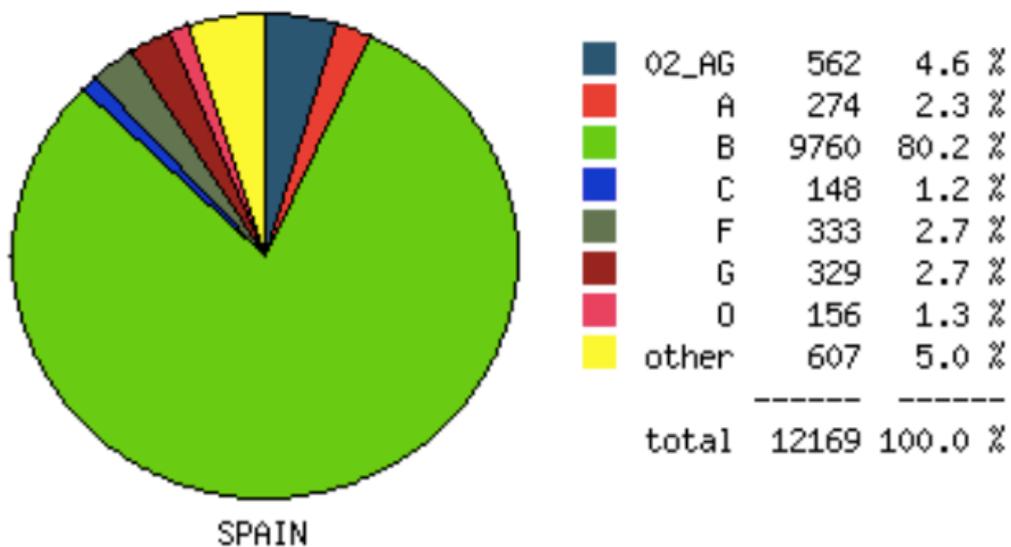


Figura 6. Distribución mundial de subtipos de VIH-1 y formas recombinantes.

La mayor prevalencia en España sigue siendo para el subtipo B (80.2%) (según datos recogidos por HIV los Alamos actualizados en 2015). La forma recombinante CRF02\_AG es la más prevalente (4.6%), seguida de otros subtipos puros como el G (2.7%), A (2.3%) o F (2.7%) (Figura 7).



**Figura 7. Prevalencia de subtipos de VIH-1 en España. Datos disponibles de HIV los Alamos (Actualizados en 21 de abril de 2015)**

La prevalencia de subtipos no-B ha aumentado en los últimos años en ciudadanos españoles. En un estudio reciente realizado en Galicia (19), de 230 personas infectadas por VIH desde 2009 a 2013 en un centro, un 37.8% fueron subtipos no-B con la siguiente distribución: F (29.6%), C (2.6%), A (2.2%), CRF02\_AG (1.7%), G (0.9%), K (0.4%) y CRF01\_AE (0.4%).

### **1.5.3 Implicaciones clínicas y epidemiológicas de los subtipos del VIH.**

La mayoría de los primeros genomas analizados a nivel molecular y fenotípico del VIH-1 fueron variantes del subtipo B, que eran las formas más prevalentes en los países desarrollados. Por este motivo, gran parte de las investigaciones sobre biología y patogénesis del VIH-1 se realizaron con cepas de este subtipo. Sin embargo, las consecuencias sobre transmisibilidad, patogenicidad, diagnóstico, tratamiento y diseño de vacunas que resultan de la alta tasa de variabilidad genética del VIH hacen esencial la caracterización de los subtipos del VIH-1 que circulan en cada área geográfica. Basándose en las diferencias de transmisibilidad entre VIH-1 y VIH-2, se ha sugerido que ciertos subtipos pueden presentar un mecanismo predominante de transmisión (homosexual, heterosexual, vertical o parenteral), pero esto no ha sido aún demostrado. Algunos trabajos han encontrado diferencias entre los subtipos y el grado de progresión a SIDA (20), al contrario que otros (21). Las diferencias entre los resultados de estos estudios podrían deberse también a factores ajenos al virus (huésped, condiciones ambientales, diseño de los estudios, tamaño de muestra, duración del estudio). Los procedimientos diagnósticos, optimizados inicialmente para el diagnóstico del subtipo B, pueden dar resultados falsamente negativos o dudosos cuando se aplican a otros subtipos. De ahí la importancia de incorporar subtipos no-B en técnicas de diagnóstico de VIH, tanto serológico como molecular.

Sin embargo, la optimización de estas técnicas para subtipos y CRFs poco frecuentes puede ser difícil, dado el escaso número de secuencias completas de las que disponemos, lo que complica su validación. Igualmente, algunas de las técnicas de cuantificación de la carga viral se basan en el uso de primers diseñados frente a secuencias del subtipo B del VIH-1, que podrían no hibridar o hibridar defectuosamente con el genoma de los aislados de subtipos no-B (22), lo que daría lugar a valores falsos de carga viral, indetectables o inferiores a los reales.

## 1.6 Generalidades del tratamiento antirretroviral

Los fármacos antirretrovirales aprobados hasta el momento se pueden clasificar en cuatro grupos principales: inhibidores de la RT, análogos de nucleósidos o nucleótidos (ITIAN), inhibidores de la RT no análogos de nucleósidos (ITINAN), inhibidores de la proteasa (IP), e inhibidores de la integrasa. Para nuestro estudio, hemos considerado exclusivamente dos grupos, los cuales se reflejan a continuación en la siguiente tabla. (Tabla 1).

INHIBIDORES DE LA TRANSCRIPTASA INVERSA	
Análogos de los nucleósidos/nucleótidos (ITIAN)	
<ul style="list-style-type: none"><li>• ZIDOVUDINA (AZT)</li><li>• DIDANOSINA (ddI)</li><li>• ZALCITABINA (ddC)</li><li>• ESTAVUDINA (d4T)</li><li>• EMTRICITABINA (FTC)</li><li>• LAMIVUDINA (3TC)</li><li>• ABACAVIR (ABC)</li><li>• TENOFOVIR DF (TDF)</li></ul>	
No análogos de los nucleósidos (ITINAN)	
<ul style="list-style-type: none"><li>• NEVIRAPINA (NVP)</li><li>• DELAVIRDINA (DLV)</li><li>• EFAVIRENZ (EFV)</li><li>• ETRAVIRINA (ETV)</li></ul>	
INHIBIDORES DE LA PROTEASA (IPs)	
<ul style="list-style-type: none"><li>• NELFINAVIR (NFV)</li><li>• SAQUINAVIR (SQV)</li><li>• INDINAVIR (IDV)</li><li>• RITONAVIR (RTV)</li><li>• LOPINAVIR (LPV)</li><li>• ATAZANAVIR (ATV)</li><li>• FOSAMPRENAVIR (FPV)</li><li>• TIPRANAVIR (TPV)</li><li>• DARUNAVIR (DRV)</li></ul>	

**Tabla 1. Fármacos aprobados para la RT y Proteasa**

### **1.6.1 Fármacos inhibidores de la RT.**

La RT, enzima fundamental en el ciclo biológico del VIH, es un heterodímero compuesto de dos subunidades, denominadas p66 y p51. En la subunidad p66 se localiza el sitio activo del enzima, formado por el dominio polimerasa unido al dominio ARNasa H. Los inhibidores de la RT se clasifican en dos grandes grupos en base a sus características moleculares, que determinan mecanismos de inhibición distintos para ambos grupos de fármacos.

#### **- Inhibidores de la RT análogos de nucleósidos / nucleótidos (ITIAN)**

Los fármacos inhibidores de la RT Análogos de Nucleósidos son estructuralmente semejantes a los nucleósidos naturales y, al igual que éstos, requieren ser trifosforilados en el citoplasma celular por una fosfoquinasa para ser biológicamente activos. Una vez fosforilados, compiten con los dNTPs en su incorporación a la cadena de ADN sintetizada por la RT interrumpiendo su proceso de síntesis, ya que carecen de grupo hidroxilo en la posición 3' del azúcar y los nuevos nucleótidos no se pueden unir a la cadena de ADN sintetizada.

Los inhibidores de la RT Análogos de Nucleótidos actúan de forma similar a los análogos de nucleósidos pero, al estar parcialmente fosforilados, necesitan menos fosforilaciones en el interior de la célula para ser activos.

#### **- Inhibidores de la RT no análogos de nucleósidos (ITINAN)**

Los ITINAN no requieren activación metabólica intracelular. A diferencia de los ITIAN no se incorporan a la cadena de ADN en formación sino que se unen de forma no competitiva a un lugar cercano al centro catalítico de la enzima, una región hidrofóbica de la subunidad p66. Esta unión conduce a un cambio conformacional en la estructura del enzima, que impide la colocación correcta del complejo molde-iniciador imposibilitando la polimerización del ADN.

### **1.6.2 Fármacos inhibidores de la proteasa**

La proteasa del VIH es un homodímero constituido por dos subunidades de 99 aminoácidos codificadas en el extremo 5' del gen pol que interaccionan con el sustrato a través de una hendidura situada en la zona de contacto entre ambas subunidades. Esta enzima actúa en la fase final del ciclo de replicación viral, participando en el procesamiento de las poliproteínas Gag y Gag-Pol. Los fármacos inhibidores de la proteasa (IPs) se unen al centro activo de esta enzima compitiendo con sus sustratos naturales (Gag y Gag-Pol) lo que resulta en la formación de partículas virales inmaduras y, por consiguiente, no infecciosas.

## 1.7 Interpretación de la resistencia del VIH-1 a los fármacos antiretrovirales.

Una de las principales causas de resistencia a los fármacos antirretrovirales es la aparición de mutaciones en la región codificante de la transcriptasa inversa y la proteasa.

Las mutaciones se representan por una letra mayúscula (Figura 8) indicativa del nombre del aminoácido en la cepa wild-type o salvaje, sin mutaciones, seguida del número que indica la posición del aminoácido y finalmente la letra del aminoácido mutado. Así, por ejemplo, M41L indicaría que en la posición 41 del virus mutado aparece una leucina (L) en vez de una metionina (M).

Nº	AMINOÁCIDO	ABREVIATURA	SÍMBOLO
1	Alanina	Ala	A
2	Cisteina	Cys	C
3	Acido aspártico	Asp	D
4	Acido glutámico	Glu	E
5	Fenilalanina	Phe	F
6	Glicina	Gly	G
7	Histidina	His	H
8	Isoleucina	Ile	I
9	Lisina	Lys	K
10	Leucina	Leu	L
11	Metionina	Met	M
12	Asparagina	Asn	N
13	Pirrolisina	Pyl	O
14	Prolina	Pro	P
15	Glutamina	Gln	Q
16	Arginina	Arg	R
17	Serina	Ser	S
18	Treonina	Thr	T
19	Selenocisteina	Sec	U
20	Valina	Val	V
21	Triptofano	Trp	W
22	Tirosina	Tyr	Y

Figura 8. Nomenclatura internacional de los diferentes aminoácidos.

Según la guía de resistencia a los antirretrovirales publicada en 2013 por la red de investigación en sida de España (RIS), se exponen las mutaciones más importantes a los fármacos antirretrovirales de primera línea (figura 9) y por consiguiente las que deberían aparecer en los informes de resistencia.

---

**ITIAN**

M A K D T K L V F Y F Q M L T K  
41 62 65 67 69 70 74 75 77 115 116 151 184 210 215 219

**ITINAN**

L A L K K V V E V Y Y V G H P F M K  
90 98 100 101 103 106 108 138 179 181 188 189 190 221 225 227 230 238

**IP**

L V L D V L K M I G I F I Q A G T L V N I N L L  
10 11 24 30 32 33 43 46 47 48 50 53 54 58 71 73 74 76 82 83 84 88 89 90

**Figura 9. Posiciones que deben aparecer recogidas en los informes de resistencia**

En la guía RIS además se establecen cuales serían las posiciones de resistencia a tener en cuenta en pacientes naïve desde el punto de vista epidemiológico, ya que estas serían las únicas mutaciones en las cuales ha sido comprobada la transmisión entre pacientes (Figura 10).

### Mutaciones que indican la transmisión de un virus resistente en pacientes naïve

#### ITIAN

M	K	D	T	K	L	V	F	Y	F	Q	M	L	T	K
41	65	67	69	70	74	75	77	115	116	151	184	210	215	219
L	R	E/G/N D		E/R	I/VA/M/T/S L			F	Y	M	I/V	W	C/D/E/F/I/N/S/V/Y E/N/Q/R	
		Ins/del												

#### ITINAN

L	K	K	V	V	Y	Y	G	P	M
100	101	103	106	179	181	188	190	225	230
I	E/P	N/S	A/M	F	C/I/V	C/H/L	A/E/S	H	L

#### IP

L	L	D	V	M	I	G	I	F	I	G	L	V	N	I	I	N	L
23	24	30	32	46	47	48	50	53	54	73	76	82	83	84	85	88	90
I	N	I	I/L	A/V	M/V	L/V	L/YA/L/M/S/T/V			A/C/S/T	VA/C/F/L/M/S/T		D A/C/V V D/S M				

Figura 10. Esta tabla recoge aquellas mutaciones que, estando presentes en pacientes naïve sin exposición previa a tratamiento antirretroviral, confirman la transmisión de un virus resistente según la última lista publicada por la Organización Mundial de la Salud (OMS). Granate: Mutaciones de reversión. Ins/del: inserciones o deleciones entre las posiciones 67 y 69.

### 1.7.1 Mutaciones de resistencia transmitidas

La terapia antirretroviral de gran actividad (TARGA) ha cambiado la historia natural de la infección VIH/SIDA, al retrasar la evolución de la enfermedad y mejorar la calidad de vida de los individuos infectados. Los países desarrollados han implementado este tipo de terapia masivamente en su población infectada en seguimiento, con resultados muy favorables. Uno de los principales problemas en la actualidad es el fracaso terapéutico y, dentro de éste, el aumento de resistencia a drogas antirretrovirales. Los virus resistentes a los fármacos también pueden transmitirse de persona a persona. En el ámbito mundial, la resistencia del VIH a los medicamentos se considera un problema creciente y en este sentido la OMS recomienda la vigilancia epidemiológica de la farmacoresistencia del VIH.

Este problema se origina por la asociación de las características propias del virus (variabilidad genética, latencia y reactivación, adaptación a vías de entrada, infección de reservorios), junto a una inadecuada supresión viral (adherencia, farmacocinética) y uso de monoterapia y biterapia previos (23). La menor susceptibilidad del virus a las drogas está asociada a mutaciones en los genes que codifican, principalmente, para las enzimas transcriptasa reversa y proteasa viral. Según el momento de adquisición de la resistencia podemos clasificar las resistencias a antirretrovirales en dos tipos (24):

- Resistencia Primaria o transmitida (TDR): Virus mutantes resistentes a uno o varios fármacos detectados en pacientes que no han sido sometidos previamente a tratamiento antirretroviral.
- Resistencia Secundaria o adquirida: Virus mutantes resistentes detectados en pacientes que han tenido experiencia a tratamiento antirretroviral.

La prevalencia de resistencias primarias ha tenido un importante aumento en todo mundo desde el inicio de la terapia antirretroviral. Las tasas de resistencia primaria han llegado a ser muy altas en Estados Unidos y Europa, probablemente debido a que han tenido acceso por más tiempo a las drogas antirretrovirales (25). La prevalencia a resistencias transmitidas por zonas del mundo oscila entre 4 y 13% (América del Norte 12,9%, Europa 10,9%, América Latina 6,3%, África 4,7% y Asia 4,2%)(26).

La transmisión de resistencias a fármacos antirretrovirales en pacientes que no han recibido nunca tratamiento es uno de los factores que limitan las opciones terapéuticas en la terapia de inicio, además de reducir su eficacia (27). El conocimiento de la prevalencia de estas resistencias primarias varía mucho de unos países a otros, observándose, como hemos dicho anteriormente, las mayores tasas en zonas donde la terapia antirretroviral está bien establecida como Europa Occidental y América del norte (28–30). Pero además este porcentaje varía en función del diseño del estudio no existiendo en la actualidad criterios concretos. Entre los factores de riesgo descritos involucrados en la transmisión de virus resistentes en pacientes naïve cabe destacar:

- La elevada carga viral
- Enfermedades de transmisión sexual
- Factores genéticos aunque no se conoce el grado de influencia

Además de estos factores, la presencia de mutaciones a más de una clase de fármacos antirretrovirales también afecta a la transmisibilidad estando disminuida en virus con resistencias primarias a dos o tres clases de fármacos. Es importante además conocer las características de los sujetos donde la tasa de resistencias es más frecuente. Aunque existen estudios contradictorios, se sugiere que la transmisión de virus con resistencia primaria es más frecuente en la etnia caucasiana, pacientes infectados con subtipo B y vía de transmisión homosexual (31,32). El que estos factores se relacionen con mayor transmisión de mutaciones de resistencia pueden ser justificados si tenemos en cuenta que los países con mayor población de raza caucasiana e infectados por subtipos B son los más desarrollados y, por tanto, los que tienen mayor acceso a la terapia antirretroviral y que los pacientes VIH homosexuales suelen ser los más concienciados a la hora de realizar la prueba de VIH. También podemos encontrar otros estudios que muestran resultados discrepantes, encontrando tendencias significativas en el aumento de TDRs en otros grupos diferentes a los anteriores (33) o no encontrando diferencias significativas al estudiar estas variables sociodemográficas (34). La alta tasa de variabilidad genética del VIH tiene también

importantes implicaciones en patogenicidad, transmisibilidad, diagnóstico y tratamiento; con un impacto potencial en la epidemiología, terapia y prevención de la enfermedad. Por todas estas implicaciones, resulta esencial la caracterización de los subtipos genéticos del VIH-1 que circulan en cada área geográfica.

## **1.8 Las herramientas de análisis filogenético en el control de la epidemia de VIH.**

La infección por el VIH ha pasado de ser una pandemia controlada en países desarrollados, a experimentar un incremento en el número de casos en países con un nivel económico y sanitario adecuado, especialmente en determinados colectivos, como en el de hombres que tienen sexo con hombres (HSH). Los cambios de comportamiento, consecuencia de vivir en una era de terapia antiviral eficiente, han contribuido al incremento de la propagación del VIH. En este contexto, el resultado de estudiar la asociación de la diversidad viral con diferentes parámetros demográficos, sugieren una epidemia muy compartimentada, determinada por las características sociales y los hábitos de comportamiento de estos pacientes (35).

La filogenia molecular es la ciencia de la estimación de la evolución basada en la comparación de secuencias genéticas. El análisis filogenético nos permite inferir relación evolutiva entre diferentes secuencias moleculares, por ello, los análisis mediante secuencias a través de herramientas para el análisis filogenético y de epidemiología molecular pueden ser usados para evaluar diferentes aspectos del virus: el origen de la epidemia de VIH-1(5,36), casos legales (37), análisis de clusters de transmisión de VIH (38), describir la relación existente entre los eventos migratorios y la diseminación del HIV a nivel local (filogeografía) (18,39–42) y entender el crecimiento o decrecimiento en determinados subtipos de VIH (filodinámica)(43–45).

La filodinámica es uno de los más emocionantes e interesantes campos en curso en el que la filogenética está contribuyendo a nuestra comprensión de la evolución del virus, en particular en el VIH. La dinámica de la población viral puede ser apoyada utilizando métodos basados en la teoría de la coalescencia, de manera que realizan una reconstrucción de la historia evolutiva sobre su distribución temporal y espacial. Además, este marco temporal se puede mejorar mediante el modelado de forma explícita de las tasas de dispersión espacial en un contexto filogeográfico(46).

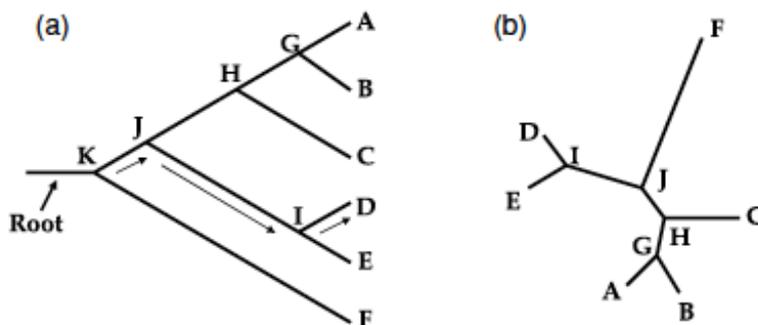
Cada vez es mayor el número de estudios que utilizan las técnicas de análisis filogenético (47) en diferentes cohortes(48), analizando la situación local, con el objetivo de poder

establecer las adecuadas medidas de contención. Por lo tanto, el uso de técnicas de análisis filogenético en grandes estudios de cohorte para el control de la propagación del virus del VIH, se plantea como una potente herramienta de alerta epidemiológica.

### 1.8.1 Introducción a los métodos de análisis filogenético

Un árbol filogenético es una estimación de la historia evolutiva de un conjunto de secuencias (49) y el medio a través del cual se puede inferir relación evolutiva entre estas secuencias. La historia evolutiva suele ser representada normalmente como una ramificación dentro del árbol filogenético.

Para una estima real de la relación filogenética entre secuencias homólogas todas las muestras deben estar bajo un mismo marco evolutivo. Las relaciones entre estas secuencias pueden ser ilustradas usando un árbol filogenético (Figura 7).



**Figura 11. Estructura de un árbol enraizado (a) y (b) no enraizado. Ambos árboles presentan la misma topología. Un árbol enraizado es representado con la raíz a la izquierda. A, B, C, D y E son los nodos externos. G, H, I, J y K son los nodos internos con K como nodo principal. Las líneas que salen de los nodos son las ramas. Las flechas indican la dirección de la evolución en el árbol enraizado, siendo desconocida en árbol de la derecha. Los taxones A, B y C forman un cluster con un mismo nodo común (H).**

Los objetos de estudio (los taxones terminales) pueden ser especies, géneros o cualquier otra categoría taxonómica y reciben el nombre de UTO (unidades taxonómicas operacionales) (en inglés, OTU, operational taxonomic units).

Un grupo de taxones (familia de organismos, individuos o virus), que comparten un mismo nodo con un mismo origen filogenético se denomina clado o cluster. El termino extrapolable en individuos infectados por el virus de la inmunodeficiencia humana y que comparten una misma cepa vírica ancestral infectiva, se denominaría “cadena o cluster de transmisión”.

La reproducibilidad, estabilidad y fiabilidad de la topología de los árboles filogenéticos generados debe analizarse mediante métodos estadísticos. El más utilizado se denomina *bootstrapping* o análisis de bootstrap, que asigna un valor (valor de bootstrap) a cada nodo, expresado en forma de porcentaje. El análisis de bootstrap es la manera más sencilla de reflejar la calidad o la precisión de nuestro árbol filogenético (50), se obtienen numerosas submuestras aleatorias de nuestra matriz de datos, se construyen árboles filogenéticos de cada una de ellas y se calcula la frecuencia con la que dos secuencias divergen a partir de un nodo en los nuevos árboles (valor de bootstrap) (51). En general los valores de bootstrap de 70% o por encima nos proporcionarían agrupaciones fiables (52).

El origen común en el tiempo de un gen o la divergencia de un grupo de taxones puede ser analizado mediante herramientas de análisis filogenético. El análisis filogenético puede ser utilizado para calcular el tiempo del más reciente común antecesor (en inglés, TMRCA, time most recent common ancestor), que es el tiempo aproximado en el que se origina un linaje, es decir, un grupo de secuencias que comparten un nodo ancestral común.

Métodos probabilísticos basados en Inferencia Bayesiana permiten establecer un marco temporal de nuestro árbol filogenético facilitándonos etiquetar a cada secuencia con datos temporales. En este contexto, nos permiten establecer el TMRCA o el origen aproximado de aparición de una epidemia y de los diferentes linajes (clusters de transmisión). Para establecer un soporte estadístico de los diferentes nodos utilizan valores de probabilidad a posteriori (pp), que son similares a los valores de bootstrap que utilizan la mayoría de métodos de construcción de árboles filogenéticos. En general valores de probabilidad a posteriori por encima de 0.8 serían fiables.

## **-Pasos previos a la construcción de árboles filogenéticos**

Los pasos previos a la obtención de filogenias moleculares se centran en la selección de regiones concretas de material genético para la realización de los árboles, amplificación y secuenciación de estas. Dichas secuencias son revisadas, alineadas en función de sus homologías y posteriormente analizadas utilizando el mejor modelo de reconstrucción filogenético que se adapte a nuestro estudio evolutivo. A partir de estos árboles filogenéticos, si el muestreo taxonómico y de regiones ha sido adecuado, se pueden realizar inferencias sobre las relaciones evolutivas, y estudiar la relación filogenética entre las diferentes secuencias. El análisis filogenético establece las relaciones entre genes o fragmentos de un individuo para inferir la historia común de un conjunto de secuencias de la misma especie, por ello, para llegar a comparar estas secuencias es imprescindible que las regiones sean homólogas.

Lo más importante para realizar una construcción filogenética de un árbol filogenético es asumir que el alineamiento de las secuencias es correcto. Hasta 1989 estos alineamientos se solían hacer manualmente, debido a que la alineación exhaustiva de más de seis u ocho secuencias era computacionalmente inviable. Los errores en el alineamiento pueden conducir a la construcción de un árbol incorrecto. Hoy en día, el alineamiento múltiple de secuencias, es realizado por el método de “aproximación progresiva”, este método construye una alineación paso a paso, empezando por las secuencias similares y añadiendo progresivamente las secuencias más divergentes(53,54). Algunos de los algoritmos más populares comienzan comparando la similaridad de todos los pares de secuencias, alineando en primer lugar las dos secuencias con la mayor similaridad. Las otras secuencias son añadidas progresivamente en función de su similaridad. Actualmente existen diversas herramientas que permiten alinear un conjunto de secuencias. Entre los más destacados se encuentra MUSCLE(55) o ClustalW(56).

Otro problema que se plantea tras el alineamiento, es cómo interpretar las inserciones o deleciones que se presentan en la secuencia (Gaps), puesto que estas inserciones o deleciones también pueden tratarse de un fallo de secuenciación. Los genes no toman generalmente inserciones y deleciones a la ligera, ya que una inserción o deleción podría hacer cambiar el marco de lectura y por lo tanto sería inviable. A menudo, es recomendable revisar el alineamiento de forma manual con programas de edición de secuencias. Muchos programas

de reconstrucción filogenética cuentan de diferentes parámetros para valorar de manera diferente a estos Gaps (57).

Cuando el alineamiento es óptimo hay varios métodos diferentes para crear una filogenia. A pesar de que existen varios métodos para la construcción de árboles filogenéticos, ninguno de ellos garantiza que el árbol filogenético inferido refleje absolutamente la verdadera reconstrucción de la historia evolutiva. A grandes rasgos, los métodos se pueden dividir en métodos basados en distancias genéticas (medida de la diferencia del material genético entre distintas especies o individuos de la misma especie) y métodos basados en caracteres.

### 1.8.2 Métodos de construcción de árboles filogenéticos

Los métodos para la construcción de arboles filogenéticos a partir de datos moleculares se clasifican en función del tipo de dato que se use:

- Caracteres discretos: Se usan las sustituciones individuales entre las secuencias para determinar las relaciones ancestrales más probables
- Matriz de distancias genéticas: Los métodos basados en distancia primero calculan la distancia total entre todos los pares de secuencias, y luego calculan un árbol en base a estas distancias.

También se clasifican en acorde a el algoritmo de aproximación que utilice el método:

- Algoritmo basado en agrupamiento o “clustering”, en el cual solo obtenemos un árbol filogenético.
- Algoritmo basado en un criterio óptimo evaluando diferentes topologías para diferentes arboles.

	Criterio óptimo de búsqueda	Agrupamiento
Carácteres	Máxima parsimonia (MP) Maxima verosimilitud o Maximum likelihood (ML) Inferencia Bayesiana	
Distancias genéticas	Fitch-Margoliash	UPGMA Neighbor-joining(NJ)

**Tabla 2. Clasificación de los métodos de análisis filogenético y sus estrategias de agrupamiento.**

Los métodos para la obtención de filogenias en VIH más utilizados actualmente para el análisis de los clusters de transmisión, son el método de Neighbor-joining (NJ) y el método de Máxima verosimilitud o Maximum likelihood (ML). Existen métodos más recientes de Inferencia Bayesiana, para para la obtención de relojes moleculares. Estas filogenias implementan un contexto temporal en sus arboles filogenéticos.

### **-Métodos basados en distancias genéticas**

Los métodos basados en una matriz de distancias genéticas, comienzan calculando la disimilitud de cada par de secuencias, para producir una matriz de distancias genéticas y de ahí inferir las relaciones filogenéticas.

Se han propuesto muchos modelos diferentes para estimar con precisión la distancia evolutiva entre secuencias. La manera más simple para calcular una distancia entre dos secuencias es contar el número de diferencias. Esta medida es llamada la distancia de Hamming. El problema es que, este método puede ser engañoso, puesto que pueden haber sitios aparentemente iguales que hayan sufrido sin embargo dos eventos evolutivos. Además pueden existir tasas de evolución no uniformes entre diferentes sitios del material genético. Modelos de distancia evolutiva tales como Jukes-Cantor, Kimura 2 parámetros, Jin-Nei, etc... pueden resolver estos problemas.

Muchas simulaciones de modelos evolutivos, muestran que la importancia de la utilización de un modelo de sustitución realista para estimar la distancia genética depende en gran medida de la divergencia de las secuencias. Si el número esperado de sustituciones por sitio es pequeño (por debajo de 0,2), el árbol resultante no cambia mucho cuando se utilizan diferentes modelos de sustitución. Si las secuencias son muy divergentes sin embargo, y especialmente cuando el número de sustituciones por sitio se eleva por encima de 1, las diferencias se hacen marcadas y la elección del modelo de sustitución es importante.

El método más utilizado en distancias genéticas es el método de Neighbor-Joining (NJ). Este es muy rápido y en general bastante bueno. El método de construcción de NJ construye gradualmente un árbol basándose secuencialmente en el agrupamiento de pares de OTUs

conectados por un único nodo. El método de agrupamiento usado por este algoritmo es un poco diferente al resto porque este no intenta agrupar a los OTUs con mayor relación, sino que intenta reducir al mínimo la longitud de todas las ramas internas y por lo tanto la longitud de todo el árbol. El algoritmo de NJ comienza asumiendo un árbol con forma de estrella que no tiene ramas internas. En el primer paso, se introduce la primera rama interna y calcula la longitud del árbol resultante. El algoritmo conecta secuencialmente cada pareja de OTUs posibles y finalmente une los OTUs que producen el árbol más corto.

### **-Métodos basados en caracteres**

Los métodos basados en caracteres examinan cada posición en las secuencias del alineamiento de forma separada e independiente por cada posición, evaluando árboles y examinando todas las topologías teóricamente posibles para un número dado de secuencias, y, de acuerdo a los resultados obtenidos, construyen el árbol que mejor represente las relaciones evolutivas. Los métodos más utilizados para este tipo de análisis son Máxima Verosimilitud (ML, del inglés, Maximum Likelihood) y el método de inferencia bayesiana (BI, del inglés, Bayesian inference).

#### *Método de Máxima verosimilitud*

El algoritmo de construcción del método de Máxima verosimilitud emplea modelos probabilísticos para seleccionar el mejor árbol, es decir aquel que tenga la probabilidad más alta (máxima verosimilitud) de reflejar el proceso evolutivo real. El método de máxima verosimilitud busca todas las posibles topologías y considera cada posición en un alineamiento (no sólo sitios informativos).

Empleando un modelo particular de sustitución de nucleótidos, calcula la verosimilitud total de las secuencias ancestro que evolucionan en nodos internos. Este método trabaja calculando la probabilidad de un determinado camino evolutivo para una secuencia particular existente. Los valores de probabilidad son determinados por un modelo de sustitución. El método utiliza todos los sitios para calcular la probabilidad para todos los árboles posibles con todas

las combinaciones posibles de secuencias ancestro en los nodos internos de acuerdo a un modelo de sustitución predefinido(49,58).

### *Método de Inferencia Bayesiana*

En el método Bayesiano de construcción de árboles filogenéticos también se emplea el concepto de verosimilitud, pero ajustándolo a una distribución de probabilidad para los diferentes árboles. Los métodos bayesianos requieren que el investigador especifique una serie de “priors” o parámetros del modelo evolutivo, los cuales son formalizados y ajustados a una distribución de probabilidad. La evidencia relativa de los datos depende de cómo de creíble a posteriori sean estos priors. Son obtenidas una serie de probabilidades a posteriori, utilizando una técnica de muestreo llamada, cadenas de Markov y Monte Carlo (MCMC, del inglés, Markov chain Monte Carlo). Este método de muestreo comienza simulando un conjunto aleatorio de parámetros y propone un nuevo estado, que es un nuevo conjunto de parámetros, cambiando los parámetros utilizando operadores al azar. En cada paso la razón de verosimilitud es calculada. La frecuencia con la que se muestrea una topología de árbol en particular es proporcional a su probabilidad a posteriori. Los resultados se presentan habitualmente como funciones resumen de las muestras. El árbol que refleja una aproximación más real de la historia evolutiva de un grupo de secuencias se define como el árbol con la máxima credibilidad de clados (MCC).

Para estimar el tiempo de divergencia de varios taxones, se asumen que las secuencias se ajustan a una hipótesis de reloj molecular. Cuando un reloj molecular se mantiene fijo, todos los linajes en el árbol han acumulado sustituciones a un ritmo similar. Sin embargo la tasa evolutiva depende de varios factores, como la tasa de mutación de la especie, el tamaño de la población o la presión selectiva. Por ello generalmente los datos reales nunca se ajustan a un reloj molecular estricto, la tasa de mutación casi nunca suele ser uniforme. Existen modelos recientes, que se ajustan a “relojes moleculares relajados”, para acomodar la tasa de variación a las diferentes fechas de muestreo de nuestros datos(59).

### **1.8.3 Métodos de análisis filogenético utilizados en VIH-1.**

En general, los métodos de análisis filogenético utilizados con mayor frecuencia para el análisis de los diferentes cluster de transmisión de VIH-1 en grandes cohortes se denominan métodos de análisis filogenético “clásico” y son:

- Método de Neighbor-Joining.
- Método de Máxima Verosimilitud

Ambos métodos ofrecen resultados similares. Estos métodos de análisis filogenético “clásico” utilizan el análisis de bootstrap para reflejar la fiabilidad de las diferentes agrupaciones filogenéticas: expresados en valores de porcentaje de 0 a 100%.

Existen métodos “avanzados”, que utilizan métodos probabilísticos (Inferencia Bayesiana), para reflejar la fiabilidad de las diferentes agrupaciones filogenéticas, estos valores son expresados en valores de probabilidad a posteriori de 0 a 1. Estos métodos son utilizados con menor frecuencia debido a su mayor complejidad. A diferencia de los métodos filogenéticos “clásicos”, estos ofrecen una escala temporal (reloj molecular) que permiten estudiar, estableciendo un marco temporal, la historia evolutiva del VIH.

En esta tesis doctoral se estudiarán los perfiles filogenéticos de VIH-1 utilizando las herramientas más utilizadas en el análisis filogenético para este virus (análisis filogenético “clásico”) e implementando métodos de análisis filogenético de mayor complejidad (análisis filogenéticos “avanzados”).

#### 1.8.4 Modelos de sustitución nucleotídica

Los modelos evolutivos o modelos de sustitución nucleotídica en filogenias moleculares describen el modo y la probabilidad de que una secuencia de nucleótidos cambie a otra secuencia de nucleótidos homóloga a lo largo del tiempo. La forma más sencilla de calcular esta divergencia genética entre dos secuencias es contar el número de sitios en los que difieren. Esta proporción de sitios no homólogos se denomina distancia genética y se expresa como el número de nucleótidos diferentes por sitio (51,58).

El ritmo de sustitución nucleotídica puede ser diferente en las diferentes regiones de la secuencia. En la tercera base del codón se fijan mutaciones mucho más rápido que en la primera base, y en ésta a su vez más que en la segunda. Este hecho se explica por la presencia de diferentes fuerzas que actuarían en cada sitio en cuestión. Asimismo, esta heterogeneidad en los ritmos de evolución entre posiciones puede jugar un papel importante en el cálculo de las distancias genéticas

Existen diversos modelos de sustitución de nucleótidos que se utilizan dependiendo de los parámetros que se estimen. El más sencillo es el de Jukes-Cantor (JC), que estima la misma probabilidad para todas las sustituciones de nucleótidos. Otros modelos más complejos asignan distintos valores a diferentes sustituciones y algunos también tienen en cuenta la frecuencia de nucleótidos y la heterogeneidad de ritmos de evolución entre posiciones. Un modelo comúnmente utilizado es el modelo Kimura 2 parámetros (K2P), que divide las sustituciones en dos categorías, transiciones y transversiones, asignándoles valores distintos, pero asumiendo iguales frecuencias de nucleótidos. Otro modelo utilizado con frecuencia es el Hasegawa-Kishino-Yano (HKY), que asume diferentes ritmos de transiciones y transversiones y frecuencias de nucleótidos variable. El modelo más complejo con una matriz de sustitución simétrica (que asume igual probabilidad de cambio entre dos nucleótidos en ambos sentidos) y por lo tanto reversible en el tiempo, es el general time reversible (GTR), en el que las 6 posibles sustituciones entre cada par de nucleótidos tienen un ritmo diferente y los nucleótidos pueden ocurrir con frecuencias diferentes.

Existen herramientas tales como Find Model de HIV Los Alamos, que analizan diferentes modelos de sustitución nucleotídica y calculan el grado de ajuste de ese modelo de sustitución nucleotídica a nuestros datos.

En resumen, la filogenia consta de tres etapas claves:

1º Alineamiento múltiple de secuencias

2º Elección de un modelo de evolución que se ajuste a nuestras secuencias

3º Construcción del árbol filogenético en base al algoritmo evolutivo elegido.

## **Capítulo 2: OBJETIVOS**



El uso de una metodología eminentemente práctica, utilizando diferentes herramientas de análisis filogenético, puede aportarnos información valiosa para el conocimiento de la infección y de la historia evolutiva del virus de VIH en Andalucía Oriental en los últimos años. En el marco de los antecedentes expuestos en el capítulo de Introducción, se ha planteado el trabajo experimental de la presente Tesis, proponiendo los siguientes objetivos:

- Estudiar la prevalencia, las características demográficas, clínicas y virológicas de los diferentes subtipos de VIH-1 en Andalucía oriental, y su heterogeneidad y las zonas geográficas donde existe una mayor recombinación y diversidad genética.
- Describir los posibles cluster de transmisión de VIH-1 y la existencia de variables clínico-epidemiológicas predictoras de agrupamiento de los pacientes en cluster de transmisión para las formas genéticas de VIH-1 más prevalentes en nuestra población.
- Analizar la presencia de mutaciones de resistencia transmitidas (TDR) en los pacientes agrupados en cluster de transmisión.
- Reconstruir la historia evolutiva de los diferentes subtipos de VIH-1 en Andalucía Oriental, estableciendo el tiempo del más reciente común antecesor (tMRCA) para los diferentes linajes y cluster de transmisión y estimando diferentes parámetros filogenéticos (tales como la tasa evolutiva, probabilidad de localización...).
- Analizar los perfiles filodinámicos del VIH-1 subtipo B.
- Analizar los patrones de migración de la epidemia de VIH, analizando la direccionalidad de los eventos de transmisión mediante el análisis filogeográfico de las diferentes formas genéticas de VIH-1.

**Capitulo 3: ANÁLISIS FILOGENÉTICO  
GLOBAL DE LA TRANSMISIÓN DE VIH EN  
ANDALUCÍA ORIENTAL.**

### 3.1 Antecedentes

El diseño de medidas de prevención para evitar la expansión del virus del VIH-1 requiere un conocimiento exhaustivo de la epidemiología del virus. Para contribuir a este conocimiento, los estudios filogenéticos se valen del creciente número de secuencias del gen *pol* del VIH-1 generadas durante la rutina clínica para los ensayos de resistencia genotípica. Debido a la gran diversidad genética del virus, estas secuencias *pol* aportan suficiente información para deducir y recrear la historia de la epidemia del virus (60).

Desde la identificación del VIH-1 a principios de los años 1980s (61), la epidemiología molecular ha contribuido a la comprensión de diferentes aspectos del virus, resultando ser fundamental tanto para el conocimiento de la heterogeneidad y la prevalencia de los diferentes subtipos de VIH-1, como en el estudio de la resistencia basal del virus a fármacos antiretrovirales(62–66). Dentro de la epidemiología molecular, el análisis filogenético, utilizando para ello, Las secuencias de proteasa y transcriptasa reversa del VIH-1 aportan información muy valiosa para el manejo de la infección por VIH, más allá de la información de resistencias a los antirretrovirales. De esta forma, y en combinación con los datos generados por la epidemiología tradicional, han sido muy utilizados para conocer fenómenos de transmisión implicados en la propagación del virus(47).

El objetivo de este capítulo fue caracterizar a nivel global las diferentes agrupaciones filogenéticas presentes dentro de nuestra cohorte. Para este estudio se utilizaron todas las secuencias de pacientes diagnosticados durante el periodo de 2005 a 2012 en diferentes hospitales de Andalucía oriental (Almería, El Ejido, Motril, Granada y Jaén). Se utilizaron técnicas de análisis filogenético “clásico” para estudiar las variables demográficas, clínicas y virológicas que estuvieran relacionadas con la formación de colectivos de riesgo y la presencia de mutaciones de resistencia transmitidas en los pacientes incluidos en estos cluster de transmisión. Como objetivo secundario analizamos las regiones con mayor diversidad genética, mediante la georeferenciación de los diferentes pacientes incluidos en los diferentes cluster de transmisión.

## **3.2 Pacientes y métodos**

### **3.2.1 Descripción de la cohorte**

Se recopilaron 693 secuencias parciales del gen *pol* del VIH-1 (posiciones 4-99 de la proteasa y 38-247 de la transcriptasa reversa), obtenidas utilizando el ensayo Trugene HIV-1 genotyping Kit (Siemens) para el análisis de resistencias basales en pacientes naïve, procedentes de pacientes de diferentes hospitales de Granada, Almería y Jaén durante 2005 a 2012. Se recogió la información demográfica (edad, sexo, fecha de diagnóstico y forma de infección por VIH-1), clínica (recuento de CD4) y virológica (carga viral) de cada uno de estos pacientes y esta información se vinculó a cada una de nuestras secuencias mediante la construcción de una base de datos anonimizada, encriptada y protegida.

Los pacientes estudiados en nuestra cohorte fueron mayoritariamente hombres jóvenes (41 años), homosexuales (51.5%), nativos españoles (65.5%) y que residen en la zona de Granada (43,7%) o Almería (36.9%). Las características demográficas clínicas y virológicas de los pacientes se representan en la tabla 3.

CARACTERÍSTICAS	Nº DE PACIENTES (%)
<b>SEXO (675)*</b>	
Mujer	147(21.2)
Hombre	528(76.2)
<b>NACIONALIDAD (642)*</b>	
Español	454(65.5)
Extranjeros	188(27.1)
<b>EDAD</b>	
<35	173(26.1)
35-45	256(36.9)
>45	223(32.2)
<b>LOCALIDAD (678)*</b>	
Almería	256(36.9)
Granada	303(43.7)
Motril	49(7.1)
Jaén	70(10.1)
<b>AÑO DE DIAGNÓSTICO (686)*</b>	
2005-2006	48(7.2)
2006-2007	50(2.9)
2007-2008	100(14.4)
2008-2009	83(12)
2009-2010	142(20.5)
2010-2011	60(8.7)
2011-2012	102(15)
2012-2013	101(15)
<b>CARGA VIRAL(646)*</b>	
<10.000	133(19.2)
10.000-100.000	277(41.8)
>100.000	236(34.1)
<b>RECuento DE CD4 (602)*</b>	
<200	182(16.5)
201-350	146(21.1)
>350	274(39.5)
<b>MECANISMO DE TRANSMISIÓN (591)*</b>	
UDVP	102(14.7)
HTX	273(39.4)
HSH	216(31.2)

**Tabla 3. Características demográficas, clínicas y virológicas de los pacientes infectados con VIH-1 durante el periodo 2005 a 2012.**

\* Datos disponibles para el nº de pacientes indicado.

### 3.2.2 Secuenciación del genoma vírico

La técnica de secuenciación de ácidos nucleicos utilizada para el estudio de resistencias basales a fármacos antiretrovirales fue TruGene™ HIV-1. Esta técnica consta de 5 etapas ( Figura 12) :

- Extracción del ácido nucleico
- Amplificación mediante RT-PCR
- Reacción de secuenciación bidireccional
- Electroforesis en gel de poliacrilamida de los secuenciados
- Alineamiento de las secuencias e interpretación de los resultados

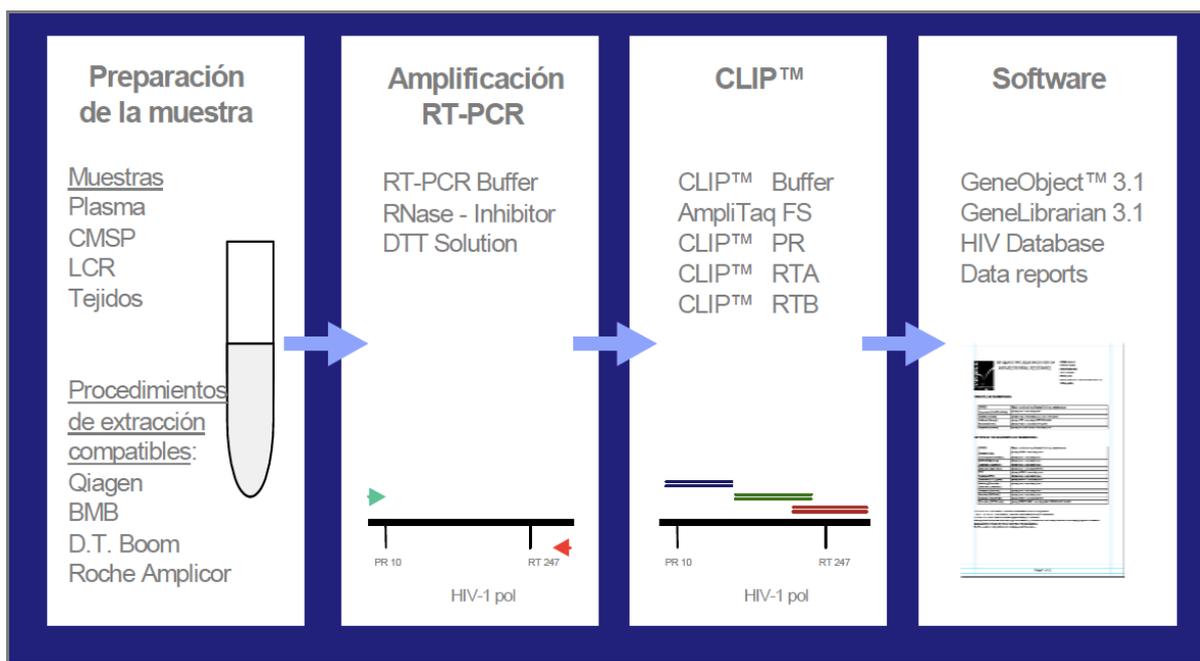


Figura 12. Fases de la técnica de secuenciación.

### *-Extracción del ácido nucleico*

Consiste en hacer accesible la diana, es decir, el ácido nucleico del VIH, que se encuentra dentro de la cápside viral. Para ello se realiza un pretratamiento del suero con el fin de alcanzar la lisis de las partículas víricas; posteriormente se procede a separar el ácido nucleico de los restos virales y de las proteínas séricas (generalmente se hace por unión a una partícula sólida o por precipitación); finalmente se lava el ácido nucleico y se purifica, quedando al final una solución que contiene el ARN viral en condiciones de poder ser utilizado para la amplificación.

### *-Amplificación del ácido nucleico: RT-PCR*

Este proceso se basa en una retrotranscripción del material extraído (ARN) con una enzima transcriptasa reversa obteniéndose su ADN copia que servirá de base para la posterior amplificación mediante una enzima polimerasa.

### *-Secuenciación bidireccional y electroforesis*

La reacción de secuenciación realizada se basa en la secuenciación bidireccional conocida como CLIPTM, la cual no es más que una modificación del método de Sanger que emplea terminadores de la secuencia de ADN en forma de didesoxinucleótidos (ddNTP). Los ddNTPs carecen del grupo -OH en el carbono 3 del azúcar que es el responsable de la unión de los nucleósidos entre sí, a través del grupo fosfato para formar el nucleótido e incorporarse a la cadena del ácido nucleico. Por lo tanto, si un enzima al elongar una cadena de ADN coloca en vez de una adenina en forma de dATP una ddATP, la cadena que se está formando se trunca y no puede ser elongada más.

En este caso los terminadores no están marcados, son los primers los que se marcan. En concreto, un primer incorpora Cy5.5 como fluoróforo en 5', y el otro primer incorpora Cy5.0 en 3'. Estos dos fluoróforos tienen la característica de emitir luz a distintas longitudes de onda (690 y 670 nm respectivamente) de modo que podremos diferenciar con un láser la emisión de los productos de secuenciación iniciados con cada uno de estos dos primers y, de

esta manera, conoceremos en un mismo tubo la secuencia *sense* y *antisense* (*forward* y *reverse*).

La electroforesis en gel de poliacrilamida la realizamos en un secuenciador semiautomático (*Long-Read Tower™*) previa polimerización del gel de acrilamida.

El secuenciador dispone de un sistema láser que hace lecturas cada 30 segundos de manera que va registrando las señales que emiten los fluoróforos: el producto de secuenciación en la dirección 5' que incorpora Cy5.5 y emite a 550 nm y en dirección 3' marcado con Cy5.0, que emite a 670 nm. El secuenciador tiene un sistema de registro de ambas longitudes de onda de modo que lee a la vez la secuencia de la dirección 5' y 3' (secuenciación bidireccional).

Para cada amplificado a secuenciar vamos a realizar tres reacciones distintas de secuenciación; CLIP PR, en la que investigaremos la secuencia de la proteasa (codones 6-95 aproximadamente), CLIP RTA en la que investigaremos una parte de la transcriptasa reversa (*RT beginning*, codones 38-140 aproximadamente) y por último, CLIP RTB, en la que investigaremos otro fragmento de la transcriptasa reversa (*RT middle*, codones 130-249 aproximadamente).

Para llevar a cabo la secuenciación, para cada una de las tres regiones (PR, RTA, RTB) se necesitan 4 tubos (A, C, G, T). En cada uno de ellos se añaden los primers de secuenciación marcados con Cy5.5 y Cy5.0, el didesoxinucleótido correspondiente (ddATP, ddCTP, ddGTP y ddTTP) y los dNTPs.

#### *-Alineamiento de las secuencias e interpretación de los resultados*

El paso final para la secuenciación consiste en el alineamiento de las secuencias. Mediante este procedimiento el software integra los 4 cromatogramas obtenidos para cada región secuenciada (PR, RTA y RTB) y para cada dirección (5' y 3'). Asimismo, enfrentamos las secuencias obtenidas con una cepa de referencia en la que no existen mutaciones de resistencia (cepa salvaje, *wild type* o WT, HIV-1 B LAV-1).

A continuación, revisamos la secuencia completa, viendo las bases en las que no exista concordancia entre las direcciones 5' y 3' o con la secuencia *wild type* y revisando todas las posiciones de resistencia.

En estos pasos debemos ser extremadamente cautelosos ya que somos nosotros los que vamos a decidir si estamos o no de acuerdo con la interpretación que el software realiza.

Finalmente, comprobamos que no existen posibles contaminaciones cruzadas con otras muestras (*fingerprinting*).

### **3.2.3 Caracterización del subtipo viral.**

Para la caracterización del subtipo viral puro o forma recombinante de cada una de las 693 secuencias recopiladas durante el periodo de 2005 a 2012, utilizamos la herramienta de subtipado rápido REGA HIV-1 Subtyping tool versión 2.0 (67).

### **3.2.4 Caracterización de los cluster de transmisión a través de análisis filogenético clásico.**

Con la finalidad de estudiar las diferentes agrupaciones filogenéticas presentes en nuestra población se construyó un árbol filogenético mediante el método de construcción de análisis filogenético “clásico” NJ. Las secuencias fueron procesadas utilizando PhyDE (disponible en <http://www.phyde.de>) y alineadas utilizando MEGA 5.2 (57). El árbol filogenético de consenso fue inferido a partir de 1000 replicas de bootstrap con una tasa de variación entre sitios modelada con una distribución gamma equivalente a 2. Las ramas pertenecientes a los clusters con un valor de bootstrap inferior al 80% de las replicas fueron eliminadas. Se utilizó FindModel (disponible en <http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>) para estudiar el modelo de sustitución nucleotídica que mejor se adaptaba en el conjunto de secuencias. El modelo de sustitución nucleotídica utilizado fue Kimura 2 parámetros. Finalmente, para la edificación del árbol filogenético se utilizó el editor FigTree v 1.4.0 (disponible en <http://tree.bio.ed.ac.uk/software/figtree>).

### **3.2.5 Análisis pormenorizado de los cluster de transmisión.**

Se realizó un estudio epidemiológico pormenorizado de los clusters con un valor de bootstrap superior al 85%. Asimismo se analizaron las diferentes características demográficas (fecha de nacimiento, año de diagnóstico de VIH, hospital de referencia, sexo, mecanismo de transmisión de VIH, país de nacimiento, y lugar de residencia actual), clínicas (recuento de CD4) y virológicas (carga viral) de los pacientes relacionados epidemiológicamente.

### **3.2.5. Georeferenciación de los pacientes incluidos en los cluster de transmisión.**

Se utilizó ArcGIS Explorer Desktop Esri©, un visor SIG (*Geographic Information System*) (disponible en <http://www.esri.com/software/arcgis/explorer>) para georeferenciar los diferentes pacientes que estaban incluidos en los diferentes cluster de transmisión. Posteriormente, se realizó un de proximidad para el estudio de zonas geográficas con mayor heterogeneidad de subtipos.

### **3.2.5 Análisis de las mutaciones de resistencia transmitidas.**

Se estudiaron todas las mutaciones de resistencia transmitidas de los pacientes agrupados en diferentes cluster de transmisión, a través de la lista de mutaciones transmitidas que establece la Organización Mundial de la Salud OMS (68). Asimismo, se interpretaron a través del algoritmo de resistencias de la Universidad de Stanford (69).

### **3.2.6 Análisis de las características clínico-epidemiológicas predictoras de agrupamiento en cluster.**

Se realizó un análisis de regresión logística para evaluar las características demográficas, clínicas y virológicas predictoras de agrupamiento de los pacientes en cluster de transmisión. El límite de significación fue establecido en 0,05 para los valores de *P*. Los análisis se realizaron con el programa SPSS versión 15 (SPSS Inc., Chicago, EE. UU.).

### **3.3 Resultados**

#### **3.3.1 Resultados obtenidos a través del árbol filogenético realizado mediante NJ.**

Como se muestra en la figura 13, nuestro árbol mostró la presencia de diferentes agrupaciones filogenéticas que se asociaron con valores de Bootstrap de 0 a 100, medida que daba fiabilidad a estas cadenas de transmisión y que establecimos en valores superiores a 85% para definir estos cluster. Asimismo, la figura 14 muestra los diferentes cluster encontrados en nuestra población coloreados en acorde a su tamaño.

De las 693 secuencias utilizadas para generar el árbol filogenético, 298 secuencias (43%) se agruparon en 77 clusters diferentes, con un valor medio de bootstrap de 97,12 %.

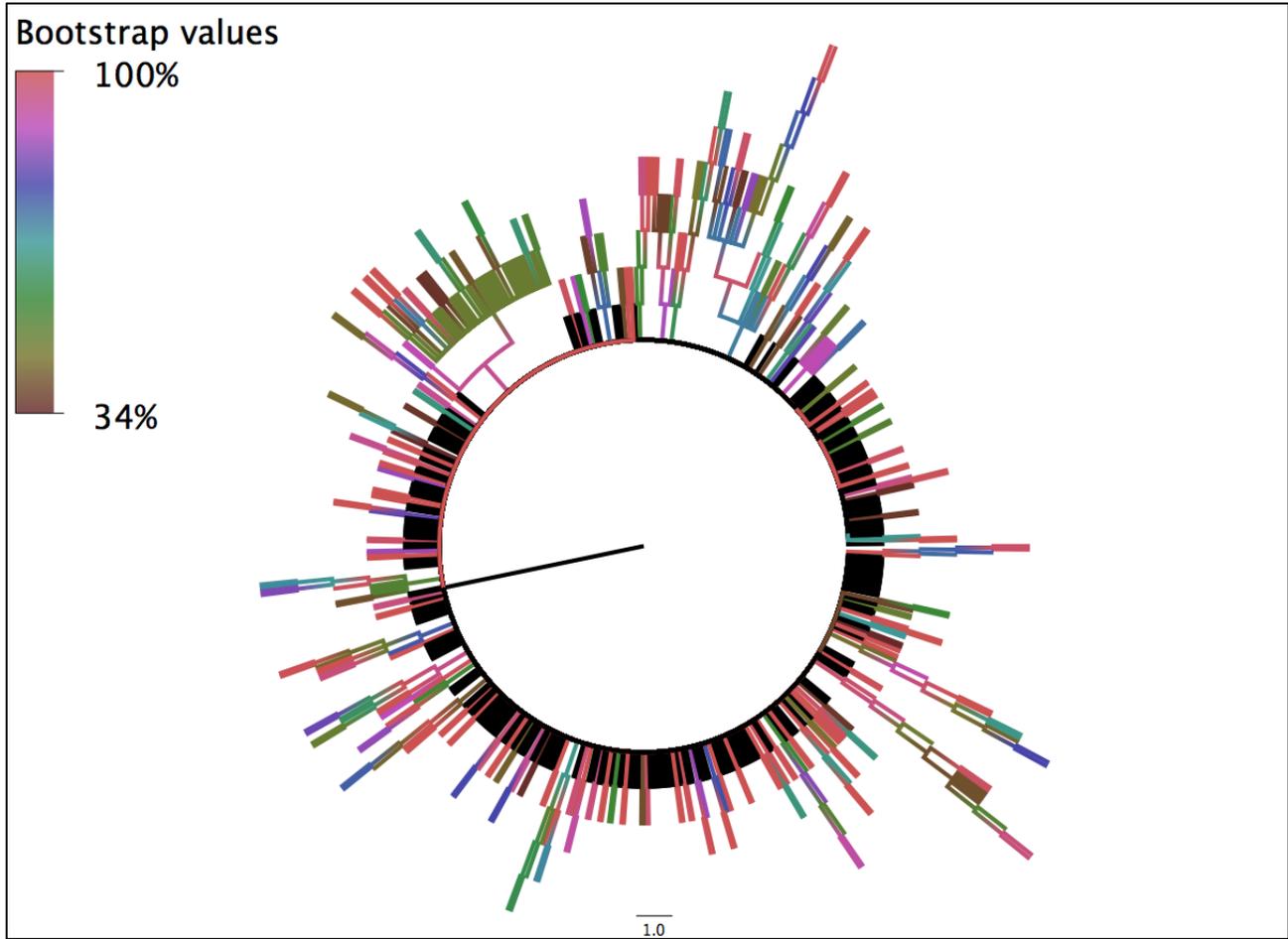
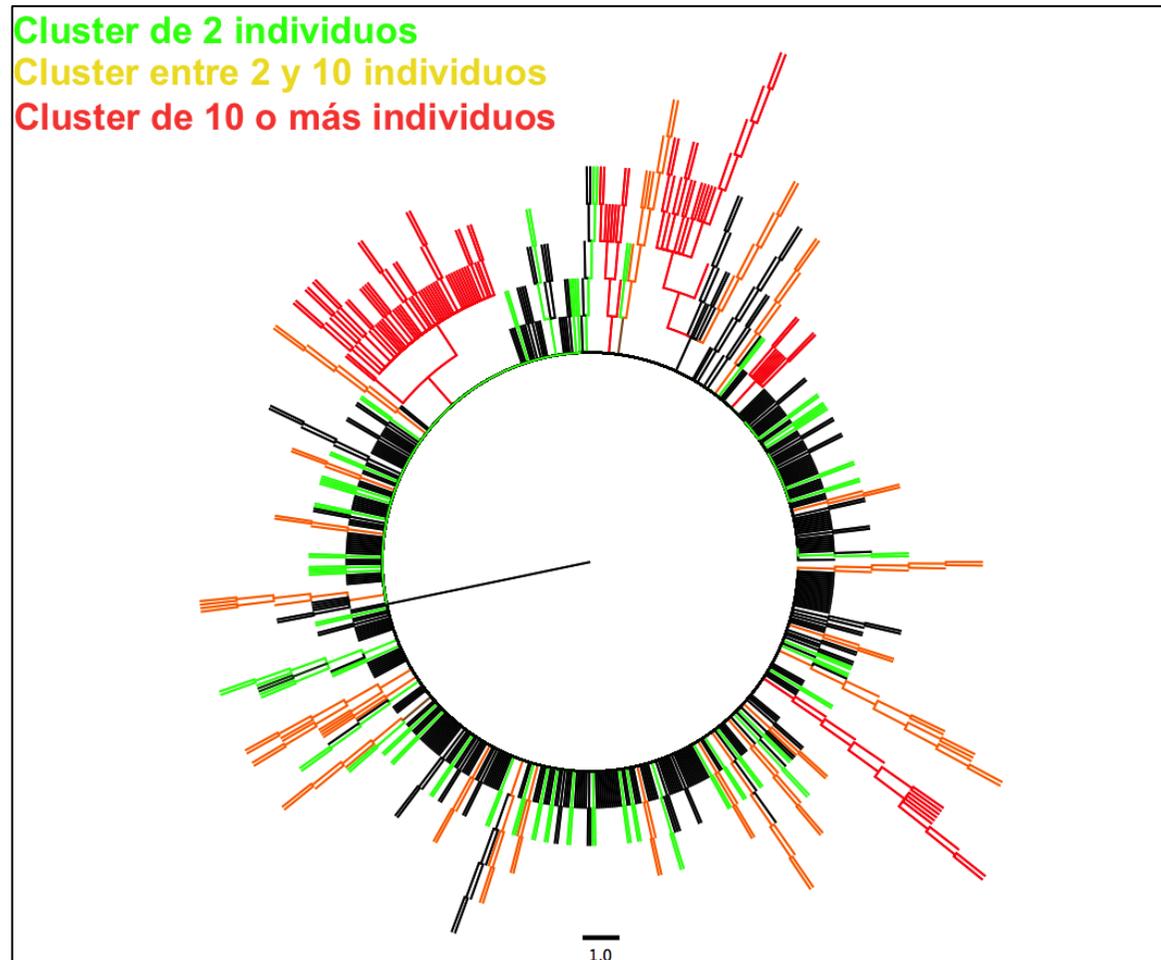


Figura 13. Arbol filogenético consenso representado por valores de Bootstrap (de 0 a 100%) para el global de pacientes (n=693) diagnosticados de VIH-1 desde 2005 a 2012. La escala de la parte superior izquierda muestra un gradiente de color para los valores de Bootstrap.



**Figura 14. Representación esquemática del árbol filogenético consenso realizado para el global de pacientes (n=693) diagnosticados de VIH-1 desde 2005 a 2012. Los cluster de transmisión con valores de Bootstrap superiores al 85% quedan coloreados en acorde a su tamaño.**

### **3.3.2 Variables predictoras de agrupamiento en cluster**

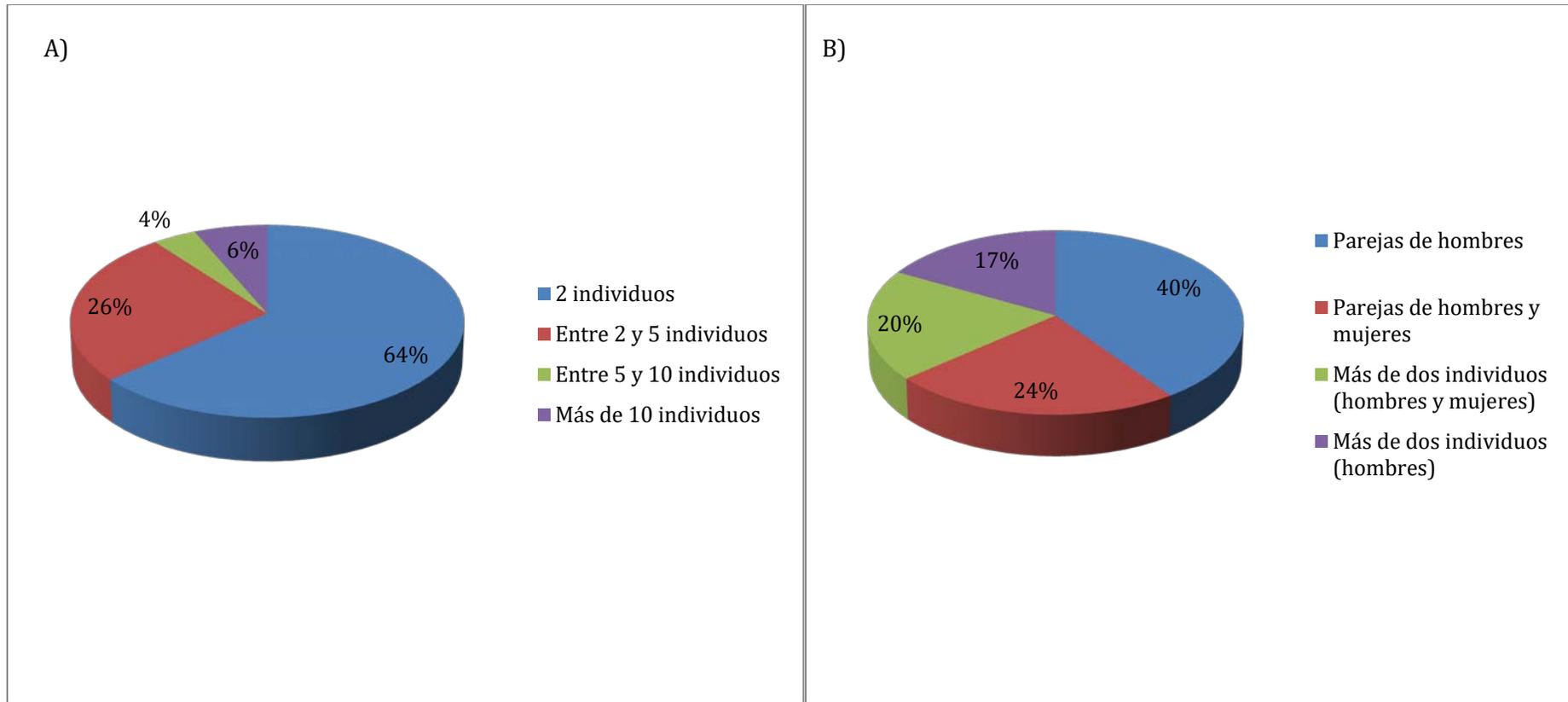
En la tabla 4 se reflejan las características de los pacientes incluidos en cluster de transmisión en comparación con los que no estuvieron incluidos en ellas. Existieron diferencias estadísticamente significativas entre los pacientes asociados en cluster y los que no lo estaban. En concreto, la asociación en cluster fue más frecuente en hombres ( $p=0.02$ ), españoles ( $p=0.01$ ), con edad menor de 35 años ( $p=0.01$ ), residentes en Granada ( $p=0.03$ ), con un recuento de CD4 mayor de 350 ( $p<0.001$ ), homosexuales ( $p<0.001$ ) y que no presentaban mutaciones de resistencia a antiretrovirales de primera línea ( $p=0.01$ ).

<b>Pacientes, n(%)</b>				
	<b>Global</b>	<b>En Cluster</b>	<b>Fuera de Cluster</b>	<b>p</b>
	693	298 (43)	395 (57)	
<b>Sexo</b>				<b>0.02</b>
Mujer	147(21.2)	51(17.1)	96(24.3)	
Hombre	528(76.2)	242(81.2)	286(72.4)	
Desconocido	18(2.6)	5(1.7)	13(3.3)	
<b>Nacionalidad</b>				<b>0.01</b>
Español	454(65.5)	220(73.8)	234(59.2)	
No-Español	188(27.1)	71(23.8)	117(29.6)	
Desconocido	51(7.4)	7(2.3)	44(11.1)	
<b>Edad</b>				<b>0.01</b>
<35	173(26.1)	101(33.9)	72(18.2)	
35-45	256(36.9)	118(39.6)	138(34.9)	
>45	223(32.2)	79(26.5)	144(36.5)	
Desconocido	41(5.9)	0	41(10.4)	
<b>Localidad</b>				<b>0.03</b>
Almería	256(36.9)	96(32.2)	160(17.1)	
Granada	303(43.7)	149(50)	154(39)	
Motril	49(7.1)	24(8)	25(6.3)	
Jaén	70(10.1)	29(9.7)	41(10.4)	
Desconocido	15(2.3)	0	15(3.8)	
<b>Carga viral</b>				<b>0.8</b>
<10.000	133(19.2)	51(17.1)	82(20.8)	
10.000-100.000	277(41.8)	129(43.3)	148(37.5)	
>100.000	236(34.1)	110(37)	126(31.9)	
Desconocido	47(6.8)	8(2.7)	39(9.9)	
<b>CD4</b>				<b>&lt;0.001</b>
<200	182(16.5)	54(18.1)	128(32.4)	
201-350	146(21.1)	63(21.1)	83(21)	
>350	274(39.5)	148(49.7)	126(31.9)	
Desconocido	91(13.1)	33(11.1)	58(14.7)	
<b>Mecanismo de transmisión</b>				<b>&lt;0.001</b>
UDVP	102(14.7)	26(8.7)	76(19.2)	
HTX	273(39.4)	123(41.3)	150(38)	
HSH	216(31.2)	146(49)	70(17.7)	
Desconocido	102(14.3)	3(1)	99(25.1)	
<b>Mutaciones de resistencia</b>				<b>0.01</b>
Si	76(11)	22(7.4)	54(13.7)	
No	617(89)	276(92.6)	341(88.6)	
<b>Subtipo</b>				<b>0.8</b>
Subtipo B	452(65.2)	202(67.8)	250(63.3)	
Subtipo no-B	211(34.8)	96(32.3)	115(36.7)	

**Tabla 4. Análisis comparativo de las características demográficas, clínicas y virológicas de los pacientes agrupados en cluster frente a los pacientes no incluidos en clust**

### **3.3.2 Características epidemiológicas de los cluster de transmisión.**

Como se representa en los gráficos 15 y 16, la gran mayoría de los clusters de transmisión que encontramos estaban formados por pares (n=49) (64%). La mayor proporción de parejas eran de HSH (53.1%) seguido por parejas HTX (37.5%) y de UDVP (9.4%). El 48% de los clusters estaba formado por nativos españoles, el 23% incluía algún extranjero entre sus integrantes y solo el 3% estaba compuesto íntegramente por extranjeros. La mayoría de los cluster estaban formados por pacientes de Granada (33%), Almería (28%) o por un conjunto de pacientes de ambos sitios (21%).



**Figura 15. Análisis comparativo de la proporción relativa de cluster de transmisión en acorde a su tamaño (A) y género de los pacientes(B).**

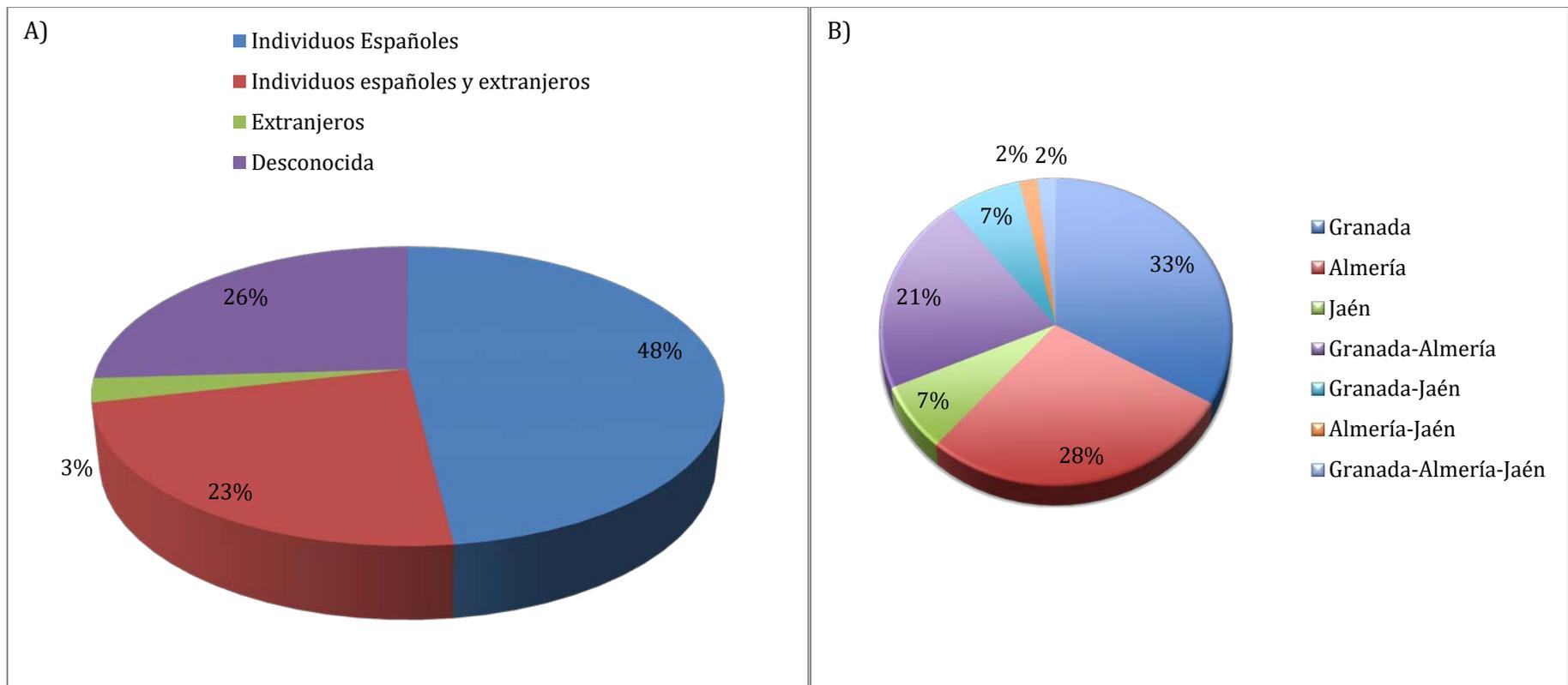


Figura 16. Análisis comparativo de la proporción relativa de cluster de transmisión en acorde a su nacionalidad (A) y localización geográfica de los pacientes (B).

### **3.3.6 Características de los cluster con mayor repercusión epidemiológica: agrupaciones filogenéticas formadas por 10 o más pacientes.**

Se detectaron 5 cadenas de transmisión en los que estaban incluidos un número igual o mayor a 10 pacientes y que representaban el 16% (n=113) de los pacientes de nuestra cohorte. En Andalucía, las agrupaciones filogenéticas con mayor repercusión epidemiológica y que involucraban extranjeros entre sus integrantes (Europa del este, América latina, África) se dieron todos entre población HTX y con subtipos no-B, en cambio las que involucraban individuos españoles fueron todos de población HSH y con subtipo B. Cabe destacar, por su número, una cadena de transmisión formada por 58 HSH de nacionalidad española que se inicia en el año 2007 e incorpora individuos hasta el año 2012, y otro de 21 pacientes HTX formado por varones de nacionalidad española y mujeres de nacionalidad rusa. Asimismo, destaca también un cluster formado por 10 pacientes de Senegal, Mali, Nigeria, Ghana y Guinea en el que hemos detectado en 3 pacientes la transmisión de M184V y V179D, con resistencia a los análogos de primera línea emtricitabina y lamivudina, y resistencia de bajo nivel a Abacavir.

<b>N° de cluster</b>	<b>n</b>	<b>Factor de Riesgo</b>	<b>Edad (mediana, IQR)</b>	<b>Nacionalidad</b>	<b>Subtipo</b>	<b>Carga Viral (mediana, IQR) Log<sub>10</sub></b>	<b>CD4 (mediana, IQR)</b>	<b>Resistencias</b>
7	12	HSH	37.5(30-44.3)	Española	B	5.17(4.81-5.98)	331(252.3-388)	-
58	58	HSH	32.5(28-39.3)	Española	B- (B-D)	4.42(4.03-5.27)	518(301-896)	-
64	10	HTX	38(27-49)	España-Brasil-Mali-Rumania	F	5.47(5.13-5.76)	103(50-157)	-
67	21	HTX	33(31-38)	España-Rusia-Rep. Dominicana	A	4.56(3.59-5.15)	624(309-867)	-
71	10	HTX	43(35-48)	Senegal-Mali-Nigeria-Ghana-Guinea	G	4.87(4.31-5.56)	173(20-364)	M184V, V179E

**Tabla 5. Características clínicas, demográficas y virológicas de las cadenas de transmisión encontradas con mayor relevancia epidemiológica ( $\geq$  a 10 pacientes relacionados epidemiológicamente) del árbol realizado mediante NJ para el total de los pacientes infectados con VIH-1.**

### **3.3.4 Diversidad genética de los cluster de transmisión**

Como cabría esperar, la mayoría de nuestros pacientes estaban infectados con subtipo B (79,3%). Los análisis de subtipaje rápido realizados a través de REGA 2.0 mostraron que la mayoría de cluster eran de subtipo B (n=51), 17 de subtipo no-B (22%) y 9 de recombinantes B-D (12%).

Como se observa en el mapa de la figura 17 que refleja al total de pacientes georeferenciados que estaban incluidos en cluster de transmisión, las áreas de mayor agregación (basándonos en la distancia entre individuos que nos facilitaba el visor SIG de ArcGIS) de pacientes infectados con VIH por Km<sup>2</sup> se dieron en la zona del área metropolitana de granada y en las zonas de costa de Almería. La figura 18 representa la diversidad genética de los pacientes incluidos en cluster de transmisión y que estuvieron infectados durante este periodo. Cabe destacar que en estas mismas zonas existió una mayor heterogeneidad de subtipos y por ende mayor recombinación.

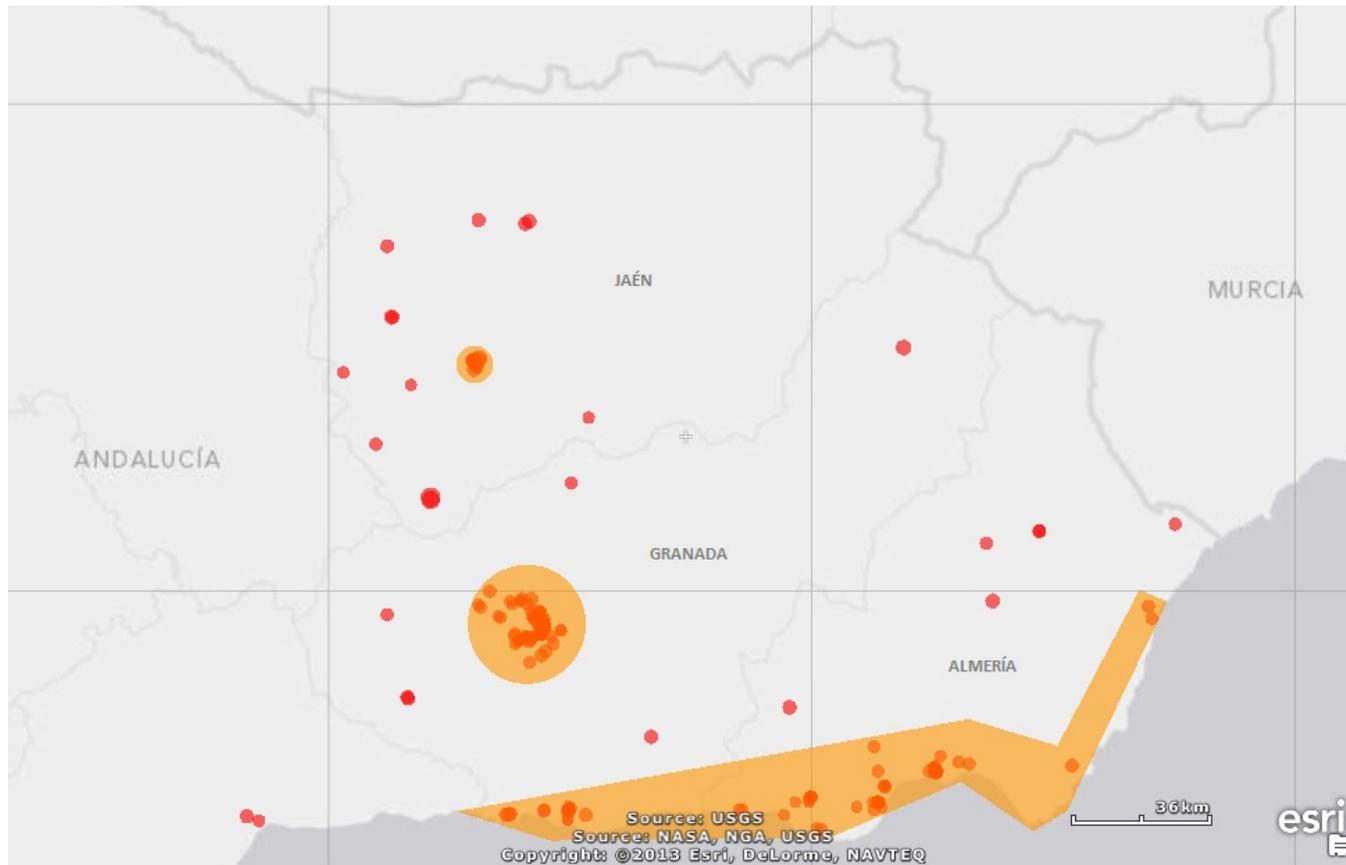


Figura 17. Georeferenciación de los pacientes incluidos en los cluster de transmisión VIH-1 durante el periodo 2005-2012. Las zonas resaltadas, representan una mayor agregación de individuos por Km<sup>2</sup>.

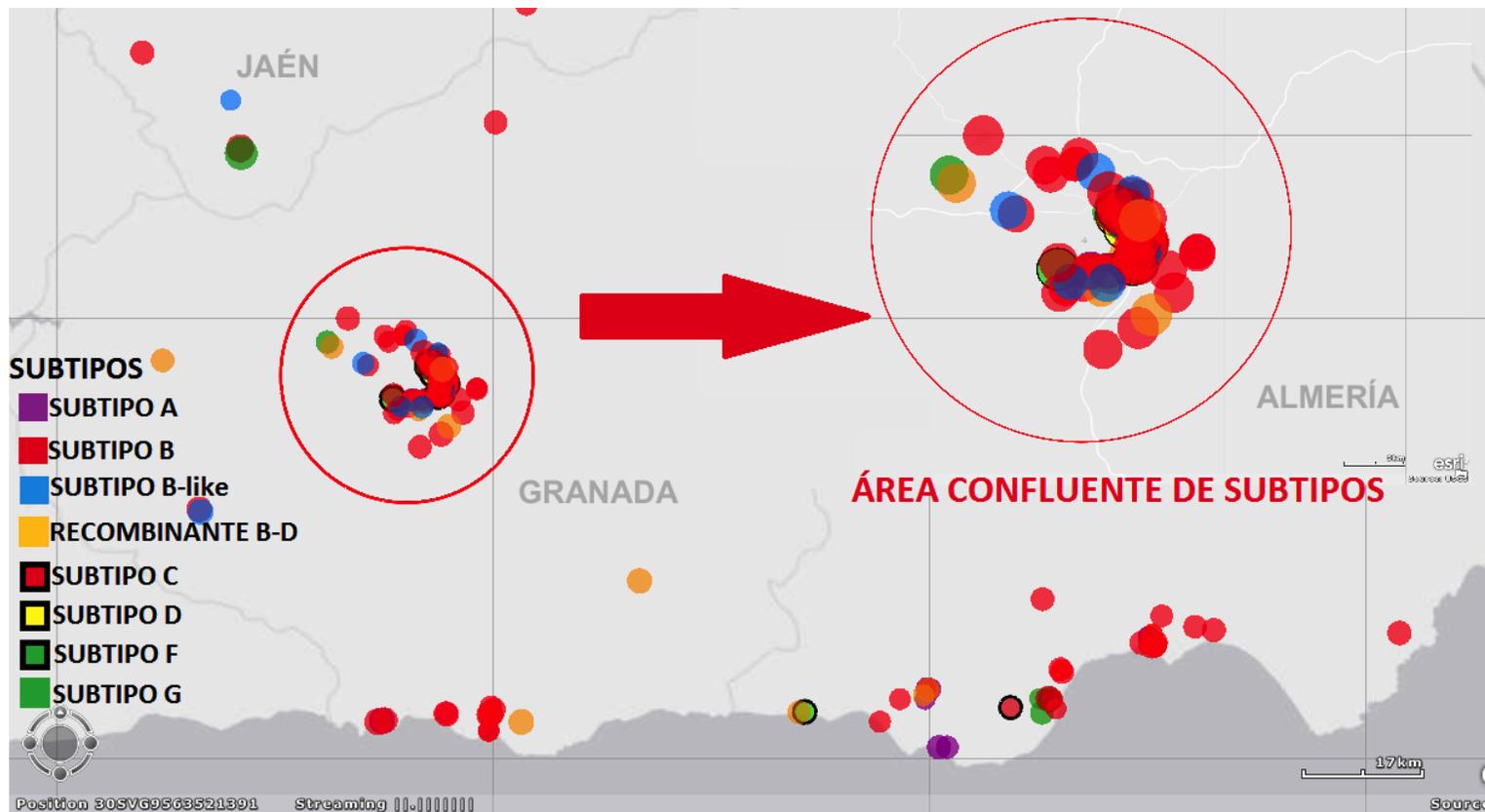


Figura 18. Georeferenciación de los pacientes incluidos en los diferentes cluster de transmisión VIH-1 durante el periodo 2005 a 2012 en función de la forma genética de infección.

### **3.3.3 Análisis de la presencia de mutaciones de resistencia transmitida.**

Como se representa en la tabla 6, se encontraron transmisión de mutaciones de resistencia para fármacos antirretrovirales de primera línea en 15 clusters diferentes, generalmente en clusters de tamaño pequeño. Para los análogos de nucleósidos, se detectaron mutaciones en 3 clusters diferentes frente a análogos de timidina (TAMS, M41L; D67N+K219Q; K70R respectivamente), en dos cambios en la posición M184 (M184I, en uno; M184V en otro), en un cluster se detectó la revertiente T215L, y en otro la combinación D67E+F77L. Las mutaciones para los NNRTIs fueron las más frecuentes, detectándose en un total de 7 clusters; su distribución fue la siguiente: K101E; V179D; P225H; V90I+Y181C+H221Y; K103N; V179E; E138A+V179D. Finalmente, 5 clusters mostraron mutaciones de resistencia en la proteasa: M46I (3 clusters); L90M (2 clusters) y 2 clusters mostraron mutaciones de resistencia frente a más de una familia de antirretrovirales.

CLUSTER (ID)	Tamaño del cluster (N° de pacientes agrupados)	N° de individuos con mutaciones de resistencia	Factor de riesgo	Subtipo	Mutaciones de resistencia a ARV
1	5	5	HSH	B	T215L, M41LM
3	2	1	HTX	B	M46L, K101E
13	2	1	HTX	B	V179DV
22	2	2	HTX	B	L90M
24	2	2	HSH	B	L90M
36	3	1	HSH	B	D76E, F77L
38	3	1	HSH	B	P225HP
46	2	2	HSH	B	D67N, T69N, K219Q
49	3	2	HTX	B	M184L, Y181C, H221Y, V90IV, Y181C, H221Y
54	3	1	HSH	B	K70R
68	5	4	HTX	A	K103N
70	2	1	HTX	CRF14_BG	M46I
71	10	3	HTX	G	M184V, V179E
73	2	1	HTX	CRF02_AG	M46IM
77	2	1	HSH	D	E138A, V179DV

**Tabla 6. Mutaciones de resistencia transmitidas presentes en los pacientes agrupados en diferentes cluster de transmisión para la población infectada por VIH-1 durante el periodo 2005 a 2012.**

### 3.4 Discusión

El primer objetivo de esta tesis doctoral fue caracterizar a nivel molecular la epidemia de VIH-1 en Andalucía oriental. Para ello se realizó un análisis filogenético clásico que permitió analizar la presencia de las diferentes cadenas de transmisión. Se encontraron 298 pacientes agrupados en 77 cadenas de transmisión, datos que demuestran que un 43% de los pacientes infectados con VIH-1 en el periodo de 2005 a 2012 estuvieron relacionados epidemiológicamente. La agregación de pacientes ocurrió con mayor frecuencia entre parejas de HSH de nacionalidad española ( $p=0.01$ ), residentes en Granada ( $p=0.03$ ) y con una edad menor a 35 años ( $p=0.01$ ). Asimismo, los datos de este estudio son similares a los descritos en otras series (70), que también comunican una clara tendencia a la agrupación de los pacientes en clusters formados entre parejas de HSH.

Encontramos grandes cadenas de transmisión que se dieron también con mayor frecuencia entre HSH. Cabe destacar por su gran tamaño el cluster 58, formado por 58 pacientes HSH y que incorporó individuos hasta el año 2012. Estos datos están en la misma línea a otros estudios y sugieren una rápida y eficaz propagación del VIH-1 en determinados colectivos: *Yebra et al* (38), demostró como en la comunidad de Madrid se ha producido en los últimos años una mayor inclusión de HSH en estas cadenas de transmisión. Otros estudios realizados en otras cohortes Europeas han demostrado resultados similares: un 29% de la población en Reino Unido se agrupó en grandes cadenas de transmisión con un número mayor de 10 individuos (71) y un 42% de los pacientes pertenecientes a la cohorte suiza (70) estaban también relacionados epidemiológicamente en cadenas de transmisión de gran tamaño, generalmente formados entre pacientes HSH.

En consonancia con el aumento de la población migrante en los últimos años en España, hemos visto como en Andalucía Oriental hasta el 23% de las cadenas de transmisión estaban formadas por españoles y pacientes de otras nacionalidades. Además, las cadenas de transmisión con mayor repercusión epidemiológica entre la población HTX estaban integradas por una gran proporción de individuos extranjeros, como es el caso del cluster 67, formado por 21 individuos de subtipo A, y con una gran proporción de mujeres de nacionalidad rusa.

Aunque hemos encontrado transmisión de resistencias a antirretrovirales de primera línea en cluster de transmisión, en la mayoría de las ocasiones se detectan en cluster de pequeño tamaño (72,73). Sólo en dos casos se han detectado en agrupaciones filogenéticas de mayor tamaño: en el subtipo G (3 de 10 pacientes) y en el subtipo A (4 de 5 pacientes). Estos datos explican en gran medida que los virus con mutaciones de resistencia presentan una desventaja biológica frente a una cepa silvestre y que por tanto este hecho disminuye su eficacia en la transmisión (74).

Aunque el subtipo de VIH-1 con mayor prevalencia sigue siendo el subtipo B, hemos demostrado la existencia de “zonas calientes” donde existe una mayor prevalencia de VIH-1 y por ende una mayor recombinación y diversidad genética. Estas zonas estuvieron localizadas fundamentalmente en la costa de Almería y en el área metropolitana de Granada.

Como limitaciones encontradas en este primer capítulo destacamos que, como es frecuente en estudios retrospectivos, no pudimos recopilar todos los datos demográficos en los pacientes estudiados. De especial importancia hubiera sido poder disponer de la proporción de individuos con infección aguda/reciente, ya que los niveles de carga viral y de CD4 de los clusters mayoritarios de HSH con subtipo B sugieren que probablemente existe un mayor número de infecciones recientes entre estos pacientes. Elegimos un valor de bootstrap del 85% para definir los clusters de transmisión, para evitar perder individuos pertenecientes a cadenas de transmisión con un tamaño muy grande, y que presentaban evidencias clínicas y demográficas de estar incluidos en ellas. En general, valores de bootstrap por encima del 70% (52) para cadenas de transmisión grandes son fiables, a pesar de ello, asumimos el hecho en este capítulo de que algunos pocos individuos pudieran quedar fuera de alguna cadena de transmisión. Asimismo, en este primer estudio no hemos incluido secuencias que están relacionadas epidemiológicamente con los virus que han infectado a los pacientes incluidos en nuestra cohorte. Por lo tanto, asumimos que debido a la gran cantidad de pacientes migrantes infectados fundamentalmente con variantes de VIH-1 no-B (infecciones que a menudo se producen en sus países de origen), se hace fundamental en posteriores capítulos incluir estas secuencias control para estudiar de forma más precisa la relación epidemiológica entre nuestros pacientes. Por último, aunque hemos caracterizado el subtipo viral mediante la herramienta automatizada de subtipado rápido REGA 2.0, se hace necesario en próximos capítulos confirmar estas formas genéticas utilizando secuencias referencia a través de análisis filogenético (75).

En resumen, en este primer capítulo presentamos datos que demuestran que el estudio epidemiológico de las diferentes cadenas de transmisión de VIH-1 mediante análisis filogenético, se presenta como una herramienta potente y de gran utilidad para estudiar los colectivos de riesgo y las diferentes características que influyen en el agrupamiento de los pacientes en diferentes cluster de transmisión.

**Capítulo 4: ESTUDIO DE LOS PERFILES  
FILOGENÉTICOS, FILODINÁMICOS Y  
FILOGEOGRAFICOS DEL VIH-1  
SUBTIPO B.**

## 4.1 Antecedentes

El VIH-1 subtipo B es el más prevalente en países desarrollados de Europa occidental y central (76,77), implicando casi un 70% del total de las nuevas infecciones en pacientes naïve(35).

A pesar de que la tasa de infección por el virus de la inmunodeficiencia humana está decreciendo en los últimos años (6), existen ciertas sub-epidemias, especialmente en el subtipo B y dentro del colectivo HSH, donde el virus se transmite con mayor eficiencia y la infección resulta desproporcionada en comparación al resto de la población (38,70,71,78).

Desde comienzos de las primeras infecciones de VIH-1 en los años 1982, la epidemia ha seguido una dinámica cambiante en la distribución de subtipos, prevalencia de las rutas de transmisión y en las características de estos clusters de transmisión. La filodinámica y la filogeografía del VIH-1, pueden revelarnos información acerca de la historia evolutiva de una población viral, y ser utilizadas para investigar el curso de la epidemia y la dinámica temporal y espacial(39,40,79,80)(39,40,78,79)(39,40,77,78) del virus. En este escenario hemos decidido investigar la historia evolutiva del VIH-1 subtipo-B en Andalucía Oriental, mediante el uso de diversas herramientas de análisis filogenético.

En este capítulo se estudiarán los perfiles filogenéticos del VIH-1 subtipo B, realizando un análisis pormenorizado de los cluster de transmisión de VIH-1 subtipo B mediante métodos de análisis filogenético “clásico”. Asimismo, se implementarán métodos de análisis filogenético avanzados de inferencia Bayesiana (como explicamos en el capítulo 1) para establecer un marco temporal de la epidemia, inferir el origen más probable de aparición de los primeros linajes ancestrales de VIH-1 subtipo B y de los cluster de transmisión para este subtipo. Por último, se estudiarán los perfiles filodinámicos y filogeográficos del VIH-1 subtipo B, es decir, la dinámica de crecimiento del virus de VIH-1 subtipo B y su propagación a lo largo del tiempo, estudiando las rutas de mayor tránsito de VIH-1 y realizando una reconstrucción espacio-temporal de la epidemia de VIH-1 subtipo B.

## **4.2 Pacientes y métodos**

### **4.2.1 Población de estudio**

Con el fin de investigar la historia evolutiva de la epidemia de VIH-1 subtipo-B, en este capítulo se recogieron todas las secuencias caracterizadas en el capítulo 3 (a través de la herramienta REGA 2.0) como secuencias de VIH-1 Subtipo B (Suntipo B=452) o secuencias similares: secuencias Recombinantes B-D (n=45), secuencias subtipo B-Like (n=17) y secuencias subtipo B-potencial recombinante (n=6). Como se describe en el capítulo 3, estas secuencias fueron asociadas a datos demográficos, clínicos y virológicos.

### **4.2.2 Caracterización genética del VIH-1 subtipo B**

Se caracterizo pormenorizadamente el subtipo viral de cada secuencia de VIH-1 subtipo B o similar, a través de la herramienta REGA v3.0 ( disponible en <http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/>) y estos subtipos fueron confirmados a través de análisis filogenético a través de Máxima Verosimilitud. Para la confirmación a través de análisis filogenético se utilizó un grupo de secuencias de referencia de subtipos puros (subtipos de la A a la K) y CRFs disponibles hasta el momento del análisis obtenido de HIV Los Alamos (<http://www.hiv.lanl.gov>).

### **4.2.3 Estudio de los cluster de transmisión de VIH-1 subtipo B**

Para el estudio de los cluster de transmisión se seleccionaron todas las secuencias que fueron confirmadas definitivamente mediante análisis filogenético como VIH-1 subtipo B. Todas las secuencias fueron editadas manualmente y alineadas a través del algoritmo Clustal W(56), disponible en la plataforma CIPRES science Gateway (81). El modelo de sustitución nucleotídica más apropiado para este análisis fue determinado utilizando FindModel. Se realizó un análisis filogenético preliminar mediante Máxima Verosimilitud (ML), utilizando RAxML (del ingles, *randomized Axelerated Maximum Likelihood*) disponible en la plataforma CIPRES science Gateway. Para este análisis utilizamos el modelo de sustitución

nucleotídica General-Time Reversible (GTR) con 1000 réplicas de Bootstrap. Se estudiaron las asociaciones entre secuencias que presentaban nodos comunes con una significación estadística superior a un límite confiable, que establecimos esta vez en 70% de valor de Bootstrap (52) y se estudiaron de manera pormenorizada los clusters de subtipo-B con mayor relevancia epidemiológica (asociaciones de 5 o más secuencias con un valor superior al 70% de soporte de Bootstrap). Con la finalidad de aportar mayor fiabilidad a nuestros cluster de transmisión, como se explica más abajo, estas agrupaciones filogenéticas fueron confirmadas utilizando valores de probabilidad a posteriori a través de Inferencia Bayesiana.

#### **4.2.4 Reconstrucción de la historia evolutiva del VIH-1 subtipo B.**

Con la finalidad de establecer un marco temporal de todas las secuencias de VIH-1 subtipo B, se realizó un árbol filogenético de máxima credibilidad de clados (MCC) mediante inferencia bayesiana y utilizando la aproximación Bayesiana de las Cadenas de Markov (MCMC) a través de la aplicación BEAST v1.7.5 (82,83).

La escala de tiempo de nuestro reloj molecular se ajustó utilizando únicamente 408 secuencias de subtipo-B puro (secuencias que se asociaban con secuencias de subtipo B referencia, con un valor de Bootstrap mayor del 70% en REGA 3.0). La escala de tiempo fue realizada utilizando el modelo de sustitución nucleotídica Hasegawa, Kishino, and Yano (HKY) con una distribución de heterogeneidad entre sitios equivalente a gamma (4 categorías). Se utilizó un reloj molecular relajado (84) para estimar las diferentes tasas evolutivas ( $\mu$ , sustituciones nucleotídicas por sitio por año, sust/sitio/año) y el tiempo del más reciente comun antecesor (tMRCA) de los diferentes linajes de VIH-1 subtipo B.

Se estudiaron los clusters con mayor relevancia epidemiológica (definidos previamente mediante ML) con una probabilidad a posteriori mayor a 0.9 ( $pp > 0.9$ ) utilizando un modelo de demográfico de crecimiento no-paramétrico denominado *Bayesian skyline plot* (BSP)(85). El análisis de las cadenas de Markov y Monte Carlo se calculó con 250 millones de estados muestreados cada 50000 y se aceptaron únicamente las trazas con un tamaño efectivo (del inglés, Effective Sample Size) mayor de 200, excluyendo un 10% de los datos iniciales a través de TRACER v1.6 (programa disponible en <http://tree.bio.ed.ac.uk/software/tracer/>). Para estimar los cambios en el tamaño efectivo de la población ( $N_e$ ), usamos el modelo BSP y realizamos un análisis demográfico de nuestra población en TRACERv1.6.

Por ultimo, estudiamos origen de la epidemia y de las rutas epidemiológicas con mayor soporte estadístico, realizando una filogeografía discreta: georeferenciando a cada secuencia con valores discretos de latitud y longitud a través de BEAUti v1.7.5. Para reconstruir la historia evolutiva de la epidemia de VIH-1 subtipo B, se utilizó *SPREAD V1.0.6* (disponible en <http://www.kuleuven.ac.be/aidslab/phylogeography/SPREAD.html>). Esta aplicación nos permitió generar un archivo .KML, que nos permitió visualizar el orden secuencial de movimientos del virus a través del tiempo en Google Earth. Asimismo, se estudiaron las rutas de VIH-1 con mayor soporte estadístico a través del Factor de Bayes (BF), estableciendo como límite estadístico, rutas epidemiológicas con un Factor de Bayes superior a 3. Todos los árboles fueron creados y editados a través de FigTree v 1.4.0.

#### **4.2.5 Análisis estadístico**

Se realizó un estudio de regresión logística para estudiar relación de las variables demográficas, clínicas y virológicas con el agrupamiento de los pacientes en cluster de transmisión. Estos estudios estadísticos se realizaron utilizando SPSS 15 (SPSS Inc., Chicago, IL, USA).

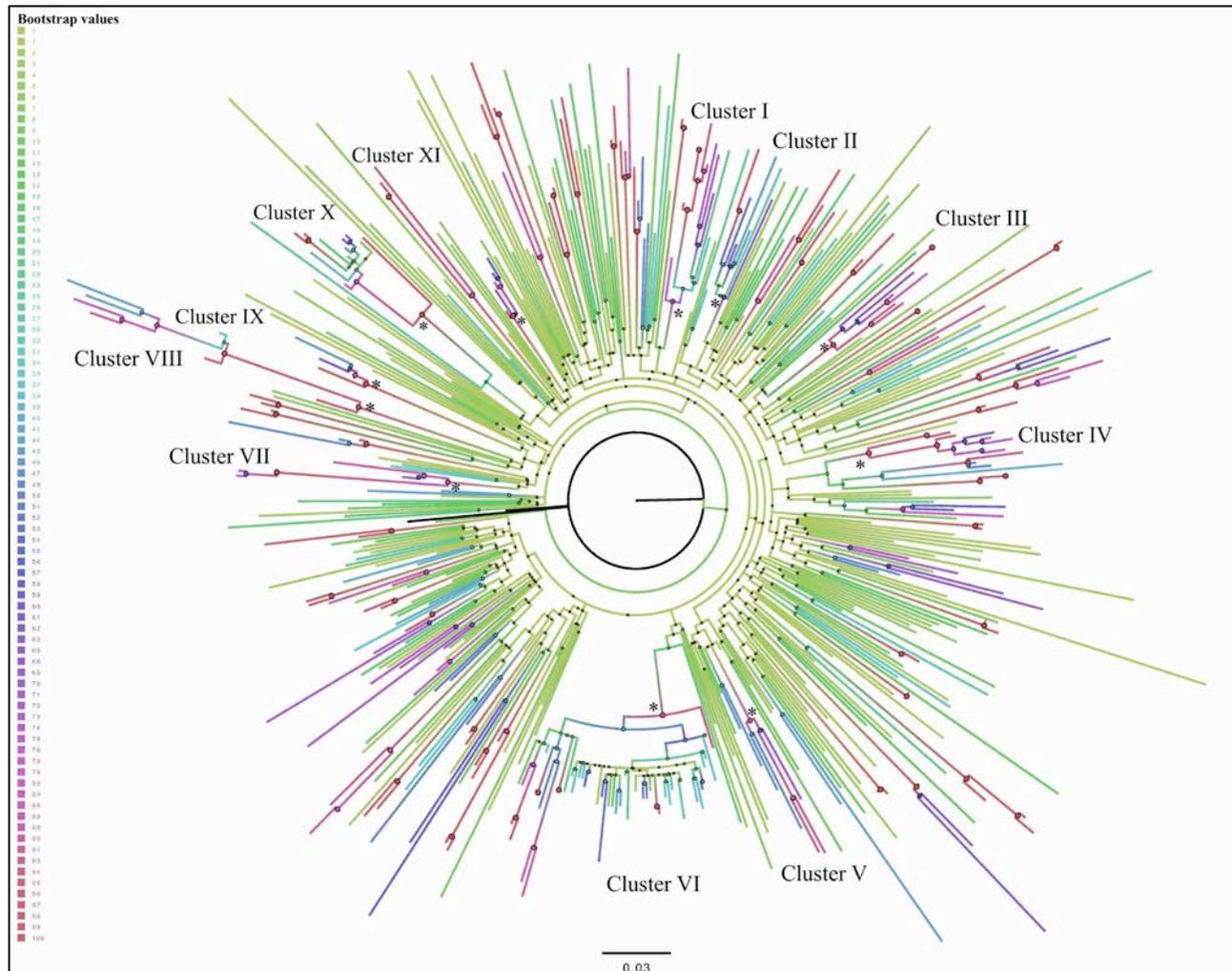
## **4.3 Resultados**

### **4.3.1 Caracterización genética del VIH-1 subtipo B**

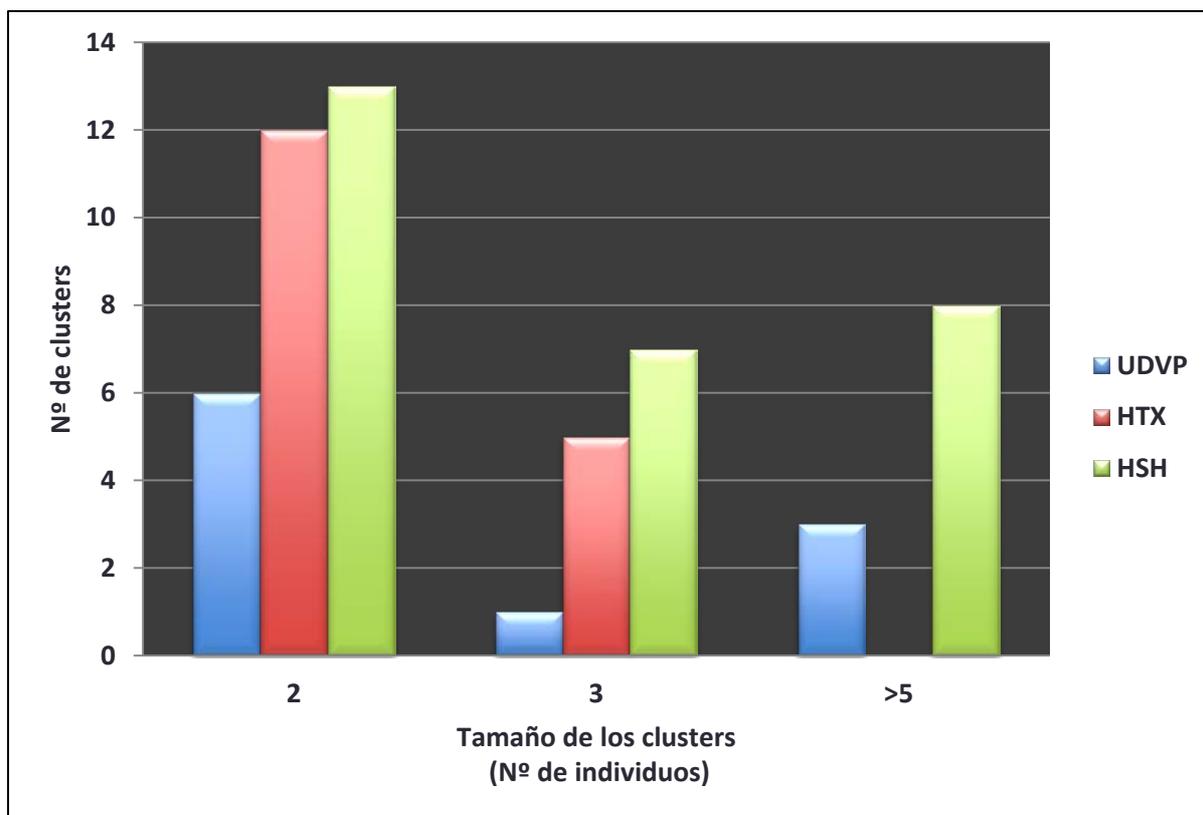
De los 693 pacientes infectados con diversas formas genéticas de VIH-1, 493 (71%) fueron confirmados definitivamente como subtipo B a través de análisis filogenético. Sin embargo, únicamente 408 secuencias fueron determinadas como VIH-1 subtipo B con un valor de Bootstrap por encima de 70% en REGA 3.0.

### **4.3.2 Estudio de los clusters de transmisión de VIH subtipo B.**

El análisis filogenético realizado sobre el total de secuencias de VIH-1 subtipo B ( N=493) a través del método de Maxima Verosimilitud (figura 19) mostró la presencia de 55 clusters que englobaban al 47% (n=234) de la población estudiada. El tamaño y vía de transmisión de los diferentes cluster se resume en la gráfica representada en la figura 20. El 43%(101/234) de los individuos se agruparon en 44 clusters de tamaño pequeño, [31 clusters, n=2 individuos, (13 HSH), (12 HTX), (6 UDVP)] y [13 clusters, n=3 individuos, (HSH=7), (HTX=5), (UDVP=1)]. El 57%(133/234) de los individuos se agruparon en clusters con 5 o más individuos, sus características demográficas, clínicas y virológicas se reflejan en la tabla 8. Como se refleja en la tabla 8, se encontraron 11 clusters de transmisión de gran tamaño, 8 clusters estaban formados por pacientes HSH y 3 por UDVP. Cabe destacar por su gran tamaño el cluster VI, formado por 58 individuos españoles HSH, de edad joven (33 años) y con un número elevado de CD4 (518).



**Figura 19.** Arbol realizado mediante Máxima Verosimilitud (ML) del global de secuencias de VIH-1 subtipo B. El árbol filogenético fue construido utilizando el modelo de sustitución nucleotídica General Time Reversible (GTR) utilizando RaxML. Los Clusters con un valor de Bootstrap confinable (>70%) y con un número igual o mayor a 5 individuos se representan con un asterisco (\*).



**Figura 20. Representación esquemática del tamaño y la vía de transmisión de los diferentes cluster de transmisión encontrados para el VIH-1 subtipo B.**

#### 4.3.4 Variables predictoras de agrupamiento de los pacientes en cluster de transmisión.

Como se muestra en la tabla 7, se encontraron asociaciones significativas entre los pacientes asociados en cluster de transmisión en comparación con los que no lo estaban. La asociación en cluster para los pacientes infectados con VIH-1 subtipo B, fue más frecuente, en hombres menores de 35 años ( $p=0.001$ ), de nacionalidad española ( $p=0.02$ ), homosexuales ( $p<0.001$ ) y con un nivel de CD4 por encima de 350 cel/ul ( $p<0.001$ ).

	Pacientes, n(%)			P valor <sup>a</sup>	OR (95% CI)	P valor <sup>b</sup>
	Total	En Cluster	Fuera de Cluster			
<b>Genero</b>	493(100)	234(100)	259(100)	<b>&lt;0.001</b>		
Hombre	404 (81,9)	209(89,3)	195(75,3)		1	
Mujer	89 (18,1)	25(10,7)	64(24,7)		0.312(0.184-0.531)	<0.001
<b>Nacionalidad (453)*</b>				<b>0.020</b>		
Español	398 (87,9)	196 (89,9)	202 (86)		1	
Extranjero	55 (12,1)	22 (10,1)	33 (14)		0.491(0.267-0.904)	0.022
<b>Edad(460)*</b>				<b>0.001</b>		
<35	123(26,7)	78(35,3)	45(18,8)		1	
35-45	167(36,3)	72(32,6)	95(39,7)		0.480(0.299-0.771)	0.002
>45	170(37)	71(32,1)	99(41,4)		0.445(0.278-0.774)	0.001
<b>Ciudad</b>				0.082		
Almería	60(12,2)	26(11,1)	34(13,1)			
El Ejido	66(13,4)	31(13,2)	35(13,5)			
Granada	233(47,3)	112(47,9)	121(46,7)			
Motril	46(9,3)	30(12,8)	16(6,2)			
Jaén	88(17,8)	35(15)	53(20,5)			
<b>Fecha de muestreo</b>				0.388		
2005-2006	40(8,1)	16(6,8)	24(9,3)			
2006-2007	41(8,3)	14(6)	27(10,4)			
2007-2008	62(12,6)	29(12,4)	33(12,7)			
2008-2009	48(9,7)	26(11,1)	22(8,5)			
2009-2010	109(22,1)	58(24,8)	51(19,7)			
2010-2011	45(9,1)	26(11,1)	19(7,3)			
2011-2012	76(15,4)	30(12,8)	46(17,8)			
2012-2013	72(14,6)	35(15)	37(14,3)			
<b>Carga Viral (456)*</b>				0.070		
<10.000	94(20,6)	38(17,4)	56(23,5)			
10.000-100.000	205(45)	101(46,3)	104(43,7)			
>100.000	157(34,4)	79(36,2)	78(32,8)			
<b>Nivel de CD4(427)*</b>				<b>&lt;0.001</b>		
<200	121(28,3)	38(19,2)	83(36,2)		1	
201-350	95(22,2)	43(21,7)	52(22,7)		2.079(1.185-3.646)	0.011
>350	211(49,4)	117(59)	94(41)		2.980(1.847-4.809)	<0.001
<b>Vía de transmisión (433)*</b>				<b>&lt;0.001</b>		
HSH	210(48,5)	156(66,7)	54(27,1)		1	
IVDU	91(21)	39(16,7)	52(26,1)		0.286(0.170-0.480)	<0.001
HTX	132(30,5)	39(16,7)	93(46,7)		0.271(0.170-0.432)	<0.001

**Tabla 7. Características Clínicas, demográficas y virológicas de los pacientes incluidos en cluster en comparación con los no incluidos en cluster de transmisión.**

<sup>a</sup>P valor para el test de chi-cuadrado

<sup>b</sup>P valor para la regresión logística univariada.

\*Datos disponibles para los pacientes indicados.

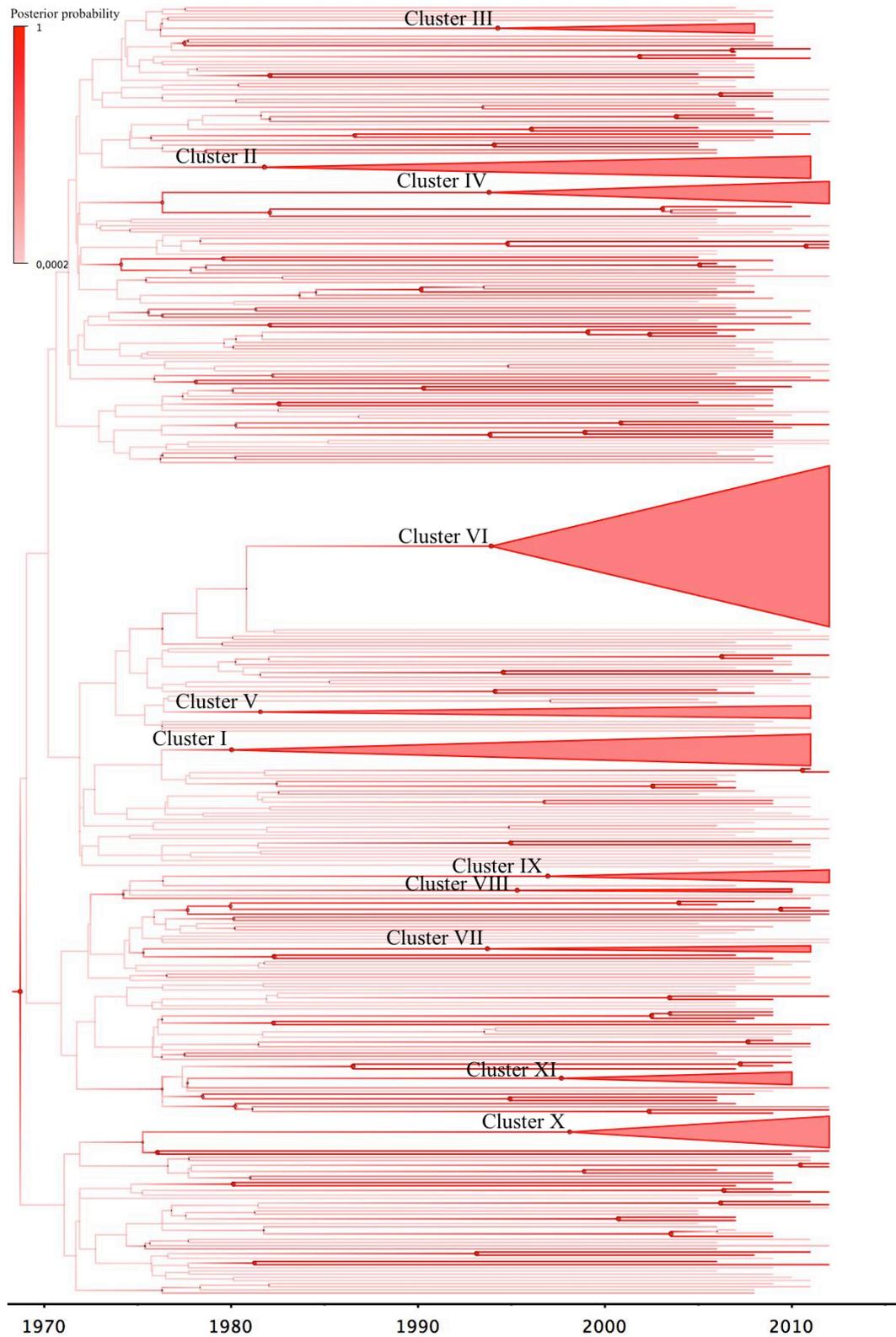
#### **4.3.5 Estudio de la historia evolutiva del VIH-1 a través de inferencia Bayesiana**

El análisis filogenético realizado sobre la región genómica *pol*, reveló una tasa evolutiva para el subtipo B estimada en 2.4 (95%IC: 1.7 - 3.1) x 10<sup>-3</sup> sustituciones/sitio/año. Como se muestra en el árbol de máxima credibilidad de clados (MCC) obtenido a través de inferencia Bayesiana (figura 21), los linajes ancestrales para el subtipo B se originaron alrededor del año 1969 (95% CI: 1960-1980). En este árbol filogenético se muestran bajo una escala temporal, el orden de aparición de los clusters con mayor relevancia epidemiológica ( $\geq 5$  pacientes) y sus características demográficas, clínicas y virológicas quedan recogidas en la tabla 8. Todos los clusters definidos previamente mediante ML fueron confirmados mostrando valores de probabilidad a posteriori superiores a 0.9.

Los diferentes parámetros filogenéticos: tMRCA de los cluster de transmisión o tasa evolutiva fueron anotados en la tabla 8.

#### **Filogeografía del VIH-1 subtipo B**

La filogeografía discreta realizada para los pacientes infectados con VIH-1 subtipo B se refleja en el árbol de MCC que aparece en la figura 22. En este árbol filogenético cada rama del árbol representa la localización más probable para cada evento de transmisión. Los nodos reflejan la probabilidad de localización de las diferentes agrupaciones filogenéticas (de 0 a 1) y los cluster de transmisión con mayor repercusión epidemiológica ( $\geq 5$  pacientes) quedan señalados bajo un asterisco. Finalmente, los datos filogenéticos inferidos de este análisis filogenético (localización y probabilidad de localización de los cluster de transmisión) se anotaron en la tabla 8.



**Figura 21.** Arbol de máxima credibilidad (MCC) realizado mediante inferencia Bayesiana en Beast v1.7.5. Las ramas son coloreadas utilizando un valor progresivo de probabilidad a posteriori de 0 a 1. Los clusters con mayor repercusión epidemiológica y con valores altos de probabilidad a posteriori ( $pp > 0.9$ ) quedan representados bajo una escala temporal y definidos con un asterisco(\*).

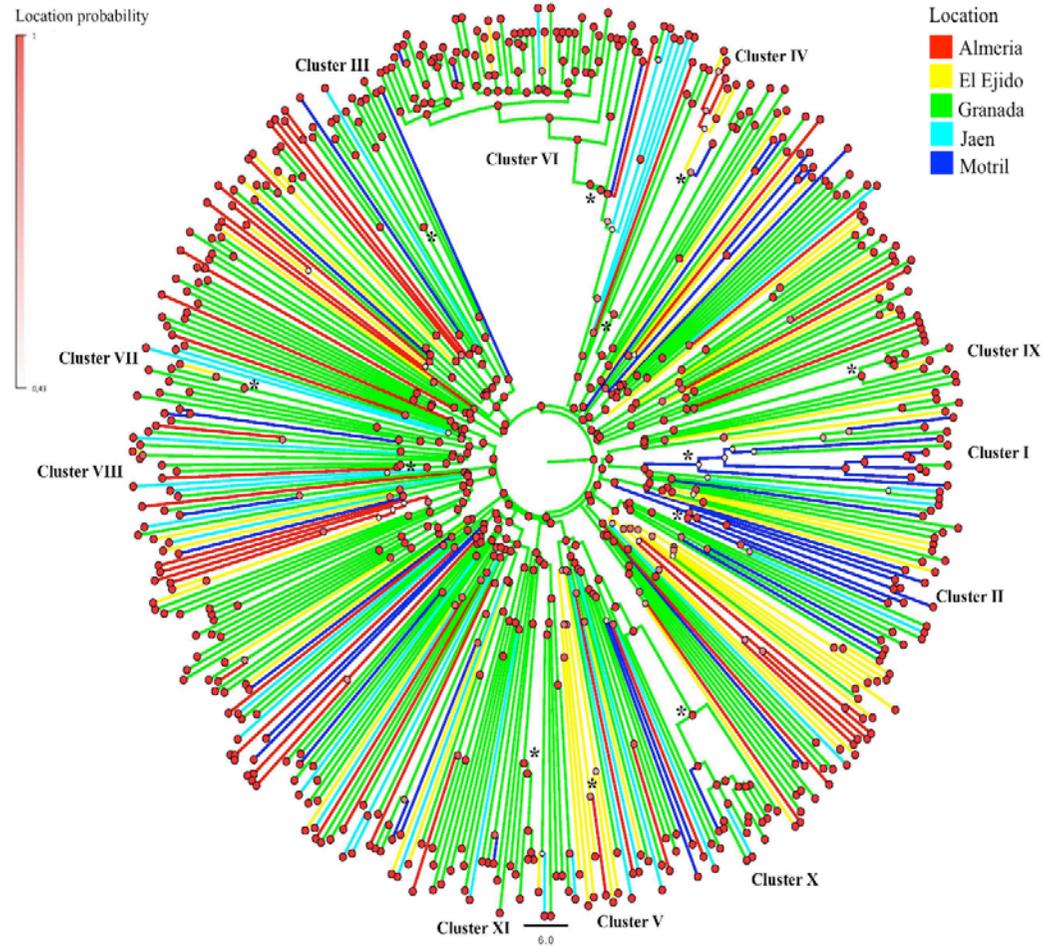


Figura 22. Filogeografía discreta del HIV-1 subtipo B realizada en Beast 1.7 de los árboles Bayesianos de máxima credibilidad (MCC). Los nodos son representados bajo un gradiente de color progresivo para la probabilidad de localización ancestral. La barra superior izquierda representa el gradiente de probabilidad progresivo de localización de 0 a 1. El color de cada rama indica la localización de cada secuencia que se refleja en la parte superior derecha. Los clusters con mayor repercusión epidemiológica quedan señalados con un asterisco(\*).

Cluster ID	N	Vía de transmisión	Fecha de muestreo	Edad (mediana, IQR)	Nacionalidad	Carga Viral (mediana, IQR) Log <sub>10</sub>	CD4 (mediana, IQR)	Valor de Bootstrap (%)	Probabilidad a posteriori	tMRCA (95% HPD)	μ(sust./sitio/año)	Localización Ancestral	Probabilidad de localización
I	11	UDVP	2005-2011	47(46-48)	Española	4.23(3.71-4.72)	330(132.5-541)	76	0.98	1980(1972-1990)	3.3 x10 <sup>-3</sup> (1.4 x10 <sup>-3</sup> -5.7 x10 <sup>-3</sup> )	Motril	0.5
II	8	UDVP	2005-2011	40(37-48.5)	Española	4.73(3.85-5.36)	323.5(142-442.8)	70	0.99	1982(1976-1994)	3.4 x10 <sup>-3</sup> (1.2 x10 <sup>-3</sup> -5.7 x10 <sup>-3</sup> )	Motril	0.94
III	7	HSH	2005-2009	49(44-54)	Española	4.84(4.44-5.22)	373.5(256.8-628.75)	96	0.99*	1994(1984-1999)*	3.0 x10 <sup>-3</sup> (1.3 x10 <sup>-3</sup> -4.8 x10 <sup>-3</sup> )*	Granada*	0.99*
IV	8	HSH	2009-2012	39.5(32-48)	Española -Rumana	5.34(3.8-5.5)	553(491-849)	94	0.97	1994(1982-1997)	2.8 x10 <sup>-3</sup> (1.2 x10 <sup>-3</sup> -4.8 x10 <sup>-3</sup> )	Granada	0.96
V	5	UDVP	2005-2011	45(37-53.5)	Española-Algelina	4.81(2.98-5.59)	295(185.5-551.5)	88	0.95*	1982(1977-1995)*	3.1 x10 <sup>-3</sup> (1.4 x10 <sup>-3</sup> -5.2 x10 <sup>-3</sup> )*	El Ejido*	0.85*
VI	58	HSH	2007-2012	32.5(28-39.3)	Española	4.42(4.03-5.27)	518(301-896)	95	0.99	1994(1984-1999)	2.8 x10 <sup>-3</sup> (1.3 x10 <sup>-3</sup> -4.7 x10 <sup>-3</sup> )	Granada	0.99
VII	7	HSH	2007-2012	39(32.3-52.5)	Española - Argentina-Ecuatoriana	5.15(4.41-6.21)	326.5(241.5-847.25)	86	0.98*	1994(1993-2006)*	2.1 x10 <sup>-3</sup> (1 x10 <sup>-3</sup> -3.4 x10 <sup>-3</sup> )*	Granada*	0.96*
VIII	7	HSH	2009-2011	38(29.5-42.3)	Española	4.46(4.28-5.04)	279.5(8.25-598)	96	0.99*	1995(1988-2004)*	3.6 x10 <sup>-3</sup> (1.7 x10 <sup>-3</sup> -5.7 x10 <sup>-3</sup> )*	Granada*	0.82*
IX	5	HSH	2006-2012	43(38-45.5)	Española	5.71(4.67-6.32)	213(86-542.5)	100	0.99	1997(1995-2004)	3.1 x10 <sup>-3</sup> (1.6 x10 <sup>-3</sup> -4.7 x10 <sup>-3</sup> )	Granada	0.99
X	12	HSH	2007-2012	37.5(30-44.3)	Española	5.17(4.81-5.98)	331(252.3-388)	99	0.93	1998(1985-1999)	3.5 x10 <sup>-3</sup> (1.6 x10 <sup>-3</sup> -5.6 x10 <sup>-3</sup> )	Granada	0.99
XI	5	HSH	2005-2010	43(34-45)	Española	4.36(3.94-5.14)	325(174.5-380)	99	0.99	1998(1993-2002)	2.3 x10 <sup>-3</sup> (1 x10 <sup>-3</sup> -3.7 x10 <sup>-3</sup> )	Granada	0.99

**Tabla 8. Características epidemiológicas y filogenéticas de los cluster con mayor repercusión epidemiológica (≥ 5 pacientes) encontrados dentro del VIH-1 subtipo B. \* Estimaciones obtenidas utilizando un menor número de pacientes debido al filtrado de subtipos B puros en REGA 3.0.**

#### **4.3.6 Interpretación de los resultados obtenidos de los árboles filogenéticos realizados mediante Inferencia Bayesiana.**

Los primeros clusters de transmisión se originaron a principios de los años 80 formados por usuarios de drogas por vía parenteral en las zonas de costa de Motril; Cluster I (location probability= 0.5) en 1980 (95% IC: 1972-1990), Cluster II (location probability= 0.94) en 1982 (95% IC: 1976-1994) y El Ejido; Cluster V (location probability=0.85) en 1982 (95% IC: 1977-1995). Asimismo, como se refleja en la tabla 8, la mayoría de los clusters con mayor repercusión epidemiológica (clusters que estaban formados todos por HSH), aparecieron en la década de los 90. Todos los clusters formados entre HSH, se originaron en Granada con un valor de probabilidad a posteriori por encima de 0.9. La tasa evolutiva para la mayoría de los clusters de transmisión de gran relevancia epidemiológica fue superior al valor medio de la población viral de VIH-1 subtipo-B (por encima de  $3 \times 10^{-3}$  sustituciones/sitio/año).

#### 4.3.7 Representación espacial de la propagación de VIH-1 subtipo B en Andalucía oriental.

La figura 23 muestra la reconstrucción espacial de la filogeografía del VIH-1 subtipo B en Andalucía oriental realizada a través de SPREAD y visualizada Google Earth. Los linajes ancestrales del virus de la inmunodeficiencia humana de subtipo B se originaron en Granada a comienzos de la década de los 70. De esta zona, la epidemia se propagó secuencialmente a lo largo del tiempo hacia otras regiones periféricas de Andalucía Oriental, en primer lugar: a las zonas de costa de El Ejido, Motril y Almería y por último, a Jaén.

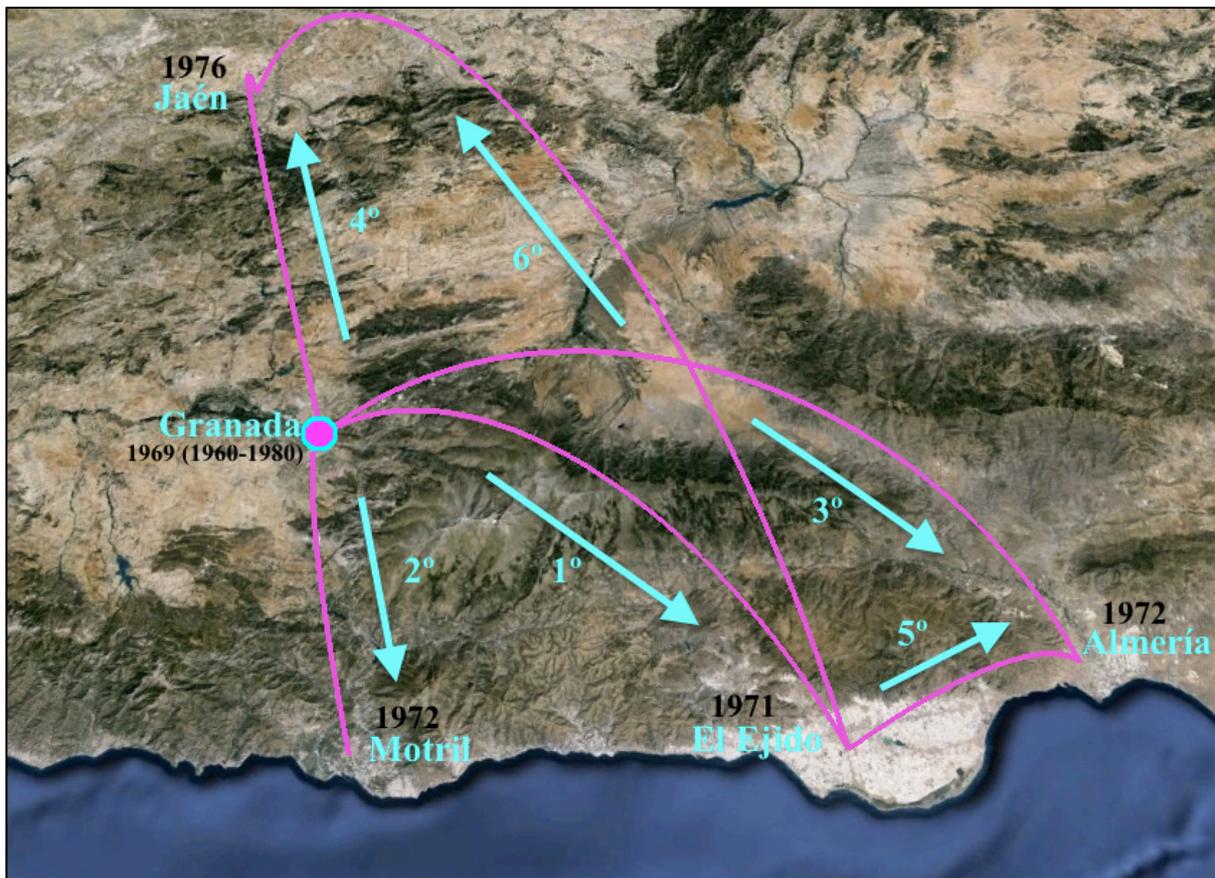


Figura 23. Reconstrucción ancestral de la propagación del VIH-1 subtipo B en Andalucía Oriental. La representación espacial de la filogeografía fue realizada en SPREADv1.0.6 y visualizada en Google Earth. Los números indican el orden secuencial de movimientos en el tiempo. Los linajes ancestrales del VIH-1 subtipo-B quedan representados en un círculo.

#### 4.3.8 Análisis de las rutas migratorias con mayor soporte estadístico

Como se muestra en la figura 24, las rutas epidemiológicas más significativas del VIH-1 subtipo-B (Factor de Bayes >3) mostraron una dirección de transmisión del VIH-1 subtipo B de las zonas de interior de Granada y Jaén, hacia las zonas periféricas de Motril y principalmente, a las zonas de costa de El Ejido y Almería.

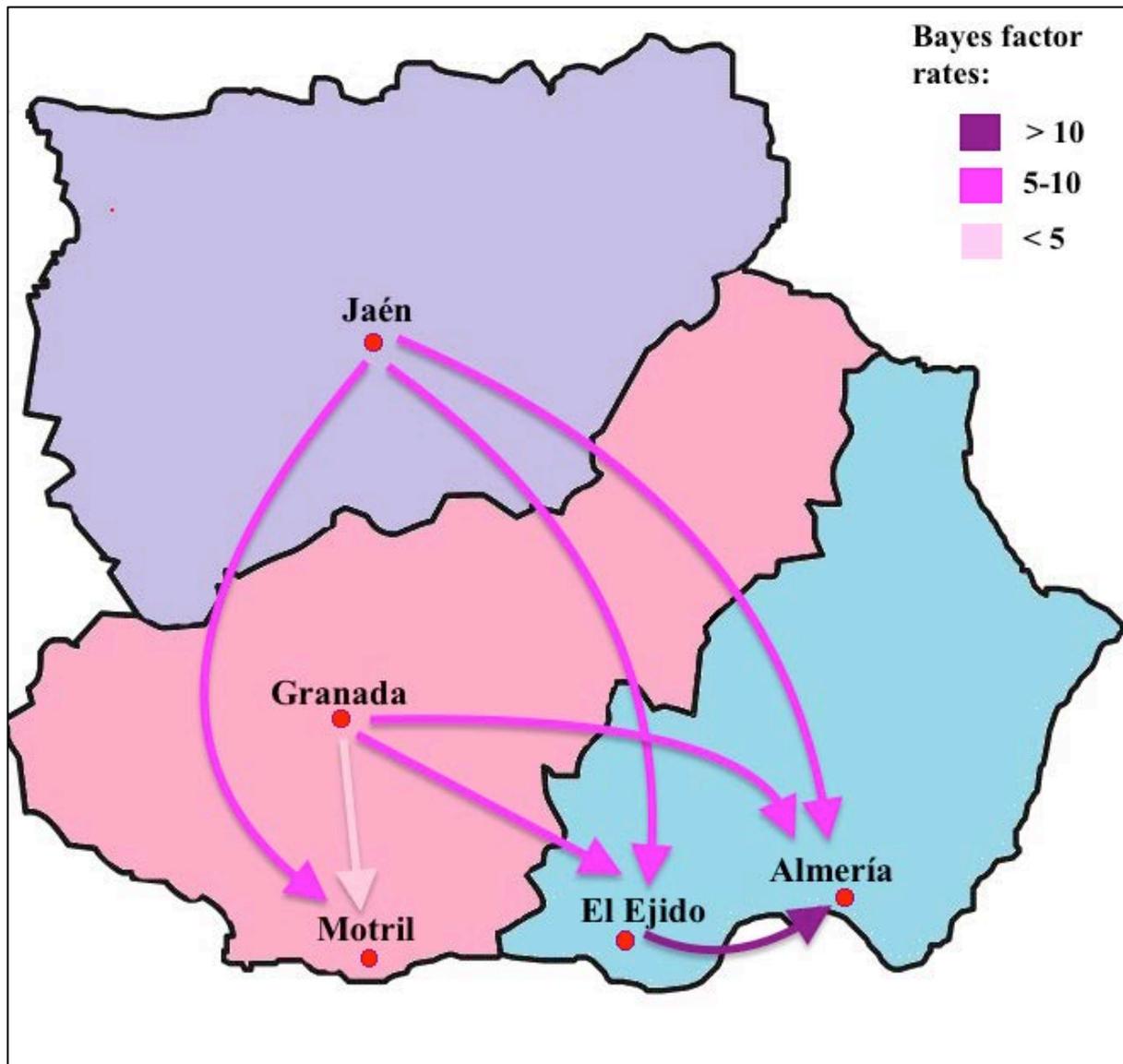
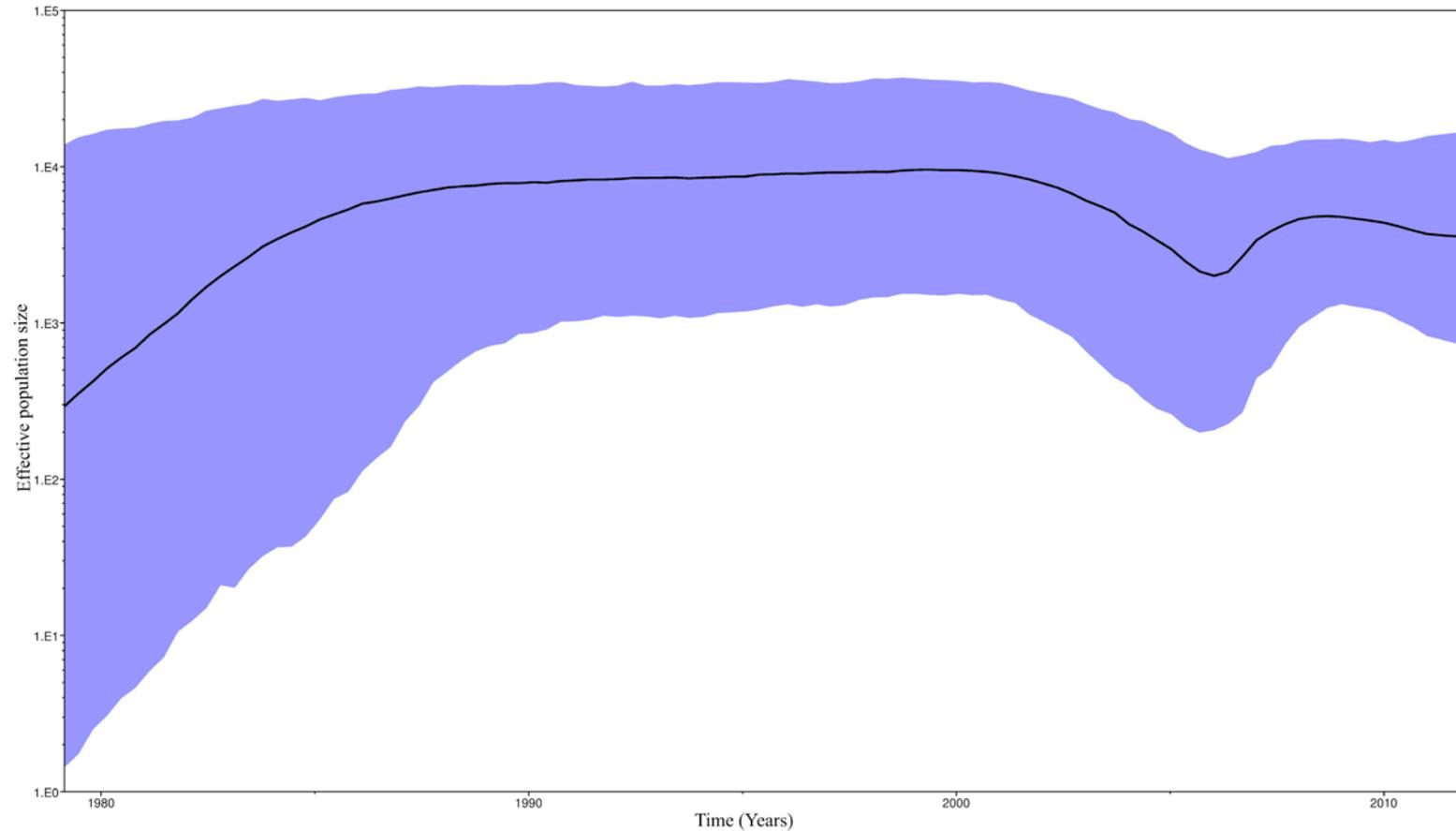


Figura 24. Representación esquemática de las rutas migratorias de VIH-1 subtipo B con mayor soporte estadístico. Solo las rutas migratorias con un valor de Factor de Bayes por encima de 3 son representadas. La leyenda superior izquierda muestra los valores de Factor de Bayes, coloreados en función del valor de soporte estadístico.

#### **4.3.9 Análisis de los perfiles filodinámicos del VIH-1 subtipo B.**

Con el fin de reflejar la dinámica de crecimiento en la población de VIH-1 subtipo-B en Andalucía Oriental, se estimaron los cambios en el tamaño efectivo de la población en las últimas tres décadas realizando un análisis demográfico en TRACERv1.6. Para este análisis se utilizó de nuevo el modelo de crecimiento demográfico no-paramétrico *Bayesian Skyline Plot* (figura 25). Así pues, existieron cambios temporales en el tamaño efectivo de la población, el análisis demográfico reveló una fase inicial de crecimiento del número efectivo de infecciones a principios de los años 80, que pasó a estabilizarse en la década de los 90 para. Asimismo, a principios del año 2000 comenzó a existir un decrecimiento del tamaño efectivo de la población viral, pero existió un nuevo rebrote en el tamaño efectivo de la población desde 2005 a 2008 que finalmente pasó a estabilizarse a principios del año 2010.



**Figura 25.** Análisis de los perfiles filodinámicos de la epidemia de VIH-1 subtipo B. El análisis demográfico de crecimiento de la población se realizó mediante Bayesian Skyline Plot. La línea negra representa la media estimada para el tamaño efectivo de la población a lo largo del tiempo, reflejada en escala logarítmica en el eje de ordenadas. El área azul refleja el intervalo que representa el 95% de intervalo de confianza (HPD, *High posterior density*) para esta estima.

#### 4.4 Discusión

En este capítulo demostramos como el VIH-1 subtipo B sigue siendo la variante de VIH-1 que predomina en España (86,87) encontrando un total de 493 (71%) pacientes infectados con este subtipo. Asimismo, El 47% de los pacientes infectados con subtipos B estaban relacionados epidemiológicamente, proporción que está dentro del rango (15-50%) de otros estudios realizados para este subtipo en Europa (88–90). Hemos demostrado como más de la mitad (57%) de los pacientes nuevos diagnósticos en Andalucía oriental infectados con VIH-1 subtipo-B se asociaron en cluster de transmisión formados por 5 individuos o más, y la mayoría de los pacientes fueron jóvenes HSH con un alto recuento de CD4. Estos datos sugieren la presencia de un número considerable de pacientes con infecciones agudas/recientes.

Nuestros resultados revelaron una tasa evolutiva para la población viral de VIH-1 subtipo-B en Andalucía Oriental alrededor de  $2.4$  (95% IC:  $1.7 - 3.1$ )  $\times 10^{-3}$  sustituciones/por sitio/por año), datos que son semejantes a los publicados en estudios realizados en otros países Europeos (91–93). Asimismo, la existencia de una tasa evolutiva superior a la media de la población viral en la mayoría de las redes de transmisión con mayor repercusión epidemiológica, podría sugerir una mayor eficiencia en la transmisión del virus en pacientes agrupados en cluster de transmisión de gran tamaño (94).

Como muestran datos obtenidos en otros países de Europa y America, los análisis de Inferencia Bayesiana han servido para datar la introducción de el VIH-1 subtipo-B a mitad de los años 60(91,95–97). Nuestros datos reflejan que el origen del VIH-1 subtipo-B en Andalucía Oriental se estimó (tMRCA inferido en 1969) años antes de la aparición de los primeros casos de VIH-1 diagnosticados en 1981(98) en España. El estudio de la historia demográfica realizado en nuestra población muestra que el tamaño efectivo de la población viral del VIH-1 subtipo B en Andalucía oriental comenzó a crecer de forma exponencial a principios del año 1980 hasta comienzos del año 1990, datos que manifiestan un patrón similar de crecimiento, al de otros estudios realizados en otros países (96,99–102) vecinos. Nuestro estudio demuestra como los primeros cluster de transmisión se originaron en la década de los 80 entre usuarios de drogas por vía parenteral en las zonas de costa: El primer cluster (Cluster I) se originó en Motril, cuyo tMRCA se estimó alrededor del año 1980, los otros dos clusters de transmisión aparecieron

dos años después en Motril (Cluster II) y en el Ejido (Cluster V). La existencia de los primeros clusters en zonas de costa podría estar relacionado con el mayor acceso a la droga en los primeros años de la infección por el VIH-1 en las zonas periféricas de costa.

El crecimiento de la epidemia de VIH-1 subtipo-B se mantuvo estable a comienzos de la década de los 90, hecho que probablemente vino acompañado de las primeras campañas de prevención realizadas en España en 1993(103). A partir de los años 90, los patrones de transmisión en España han ido evolucionando gradualmente de UDVP a HSH. Esto explica en parte por qué hemos detectado la emergencia de la mayoría de cluster de transmisión entre HSH en estas fechas, y está en la línea con el aumento de los comportamientos de alto riesgo entre HSH reportados en Europa durante estos años (104–106). En este periodo, la introducción del tratamiento antiretroviral de alta eficacia (TAR) en 1996 en España ha ayudado a reducir el tamaño efectivo de la población en los pacientes infectados con el VIH-1(107–110). Sin embargo, en 2005 la epidemia de VIH-1 subtipo-B experimentó un nuevo crecimiento exponencial, lo cual indica que las medidas preventivas para evitar el crecimiento de la epidemia deben ser implementadas y mantenidas en el tiempo. De este modo, en este capítulo demostramos que, el análisis filodinámico podría ayudar a evitar la aparición de nuevos brotes del virus de la inmunodeficiencia humana, ya que como ha sido demostrado en otros trabajos, cambios en el tamaño efectivo de una población viral pueden predecir futuros aumentos en la incidencia de VIH-1 en un área determinada(111).

De especial importancia en nuestra población es la presencia del cluster VI, formado por 58 individuos, cuyo tMRCA se estima en 1994(95% CI: 1984-1999). En esta red de transmisión se siguieron diagnosticando casos de pacientes infectados con el virus hasta el final de nuestro estudio. La existencia en este colectivo de HSH, de un alto nivel de CD4, la edad joven de los pacientes y la poca distancia genética entre las diferentes ramas de este cluster de transmisión, sugieren la presencia de una alta proporción de individuos con infección aguda/reciente.

El análisis filogeográfico realizado en este capítulo muestra como el VIH-1 subtipo-B se está propagando con mayor eficiencia de las zonas de interior a las zonas de costa. Este hecho puede explicarse parcialmente a consecuencia de la visita de muchos turistas a estas áreas. Asimismo, dado la proximidad geográfica entre El Ejido y Almería se produce también un gran tránsito de VIH entre estas dos zonas.

Como limitaciones encontradas en este capítulo destacamos que, aunque hemos incluido todas las secuencias posibles de VIH-1 subtipo-B en esta zona, gracias a los estudios de resistencia a antiretrovirales durante este periodo, la presencia de sujetos infectados de VIH no diagnosticados, nos ha impedido la total detección de todos los cluster de transmisión que circulan en Andalucía oriental. Asimismo, cabe destacar que aunque las guías recomiendan el estudio de resistencias a fármacos antiretrovirales de primera línea en todos los pacientes nuevos diagnósticos, esto no es siempre así. Por otra parte, en este estudio no tuvimos en cuenta la migración (incorporando secuencias similares en bases de datos), porque el análisis de un conjunto de datos demasiado grande en BEAST habría sido computacionalmente imposible. Este hecho nos llevó a seleccionar un conjunto de secuencias menor (n=408) con respecto al total de pacientes infectados con virus de VIH-1 subtipo-B en Andalucía Oriental.

En resumen en este capítulo presentamos el primer estudio sobre la historia evolutiva del VIH-1 subtipo-B (mediante el estudio de los perfiles filodinámicos y filogeográficos) en España, concretamente en Andalucía oriental. Asimismo, presentamos datos de suma importancia para el desarrollo de estrategias de prevención en nuestra área.

**Capítulo 5: ANÁLISIS DE LOS  
PERFILES FILOGENÉTICOS Y  
FILOGEOGRÁFICOS DE LAS FORMAS  
GENÉTICAS DE VIH-1 SUBTIPO NO-B.**

## 5.1 Antecedentes

La variabilidad genética y la capacidad de recombinación existente en el virus de la inmunodeficiencia humana (VIH) determina la existencia de una amplia diversidad de subtipos y formas recombinantes en el mundo(112–115). Los virus de subtipo no-B causan cerca del 90% de infecciones a nivel global, predominando en la población Africana o en países de Europa del Este como formas endémicas(18). Sólo los subtipos C y A y las formas CRF01\_AE, CRF02\_AG son responsables del 70% de las infecciones a nivel mundial (18). En la actualidad las infecciones por variantes de VIH-1 subtipo no-B en España oscilan entre el 12-15% (86,87), sin embargo la alta prevalencia del VIH-1 subtipo B en países donde existe una mayor cobertura en la terapia antiretroviral, hace que este subtipo sea el más estudiado desde el punto de vista genético, biológico y terapéutico.

Aún no se comprende completamente el significado o la importancia biológica de la variabilidad genética del VIH-1. Sin embargo, se han descrito diferencias importantes entre los subtipos genéticos en propiedades biológicas, como: virulencia, tropismo y transmisibilidad (116,117), utilización de los correceptores de quimiocinas (118), progresión en la enfermedad (119,120) susceptibilidad a algunas drogas antirretrovirales (121,122), sensibilidad de algunos métodos de cuantificación de carga viral (123,124) y correcta detección (125). Asimismo, diferentes estudios han comparado la población según el subtipo viral, señalando como factor clave de respuesta al tratamiento antiretroviral las características socio-culturales, por encima incluso de las características étnicas o biológicas de las formas genéticas de VIH-1(126,127). Sin embargo, otros estudios han encontrado una peor respuesta al tratamiento (expresada en probabilidad alcanzar CV plasmática indetectable) en pacientes de raza negra, frente a caucásicos (128). Por este motivo, estos hallazgos hacen evidente para la comunidad científica la importancia de la información epidemiológica referente a los diferentes subtipos.

La migración ha permitido la introducción en determinadas regiones de variantes del virus poco frecuentes (129,130), hecho que está cambiando la epidemiología molecular del VIH-1 en los países de acogida. En Europa se ha descrito una prevalencia de variantes de VIH-1 no-B variable que oscila entre un 15% y un 50% en función del escenario epidemiológico (131–135). Asimismo, en España estamos observando en algunas regiones un cambio en la prevalencia de algunas formas genéticas de VIH-1 subtipo no-B (130,136). En este contexto, los estudios de filogeografía en los que se analizan las relaciones filogenéticas de las diferentes secuencias

encontradas en una determinada region, pueden ser claves para analizar el marco epidemiológico en el que se encuentra una determinada población.

Dada la cercanía geográfica de Andalucía Oriental al continente Africano, esta región se ha caracterizado en la última década, por ser receptora de corrientes migratorias extranjeras, ocupando la cuarta posición de comunidades con mayor presencia de población extranjera, solo por detrás de Cataluña, Madrid y la Comunidad Valenciana. La principal fuente de migración en Andalucía Oriental, se debe a la agricultura intensiva, siendo Almería una de las provincias Andaluzas con mayor proporción de migrantes (uno de cada cuatro ciudadanos empadronado en Almería es migrante) estando localizada principalmente en la localidad de El Ejido. Actualmente, la economía Ejidense cimienta sus bases en la agricultura mediante invernaderos, en torno a la cual se ha generado una importante industria gracias a la mano de obra procedente de la migración, fundamentalmente de población Africana. En este contexto, la población migrante se enfrenta a grandes obstáculos para recibir atención médica debido a las grandes barreras lingüísticas o culturales y a las dificultades económicas y sociales, influyendo estos factores, en el retraso en su diagnóstico clínico, presentando en muchos casos infecciones tardías.

En toda reconstrucción filogenética se asume que no todos los eventos de transmisión se han muestreado y que por ende se detectan. Estos hechos pueden causar una subestimación de la proporción real de pacientes relacionados entre si especialmente en las infecciones importadas, ya que la probabilidad de reconstruir la verdadera historia evolutiva y de incluir miembros de una cadena de transmisión es muy baja. En estos casos quedarían por detectar fenómenos de transmisión que se han producido en los países de origen de los pacientes. Debido a la gran proporción de pacientes migrantes en nuestra población infectados con formas genéticas de VIH-1 subtipo no-B, se hace fundamental para reconstruir una historia evolutiva acertada, la incorporación en nuestros análisis filogenéticos de secuencias control similares a los virus que circulan en sus países de origen.

En este capítulo se estudian detalladamente las variables clínicas, demográficas y virológicas de los pacientes infectados con las principales variantes de VIH-1 no-B, la relación filogenética de los pacientes infectados en Andalucía Oriental, con virus circulantes en otros países y la presencia de cluster de transmisión entre la población autóctona. Asimismo, para aportar una mayor solidez a

nuestros estudio, como en el capítulo anterior, utilizaremos métodos avanzados de análisis filogenéticos basados en Inferencia Bayesiana.

## **5.2 Pacientes y métodos**

### **5.2.1 Población de estudio y caracterización genética de las variantes de VIH-1 no-B.**

En este capítulo se recopilaron todas las secuencias caracterizadas como variantes de VIH-1 no-B a través de la herramienta de subtipado rápido REGA 2.0 en el capítulo 3 (n=211). Todas estas variantes de VIH-1 subtipo no-B fueron caracterizadas de nuevo a través de REGA 3.0 y estas secuencias fueron confirmadas mediante análisis filogenético a través de Máxima Verosimilitud. Para la confirmación a través de análisis filogenético se utilizó un grupo de secuencias de referencia de subtipos puros (subtipos de la A a la K) y CRFs disponibles hasta el momento del análisis obtenido de HIV Los Alamos. Se estableció como límite de confianza una asociación con un valor de Bootstrap superior a 70% y todas las formas genéticas que no se asociaron con subtipos/CRFs conocidos, fueron caracterizados como formas recombinantes únicas (URFs) y su patrón de recombinación se estudió mediante análisis de Bootscan en REGA 3.0.

### **5.2.2 Búsqueda de secuencias similares a través de HIV-BLAST.**

Una vez caracterizadas todas las secuencias de VIH-1 no-B, se realizó una búsqueda de secuencias similares para las variantes de VIH-1 no-B con mayor relevancia epidemiológica (variantes encontradas en 5 pacientes o más) a través de HIV-BLAST (disponible en [http://www.hiv.lanl.gov/content/sequence/BASIC\\_BLAST/basic\\_blast.html](http://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html)). Estas secuencias similares fueron incorporadas a cada grupo de variantes de VIH-1 no-B, y estos grupos de secuencias fueron realineados a través de ClustalW(56). Como única excepción, se descargaron todas las secuencias disponibles para las variantes genéticas con un número poco representativo de secuencias en HIV Los Alamos.

### **5.2.3 Análisis de los perfiles filogenéticos de las variantes de VIH-1 no-B.**

#### **Análisis mediante Máxima Verosimilitud**

En primer lugar se realizó un análisis filogenético para cada grupo mayoritario de secuencias mediante Máxima Verosimilitud en RAxML a través de la plataforma CIPRES science Gateway(81). Para estos análisis utilizamos el modelo de sustitución nucleotídica GTR y 1000 réplicas de Bootstrap.

## **Análisis filogenético mediante Inferencia Bayesiana**

En segundo lugar se realizó una aproximación Bayesiana utilizando el método probabilístico de las cadenas de Markov y Montecarlo (MCMC) a través de BEAST v1.7.5. La escala evolutiva en los árboles filogenéticos fue realizada usando el modelo SRD06, adaptado al modelo de sustitución nucleotídica HKY. Se usó un reloj molecular relajado, utilizando el modelo de crecimiento BSP. Las Cadenas de Markov y Montecarlo fueron utilizadas para 250 millones de estados y muestreadas cada 50.000. Para este análisis se utilizó el programa TRACER v1.6 y se aceptaron únicamente las trazas con un tamaño efectivo de la muestra mayor a 200, después de excluir un 10% de los datos iniciales. Se estimó la tasa evolutiva ( $\mu$ , nucleotidos/sitio/año) y el tiempo del más reciente común antecesor (tMRCA) para las diferentes agrupaciones filogenéticas.

### **5.2.4 Estudio de las relaciones filogeográficas entre las variantes de VIH-1 subtipo no-B.**

Con la finalidad de caracterizar las relaciones filogenéticas existentes entre las variantes de VIH-1 no-B mayoritarias, estudiamos la presencia de linajes (grupos de secuencias que estuvieran relacionadas epidemiológicamente procedentes de diferentes países) y cluster de transmisión (agrupaciones filogenéticas encontradas exclusivamente entre pacientes de nuestra cohorte). Para ello, en el análisis filogenético mediante Máxima verosimilitud, establecimos como límite de confianza, secuencias relacionadas entre sí mediante nodos comunes con valores de Bootstrap confiables (>70%). Asimismo, estas agrupaciones filogenéticas fueron confirmadas a través de los árboles de Máxima Credibilidad de Clados (MCC) realizados mediante inferencia Bayesiana y estableciendo como límite de confianza, secuencias relacionadas con valores de probabilidad a posteriori significativos ( $pp > 0.9$ ).

### **5.2.5 Edición de los árboles filogenéticos**

Todos los árboles fueron visualizados y editados usando el programa FigTree v 1.4.0. En estos árboles, los cluster y linajes previamente definidos mediante ML y confirmados con valores de probabilidad a posteriori mayores de 0.9 fueron estudiados.

### **5.2.6 Estudio de mutaciones de Resistencia incluidas en los cluster de transmisión**

Se estudiaron todas las mutaciones presentes para los diferentes fármacos antiretrovirales según la lista de mutaciones de resistencia que establece la Organización Mundial de la Salud OMS para todas las secuencias analizadas en este capítulo y se interpretaron a través del algoritmo de resistencias de la Universidad de Stanford. Asimismo, se realizó un estudio pormenorizado de las mutaciones de resistencia incluidas dentro de los diferentes clusters de transmisión.

### **5.2.7 Análisis estadísticos**

Se realizó un análisis de regresión logística multivariante para determinar el efecto que las características demográficas, clínicas y virológicas tuvieron en la predicción de las diferentes formas genéticas. La significación estadística de estas variables clínico-epidemiológicas con respecto a la proporción total de los pacientes infectados, se estudió a través de un contraste de hipótesis mediante z-test. El análisis estadístico se realizó con el programa SPSS 22.0.

## 5.3 Resultados

### 5.3.4 Distribución epidemiológica de las variantes de VIH-1 subtipo B

De las 211 secuencias previamente caracterizados como formas genéticas VIH-1 subtipo no-B en REGA 2.0, 165 (23.8%) secuencias fueron caracterizados definitivamente como formas genéticas de VIH-1 subtipo no-B mediante análisis filogenético. La mayoría (n=104, [63%]) de estos pacientes estuvieron infectados con virus recombinantes en *pol*: 95 (57.6%) con 12 diferentes CRFs y 9 (5.5%) con URFs. El resto de pacientes (n=61, [37%]) fueron infectados con 5 subtipos puros (A1, D, C, F1 and G) (Figura 26).

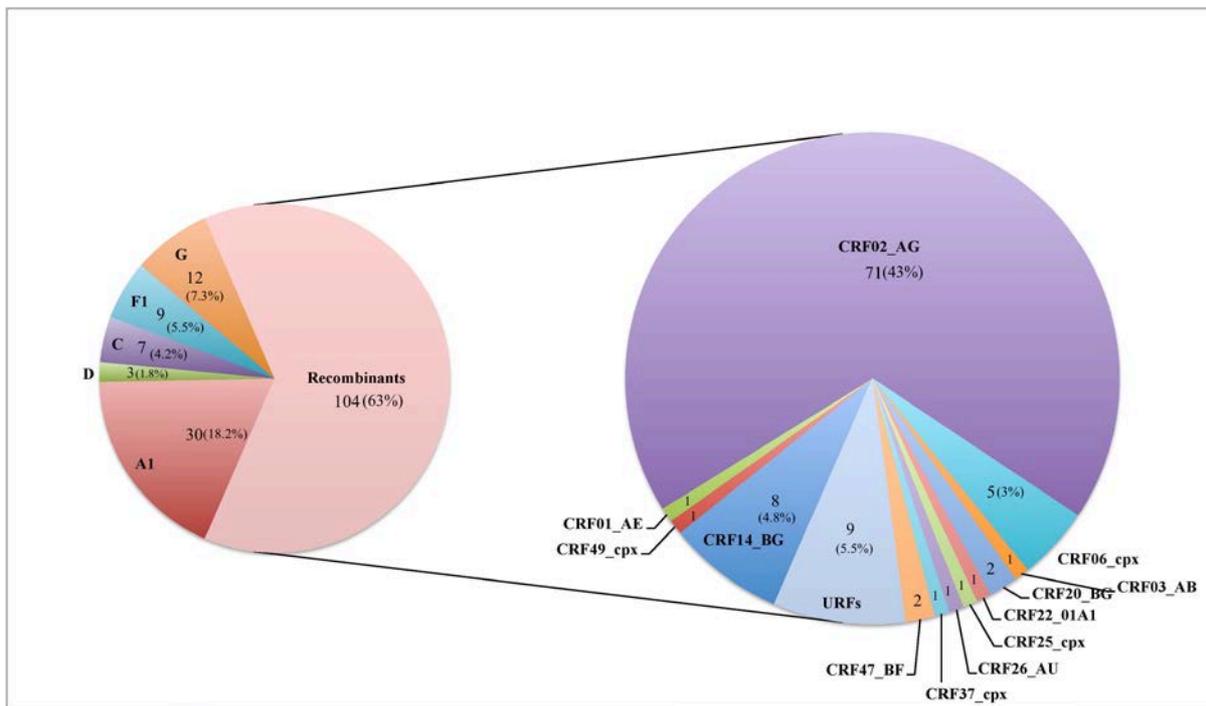


Figura 26. Distribución de las diferentes formas genéticas de VIH-1 subtipo no-B entre los pacientes infectados en Andalucía oriental durante el periodo 2005-2012.

Las características demográficas, clínicas y virológicas de los diferentes grupos de formas genéticas de VIH-1 no-B quedan reflejados en la tabla 9. La mayoría de los pacientes infectados con variantes no-B fueron hombres (63%,  $p < 0.001$ ), mayores de 35 años (73.5%,  $p < 0.001$ ), heterosexuales (92.2%,  $p < 0.001$ ) de nacionalidad Africana (58.2%,  $p < 0.001$ ) y que residían en la zona de El Ejido (62.4%,  $p < 0.001$ ). La mayoría de los pacientes infectados por formas no-B fueron extranjeros ( $n=126$ , [76.3%]) y procedentes de diferentes países: Argentina ( $n=1$ ), Brasil ( $n=4$ ), Burkina Faso ( $n=1$ ), Camerun ( $n=1$ ), Colombia ( $n=1$ ), Costa de Marfil ( $n=2$ ), Cuba( $n=1$ ), Ecuador( $n=1$ ), Gambia( $n=3$ ), Ghana( $n=12$ ), Guinea( $n=13$ ), Guinea-Bissau( $n=8$ ), Guinea-ecuatorial ( $n=7$ ), Lituania ( $n=1$ ), Mali( $n=11$ ), Marruecos( $n=2$ ), Mauritania ( $n=1$ ), Nigeria ( $n=17$ ), Republica-dominicana( $n=1$ ), Rumania ( $n=6$ ), Rusia( $n=14$ ), Senegal( $n=14$ ), Sierra Leona( $n=1$ ). El resto ( $n=37$ , [22.4%]) fueron pacientes de nacionalidad española.

Los análisis realizados mediante regresión logística multivariante resultaron únicamente significativos para el subtipo A: ser de género femenino (OR = 6.17,  $p = 0.026$ ) y de nacionalidad diferente a la Africana fueron características predictoras de presentar este subtipo (OR = 0.08,  $p = 0.008$ ).

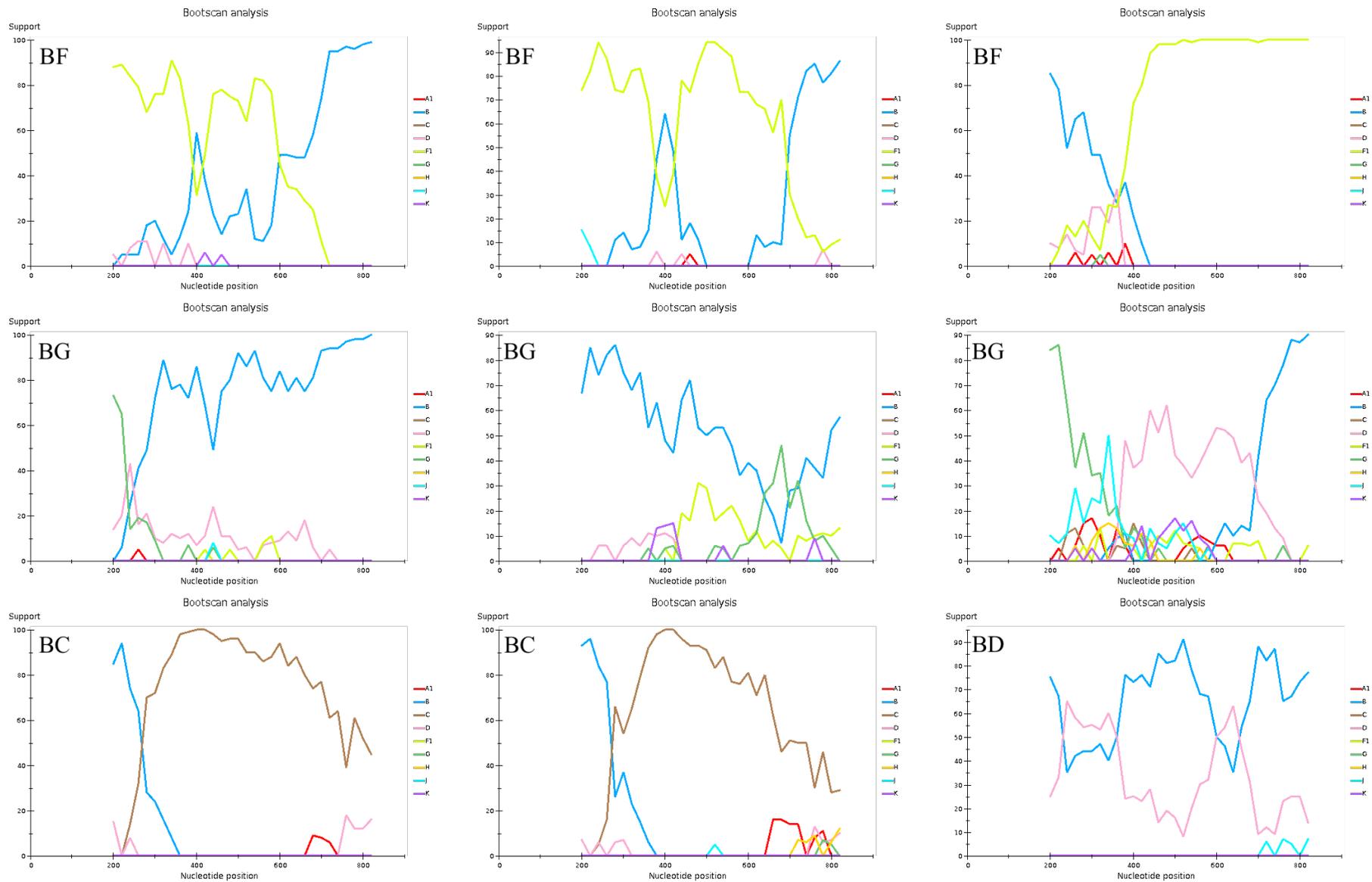
Hemos encontrado 23 casos de variantes del VIH-1 poco frecuentes en nuestra población (<3 pacientes infectados por forma genética) de las cuales 10 (44.4%) se encontraron entre pacientes de nacionalidad Española (tabla 10). Los patrones de recombinación para los diferentes URFs obtenidos mediante análisis de Bootscan se presentan en la figura 27.

CARACTERÍSTICAS	SUBTIPO PURO, n (%)				FORMAS RECOMBINANTES, n (%)				TOTAL, n (%)
	A1	C	F1	G	CRF02_AG	CRF14_BG	CRF06_cpx	Otros	
<b>Género</b>									
Hombre	10 (33.3)	4 (57.1)	8 (88.9)	9 (75)	48 (69)	6 (75)	3 (60)	15 (65.2)	104 (63)
Mujer	20 (66.7)	3 (42.9)	1 (11.1)	3 (25)	23 (31)	2 (25)	2 (40)	8 (34.8)	61 (37)
<b>Edad</b>									
<35	12 (40)	3 (42.9)	4 (50)	3 (25)	16 (23.2)		1 (20)	3 (15)	42 (26.4)
35-45	12 (40)	3 (42.9)	2 (25)	4 (33.3)	31 (44.9)	5 (62.5)	4 (80)	13 (65)	74 (46.5)
>45	6 (20)	1 (14.3)	2 (25)	5 (41.7)	22 (31.9)	3 (37.5)		4 (20)	43 (27)
<b>Factor de riesgo</b>									
HTX	30 (100)	7 (100)	8 (88.9)	12 (100)	61 (95.3)	6 (75)	3 (75)	14 (73.7)	141 (92.2)
HSH					1 (1.6)	1 (12.5)	1 (25)	5 (26.3)	8 (5.2)
UDVP			1 (11.1)		2 (3.1)	1 (12.5)			4 (2.6)
<b>Nacionalidad</b>									
Española	11 (36.7)		4 (44.4)	1 (8.3)	7 (9.9)	4 (50)		10 (47.6)	37 (22.7)
Europa del Este	15 (50)	2 (28.6)	2 (22.2)					2 (9.5)	21 (12.9)
Africa del norte					3 (4.2)				3 (1.8)
Africa occidental	2 (6.7)	3 (42.9)	1 (11.1)	11 (91.7)	55 (77.5)	4 (50)	5 (100)	3 (14.3)	84 (51.5)
Africa central	1 (3.3)				5 (7)			2 (9.5)	8 (4.9)
América central y sur	1 (3.3)	2 (28.6)	2 (22.2)		1 (1.4)			4 (19)	10 (6.1)
<b>Carga viral</b>									
<10000	8 (26.7)	1 (14.3)	1 (12.5)	1 (8.3)	15 (22.1)	1 (12.5)		5 (23.8)	32 (20.3)
10000-100000	10 (33.3)	2 (28.6)	1 (12.5)	6 (50)	24 (35.3)	5 (62.5)	2 (50)	9 (42.9)	59 (37.3)
>100000	12 (40)	4 (57.1)	6 (75)	5 (41.7)	29 (42.6)	2 (25)	2 (50)	7 (33.3)	67 (42.4)
<b>Recuento de CD4</b>									
<200	4 (16)		4 (50)	6 (50)	35 (52.2)	2 (25)	1 (20)	1 (5.3)	53 (35.6)
201-350	6 (24)	3 (60)	3 (37.5)		19 (28.4)	3 (37.5)	2 (40)	4 (21.1)	43 (28.9)
>350	15 (60)	2 (40)	1 (12.5)	3 (25)	13 (19.4)	3 (37.5)	2 (40)	14 (73.7)	53 (35.6)
<b>Fecha de diagnóstico</b>									
2005-2007	4 (13.3)	3 (42.9)		6 (50)	11 (15.5)	1 (12.5)		3 (13)	28 (17)
2007-2009	3 (10)	1 (14.3)	1 (11.1)	2 (16.7)	30 (42.3)	5 (62.5)	4 (80)	3 (13)	49 (29.7)
2009-2011	10 (33.3)	2 (28.6)	1 (11.1)	3 (25)	12 (16.9)	1 (12.5)	1 (20)	8 (34.8)	38 (23)
2011-2013	13 (43.3)	1 (14.3)	7 (77.8)	1 (8.3)	18 (25.4)	1 (12.5)		9 (39.1)	50 (30.3)

**Tabla 9. Características demográficas, clínicas y virológicas de los diferentes pacientes infectados con variantes de VIH-1 subtipo no-B en Andalucía oriental de 2005 a 2012.**

Paciente	Forma genética	Género (M/F)	Factor de riesgo	Edad	Fecha de diagnóstico	Recuento de CD4	Carga Viral	País de origen
1	CRF01_AE	M	HTX	38	2008	764	421	España
2	CRF03_AB	M	HSH	28	2011	715	404000	España
3	CRF20_BG	M	HTX	65	2005	521	5910	España
4	CRF20_BG	F	HTX	40	2005	476	77700	Cuba
5	CRF22_01A1	M	HTX	37	2009	474	11700	Guinea Equatorial
6	CRF25_CPX	F	ND	ND	2012	ND	36300	España
7	CRF26_AU	F	HTX	25	2012	15	1180000	España
8	CRF37_CPX	M	HTX	39	2007	405	135000	Guinea Ecuatorial
9	CRF47_BF	M	HTX	39	2009	389	12279	España
10	CRF47_BF	F	HTX	32	2011	392	4122	España
11	CRF49_CPX	M	HTX	36	2009	288	2250000	Senegal
12	D	F	HTX	39	2011	443	22233	Guinea Equatorial
13	D	M	HSH	48	2009	526	55723	España
14	D	M	HSH	41	2009	416	417477	Colombia
15	URF BC	F	HTX	53	2011	723	27911	Rumanía
16	URF BC	M	HTX	40	2011	450	211933	Rumanía
17	URF BF	M	HSH	38	2011	246	143039	España
18	URF BF	F	HTX	38	2012	ND	44100	Brasil
19	URF BF	M	HTX	52	2006	462	3200	Argentina
20	URF BG	F	ND	ND	2008	ND	ND	ND
21	URF BG	M	ND	ND	2009	ND	ND	ND
22	URF BG	M	HTX	38	2009	ND	ND	Guinea Bissau
23	URF BD	M	ND	49	2009	259	2800	España

**Tabla 10. Características demográficas, clínicas y virológicas de los pacientes infectados por formas genéticas poco frecuentes de VIH-1 subtipo no-B durante el periodo 2005-2012. ND: No descrito; M:Masculino; F:Femenino.**



**Figura 27. Análisis de los patrones de recombinación a través de Bootsca de los diferentes URFs encontrados en Andalucía oriental durante 2005 a 2012.**

### 5.3.5 Distribución geográfica de las diferentes formas genéticas de VIH-1 no-B.

La ubicación de los diferentes subtipos y formas recombinantes de VIH-1 subtipo no-B y su prevalencia durante el periodo 2005 a 2012 se representa en la figura 28. Como se muestra, la mayoría de los pacientes infectados con variantes de VIH-1 no-B residían en el área de El Ejido (62.4%) o Granada (22.4%). Encontramos muy poca representación de pacientes infectados con formas de VIH-1 no-B en Almería (10.9%), Jaén (2.4%) o Motril (1.8%).

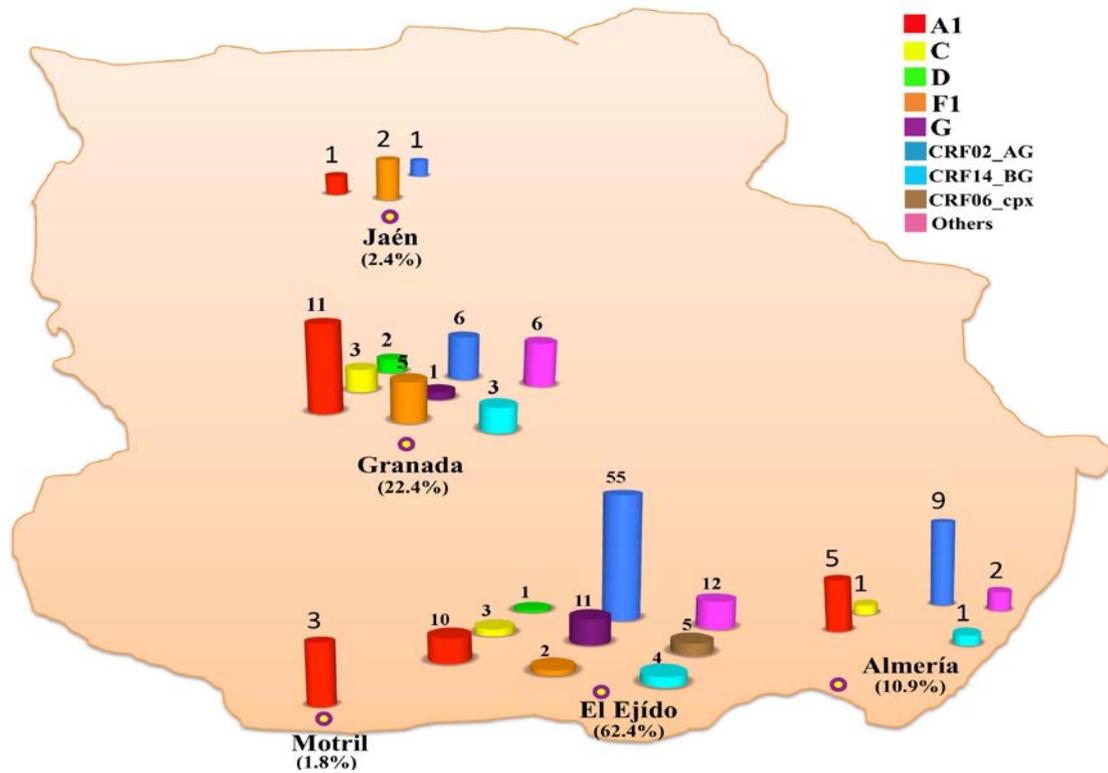


Figura 28. Distribución geográfica de los pacientes infectados con diferentes formas genéticas de VIH-1 subtipo no-B en Andalucía oriental desde 2005 a 2012.

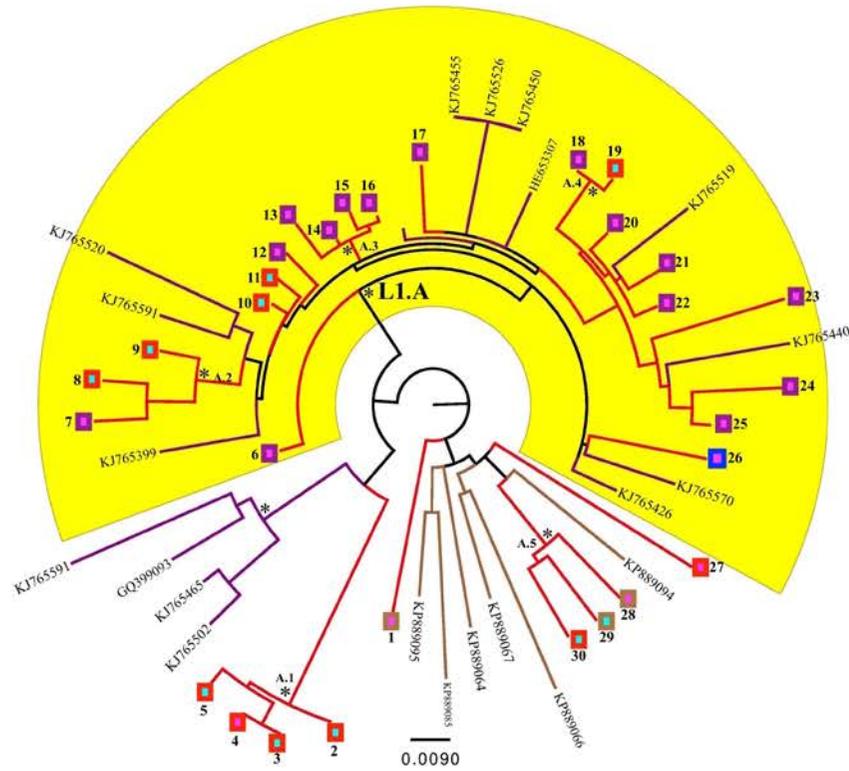
### **5.3.6 Caracterización filogenética de las las principales formas genéticas de VIH-1 subtipo no-B.**

Los análisis realizados a través de Máxima Versimilitud e Inferencia Bayesiana con el fin de caracterizar la relación filogenética de los pacientes infectados con las principales formas genéticas de variantes de VIH-1 no-B, mostraron la existencia en Andalucía Oriental de 6 linajes diferentes de virus circulantes en otros países (Tabla 12) y un total de 16 cluster de transmisión (Tabla 13) entre pacientes de nuestra cohorte, cluster generalmente de tamaño pequeño. La mayor proporción de pacientes que se agruparon en cluster se encontró entre los pacientes infectados con virus del subtipo A1 (53.3%, [16/33]) seguido por el CRF14\_BG ([4/8], 50%), C ([3/7], 43%), G ([5/12], 42%) y muy de lejos para la forma recombinante CRF02\_AG ([11/71], 14%). Como se muestra en la tabla 13, estos cluster se originaron en su mayoría en la década de los 2000, integrados principalmente por pacientes que residían en El Ejido o Granada. Asimismo, el bajo recuento de CD4 de los pacientes incluidos en la mayoría de estas redes de transmisión sugieren la presencia de una alta proporción de pacientes con un diagnóstico tardío de VIH.

### **Subtipo A1:**

Los análisis filogenéticos mediante Máxima Verosimilitud e inferencia Bayesiana realizados en el subtipo A1 (Figura 28), mostraron relación epidemiológica con secuencias recogidas en Europa del Este (Bulgaria) y Africa central (República democrática del Congo). Se encontró un linaje de especial importancia en nuestra población (L1.A). Este linaje relacionado con virus circulantes en Europa del este se originó en el año 1997.4 (95%CI: 1990.5-1999.3) e incluía 21 pacientes: 15 mujeres de Europa del Este (14 mujeres de Rusia y 1 de lituania) y 5 hombres españoles que se agruparon con 9 secuencias recogidas en Europa del Este. Como dato interesante, remarcamos la presencia del cluster (A.3) dentro de este linaje, cluster formado únicamente por mujeres de nacionalidad rusa, dato que podría sugerir la existencia de una red de transmisión de mayor tamaño, en la que probablemente estarían incluidos hombres no diagnosticados. Cuatro secuencias de pacientes españoles, residentes en el área de Granada se agruparon en un cluster (A.4) cuyo origen se estimó en 2008.5 (95%CI: 2006.6-2010.3) en el que hemos detectado la mutación K103N. Este cluster tuvo también relación epidemiológica con virus circulantes en Europa del Este. Asimismo, a diferencia de la mayoría de cluster de variantes de VIH-1 no-B, estos pacientes presentaron un elevado recuento de CD4 [590 (534-701)]. La corta longitud entre las ramas del cluster de transmisión indicó un intervalo corto de tiempo entre las infecciones producidas en los diferentes pacientes, pacientes que fueron diagnosticados todos en 2011. Finalmente, 5 secuencias: 2 pacientes de Africa occidental (Mali), 1 paciente de Africa central (Guinea Ecuatorial) y 2 pacientes españoles, se agruparon con secuencias procedentes de Africa.

A)



B)

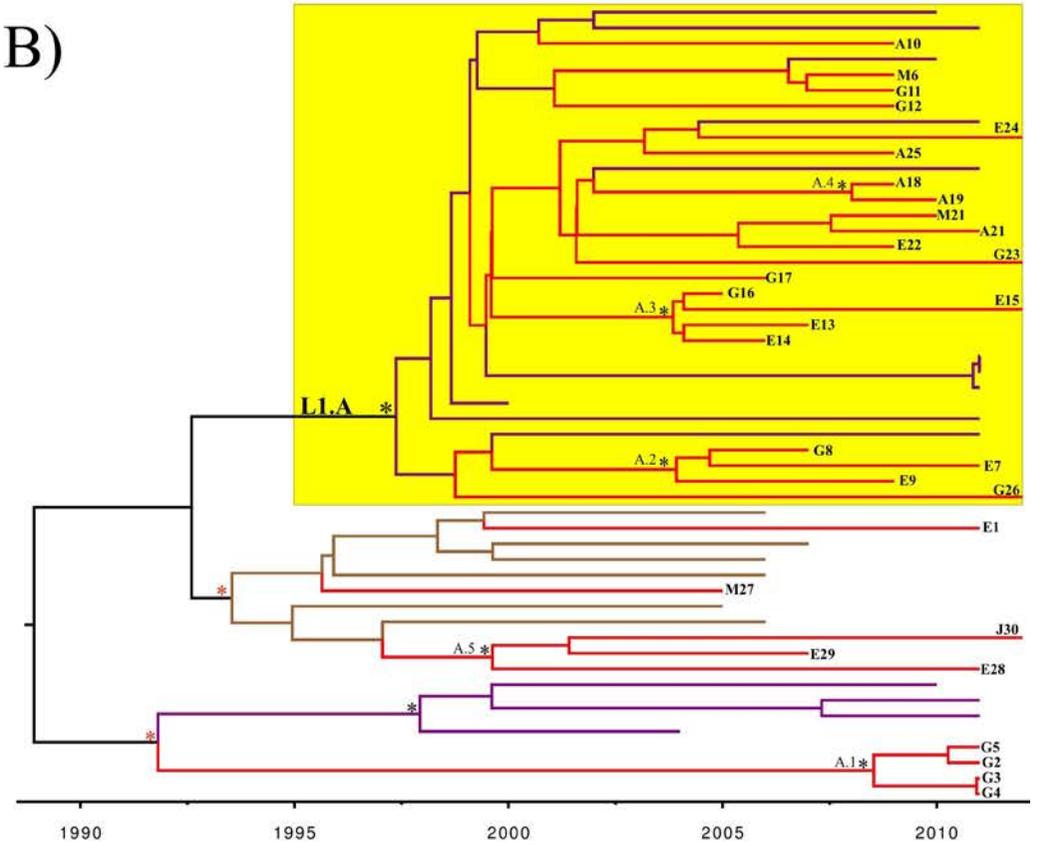
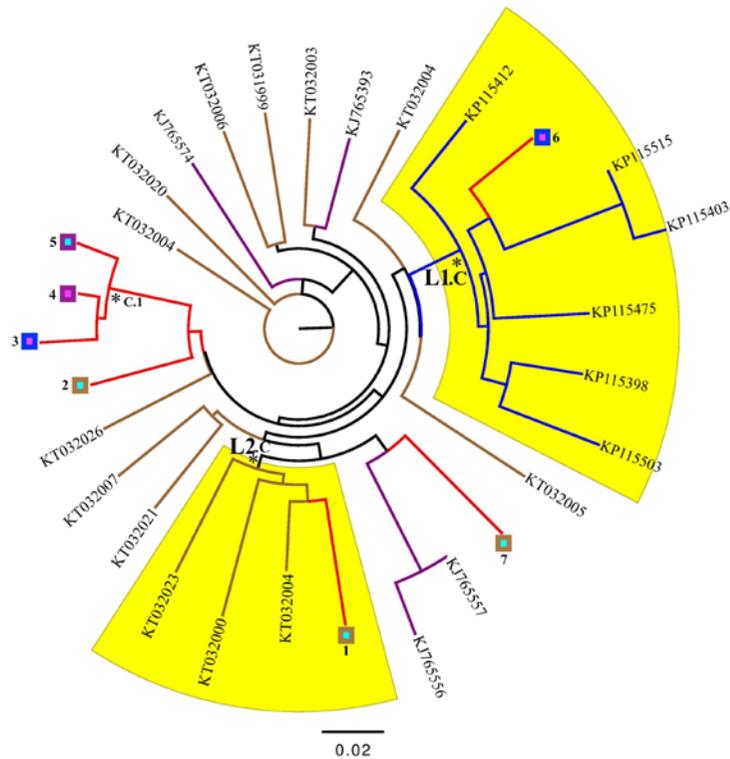


Figura 28. Árboles de Máxima Verosimilitud (ML) (A) y de Máxima Credibilidad de Clados (MCC) (B) realizados para las secuencias de subtipo A. Las ramas de color rojo corresponden con secuencias muestreadas en Andalucía. Las ramas violetas y marrones corresponden con secuencias de Europa del Este y África respectivamente. Los nodos con alto soporte estadístico quedan indicados bajo un asterisco negro y los soportados para una sola topología de árbol filogenético en rojo. Los linajes encontrados quedan resaltados en amarillo. Las etiquetas de las ramas del árbol de ML (A) muestran en la porción interna el género (Azul: Masculino y Rosa: Femenino) y en la porción externa la nacionalidad de los pacientes (Azul: America del sur, Rojo: Española, Marrón: Africana y Violeta: Europa del Este). La localización geográfica de los pacientes se muestra en las etiquetas de las ramas del árbol de MCC (B) ( A: Almería, E: El Ejido, M: Motril, G:Granada, J:Jaén). Todas las secuencias están identificadas mediante un código numérico interno y las secuencias de otros países con código de acceso en GenBank (Accesion ID).

### **Subtipo C:**

Encontramos dos vías de entrada para el subtipo C en nuestra zona: Sudáfrica y Brasil (Figura 29). Un paciente de nacionalidad brasileña se agrupó con secuencias procedentes de Brasil (Linaje L1.C) y otro paciente de nacionalidad sudafricana con secuencias procedentes de Sudafrica y Somalia (Linaje L2.C). Como nota de interés, encontramos en la literatura que algunas secuencias procedentes de Europa del Este (Bulgaria), y que agrupaban con el linaje Africano, estaban atribuidas a pacientes con una posible infección en Sudáfrica (137).

A)



B)

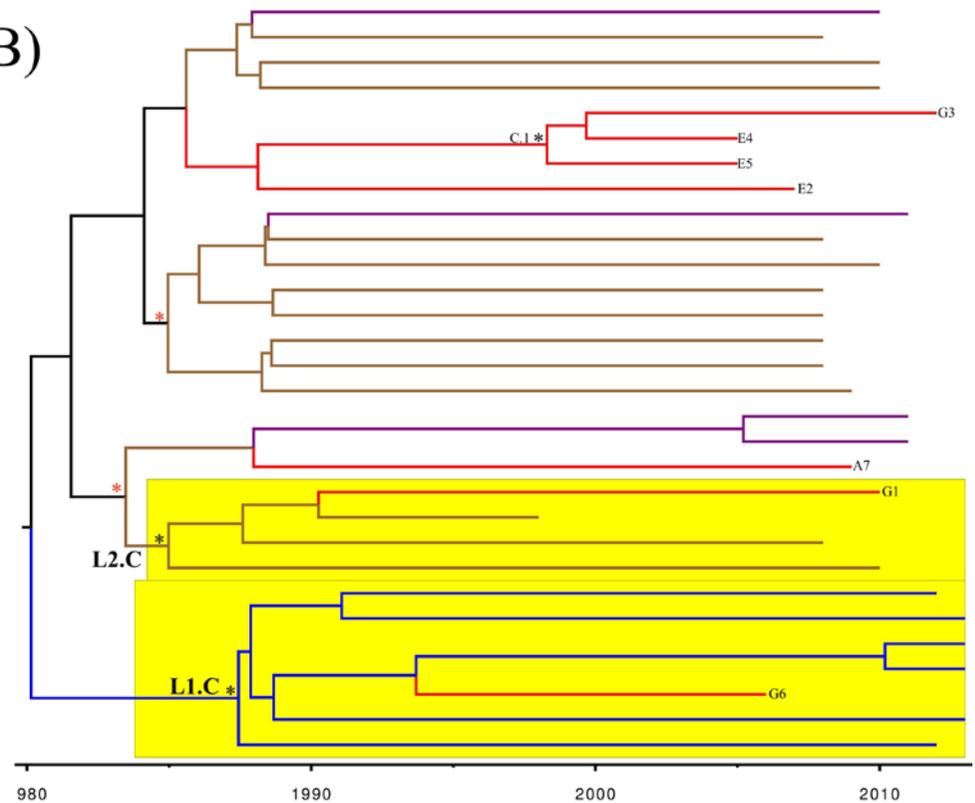


Figura 29. Árboles de Máxima Verosimilitud (ML) (A) y de Máxima Credibilidad de Clados (MCC) (B) realizados para las secuencias de subtipo C. Las ramas de color rojo corresponden con secuencias muestreadas en Andalucía. Las ramas violetas, marrones y azules corresponden con secuencias de Europa del Este, Africa y America del sur respectivamente. Los nodos con alto soporte estadístico quedan indicados bajo un asterisco negro y los soportados para una sola topología de árbol filogenético en rojo. Los linajes encontrados quedan resaltados en amarillo. Las etiquetas de las ramas del árbol de ML (A) muestran en la porción interna el género (Azul: Masculino y Rosa: Femenino) y en la porción externa la nacionalidad de los pacientes (Azul: America del sur, Rojo: Española, Marrón: Africana y Violeta: Europa del Este). La localización geográfica de los pacientes se muestra en las etiquetas de las ramas del árbol de MCC (B) (A: Almería, E: El Ejido, M: Motril, G:Granada, J:Jaén). Todas las secuencias están identificadas mediante un código numérico interno y las secuencias de otros países con código de acceso en GenBank (Accesion ID).

## **Subtipo F:**

El subtipo F1 mostró similitud genética con secuencias recogidas en Brasil, África Central (República democrática del Congo) y Europa del Este (Bulgaria) (Figura 6). No encontramos grandes asociaciones entre las secuencias de este subtipo, únicamente 2 parejas: una pareja heterosexual (cluster F.1) de pacientes de nacionalidad rumana, que se originó en 1997.4 (95%CI: 1987-2003.4) y que se agrupó con secuencias de Europa del Este y el cluster F.2 originado en 2010.2 (95%CI: 2010-2011) formado por dos hombres usuarios de drogas por vía parenteral, de nacionalidad Brasileña y Española que se agrupó con secuencias procedentes de Brasil.

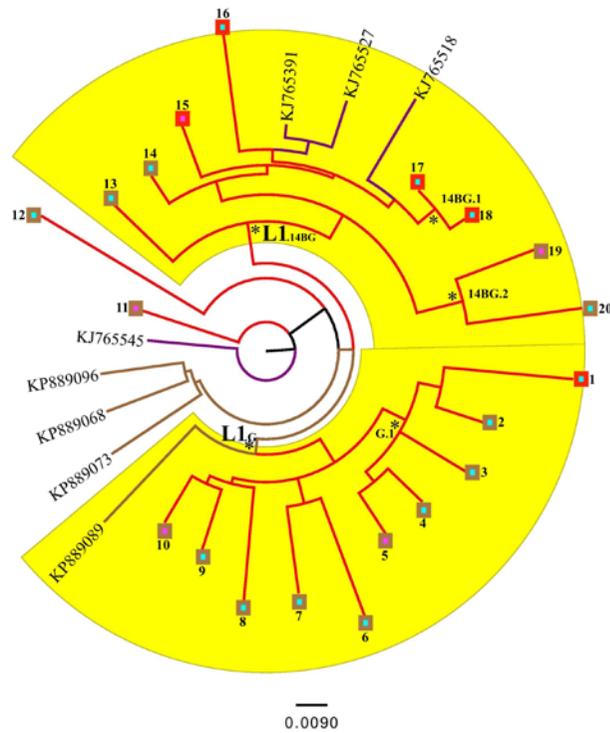


### **Subtipos G y CRF14\_BG:**

El subtipo G se relacionó con virus circulantes en África Central (República democrática del Congo) (Figura 31). Doce secuencias de subtipo G se agruparon con una secuencias de África Central (República democrática del Congo) y fueron encontradas entre pacientes procedentes de países vecinos (Senegal, Mali, Nigeria, Ghana y Guinea). Diez de estos pacientes se agruparon en un linaje (L1.G) que se originó en 1989 (95%CI: 1986-1998.5) y cinco de los pacientes infectados con virus procedentes de este linaje formaron un cluster (G.1) en 1996.2 (95%CI: 1995-2001.7) en el que se detectó la mutación M184V y M179E.

Ocho pacientes se agruparon junto a secuencias de la variante CRF14\_BG muestreadas en Bulgaria para el fragmento de subtipo G en *pol*. Cabe destacar que probablemente estos pacientes habrían adquirido la infección en España (137). Asimismo, a pesar de que los análisis filogenéticos mediante ML no fueron tan esclarecedores, se observaron dos agrupaciones filogenéticas a través de los análisis realizados mediante Inferencia Bayesiana; el clado I (cluster 14BG.2) que se originó en 1991 (1995.3-2003.8) que incluía a dos pacientes de nacionalidad guineana y el Clado II en 1995.1 (95% CI:1993.5-2001.3) formado principalmente por pacientes españoles. Estos datos sugieren que la variante CRF14\_BG fue introducida probablemente en nuestra región a través de pacientes procedentes de Guinea.

A)



B)

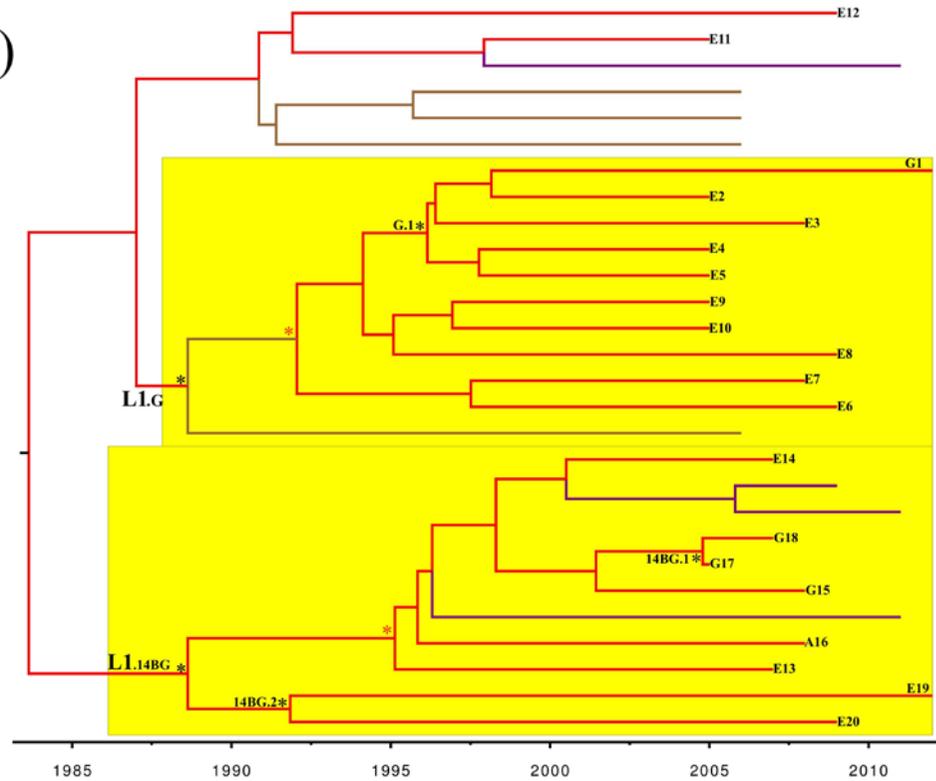


Figura 31. Árboles de Máxima Verosimilitud (ML) (A) y de Máxima Credibilidad de Clados (MCC) (B) realizados para las secuencias de subtipo G y CRF14\_BG. Las ramas de color rojo corresponden con secuencias muestreadas en Andalucía. Las ramas violetas y marrones corresponden con secuencias de Europa del Este y Africa respectivamente. Los nodos con alto soporte estadístico quedan indicados bajo un asterisco negro y los soportados para una sola topología de árbol filogenético en rojo. Los linajes encontrados quedan resaltados en amarillo. Las etiquetas de las ramas del árbol de ML (A) muestran en la porción interna el género (Azul: Masculino y Rosa: Femenino) y en la porción externa la nacionalidad de los pacientes (Azul: America del sur, Rojo: Española, Marrón: Africana y Violeta: Europa del Este). La localización geográfica de los pacientes se muestra las etiquetas de las ramas del árbol de MCC (B) ( A: Almería, E: El Ejido, M: Motril, G:Granada, J:Jaén). Todas las secuencias están identificadas mediante un código numérico interno y las secuencias de otros países con código de acceso en GenBank (Accesion ID).

### **Subtipos CRF02\_AG y CRF06\_cpx:**

No encontramos relaciones epidemiológicas relevantes entre los pacientes infectados con las formas recombinantes CRF02\_AG y CRF06\_cpx (Figura 32 y 33). En el caso de la forma CRF02\_AG encontramos 5 cluster de tamaño pequeño. Asimismo, estas secuencias mostraron similitud genética con otras recogidas en República democrática del Congo y Bulgaria, pacientes en su mayoría procedentes de diferentes países de Africa Occidental. Por otro lado, seis pacientes infectados con la variante CRF06\_cpx que procedían también de diferentes países de Africa Occidental (Nigeria, Ghana y Senegal) se agruparon con otras secuencias de países vecinos (Burkina Faso, Togo y Nigeria).

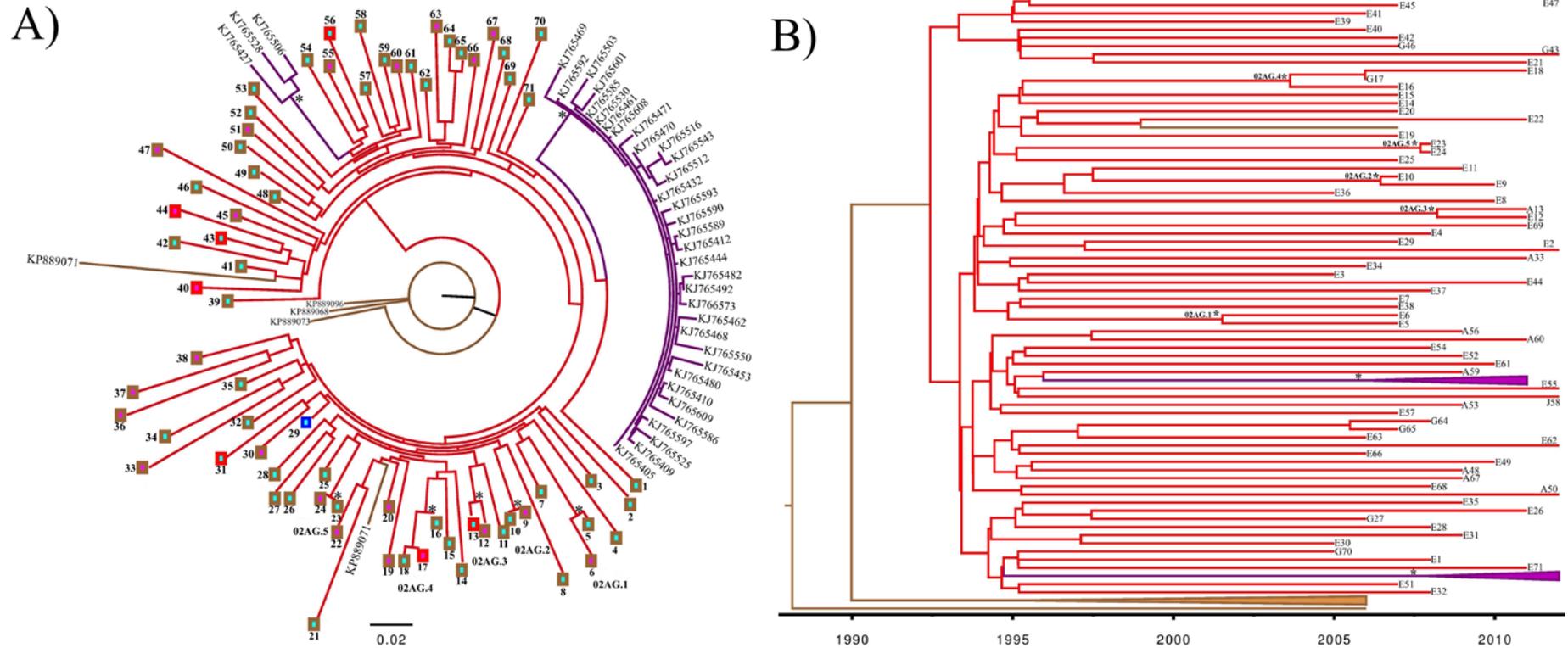


Figura 32. Árboles de Máxima Verosimilitud (ML) (A) y de Máxima Credibilidad de Clados (MCC) (B) realizados para las secuencias de subtipo G y CRF14\_BG. Las ramas de color rojo corresponden con secuencias muestreadas en Andalucía. Las ramas violetas y marrones corresponden con secuencias de Europa del Este y Africa respectivamente. Los nodos con alto soporte estadístico quedan indicados bajo un asterisco negro y los soportados para una sola topología de árbol filogenético en rojo. Las etiquetas de las ramas del árbol de ML (A) muestran en la porción interna el género (Azul: Masculino y Rosa: Femenino) y en la porción externa la nacionalidad de los pacientes (Azul: America del sur, Rojo: Española, Marrón: Africana y Violeta: Europa del Este). La localización geográfica de los pacientes se muestra en las etiquetas de las ramas del árbol de MCC (B) (A: Almería, E: El Ejido, M: Motril, G:Granada, J:Jaén). Todas las secuencias están identificadas mediante un código numérico interno y las secuencias de otros países con código de acceso en GenBank (Accesion ID).



Linajes	Secuencias (ES;OP)	Nacionalidad (Nº de pacientes)	Intervalo de muestreo	Soporte filogenético (BT:PP)	Países con secuencias similares (Código de acceso en GenBank)	tMRCA (95% HPD)	Tasa evolutiva (x 10 <sup>-3</sup> )
<b>Subtipo A</b>							
L1.A	21;12	Española (5), Rusa (14), Lituana(1) Dominicana(1)	2005-2012	99;0,9	Bulgaria (KJ765399,KJ765591,KJ765520, KJ765455,KJ765526,KJ765450,KJ765519, KJ765440, KJ765570,KJ765426) Polonia (KM057339) Croacia (HE653307)	1997.4(1990.5-1999.3)	3.2(0.7-6.7)
<b>Subtipo C</b>							
L1.C	1;6	Brasileña (1)	2006-2013	99;0,9	Brasil (KP115412, KP115515, KP115403, KP115475, KP115398, KP115503).	1987.4(1972-1995.7)	1.9 (1-3)
L2.C	1;3	Sudafricana (1)	1998-2010	75;0,9	Sudafrica (KT032000, KT032023) Somalia (AF396682).	1985(1973-1993)	1.9 (1-2.8)
<b>Subtipo F</b>							
L1.F	1;3	Española (1)	2011-2014	70;0,9	Brasil(KP115536, KP115390, KP115546)	1993.9(1983.5-2002.8)	2.2 (1.1-3.3)
<b>Subtipo G</b>							
L1.G	10;1	Española (1), Nigeriana (5), Mali(1), Guinea Bissau(1), Ghana(2).	2005-2012	70;0,9	Congo (KP889089)	1989(1986-1998.5)	3.1(1.7-4.6)
<b>Forma CRF14_BG</b>							
L1.14BG	8;3	Spañola (4), Guineana (4)	2005-2012	98;0,9	Bulgaria (KJ765391, KJ765518, KJ765527)	1989(1986-1996)	3.1(1.7-4.5)

**Tabla 12. Diferentes linajes de virus circulantes encontrados en Andalucía oriental durante el periodo 2005-2012. SP: Secuencias muestreadas en España; OP; secuencias muestreadas en otros países.**

Cluster (ID)	Nº de pacientes	Intervalo de muestreo	Soporte filogenético (BT:PP)	Factor de riesgo	Nacionalidad de los pacientes	Carga viral (mediana, IQR) Log <sub>10</sub>	Recuento de CD4 (mediana, IQR)	Área geográfica	tMRCa (95% HPD)	Mutaciones de resistencia
<b>Subtipo A</b>										
<b>A.1</b>	4	2011	100;0.9	HTX	Española	5 (4.6-6)	590 (534-701)	Granada	2008.5 (2006.6-2010.3)	K103N
<b>A.2</b>	3	2007-2011	96;0.9	HTX	Española-Rusa	4.7 (4.4-5.3)	301 (152-517)	Granada-El Ejido	2003.9 (1998.4-2004.7)	-
<b>A.3</b>	4	2005-2012	81;0.9	HTX	Rusa	4.5 (4.4-4.7)	383 (344-446)	Granada-El Ejido	2003.8 (2001.4-2004.8)	-
<b>A.4</b>	2	2009-2010	100;0.9	HTX	Española-Rusa	4.9 (4-5)	391 (325-456)	El Ejido	2008 (2007.4-2009)	-
<b>A.5</b>	3	2007-2012	98;0.9	HTX	Española-Mali	5.7 (5.5-5.8)	23 (15-125)	El Ejido-Jaén	1999.6 (1995.3-2004.4)	-
<b>Ssubtipo C</b>										
<b>C.1</b>	3	2005-2012	100;0.9	HTX	Rumana-Brasileña	5.2 (5.1-5.3)	499 (253-739)	Granada-Ejido	1998.3 (1992.3-2001.9)	-
<b>Subtipo F</b>										
<b>F.1</b>	2	2007-2010	100;0.9	UDVP	Rumana	5.3(5.1-5.4)	179 (157-200)	Granada-Almería	1997.4 (1987-2003.4)	-
<b>F.2</b>	2	2011-2012	90;0.9	HTX	Brasileña	5.7(5.4-5.9)	62 (50-73)	Jaén	2010.2 (2010-2011)	-
<b>Subtipo G</b>										
<b>G.1</b>	5	2005-2012	88;0.9	HTX	Española-Nigeriana-Ghana	4.5(4.3-5.3)	244 (173-364)	Granada-El Ejido	1996.2 (1995-2001.7)	M184V, M179E
<b>Forma CRF14_BG</b>										
<b>14BG.1</b>	2	2005-2007	100;0.9	HTX	Española	5	425 (267-583)	Granada	2004.4 (2003.8-2005)	-
<b>14BG.2</b>	2	2009-2012	97;0.9	HTX	Guineana	4.7(2-5)	186 (138-234)	El Ejido	1991.8 (1995.3-2003.8)	-
<b>Forma CRF02_AG</b>										
<b>02AG.1</b>	2	2007	100;0.9	HTX	Guineana	5.6(4-6)	309 (180-438)	El Ejido	2003.8 (2001-2006)	-
<b>02AG.2</b>	2	2007-2010	100;0.9	HTX	Senegalesa	3(4-5)	391 (381-400)	El Ejido	2006.4 (2006-2007)	-
<b>02AG.3</b>	2	2011	99;0.9	HTX	Española-Marroquí	5.9(3-6)	326 (100-551)	Granada-Almería	2008.2 (2007.2-2009.9)	-
<b>02AG.4</b>	3	2006-2011	98;0.9	HTX	Española-Mali	5.5(5.3-6.2)	140 (78-306)	El Ejido	2003.6 (2000.3-2005.3)	-
<b>02AG.5</b>	2	2008	100;0.9	HTX	Ghana	6	126 (45-200)	El Ejido	2007.7 (2007.2-2008)	-

**Tabla 13. Características demográficas, clínicas y virológicas de los pacientes encontrados en los diferentes cluster de transmisión de las principales formas genéticas de VIH-1 subtipo no-B.**

## 5.4 Discusión

En Andalucía Oriental, la mayoría de las formas genéticas de VIH-1 subtipo no-B, se encuentran entre la población migrante heterosexual, fundamentalmente en hombres africanos, o mujeres de países de Europa del Este. Estos pacientes residen generalmente en El Ejido, actuando esta zona como puerta de entrada para la mayor parte de subtipos a otras zonas de Andalucía oriental.

En este capítulo hemos investigado la presencia de cadenas de transmisión de VIH-1 subtipo no-B añadiendo secuencias similares, y procedentes de otros países. Como en otros estudios (138), demostramos la necesidad de incorporar estas secuencias para el estudio de la transmisión entre pacientes que con alta probabilidad podrían no encontrarse relacionados epidemiológicamente ya que estos datos apoyarían grupos monofiléticos procedentes de pacientes de diferentes áreas geográficas. En este contexto, el análisis filogeográfico realizado para las variantes de VIH-1 subtipo no-B en nuestra región, muestra la presencia de una alta proporción de pacientes migrantes que se relacionan con virus circulantes próximos a su país de procedencia, datos que sugieren un gran número de pacientes infectados en su país de origen. Asimismo, como muestran otros estudios realizados para la variante CRF14\_BG, su relación epidemiológica se vincula a virus circulantes en Portugal y España (139,140). Debido a que no hemos encontrado secuencias similares procedentes de estos países, estos datos podrían sugerir una vía de entrada diferente en nuestra zona y una posterior propagación de esta forma genética entre pacientes Españoles.

La prevalencia de variantes de VIH-1 no-B en Andalucía oriental, se ha mantenido estable ( $\approx 23\%$ ) en los últimos años (129) y está en la línea a la recogida en otros estudios realizados en regiones próximas (141) aunque sigue siendo mucho mayor a la reportada en otros estudios realizados en otras zonas de España (86,87). Asimismo, hemos demostrado un aumento tanto de la prevalencia de las variantes no-B de VIH-1 como de su diversidad genética: un 22% de los pacientes nativos estuvieron infectados con formas genéticas no-B entre 2005 y 2012 frente a un 12.8% reportado en otros estudios previos entre 1983 y 2001 (129) en esta misma zona. Además, hemos detectado 12 CRFs diferentes y 9 URFs en esta

zona, posiblemente relacionado con el aumento de la tasa de migración en el sur de España en la última década (141,142). Sin embargo, estas variantes no-B también fueron detectadas en pacientes españoles, que fueron incluidos en un alto número de clusters no-B (8 de 16 cluster de transmisión). Asimismo, 10 de los 23 pacientes (43%) infectados por variantes no-B poco frecuentes eran de nacionalidad Española.

La mayor prevalencia en España de variantes de VIH-1 diferentes al subtipo B sigue siendo para la forma recombinante CRF02\_AG(87), sin embargo existió una baja proporción de cluster de transmisión entre pacientes que estaban infectados con esta forma genética, datos que irían en la misma línea a los reportados en otras regiones de España, que indican la emergencia de cluster de transmisión en otros subtipos diferentes a el CRF02\_AG y por consiguiente un aumento en la prevalencia de estos subtipos en los últimos años (19,143,144). La presencia de grandes linajes de VIH-1 no-B (portadores a su vez de mutaciones de resistencia), sugieren ciertos comportamientos de alto riesgo: dentro del subtipo A1, encontramos un colectivo de riesgo especialmente vulnerable de mujeres trabajadoras del sexo y sus clientes (linaje L1.A1) que aparece a finales de la década de los noventa, formado principalmente por mujeres migrantes de nacionalidad rusa y en menor proporción por hombres españoles, que provocaron el origen de diferentes redes de transmisión durante la década de los 2000. Además, en este subtipo encontramos el cluster A.4 en el que se detectó resistencia de alto nivel a nevirapina y efavirenz. En segundo lugar, dentro del subtipo G, demostramos la existencia de un linaje (L1.G1), en el que detectamos un cluster de 5 pacientes (G.1), algunos de estos pacientes a su vez presentaban resistencia a los fármacos antiretrovirales de primera línea emtricitabina y lamivudina, y resistencia de bajo nivel a abacavir.

Como se demuestra en el capítulo anterior y en otros estudios similares realizados en países vecinos (88,145), existen diferencias epidemiológicas importantes entre los pacientes infectados con formas genéticas de VIH-1 subtipo B y los pacientes infectados con variantes no-B. La mayoría de los cluster de transmisión de VIH-1 se originaron en la década de los 90 entre pacientes homosexuales infectados con el subtipo B. En este contexto, la epidemia se mantuvo estable gracias a la aparición de los primeros fármacos antiretrovirales de alta eficacia (TAR) en 1996 en España. En contraposición, aportamos datos que demuestran una historia evolutiva diferente en las formas genéticas de VIH-1 no-B en nuestra población. La mayoría de estos cluster de subtipos no-B se originaron durante la década de los 2000 entre la

población migrante y heterosexual, década en la que la migración en España aumentó considerablemente. Estos pacientes presentaron signos clínicos y virológicos de infecciones tardías, por lo que estos datos, podrían sugerir un aislamiento de este colectivo social y culturalmente muy heterogéneo, tanto de las primeras campañas de prevención realizadas en España, como de la terapia antiretroviral.

A diferencia de otros estudios publicados en España, no hemos encontrado cluster de variantes no-B entre HSH. La emergencia en España de cluster de transmisión entre HSH sugiere en un futuro, un cambio importante en la prevalencia para determinados subtipos: Thomsom et al (136) describen recientemente la existencia de un cluster de rápida propagación entre hombres que tienen sexo con hombres para el subtipo F en Galicia y Delgado et al (146) describen un cluster de subtipo A, que se está transmitiendo entre individuos de diversas áreas de España. Finalmente, Patiño et al (147) alertan por primera vez de la emergencia de la variante CRF19\_cpx entre la población HSH

En definitiva, creemos que dado que el tamaño de los cluster parece estar relacionado con su expansión en el tiempo (90) cabría esperar un mayor tamaño de algunas cadenas de transmisión de las variantes no-B en los próximos años y una mayor inclusión de individuos españoles entre sus integrantes. En este contexto, la existencia de una vigilancia epidemiológica continua y las herramientas de análisis filogenético en nuestra población, son especialmente importantes en la prevención de futuros brotes de variantes genéticas de VIH-1 subtipo no-B.

## **Capitulo 6: DISCUSIÓN FINAL**



El trabajo realizado en esta tesis doctoral ha sido enfocado al estudio de la historia evolutiva del VIH-1 en Andalucía oriental a través de diferentes técnicas de análisis filogenético. Aprovechando el nivel elevado de cambios evolutivos que se produce en el VIH y la gran fuente de información epidemiológica que proporcionan las secuencias de gen *pol* para el estudio de resistencias a antiretrovirales, las herramientas de análisis filogenético han demostrado tener especial importancia en la obtención de información acerca del origen del VIH, su historia epidemiológica, los patrones de migración del VIH y las rutas de infección y de transmisión del VIH entre pacientes relacionados epidemiológicamente.

Dado que el tamaño de la secuencia genética está directamente relacionado con el apoyo filogenético (148), siendo recomendable un tamaño superior a 800pb para dar una suficiente señal filogenética y desarrollar una topología acertada de árbol filogenético. En este trabajo hemos demostrado que el uso de las secuencias derivadas de los estudios de resistencia a fármacos antiretrovirales aporta suficiente información para la realización de diversas filogenias robustas y el estudio de la historia evolutiva del VIH-1 en Andalucía oriental.

En esta tesis doctoral hemos utilizado diferentes métodos de análisis filogenético dado que la filogenia molecular de una población no puede ser reducida a la reconstrucción simple de una topología de un árbol filogenético, sino que un grupo de hipótesis deben ser evaluadas (149) y la comparación de diferentes topologías puede ayudarnos a recrear una historia evolutiva acertada. Asimismo, los análisis filogenéticos complejos mediante Inferencia Bayesiana han aportado información complementaria a los análisis filogenéticos clásicos siendo fundamentales para enmarcar en un contexto temporal la epidemia de VIH.

Hemos demostrado la importancia de algunas variables clínicas como variables predictoras del agrupamiento de los pacientes en cluster de transmisión (150): los pacientes que presentaron un recuento elevado de CD4 demostraron una tendencia significativa al agrupamiento en cluster de transmisión de gran tamaño, sin embargo, los pacientes con bajos recuentos de CD4 formaron cluster de menor relevancia epidemiológica. Estos datos sugieren que en Andalucía Oriental existe un importante número de pacientes con infecciones agudas/recientes con prácticas de alto riesgo, especialmente dentro del colectivo homosexual.

Asimismo, el empleo de técnicas de análisis filogenético para estudiar la filogeografía de una determinada región fue especialmente importante en un escenario donde existió una alta tasa

de migración, debido a que ésta afectó directamente en la epidemiología local introduciendo nuevas variantes de subtipos no-B(42). Asimismo, creemos que el estudio de los perfiles filodinámicos de los subtipos de VIH-1 podrían ayudarnos en la predicción de futuros brotes epidemiológicos de VIH (111).

En los últimos años están apareciendo una serie de recomendaciones de cómo estos estudios deben ser realizados e informados (47,151,152) por lo que la vigilancia de los eventos de transmisión, y la contribución a la epidemiología molecular por parte de los centros asistenciales en los que se determinan resistencias a los antirretrovirales, son fundamentales para alertar y generar esfuerzos colaborativos con los servicios de Medicina Preventiva y Enfermedades Infecciosas, y con las autoridades sanitarias, para elevar las medidas de contención en la propagación del virus. Asimismo, han sido desarrolladas un amplio abanico de propuestas basadas en el uso de diferentes modelos intervencionistas de comportamiento, principalmente en el colectivo de las mujeres trabajadoras del sexo y sus clientes(153) por lo que creemos que las estrategias de prevención para reducir el contagio por el virus del VIH-1 deberían ir orientadas a ciertos colectivos de alto riesgo(154).

A la luz de todos estos hechos, se hace de gran utilidad la información epidemiológica sobre la distribución de los diferentes subtipos de VIH-1 en las diferentes regiones geográficas del mundo, siendo necesarias para orientar a las autoridades sanitarias y a la comunidad científica internacional para que se establezcan unos adecuados programas de prevención del SIDA, y de desarrollo de vacunas efectivas contra el VIH-1.

En conclusión, en esta tesis presentamos datos que demuestran que el estudio epidemiológico a través diversas herramientas de análisis filogenético, la filogeografía y la filodinámica de los diferentes subtipos de VIH-1 en Andalucía Oriental se presenta como una fuente de información de gran utilidad para la vigilancia y control epidemiológico de la propagación del VIH, y que su utilización ayudaría a reforzar e implantar actuaciones eficaces para prevenir la diseminación del VIH.

## **Capítulo 7: CONCLUSIONES**

1. En Andalucía oriental, la mayoría de cluster de transmisión se dan entre parejas, fundamentalmente entre parejas de hombres que tienen sexo con hombres.
2. La presencia de resistencia a fármacos antiretrovirales de primera línea en pacientes que forman cluster de transmisión es poco frecuente y se da fundamentalmente entre pacientes que forman cluster de tamaño pequeño.
3. Los primeros cluster de transmisión de gran tamaño aparecieron entre usuarios de drogas por vía parenteral en las zonas de costa durante la década de los 80.
4. Los cluster de transmisión con mayor repercusión epidemiológica se han encontrado dentro del subtipo B, originándose todos durante la década de los 90 en Granada. Estos cluster estaban formados fundamentalmente por pacientes menores de 35 años, españoles, homosexuales y con un alto recuento de CD4.
5. De especial importancia en nuestra población, es la presencia de un cluster que ha incluido a 58 hombres que tienen sexo con hombres desde 2007 hasta el final de este estudio. Los datos clínicos de estos individuos sugieren una alta proporción de pacientes con infección aguda/reciente.
6. Los cluster de pacientes infectados con formas genéticas de VIH-1 subtipo no-B fueron de menor tamaño en comparación a los VIH-1 subtipo B y se originaron fundamentalmente en la década de los 2000 entre pacientes heterosexuales, extranjeros y con un recuento menor de CD4, década en la que la migración en el sur de España aumentó considerablemente.
7. El análisis de los perfiles filogenéticos del VIH-1 subtipo B muestra que los linajes ancestrales de este subtipo se originaron en Granada años antes de los primeros casos con

infección por VIH descritos en España. La epidemia se propagó hacia zonas de costa creciendo exponencial durante la década de los 80. Asimismo esta epidemia se estabilizó en la década de los 90, existiendo un nuevo crecimiento exponencial desde 2005 a 2008.

8. La situación geográfica de Andalucía oriental, la cercanía con el continente Africano y la alta tasa de migración en los últimos años en la zona de El Ejido, han desencadenado una alta prevalencia y diversidad genética de variantes de VIH-1 subtipo no-B en comparación con otras regiones de España. La mayoría de los pacientes infectados con variantes de VIH-1 no-B fueron hombres, mayores de 35 años, heterosexuales, de nacionalidad Africana y que residían en la zona de El Ejido .
9. La presencia de una alta proporción de pacientes migrantes que se relacionan con virus circulantes próximos a su país de origen sugieren un gran número de infecciones importadas.
10. Hemos encontrado procedencia importada en todas las variantes de VIH-1 subtipo no-B. El subtipo A fue importado principalmente a través de mujeres trabajadoras del sexo de Europa del Este. El subtipo C de virus circulantes en Sudáfrica y Brasil. El subtipo F1 de virus circulantes en Brasil, Africa central y Europa del Este. Finalmente las variantes G, CRF02AG y CRF06\_cpx fueron importadas a través de virus circulantes de Africa central y occidental.
11. Los análisis filogenéticos son herramientas valiosas y fundamentales en el diseño de campañas de prevención del VIH y en el desarrollo de actuaciones eficaces para el control de la propagación del virus.

## **ANEXO I: RECURSOS DE INTERNET**

### **-Listado de CFR de LANL**

<http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>

### **-Herramienta de HIV BLAST de LANL**

[http://www.hiv.lanl.gov/content/sequence/BASIC\\_BLAST/basic\\_blast.html](http://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html)

### **-Plataforma Science Gateway V. 3.3**

<https://www.phylo.org>

### **-Editor de secuencias PhyDE**

<http://www.phyde.de>

### **-Beast**

<http://tree.bio.ed.ac.uk/software/figtree/>

### **-Spread**

<http://www.kuleuven.be/aidslab/phylogeography/SPREAD.html>

### **-Google Earth**

<http://www.google.es/intl/es/earth/>

### **-Rega 2.0**

<http://www.bioafrica.net/rega-genotype/html/>

### **-Rega 3.0**

<http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/>

### **-HIVdb (Universidad de Stanford, EE. UU.)**

<http://hivdb.stanford.edu/>

### **-RIS (Red de investigación de SIDA en España)**

<http://www.retic-ris.net/>

### **-Find Model**

<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>



## ANEXO III: PARTICIPACIÓN EN CONGRESOS



El Comité Científico del XV Congreso de la Sociedad Andaluza de Enfermedades Infecciosas, celebrado en Jaén del 12 al 14 de diciembre de 2013, certifica que la comunicación titulada:

***CO-09 Canales de transmisión de VIH-1 en Andalucía oriental: aportaciones de la filogeografía***

Fue presentada en el mencionado congreso y en formato oral por los siguientes autores:

***Pérez-Parra S, Chueca N, Álvarez M, López-Ruz MA, Lozano A, Gálvez MC, Gea I, Martínez-Pérez MA, Hernández Quero J, García F.***

Y para que conste se expide el presente certificado en Jaén, a 14 de diciembre de 2013.



Sociedad Andaluza de  
Enfermedades Infecciosas  
Jesús Rodríguez Baño  
Presidente del Comité Científico

**CERTIFICA QUE:****S.PEREZ, N.CHUECA, M.ALVAREZ,  
T. CABEZAS, F.GARCIA****Han presentado en dicho congreso el  
POSTER con título:****“Filodinámica de VIH-1 durante el periodo  
2005-2012 en Andalucía Oriental”**Para que surta a los efectos oportunos, se firma el presente certificado  
en Córdoba, 16 de octubre de 2014

El Presidente del Comité Organizador



El Comité Científico del XVI Congreso de la Sociedad Andaluza de Enfermedades Infecciosas, celebrado en Almería del 11 al 13 de diciembre de 2014, certifica que la comunicación titulada:

**CO-07 Historia evolutiva y direccionalidad de la transmisión de VIH-1 en Andalucía Oriental.**

Fue presentada en el mencionado congreso y en formato oral por los siguientes autores:

**Pérez-Parra S, Chueca N, Álvarez M, Hernández Quero J, López Ruz M, Lozano AB, Collado A, Omar M, Palomares J, Delgado C, García F.**

Y para que conste se expide el presente certificado en Almería, a 13 de diciembre de 2014.



Sociedad Andaluza de  
Enfermedades Infecciosas

Antonio Rivero Román  
Presidente del Comité Científico



## CERTIFICADO DE PARTICIPACIÓN

El Presidente del Comité Organizador del **XIX Congreso SEIMC 2015**, en nombre de la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica

### CERTIFICA QUE:

**S. Pérez, N. Chueca, M. Alvarez, A. Collado, A.B. Lozano, D. Vinuesa, M.A. Lopez, M. Omar, J. Palomares, C. Delgado, F. García**

han realizado la presentación ORAL de la comunicación con título:

**“DESCRIPCIÓN DE UN CLUSTER VIH-1 SUBTIPO A1 DE MUJERES TRABAJADORAS DEL SEXO INMIGRANTES RUSAS Y HOMBRES ESPAÑOLES. ”**

en el XIX Congreso de la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica, celebrado en Sevilla, del 28 al 30 de mayo de 2015.

Y para que conste se expide el presente certificado en Sevilla a 30 de mayo de 2015.

**Dr. Álvaro Pascual Hernández**  
Presidente del Comité Organizador

**XXVIII REUNION DE LA SOCIEDAD ANDALUZA DE  
MICROBIOLOGIA Y PARASITOLOGIA CLINICA  
GRANADA, 5 Y 6 DE NOVIEMBRE 2015**

Certificamos que

Los Firmantes :

Santiago Pérez-Parra, Natalia Chueca-Porcuna, Marta Álvarez-Estevez, Josefa Lopez-Bueno,  
Antonio Sánchez, Maria Dolores Mérida, Jose-Angel Fernández-Caballero, Ana Belén Pérez,  
Federico García  
Servicio de Microbiología. Complejo Hospitalario Granada. Hospital Universitario San Cecilio.  
Instituto de Investigación Ibs.Granada.

Han presentado la Comunicación Poster Oral :

**Estudio de los perfiles filodinámicos y filogeográficos de la epidemia  
HIV-1 subtipo NO-B en Andalucía durante el periodo 2005 a 2012**

En la :

**XXVIII REUNION DE LA SOCIEDAD ANDALUZA DE MICROBIOLOGÍA Y  
PARASITOLOGÍA CLÍNICA.**

**RECONOCIDA DE INTERES CIENTIFICO SANITARIO POR LA CONSEJERIA DE SALUD DE  
LA JUNTA DE ANDALUCIA.**

**CELEBRADA EN EL HOTEL ABADES NEVADA PALACE DE GRANADA  
LOS DIAS 5 Y 6 DE NOVIEMBRE DE 2015**

Fdo. Dr. Federico García García  
Presidente Comité Organizador



VII CONGRESO NACIONAL **GeSIDA** 1-4 DICIEMBRE  
GRUPO DE ESTUDIO DEL SIDA-SEIMC **MADRID 2015**  
Y 9ª REUNIÓN DOCENTE DE LA RED DE INVESTIGACIÓN EN SIDA



## CERTIFICADO DE PARTICIPACIÓN

Los presidentes del Comité Organizador del  
**VII CONGRESO NACIONAL GeSIDA**

**CERTIFICAN QUE:**

**S. Pérez , N. Chueca, M. Álvarez, J. Pasquau, M. Omar, A. Collado, D.  
Vínuesa, A.B. Lozano, F. Garcia**

han presentado la comunicación **POSTER** con título:

**EPIDEMIOLOGÍA MOLECULAR, FILOGEOGRAFÍA Y FILODINÁMICA  
DEL VIH-1 SUBTIPO NO-B EN ANDALUCÍA ORIENTAL.**

En el **VII Congreso Nacional de GeSIDA**,  
celebrado en Madrid del 1 al 4 de diciembre de 2015.

Y para que conste se expide el presente certificado  
en Madrid, a 4 de diciembre de 2015.

A handwritten signature in blue ink, reading "J. BERENGUER", written over a horizontal line.

**Dr. Juan Berenguer**

Presidente VII Congreso Nacional Gesida

A handwritten signature in blue ink, reading "JER", written over a horizontal line.

**Dr. José Ramón Arribas**

Co-Presidente VII Congreso Nacional Gesida

Código Presentación: **P-091**

VII CONGRESO NACIONAL **GeSIDA** 1-4 DICIEMBRE  
GRUPO DE ESTUDIO DEL SIDA-SEIMC **MADRID 2015**  
Y 9ª REUNIÓN DOCENTE DE LA RED DE INVESTIGACIÓN EN SIDA



## CERTIFICADO DE PARTICIPACIÓN

Los presidentes del Comité Organizador del  
**VII CONGRESO NACIONAL GeSIDA**

**CERTIFICAN QUE:**

**S. Pérez, N. Chueca, M. Álvarez, M.Á. López, M. Omar, M.D.C. Gálvez, M.Á. Martínez, A.B. Lozano, C. Delgado, S. Hernández, J. Palomares, F. García**

han presentado la comunicación **POSTER** con título:

**ESTUDIO DE LOS PERFILES FILODINÁMICOS Y FILOGEOGRÁFICOS  
DE LA EPIDEMIA VIH-1 SUBTIPO B EN EL SUR DE ESPAÑA**

En el **VII Congreso Nacional de GeSIDA**,  
celebrado en Madrid del 1 al 4 de diciembre de 2015.

Y para que conste se expide el presente certificado  
en Madrid, a 4 de diciembre de 2015.

**Dr. Juan Berenguer**

Presidente VII Congreso Nacional GeSIDA

**Dr. José Ramón Arribas**

Co-Presidente VII Congreso Nacional GeSIDA

Código Presentación: **P-092**

## **ANEXO IV: PUBLICACIONES SURGIDAS DE ESTA TESIS**

**Pérez-Parra S, Chueca-Porcuna N, Álvarez-Estevez M, Pasquau J, Omar M, Collado A, et al. Los estudios de resistencias a antirretrovirales como herramienta para el análisis de los clusters de transmisión del virus de la inmunodeficiencia humana. Enfermedades Infecc Microbiol Clínica. 2015 Nov;33(9):603–8.**

**Pérez-Parra S, Chueca-Porcuna N, Álvarez-Estevez M, Pasquau J, Omar M, Collado A, Vinuesa S, Lozano AB, García F. Phylodynamic and Phylogeographic profiles of Subtype B HIV-1 Epidemics in South Spain. Plos One. 2016. [Enviado].**

**Pérez-Parra S, Chueca-Porcuna N, Álvarez-Estevez M, Pasquau J, Omar M, Collado A, Vinuesa S, Lozano AB, García F. High prevalence and diversity of HIV-1 non-B genetic forms due to immigration in southern Spain: a phylogeographical approach. [En redacción].**

**Pérez-Parra S, Chueca-Porcuna N, Álvarez-Estevez M, Pasquau J, Omar M, Collado A, Vinuesa S, Lozano AB, García F. Emergence of the CRF19 cpx Variant in Spain. [En redacción]**

## BIBLIOGRAFÍA

1. Information NC for B, Pike USNL of M 8600 R, MD B, Usa 20894. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. - PubMed - NCBI [Internet]. [cited 2015 Apr 8]. Available from: [http://www.ncbi.nlm.nih.gov/pubmed/?term=Centers+for+Disease+Control+\(CD+C\).+Kaposi%E2%80%99s+sarcoma+and+Pneumocystis+pneumonia+among+homosexual+men--New+York+City+and+California.+MMWR+Morb.+Mortal.+Wkly.+Rep.+30%2C+305%E2%80%93308+\(1981\)](http://www.ncbi.nlm.nih.gov/pubmed/?term=Centers+for+Disease+Control+(CD+C).+Kaposi%E2%80%99s+sarcoma+and+Pneumocystis+pneumonia+among+homosexual+men--New+York+City+and+California.+MMWR+Morb.+Mortal.+Wkly.+Rep.+30%2C+305%E2%80%93308+(1981))
2. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 1983 May 20;220(4599):868–71.
3. Ascher MS, Sheppard HW, Winkelstein W, Vittinghoff E. Does drug use cause AIDS? *Nature*. 1993 Mar 11;362(6416):103–4.
4. Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, et al. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science*. 1983 May 20;220(4599):865–7.
5. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in the chimpanzee *Pan troglodytes*. *Nature*. 1999 Feb 4;397(6718):436–41.
6. Maartens G, Celum C, Lewin SR. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet Lond Engl*. 2014 Jul 19;384(9939):258–71.
7. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature*. 1989 Jun 1;339(6223):389–92.
8. Popper SJ, Sarr AD, Guèye-Ndiaye A, Mboup S, Essex ME, Kanki PJ. Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. *J Virol*. 2000 Feb;74(3):1554–7.
9. Popper SJ, Sarr AD, Travers KU, Guèye-Ndiaye A, Mboup S, Essex ME, et al. Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J Infect Dis*. 1999 Oct;180(4):1116–21.
10. Mellors JW, Muñoz A, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann Intern Med*. 1997 Jun 15;126(12):946–54.
11. Rodríguez B, Sethi AK, Cheruvu VK, Mackay W, Bosch RJ, Kitahata M, et al. Predictive value of plasma HIV RNA level on rate of CD4 T-cell decline in untreated HIV infection. *JAMA*. 2006 Sep 27;296(12):1498–506.

12. Greene WC. The molecular biology of human immunodeficiency virus type 1 infection. *N Engl J Med*. 1991 Jan 31;324(5):308–17.
13. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CCR5* structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science*. 1996 Sep 27;273(5283):1856–62.
14. Varmus H. Retroviruses. *Science*. 1988 Jun 10;240(4858):1427–35.
15. Zack JA, Arrigo SJ, Weitsman SR, Go AS, Haislip A, Chen IS. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell*. 1990 Apr 20;61(2):213–22.
16. Cullen BR. Does HIV-1 Tat induce a change in viral initiation rights? *Cell*. 1993 May 7;73(3):417–20.
17. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*. 1996 Mar 15;271(5255):1582–6.
18. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS Lond Engl*. 2011 Mar 13;25(5):679–89.
19. Pernas B, Grandal M, Mena A, Castro-Iglesias A, Cañizares A, Wyles DL, et al. High prevalence of subtype F in newly diagnosed HIV-1 persons in northwest Spain and evidence for impaired treatment response. *AIDS Lond Engl*. 2014 Jul 31;28(12):1837–40.
20. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, et al. Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2009 May 1;48(9):1296–305.
21. de Arellano ER, Benito JM, Soriano V, López M, Holguín A. Impact of ethnicity and HIV type 1 subtype on response to first-line antiretroviral therapy. *AIDS Res Hum Retroviruses*. 2007 Jul;23(7):891–4.
22. Karasi JC, Dziezuk F, Quennery L, Förster S, Reischl U, Colucci G, et al. High correlation between the Roche COBAS® AmpliPrep/COBAS® TaqMan® HIV-1, v2.0 and the Abbott m2000 RealTime HIV-1 assays for quantification of viral load in HIV-1 B and non-B subtypes. *J Clin Virol Off Publ Pan Am Soc Clin Virol*. 2011 Nov;52(3):181–6.
23. Weinstock HS, Zaidi I, Heneine W, Bennett D, Garcia-Lerma JG, Douglas JM, et al. The epidemiology of antiretroviral drug resistance among drug-naive HIV-1-infected persons in 10 US cities. *J Infect Dis*. 2004 Jun 15;189(12):2174–80.

24. L. Ruiz. Guía para el manejo de las resistencias virales, de los aspectos farmacocinéticos del tratamiento y de las hepatitis virales en los pacientes infectados por VIH Resistencia a los fármacos antirretrovirales.
25. Wensing AMJ, van de Vijver DA, Angarano G, Asjö B, Balotta C, Boeri E, et al. Prevalence of drug-resistant HIV-1 variants in untreated individuals in Europe: implications for clinical management. *J Infect Dis.* 2005 Sep 15;192(6):958–66.
26. Frentz D, Boucher CAB, van de Vijver DAMC. Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. *AIDS Rev.* 2012 Mar;14(1):17–27.
27. Green H, Tilston P, Fearnhill E, Pillay D, Dunn DT, UK Collaborative Group on HIV Drug Resistance. The impact of different definitions on the estimated rate of transmitted HIV drug resistance in the United Kingdom. *J Acquir Immune Defic Syndr* 1999. 2008 Oct 1;49(2):196–204.
28. Zu Knyphausen F, Scheufele R, Kücherer C, Jansen K, Somogyi S, Dupke S, et al. First line treatment response in patients with transmitted HIV drug resistance and well defined time point of HIV infection: updated results from the German HIV-1 seroconverter study. *PloS One.* 2014;9(5):e95956.
29. Karlsson A, Björkman P, Bratt G, Ekvall H, Gisslén M, Sönnnerborg A, et al. Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010. *PloS One.* 2012;7(3):e33484.
30. Grgic I, Lepej SZ, Lunar MM, Poljak M, Vince A, Vrakela IB, et al. The prevalence of transmitted drug resistance in newly diagnosed HIV-infected individuals in Croatia: the role of transmission clusters of men who have sex with men carrying the T215S surveillance drug resistance mutation. *AIDS Res Hum Retroviruses.* 2013 Feb;29(2):329–36.
31. Bracciale L, Colafigli M, Zazzi M, Corsi P, Meraviglia P, Micheli V, et al. Prevalence of transmitted HIV-1 drug resistance in HIV-1-infected patients in Italy: evolution over 12 years and predictors. *J Antimicrob Chemother.* 2009 Sep;64(3):607–15.
32. Frentz D, van de Vijver D, Abecasis A, Albert J, Hamouda O, Jørgensen L, et al. Patterns of transmitted HIV drug resistance in Europe vary by risk group. *PloS One.* 2014;9(4):e94495.
33. Sagir A, Oette M, Kaiser R, Däumer M, Fätkenheuer G, Rockstroh JK, et al. Trends of prevalence of primary HIV drug resistance in Germany. *J Antimicrob Chemother.* 2007 Oct;60(4):843–8.
34. Riva C, Lai A, Caramma I, Corvasce S, Violin M, Dehò L, et al. Transmitted HIV Type 1 drug resistance and Non-B subtypes prevalence among seroconverters and newly diagnosed patients from 1992 to 2005 in Italy. *AIDS Res Hum Retroviruses.* 2010 Jan;26(1):41–9.
35. Abecasis AB, Wensing AMJ, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DAMC, et al. HIV-1 subtype distribution and its demographic determinants in

- newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology*. 2013;10:7.
36. Sharp PM, Bailes E, Gao F, Beer BE, Hirsch VM, Hahn BH. Origins and evolution of AIDS viruses: estimating the time-scale. *Biochem Soc Trans*. 2000 Feb;28(2):275–82.
  37. Bernard E, Azad Y, Vandamme A, Weait M, Geretti A. HIV forensics: pitfalls and acceptable standards in the use of phylogenetic analysis as evidence in criminal investigations of HIV transmission\*. *HIV Med*. 2007 Sep 1;8(6):382–7.
  38. Yebra G, Holguín A, Pillay D, Hué S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2013 Mar;14:232–9.
  39. Gray RR, Tatem AJ, Lamers S, Hou W, Laeyendecker O, Serwadda D, et al. Spatial phylodynamics of HIV-1 epidemic emergence in east Africa. *AIDS Lond Engl*. 2009 Sep 10;23(14):F9–17.
  40. Dalai SC, de Oliveira T, Harkins GW, Kassaye SG, Lint J, Manasa J, et al. Evolution and molecular epidemiology of subtype C HIV-1 in Zimbabwe. *AIDS Lond Engl*. 2009 Nov 27;23(18):2523–32.
  41. Esbjörnsson J, Mild M, Månsson F, Norrgren H, Medstrand P. HIV-1 molecular epidemiology in Guinea-Bissau, West Africa: origin, demography and migrations. *PloS One*. 2011;6(2):e17025.
  42. González-Alba JM, Holguín Á, Garcia R, García-Bujalance S, Alonso R, Suárez A, et al. Molecular Surveillance of HIV-1 in Madrid, Spain: a Phylogeographic Analysis ▽ . *J Virol*. 2011 Oct;85(20):10755–63.
  43. Delatorre EO, Bello G. Phylodynamics of HIV-1 subtype C epidemic in east Africa. *PloS One*. 2012;7(7):e41904.
  44. Faria NR, Suchard MA, Abecasis A, Sousa JD, Ndembi N, Bonfim I, et al. Phylodynamics of the HIV-1 CRF02\_AG clade in Cameroon. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2012 Mar;12(2):453–60.
  45. de Silva TI, van Tienen C, Onyango C, Jabang A, Vincent T, Loeff MFS van der, et al. Population dynamics of HIV-2 in rural West Africa: comparison with HIV-1 and ongoing transmission at the heart of the epidemic. *AIDS Lond Engl*. 2013 Jan 2;27(1):125–34.
  46. Lemey P, Rambaut A, Welch JJ, Suchard MA. Phylogeography takes a relaxed random walk in continuous space and time. *Mol Biol Evol*. 2010 Aug;27(8):1877–85.
  47. Dennis AM, Herbeck JT, Brown AL, Kellam P, de Oliveira T, Pillay D, et al. Phylogenetic studies of transmission dynamics in generalized HIV epidemics: an essential tool where the burden is greatest? *J Acquir Immune Defic Syndr* 1999. 2014 Oct 1;67(2):181–95.

48. Sobrino-Vegas P, Gutiérrez F, Berenguer J, Labarga P, García F, Alejos-Ferreras B, et al. [The Cohort of the Spanish HIV Research Network (CoRIS) and its associated biobank; organizational issues, main findings and losses to follow-up]. *Enfermedades Infecc Microbiol Clínica*. 2011 Nov;29(9):645–53.
49. Hall BG, Barlow M. Phylogenetic Analysis as a Tool in Molecular Epidemiology of Infectious Diseases. *Ann Epidemiol*. 2006 Mar 1;16(3):157–69.
50. Zharkikh A, Li WH. Estimation of confidence in phylogeny: the complete-and-partial bootstrap technique. *Mol Phylogenet Evol*. 1995 Mar;4(1):44–63.
51. Baldauf SL. Phylogeny for the faint of heart: a tutorial. *Trends Genet TIG*. 2003 Jun;19(6):345–51.
52. Hillis DM, Bull JJ. An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. *Syst Biol*. 1993 Jun 1;42(2):182–92.
53. Feng DF, Doolittle RF. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J Mol Evol*. 1987;25(4):351–60.
54. Higgins DG, Sharp PM. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*. 1988 Dec 15;73(1):237–44.
55. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004 Aug 19;5:113.
56. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinforma Oxf Engl*. 2007 Nov 1;23(21):2947–8.
57. Hall BG. Building phylogenetic trees from molecular data with MEGA. *Mol Biol Evol*. 2013 May;30(5):1229–35.
58. Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. *Nat Rev Genet*. 2003 Apr;4(4):275–84.
59. Baele G, Li WLS, Drummond AJ, Suchard MA, Lemey P. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Mol Biol Evol*. 2013 Feb;30(2):239–43.
60. Hué S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS Lond Engl*. 2004 Mar 26;18(5):719–28.
61. Montagnier L. A History of HIV Discovery. *Science*. 2002 Nov 29;298(5599):1727–8.
62. García F, Pérez-Cachafeiro S, Guillot V, Alvarez M, Pérez-Romero P, Pérez-Elías MJ, et al. Transmission of HIV drug resistance and non-B subtype distribution in the Spanish cohort of antiretroviral treatment naïve HIV-infected individuals (CoRIS). *Antiviral Res*. 2011 Aug;91(2):150–3.

63. Fox J, Castro H, Kaye S, McClure M, Weber JN, Fidler S, et al. Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK. *AIDS Lond Engl*. 2010 Sep 24;24(15):2397–401.
64. Lai A, Riva C, Marconi A, Balestrieri M, Razzolini F, Meini G, et al. Changing patterns in HIV-1 non-B clade prevalence and diversity in Italy over three decades. *HIV Med*. 2010 Oct 1;11(9):593–602.
65. Palma AC, Araújo F, Duque V, Borges F, Paixão MT, Camacho R, et al. Molecular epidemiology and prevalence of drug resistance-associated mutations in newly diagnosed HIV-1 patients in Portugal. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2007 Jun;7(3):391–8.
66. Recordon-Pinson P, Anies G, Bruyand M, Neau D, Morlat P, Pellegrin J-L, et al. HIV type-1 transmission dynamics in recent seroconverters: relationship with transmission of drug resistance and viral diversity. *Antivir Ther*. 2009;14(4):551–6.
67. Alcantara LCJ, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. *Nucleic Acids Res*. 2009 Jul;37(Web Server issue):W634–42.
68. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PloS One*. 2009;4(3):e4724.
69. HIV drug resistance database. Stanford university. <http://hivdb.stanford.edu>.
70. Ambrosioni J, Junier T, Delhumeau C, Calmy A, Hirschel B, Zdobnov E, et al. Impact of highly active antiretroviral therapy on the molecular epidemiology of newly diagnosed HIV infections. *AIDS Lond Engl*. 2012 Oct 23;26(16):2079–86.
71. Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT, et al. Transmission network parameters estimated from HIV sequences for a nationwide epidemic. *J Infect Dis*. 2011 Nov;204(9):1463–9.
72. Novitsky V, Bussmann H, Logan A, Moyo S, van Widenfelt E, Okui L, et al. Phylogenetic Relatedness of Circulating HIV-1C Variants in Mochudi, Botswana. *PLoS ONE*. 2013 Diciembre;8(12):e80589.
73. Aldous JL, Pond SK, Poon A, Jain S, Qin H, Kahn JS, et al. Characterizing HIV transmission networks across the United States. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2012 Oct;55(8):1135–43.
74. Leigh Brown AJ, Frost SDW, Mathews WC, Dawson K, Hellmann NS, Daar ES, et al. Transmission fitness of drug-resistant human immunodeficiency virus and the prevalence of resistance in the antiretroviral-treated population. *J Infect Dis*. 2003 Feb 15;187(4):683–6.

75. Yebra G, de Mulder M, Martín L, Pérez-Cachafeiro S, Rodríguez C, Labarga P, et al. Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naïve HIV-infected patients (CoRIS). *Antiviral Res.* 2011 Jan;89(1):19–25.
76. Vercauteren J, Wensing AMJ, van de Vijver DAMC, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. *J Infect Dis.* 2009 Nov 15;200(10):1503–8.
77. Bannister WP, Ruiz L, Loveday C, Vella S, Zilmer K, Kjaer J, et al. HIV-1 subtypes and response to combination antiretroviral therapy in Europe. *Antivir Ther.* 2006;11(6):707–15.
78. Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review. *PLoS Med.* 2007 Dec;4(12):e339.
79. Mbisa JL, Hué S, Buckton AJ, Myers RE, Duiculescu D, Ene L, et al. Phylodynamic and phylogeographic patterns of the HIV type 1 subtype F1 parenteral epidemic in Romania. *AIDS Res Hum Retroviruses.* 2012 Sep;28(9):1161–6.
80. Mehta SR, Wertheim JO, Delpont W, Ene L, Tardei G, Duiculescu D, et al. Using phylogeography to characterize the origins of the HIV-1 subtype F epidemic in Romania. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis.* 2011 Jul;11(5):975–9.
81. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Gateway Computing Environments Workshop (GCE), 2010.* 2010. p. 1–8.
82. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics.* 2002 Jul;161(3):1307–20.
83. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 2007;7:214.
84. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed Phylogenetics and Dating with Confidence. *PLoS Biol.* 2006 Mar 14;4(5):e88.
85. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol.* 2005 May;22(5):1185–92.
86. Yebra G, de Mulder M, Martín L, Rodríguez C, Labarga P, Viciano I, et al. Most HIV type 1 non-B infections in the Spanish cohort of antiretroviral treatment-naïve HIV-infected patients (CoRIS) are due to recombinant viruses. *J Clin Microbiol.* 2012 Feb;50(2):407–13.
87. García F, Pérez-Cachafeiro S, Guillot V, Alvarez M, Pérez-Romero P, Pérez-Elías MJ, et al. Transmission of HIV drug resistance and non-B subtype distribution in the

- Spanish cohort of antiretroviral treatment naïve HIV-infected individuals (CoRIS). *Antiviral Res.* 2011 Aug;91(2):150–3.
88. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ. Episodic Sexual Transmission of HIV Revealed by Molecular Phylodynamics. *PLoS Med.* 2008 Mar 18;5(3):e50.
  89. Fisher M, Pao D, Brown AE, Sudarshi D, Gill ON, Cane P, et al. Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. *AIDS Lond Engl.* 2010 Jul 17;24(11):1739–47.
  90. Brenner BG, Roger M, Stephens D, Moisi D, Hardy I, Weinberg J, et al. Transmission Clustering Drives the Onward Spread of the HIV Epidemic Among Men Who Have Sex With Men in Quebec. *J Infect Dis.* 2011 Oct 1;204(7):1115–9.
  91. Wertheim JO, Fourment M, Pond SLK. Inconsistencies in Estimating the Age of HIV-1 Subtypes Due to Heterotachy. *Mol Biol Evol.* 2012 Feb 1;29(2):451–6.
  92. Abecasis AB, Vandamme A-M, Lemey P. Quantifying Differences in the Tempo of Human Immunodeficiency Virus Type 1 Subtype Evolution. *J Virol.* 2009 Dec 15;83(24):12917–24.
  93. Hué S, Pillay D, Clewley JP, Pybus OG. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci U S A.* 2005 Mar 22;102(12):4425–9.
  94. Vrancken B, Baele G, Vandamme A-M, van Laethem K, Suchard MA, Lemey P. Disentangling the impact of within-host evolution and transmission dynamics on the tempo of HIV-1 evolution. *AIDS Lond Engl.* 2015 Jul 31;29(12):1549–56.
  95. Ciccozzi M, Lai A, Ebranati E, Gabanelli E, Galli M, Mugosa B, et al. Phylogeographic Reconstruction of HIV Type 1B in Montenegro and the Balkan Region. *AIDS Res Hum Retroviruses.* 2012 Feb 25;28(10):1280–4.
  96. Roy CN, Khandaker I, Oshitani H. Evolutionary Dynamics of Tat in HIV-1 Subtypes B and C. *PLoS ONE [Internet].* 2015 Jun 18 [cited 2015 Aug 12];10(6). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4472691/>
  97. Walker PR, Pybus OG, Rambaut A, Holmes EC. Comparative population dynamics of HIV-1 subtypes B and C: subtype-specific differences in patterns of epidemic growth. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis.* 2005 Apr;5(3):199–208.
  98. National AIDS Case Registry. Epidemiological Surveillance of AIDS in Spain [Internet]. 2010. Available from: [http://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/Registro\\_sida\\_junio2010\\_ENG\\_REV.pdf](http://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/Registro_sida_junio2010_ENG_REV.pdf)
  99. Kim GJ, Yun M-R, Koo MJ, Shin B-G, Lee J-S, Kim SS. Estimating the Origin and Evolution Characteristics for Korean HIV Type 1 Subtype B Using Bayesian Phylogenetic Analysis. *AIDS Res Hum Retroviruses.* 2012 Aug;28(8):880–4.

100. Ng KT, Ong LY, Lim SH, Takebe Y, Kamarulzaman A, Tee KK. Evolutionary History of HIV-1 Subtype B and CRF01\_AE Transmission Clusters among Men Who Have Sex with Men (MSM) in Kuala Lumpur, Malaysia. *PLoS ONE* [Internet]. 2013 Jun 20 [cited 2015 Aug 12];8(6). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3688664/>
101. Castro-Nallar E, Pérez-Losada M, Burton GF, Crandall KA. The evolution of HIV: Inferences using phylogenetics. *Mol Phylogenet Evol.* 2012 Feb;62(2):777–92.
102. Volz EM, Pong SLK, Ward MJ, Brown AJL, Frost SDW. Phylodynamics of Infectious Disease Epidemics. *Genetics.* 2009 Dec 1;183(4):1421–30.
103. Dávila PF. Conductas de Riesgo, Prueba del VIH y Seroestatus en HSH de España: Un compendio de tres artículos. *Editorial Académica Española*; 2012. 64 p.
104. Zehender G, Ebranati E, Lai A, Santoro MM, Alteri C, Giuliani M, et al. Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. *J Acquir Immune Defic Syndr* 1999. 2010 Oct;55(2):156–60.
105. Lunar MM, Vandamme A-M, Tomažič J, Karner P, Vovko TD, Pečavar B, et al. Bridging epidemiology with population genetics in a low incidence MSM-driven HIV-1 subtype B epidemic in Central Europe. *BMC Infect Dis* [Internet]. 2015 Feb 15 [cited 2015 Aug 12];15. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4345027/>
106. Marina Siljic DS. Tracing the origin and dynamics of the HIV-1 epidemic in Serbia. *Arch Biol Sci.* 2014;66(2):507–15.
107. Bezemer D, Jurriaans S, Prins M, van der Hoek L, Prins JM, de Wolf F, et al. Declining trend in transmission of drug-resistant HIV-1 in Amsterdam. *AIDS Lond Engl.* 2004 Jul 23;18(11):1571–7.
108. Yerly S, von Wyl V, Ledergerber B, Böni J, Schüpbach J, Bürgisser P, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. *AIDS Lond Engl.* 2007 Oct 18;21(16):2223–9.
109. Lohse N, Obel N, Kronborg G, Jørgensen LB, Pedersen C, Larsen CS, et al. Declining prevalence of HIV-infected individuals at risk of transmitting drug-resistant HIV in Denmark during 1997-2004. *Antivir Ther.* 2006;11(5):591–600.
110. UK Collaborative Group on HIV Drug Resistance, UK Collaborative HIV Cohort Study, UK Register of HIV Seroconverters. Evidence of a decline in transmitted HIV-1 drug resistance in the United Kingdom. *AIDS Lond Engl.* 2007 May 11;21(8):1035–9.
111. Frost SDW, Volz EM. Viral phylodynamics and the search for an “effective number of infections.” *Philos Trans R Soc Lond B Biol Sci.* 2010 Jun 27;365(1548):1879–90.
112. Vidal N, Mulanga C, Bazepeo SE, Mwamba JK, Tshimpaka J-W, Kashi M, et al. Distribution of HIV-1 variants in the Democratic Republic of Congo suggests

- increase of subtype C in Kinshasa between 1997 and 2002. *J Acquir Immune Defic Syndr* 1999. 2005 Dec 1;40(4):456–62.
113. Taniguchi Y, Takehisa J, Bikandou B, Mboudjeka I, N'Doundou-N'Kodia M-Y, Obengui null, et al. Genetic subtypes of HIV type 1 based on the vpu/env sequences in the Republic of Congo. *AIDS Res Hum Retroviruses*. 2002 Jan 1;18(1):79–83.
  114. Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, et al. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol*. 1999 Dec;73(12):10489–502.
  115. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 nomenclature proposal. *Science*. 2000 Apr 7;288(5463):55–6.
  116. Artenstein AW, VanCott TC, Mascola JR, Carr JK, Hegerich PA, Gaywee J, et al. Dual infection with human immunodeficiency virus type 1 of distinct envelope subtypes in humans. *J Infect Dis*. 1995 Apr;171(4):805–10.
  117. van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson AL, Williamson C. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. *AIDS Lond Engl*. 1997 Jan;11(1):81–7.
  118. Tscherning C, Alaeus A, Fredriksson R, Björndal A, Deng H, Littman DR, et al. Differences in chemokine coreceptor usage between genetic subtypes of HIV-1. *Virology*. 1998 Feb 15;241(2):181–8.
  119. Kanki PJ, Hamel DJ, Sankalé JL, Hsieh C c, Thior I, Barin F, et al. Human immunodeficiency virus type 1 subtypes differ in disease progression. *J Infect Dis*. 1999 Jan;179(1):68–73.
  120. Kouri V, Khouri R, Alemán Y, Abrahantes Y, Vercauteren J, Pineda-Peña A-C, et al. CRF19\_cpx is an Evolutionary fit HIV-1 Variant Strongly Associated With Rapid Progression to AIDS in Cuba. *EBioMedicine*. 2015 Mar;2(3):244–54.
  121. Apetrei C, Descamps D, Collin G, LouSSERT-Ajaka I, Damond F, Duca M, et al. Human immunodeficiency virus type 1 subtype F reverse transcriptase sequence and drug susceptibility. *J Virol*. 1998 May;72(5):3534–8.
  122. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med*. 2008 Apr 10;358(15):1590–602.
  123. Alaeus A, Lidman K, Sönnnerborg A, Albert J. Subtype-specific problems with quantification of plasma HIV-1 RNA. *AIDS Lond Engl*. 1997 Jun;11(7):859–65.
  124. Parekh B, Phillips S, Granade TC, Baggs J, Hu DJ, Respass R. Impact of HIV type 1 subtype variation on viral RNA quantitation. *AIDS Res Hum Retroviruses*. 1999 Jan 20;15(2):133–42.

125. Brennan CA, Bodelle P, Coffey R, Harris B, Holzmayer V, Luk K-C, et al. HIV global surveillance: foundation for retroviral discovery and assay development. *J Med Virol.* 2006;78 Suppl 1:S24–9.
126. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, et al. Effect of HIV-1 Subtype on Virologic and Immunologic Response to Starting Highly Active Antiretroviral Therapy. *Clin Infect Dis.* 2009 May 1;48(9):1296–305.
127. Olivier Bouchaud VLM. Similar Short-Term Efficacy of Antiretroviral Therapy in Patients Infected With HIV B and Non-B Subtype Strains in France. *J Acquir Immune Defic Syndr* 1999. 2011;56(2):e67–9.
128. Weintrob AC, Grandits GA, Agan BK, Ganesan A, Landrum ML, Crum-Cianflone NF, et al. Virologic response differences between African Americans and European Americans initiating highly active antiretroviral therapy with equal access to care. *J Acquir Immune Defic Syndr* 1999. 2009 Dec;52(5):574–80.
129. Alvarez M, García F, Martínez NM, García F, Bernal C, Vela CM, et al. Introduction of HIV type 1 non-B subtypes into Eastern Andalusia through immigration. *J Med Virol.* 2003 May;70(1):10–3.
130. Holguín A, de Mulder M, Yebra G, López M, Soriano V. Increase of non-B subtypes and recombinants among newly diagnosed HIV-1 native Spaniards and immigrants in Spain. *Curr HIV Res.* 2008 Jun;6(4):327–34.
131. Gifford RJ, Oliveira T de, Rambaut A, Pybus OG, Dunn D, Vandamme A-M, et al. Phylogenetic Surveillance of Viral Genetic Diversity and the Evolving Molecular Epidemiology of Human Immunodeficiency Virus Type 1. *J Virol.* 2007 Dec 1;81(23):13050–6.
132. Yerly S, von Wyl V, Ledergerber B, Böni J, Schüpbach J, Bürgisser P, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. *AIDS Lond Engl.* 2007 Oct 18;21(16):2223–9.
133. Frange P, Galimand J, Vidal N, Goujard C, Deveau C, Souala F, et al. New and old complex recombinant HIV-1 strains among patients with primary infection in 1996-2006 in France: the French ANRS CO06 primo cohort study. *Retrovirology.* 2008;5:69.
134. Brennan CA, Yamaguchi J, Devare SG, Foster GA, Stramer SL. Expanded evaluation of blood donors in the United States for human immunodeficiency virus type 1 non-B subtypes and antiretroviral drug-resistant strains: 2005 through 2007. *Transfusion (Paris).* 2010 Dec;50(12):2707–12.
135. von Wyl V, Kouyos RD, Yerly S, Böni J, Shah C, Bürgisser P, et al. The role of migration and domestic transmission in the spread of HIV-1 non-B subtypes in Switzerland. *J Infect Dis.* 2011 Oct 1;204(7):1095–103.
136. Thomson MM, Fernández-García A, Delgado E, Vega Y, Díez-Fuertes F, Sánchez-Martínez M, et al. Rapid expansion of a HIV-1 subtype F cluster of recent origin

- among men who have sex with men in Galicia, Spain. *J Acquir Immune Defic Syndr* 1999. 2012 Mar 1;59(3):e49–51.
137. Alexiev I, Shankar A, Wensing AMJ, Beshkov D, Elenkov I, Stoycheva M, et al. Low HIV-1 transmitted drug resistance in Bulgaria against a background of high clade diversity. *J Antimicrob Chemother*. 2015;70(6):1874–80.
  138. González-Candelas F, Bracho MA, Moya A. Molecular epidemiology and forensic genetics: application to a hepatitis C virus transmission event at a hemodialysis unit. *J Infect Dis*. 2003 Feb 1;187(3):352–8.
  139. Bártolo I, Abecasis AB, Borrego P, Barroso H, McCutchan F, Gomes P, et al. Origin and Epidemiological History of HIV-1 CRF14\_BG. *PLOS ONE*. 2011 Sep 28;6(9):e24130.
  140. de Pina-Araujo IIM, Delatorre E, Guimarães ML, Morgado MG, Bello G. Origin and Population Dynamics of a Novel HIV-1 Subtype G Clade Circulating in Cape Verde and Portugal. *PloS One*. 2015;10(5):e0127384.
  141. de Felipe B, Pérez-Romero P, Abad-Fernández M, Fernandez-Cuenca F, Martinez-Fernandez FJ, Trastoy M, et al. Prevalence and resistance mutations of non-B HIV-1 subtypes among immigrants in Southern Spain along the decade 2000-2010. *Virology*. 2011;8:416.
  142. Instituto Nacional de Estadística (INE) [Internet]. Available from: [www.ine.es](http://www.ine.es)
  143. Paraskevis D, Kostaki E, Beloukas A, Cañizares A, Aguilera A, Rodríguez J, et al. Molecular characterization of HIV-1 infection in Northwest Spain (2009-2013): Investigation of the subtype F outbreak. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2015 Mar;30:96–101.
  144. Pernas B, Mena A, Cañizares A, Grandal M, Castro-Iglesias A, Pértega S, et al. Trends on epidemiological, virological, and clinical features among newly diagnosed HIV-1 persons in Northwest Spain over the last 10 years. *J Med Virol*. 2015 Mar 16;
  145. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect Dis*. 2010 Sep 7;10(1):262.
  146. VI Congreso Nacional de GESIDA y 8ª Reunión docente de la RIS.
  147. Patiño Galindo JA, Torres-Puente M, Gimeno C, Ortega E, Navarro D, Galindo MJ, et al. Expansion of the CRF19\_cpx Variant in Spain. *J Clin Virol Off Publ Pan Am Soc Clin Virol*. 2015 Aug;69:146–9.
  148. Lemey P, Vandamme A-M. Exploring full-genome sequences for phylogenetic support of HIV-1 transmission events. *AIDS Lond Engl*. 2005 Sep 23;19(14):1551–2.

149. Lemey P, Van Dooren S, Van Laethem K, Schrooten Y, Derdelinckx I, Goubau P, et al. Molecular testing of multiple HIV-1 transmissions in a criminal case. *AIDS Lond Engl*. 2005 Oct 14;19(15):1649–58.
150. Volz EM, Koopman JS, Ward MJ, Brown AL, Frost SDW. Simple Epidemiological Dynamics Explain Phylogenetic Clustering of HIV from Patients with Recent Infection. *PLoS Comput Biol*. 2012 Jun 28;8(6):e1002552.
151. Field N, Cohen T, Struelens MJ, Palm D, Cookson B, Glynn JR, et al. Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement. *Lancet Infect Dis*. 2014 Apr;14(4):341–52.
152. Stack JC, Welch JD, Ferrari MJ, Shapiro BU, Grenfell BT. Protocols for sampling viral sequences to study epidemic dynamics. *J R Soc Interface*. 2010 Feb 10;rsif20090530.
153. Kerrigan D, Moreno L, Rosario S, Gomez B, Jerez H, Barrington C, et al. Environmental-structural interventions to reduce HIV/STI risk among female sex workers in the Dominican Republic. *Am J Public Health*. 2006 Jan;96(1):120–5.
154. Caro-Murillo AM, Gil Luciano A, Navarro Rubio G, Leal Noval M, Blanco Ramos JR, Cohorte de la Red de Investigación en Sida (CoRIS). [HIV infection in different age groups: Potential implications for prevention. CoRIS Cohort, Spain, 2004-2008]. *Med Clínica*. 2010 Apr 24;134(12):521–7.