

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

Departamento de Microbiología II



**VARIACIÓN DE LOS GENES ACCESORIOS Y SU
IMPLICACIÓN EN LA BIOLOGÍA DEL VIH-1**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Víctor Manuel Sánchez Merino

Bajo la dirección del doctor

Cecilio López Galíndez

Madrid, 2001

ISBN: 84-669-2018-8

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Farmacia

Departamento de Microbiología II

**VARIACIÓN DE LOS GENES ACCESORIOS
Y SU IMPLICACIÓN EN LA BIOLOGÍA DEL
VIH-1**

Memoria presentada para optar al grado de Doctor en Farmacia por:

Víctor Manuel Sánchez Merino

Director de la Tesis:

Dr. Cecilio López Galíndez

Centro Nacional de Biología Fundamental

Instituto de Salud Carlos III

Madrid, Junio de 2001

En la elaboración de una tesis doctoral es indudable que, aparte del esfuerzo personal que suponen los años de investigación e incluso la propia redacción del texto, existen otras personas que sin su ayuda no habría sido posible la presentación de éste trabajo. Por lo tanto, quisiera expresar mi más sentido agradecimiento a todas aquellas personas que, de un modo u otro han contribuido a que todo esto saliese adelante.

En primer lugar me gustaría agradecer al Dr. Cecilio López Galíndez su ayuda y la excelente dirección de ésta tesis. Él ha sido el responsable directo de que me embarcase en ésta aventura y por lo tanto, le agradezco que me brindase la oportunidad de realizar una tesis doctoral.

También agradezco a la Dra. Mercedes Domínguez la ayuda recibida, puesto que ha tenido parte de responsabilidad en que me dedicase a la investigación.

Agradezco al Instituto de Salud Carlos III las facilidades dadas para realizar el trabajo, tanto en medios materiales como humanos. Mi agradecimiento al Plan Nacional del Sida, a ONUSIDA, al Fondo de Investigación Sanitaria y de nuevo al Instituto de Salud Carlos III por la financiación y la ayuda recibida durante todos estos años.

Sin lugar a dudas, una de las cosas más importantes en la elaboración de cualquier tesis, es que se respire un buen ambiente en el laboratorio. Tengo que agradecer a todos mis compañeros del laboratorio de Virología Molecular que ésto haya sido así, y que el trabajo se hiciese menos duro. Su ayuda nunca me ha faltado cuando lo he necesitado y han hecho que todo fuese mucho más sencillo. Mi más sincero agradecimiento tanto a los que permanecen (Isabel, Concha, María, María José, Gonzalo, Tamara y Beatriz) como a los que recientemente nos han abandonado (Loles y María José). También mi agradecimiento a los que han pasado temporadas más o menos largas en nuestro laboratorio (Marga, Imelda y Juan).

Agradezco a Lourdes todo el trabajo que ha realizado en el establecimiento de las infecciones persistentes, puesto que, sin su trabajo ésta tesis no habría sido posible.

He tenido la suerte de conocer durante éste tiempo a un gran número de personas del Instituto que me han brindado su amistad desinteresada. Muchas gracias Paz, Fali, Pepe, Mónica, Bea, Beatriz Gil, Elena, Víctor, Mila, Isidoro, Tito, Luis, Esteban Marta y Yolanda.

También agradezco a Jose, Juan, Elia, Carlos y Almudena que, a pesar de no dedicarse al mundo de la ciencia, me han apoyado y ayudado a llevarlo lo mejor posible.

Muchas gracias Eloísa por estar a mi lado en todo momento, y sobre todo en los momentos más difíciles. No se como agradecerle todo lo que has hecho por mí. Gracias por tragarte otra tesis y darme tu cariño y apoyo.

Por último, quisiera agradecer especialmente el cariño y apoyo que me ha dado mi familia, y especialmente mis padres.

A mis padres y hermanos

A Eloísa

Abreviaturas

ADN	Ácido desoxirribonucleico
Ampli Taq	ADN polimerasa de <i>Thermophilus Aquaticus</i>
ARN	Ácido ribonucleico
AZT	3'-azido, 3'-desoxitimidina, zidovudina, Retrovir
Ci	Curio
CMSP	Células mononucleares humanas de sangre periférica
dATP	2'-desoxiadenosina-5'-trifosfato
ddNTP	2', 3'-dideoxynucleótido-5'-trifosfato
dNTP	2'-desoxynucleótido-5'-trifosfato
DI	Dosis infectiva
DTT	Ditiotreitol
IS	Aislados inductores de sincitios
KDa	Kilodaltons
LTR	Región terminal repetida
NIS	Aislados no inductores de sincitios
NJ	Análisis filogenético por el método del vecino más próximo
ONUSIDA	Programa de las Naciones Unidas para el SIDA
PBS	Tampón fosfato sódico-potásico
PCR	Reacción en cadena de la polimerasa
pb	Pares de bases
r.p.m.	Revoluciones por minuto
RNasa H	Ribonucleasa H
RRE	Elemento de respuesta a <i>rev</i>
s/nt	Sustituciones por nucleótido
RT	Retrotranscriptasa
RT-PCR	Reacción en cadena de la polimerasa a partir de ARN
³⁵ S	Isótopo azufre 35
SBF	Suero bovino fetal
SDS	Sodio dodecil sulfato
SIDA	Síndrome de inmunodeficiencia adquirida
TAE	Tampón Tris-acetato-EDTA

TAR	Región de la respuesta transactivadora
TBE	Tampón Tris-borato-EDTA
T CD4 ⁺	Linfocitos T que expresan la molécula CD4 en su superficie
T CD8 ⁺	Linfocitos T que expresan la molécula CD8 en su superficie
UDVP	Usuarios de drogas por vía parenteral
VAEC	Virus de la artritis encefalitis equina
VAEI	Virus de la anemia infecciosa equina
VFA	Virus de la fiebre aftosa
VIB	Virus de la inmunodeficiencia bovina
VIF	Virus de la inmunodeficiencia felina
VIH-1	Virus de la inmunodeficiencia humana tipo 1
VIH-2	Virus de la inmunodeficiencia humana tipo 2
VIS	Virus de la inmunodeficiencia de simio
VLB	Virus de la leucemia bovina
VLFe	Virus de la leucemia felina
VLMMo	Virus de la leucemia murina de Moloney
VLTH	Virus linfotrópicos T humanos
VSV	Virus de la estomatitis vesicular

ÍNDICE

1. Introducción	1
1.1. Virus de la Inmunodeficiencia Humana tipo I (VIH-1)	2
1.1.1. Familia retroviridae	3
1.1.2. Estructura del virión	5
1.1.1. Estructura del genoma	5
1.1.3.1. Genes estructurales	6
1.1.3.2. Genes reguladores	6
1.1.3.3. Genes accesorios	7
1.1.3.4. Regiones Terminales Repetidas (LTRs)	10
1.2. Ciclo biológico	14
1.3. Patogénesis de la infección por el VIH-1	16
1.4. Persistencia en virología	19
1.4.1. Persistencia en virus ARN	19
1.4.2. Persistencia en el VIH-1	20
1.5. Variabilidad en virus ARN	22
1.5.1. Variabilidad en el VIH-1	23
1.5.1.1. Variabilidad genética	23
1.5.1.2. Variabilidad fenotípica	24
1.6. Epidemiología molecular de la infección por el VIH-1	25
1.6.1. Clasificación y distribución de subtipos	26
1.6.2. Situación actual a nivel mundial de la epidemia de VIH/SIDA	29
1.6.3. Situación de la infección VIH/SIDA en España	30
2. Objetivos	31
3. Materiales y métodos	33
3.1. Cultivos celulares	34
3.1.1. Células	34
3.1.2. Medios de cultivo	34
3.1.3. Cepas de virus utilizadas	35
3.2. Muestras del VIH-1	
3.2.1. Aislados naturales	35
3.2.2. Provirus obtenidos de células persistentemente infectadas por el aislado s61	36

3.3. Obtención de ácidos nucleicos.....	36
3.3.1. Ácido desoxirribonucleico (ADN).....	36
3.3.2. Ácido ribonucleico (ARN).....	37
3.4. Reacción en cadena de la polimerasa (PCR).....	37
3.4.1. A partir de ADN.....	37
3.4.2. A partir de ARN.....	39
3.4.3. Iniciadores empleados.....	39
3.5. Purificación de productos de PCR.....	42
3.6. Clonaje molecular.....	42
3.6.1. Preparación del inserto y ligación.....	42
3.6.2. Transformación de bacterias competentes.....	43
3.7. Obtención de ADN de plásmidos.....	43
3.7.1. En pequeña escala.....	43
3.7.2. En gran escala.....	44
3.8. Secuenciación.....	44
3.8.1. Secuenciación manual.....	44
3.8.2. Secuenciación automática.....	45
3.9. Análisis de secuencias.....	45
3.10. Análisis filogenéticos.....	46
4. Resultados.....	47
4.1. Análisis de muestras de ONUSIDA.....	48
4.1.1. Análisis filogenético de muestras ONUSIDA cocultivo negativo en los genes <i>gag</i> , <i>env</i> y <i>nef</i>	48
4.1.2. Análisis genético en las regiones p17 y C2-V3-C3 de las muestras.....	55
4.1.3. Análisis del gen <i>nef</i>	59
4.2. Estudio del proceso de persistencia a partir de un mismo aislado del VIH-1 en tres líneas celulares.....	62
4.2.1. Criterios de infección persistente.....	62
4.2.2. Establecimiento de la infección persistente en la línea celular H9.....	64
4.2.3. Establecimiento de la infección persistente en la línea celular U-937 clon 2...65	

4.2.4. Establecimiento de la infección persistente en la línea celular MT-4.....	66
4.2.5. Comparación del proceso de persistencia entre las líneas celulares.....	67
4.2.6. Secuenciación del genoma completo del provirus del aislado inicial s61 y de los provirus VH61, VM61 y VU61.....	68
4.2.6.1. Estrategia de amplificación y secuenciación.....	69
4.2.6.2. Descripción de las mutaciones encontradas.....	70
4.2.6.2.1. Frecuencias de mutación y tipos de mutaciones encontradas.	
4.2.6.2.2. Relación entre mutaciones sinónimas y no sinónimas.....	80
4.2.6.2.3. Descripción y localización de las mutaciones no sinónimas en los tres provirus.....	81
4.2.6.2.4. Relaciones filogenéticas de los genomas completos secuenciados.....	85
4.2.6.2.5. Mutaciones comunes en los dos provirus VH61 y VM61 obtenidos de las líneas celulares T CD4 ⁺	87
4.2.6.2.6. Mutaciones comunes a los tres provirus.....	87
4.2.6.2.7. Análisis de la cuasiespecie vírica en los tres provirus.	89
5. Discusión.....	91
5.1. Análisis filogenético de las muestras cocultivo negativo de ONUSIDA.....	92
5.2. Establecimiento de las infecciones persistentes in vitro: mutaciones asociadas.....	94
5.2.1. Frecuencias de mutación observadas.....	95
5.2.2. Distribución de las mutaciones.....	96
5.3. Relación filogenética entre los virus persistentes en líneas T y en U-937.....	97
5.4. Mutaciones en los genes accesorios posiblemente relacionadas con el fenotipo de persistencia.....	98
5.4.1. Persistencia en la línea celular H9.....	98
5.4.2. Persistencia en la línea celular MT-4.....	99
5.5. Persistencia en la línea celular U-937 clon 2.....	100
5.6. Otras mutaciones posiblemente relacionadas con el fenómeno de la persistencia.....	101
5.7. Proceso de establecimiento de persistencia reflejado en la cuasiespecie vírica.....	101
5.8. Modelos de persistencia.....	102
6. Conclusiones.....	104
7. Bibliografía.....	108

8. Anexo.....127

1. INTRODUCCIÓN

1.1. Virus de la inmunodeficiencia Humana tipo 1 (VIH-1).

El VIH es el agente etiológico causal del síndrome de inmunodeficiencia adquirida (SIDA). Esta enfermedad se describió por primera vez en 1981 en varones jóvenes homosexuales, que padecían sarcoma de Kaposi y/o neumonía por *Pneumocistis carinii*. El estudio de estos pacientes reveló que presentaban un cuadro de inmunodeficiencia caracterizado por la disminución de linfocitos T CD4⁺. Los estudios epidemiológicos posteriores implicaron en la enfermedad a un agente infeccioso transmisible por sangre, por productos sanguíneos, contacto sexual, uso de drogas por vía intravenosa y verticalmente de madres a hijos (Luciw, 1996). La búsqueda de este agente infeccioso tuvo sus frutos en 1983 cuando Barre-Sinoussi, Chermann y Montagnier, del Instituto Pasteur de París, aislaron un retrovirus a partir de nódulos linfáticos de un paciente con linfadenopatía. Este retrovirus, denominado virus de la linfadenopatía (LAV), era capaz de replicar y causar efectos citopáticos en cultivos de células mononucleares humanas de sangre periférica (CMSP) (Barre-Sinoussi y cols., 1983). Pocos meses después, Gallo y colaboradores (Popovic y cols., 1984) y Levy y colaboradores (Levy y cols., 1984) aislaron un retrovirus a partir de muestras de pacientes con SIDA al que denominaron HTLV-III. Finalmente se decidió denominar al virus que causaba el SIDA como virus de la inmunodeficiencia humana (VIH). Dos años más tarde fue aislado un segundo tipo de virus que producía SIDA en pacientes procedentes de Africa Occidental (Clavel y cols., 1986). Actualmente se denomina VIH-1 a los virus genéticamente relacionados con los virus que circulan en Asia, Europa, Oceanía, América y ciertas regiones de Africa y VIH-2 al conjunto de virus relacionados con el virus descrito por Clavel que es prevalente en Africa Central y Occidental. Desde su identificación hasta la actualidad, el VIH se ha convertido posiblemente en el agente infeccioso que se ha estudiado más rápidamente y más trabajos ha generado en toda la historia de la biomedicina.

1.1.1. Familia retroviridae.

El VIH-1 se clasifica dentro del género de los lentivirus, perteneciente a la familia de los retrovirus (retroviridae). Actualmente, la familia retroviridae se clasifica en siete géneros (Tabla 1).

Los retrovirus forman un grupo muy heterogéneo de virus. Se caracterizan por tener como material genético ARN de polaridad positiva y un intermediario replicativo de ADN de doble banda, capaz de integrarse en el ADN genómico de la célula huésped. El nombre de retrovirus viene del hecho de que poseen una replicación característica y diferente al resto de los organismos vivos. Esta replicación conlleva un proceso de transcripción inversa, es decir una conversión de ARN en ADN, catalizada por una ADN polimerasa denominada retrotranscriptasa (RT) (Baltimore, 1970, Mizutani y cols., 1970). Otras características importantes que engloban a todos los retrovirus es la de poseer una organización genómica común, con tres genes estructurales *gag*, *pol* y *env*, que codifican para las proteínas estructurales y un ciclo biológico común.

Tabla 1. Clasificación de la familia retroviridae (Murphy y cols., 1995).

Género	Especie Tipo	Especies (ejemplos)
Tipo B de mamíferos	Virus del tumor mamario de ratón	Virus del tumor mamario de ratón
Tipo C de mamíferos	Virus de la leucemia murina	Tipo C de mamífero (VLMMo, VLFe) Tipo C de reptiles (Retrovirus de víbora) Virus de la retículoendoteliosis
Tipo C de aves	Virus de la leucosis aviar	Virus del carcinoma aviar Virus de la leucosis aviar Virus de la mieloblastosis aviar Virus del sarcoma de Rous
Tipo D	Virus de Mason-Pfizer	Virus de Manson-Pfizer Virus de Langur
Retrovirus VBL-VLTH	Virus de la leucemia bovina	Virus de la leucemia bovina Virus linfotrópico-T humano tipo1 Virus linfotrópico-T humano tipo 2 Virus Linfotrópico-T de simio
Lentivirus	Virus de la inmunodeficiencia humana tipo1	Lentivirus bovinos (VIB) Lentivirus equinos (VAIE) Lentivirus felinos (VIF) Lentivirus ovinos y caprino (VAEC) Lentivirus de primates (VIH-1, VIH-2, VIS)
Espumavirus	Espumavirus humanos	Virus sincitial bovino Virus sincitial felino Espumavirus humano Espumavirus simio

1.1.2. Estructura del virión.

El virión es una partícula esférica de entre 80 y 130 nm de diámetro constituida por tres capas concéntricas: interna, intermedia y externa (ver figura 1B). La capa interna, denominada nucleoide o core, tiene forma de cono truncado y está constituida por la proteína de la cápside (p24). En su interior contiene dos copias de ARN genómico, de cadena sencilla, polaridad positiva y de aproximadamente 9.8 Kb, que forman un complejo ribonucleoproteico con las enzimas y las proteínas de la nucleocápside. La capa intermedia, denominada matriz, es de forma icosaédrica y está constituida por la proteína matriz (p17). La capa externa o envuelta está constituida por una membrana lipídica, derivada de la célula que ha infectado y en su superficie se insertan las glicoproteínas gp120 y gp41.

Estas proteínas son traducidas a partir de diez transcritos distintos, generándose un total de 15 proteínas víricas. Las proteínas del VIH-1 son traducidas por la maquinaria celular. Tras la salida del virión inmaduro se produce el procesamiento de los precursores de la nucleocápsida por la proteasa vírica dando lugar al virión infeccioso maduro (ver figura 1C).

1.1.3. Estructura del genoma

Todos los retrovirus contienen 3 genes estructurales denominados *gag*, *pol* y *env* que codifican los componentes estructurales y enzimáticos de las partículas víricas. Además, el genoma del VIH contiene otros 6 fragmentos de lectura abierta solapantes llamados *tat* y *rev* (genes reguladores) y *vif*, *nef*, *vpr* y *vpu* (genes accesorios). En ambos extremos del provirus integrado se localizan dos zonas idénticas denominadas regiones terminales repetidas (LTRs). La unión de diversas proteínas celulares a secuencias de ADN pertenecientes a las LTRs, intervienen en la regulación de la expresión de los genes del provirus.

1.1.3.1. Genes estructurales

gag: la traducción del gen *gag* da lugar a la poliproteína precursora Gag (p55), que una vez procesada, da lugar a la proteína matriz (p17), la proteína de la cápsida (p24) y las proteínas de la nucleocápsida (p6 y p7).

pol: el precursor pol se lee en la fase -1 respecto al precursor de las proteínas de Gag. El precursor pol da lugar a la proteasa (p11), la retrotranscriptasa (RT) (p66/51) y la integrasa (p32).

Durante el proceso de maduración morfológica de la partícula vírica, se producen las rupturas proteolíticas de los precursores de Gag y Pol catalizadas por la proteasa. Por otra parte, la retrotranscriptasa posee actividad ADN polimerasa dependiente tanto de ARN como de ADN y actividad endonucleasa (Rnasa H). Finalmente, la integrasa tiene como función fundamental la de integrar el provirus en el ADN celular actuando como una endonucleasa.

env: el producto primario del gen *env* es una proteína altamente glicosilada denominada gp160. Esta proteína da lugar a la gp120 (proteína de superficie) y a la gp41 (proteína transmembrana), como consecuencia de su procesamiento por proteasas celulares. Ambas proteínas se encuentran unidas en la superficie del virión por interacciones no covalentes y son las responsables de que el virus pueda unirse a la membrana de la célula hospedadora a través de la unión al receptor CD4 y correceptores R5 y X4. En la glicoproteína gp120, se encuentran los sitios de unión de a los receptores celulares (Wyatt & Sodroski, 1998) y la glicoproteína gp41 actúa como anclaje a la membrana lipídica durante el proceso de entrada del virus en la célula.

1.1.3.2. Genes reguladores

Los genes reguladores *tat* y *rev* tienen una gran importancia en la estructura genómica del VIH-1 puesto que las proteínas que codifican, a pesar de su pequeño tamaño, son esenciales para la replicación viral en todas las células a las que infecta.

tat: este gen codifica una proteína de 14 KDa cuya función fundamental es la de aumentar extraordinariamente la replicación del virus al unirse al elemento TAR localizado en el extremo 5' de todos los transcritos del VIH-1 y posteriormente interactuar con factores celulares.

rev: codifica una proteína de 19 KDa que facilita el transporte al citoplasma de los ARN mensajeros mediante su unión a secuencias localizadas dentro del gen *env* denominadas RRE (Rev Responsive Element).

1.1.3.3. Genes accesorios.

Cuando los genomas de los lentivirus de primates fueron caracterizados por primera vez se definieron a los genes *nef*, *vif*, *vpr* y *vpu* (en el VIH-1) o *vpx* (en HIV-2 y SIV) como auxiliares o accesorios puesto que parecían ser innecesarios para la replicación viral en la mayoría de los cultivos celulares estudiados. Actualmente, los denominados genes accesorios, han sido reconocidos como esenciales no solo como factores importantes en la virulencia sino que también juegan un importante papel en las interacciones entre los virus y su hospedador (Trono, 1998).

nef: dentro de los retrovirus, este gen se encuentra solamente en el VIH y en el VIS. Codifica una proteína miristilada de entre 27 y 35 KDa y es producida en grandes cantidades en la fase temprana del ciclo biológico. Se han realizado numerosos estudios que ponen de manifiesto la importancia de Nef en la replicación y en la patogénesis del VIH *in vivo* pudiéndose considerar que Nef es una proteína multifuncional con distintos dominios (Cullen, 1998). Aunque los mecanismos responsables de su importancia *in vivo* no han sido establecidos claramente, si se han descrito un gran número de funciones de Nef *in vitro*:

- Regulación negativa de los receptores CD4 y de la expresión de las moléculas del complejo mayor de histocompatibilidad de clase I (MHC I) en la superficie celular. Estas dos acciones tienen una gran importancia en la patogénesis del VIH-1. La regulación negativa de las moléculas MHC I protege a las células infectadas del ataque de los linfocitos T citotóxicos (CTLs) y de las células natural killer (NK) (Cohen y cols., 1999, Collins y cols., 1998). Mientras que la modulación de los

receptores CD4 contribuye a la inducción del aumento en la infectividad de los viriones (Lama y cols., 1999, Ross y cols., 1999). El sitio de unión a los receptores CD4 se encuentra alrededor de los residuos 57-59, extendiéndose sobre la región proximal del core de Nef.

■ Aumento de la infectividad de las partículas virales.

La participación del gen *nef* en el aumento de la infectividad de los viriones se ha puesto de manifiesto con la comprobación de que los virus con el gen deletado son menos infecciosos (Aiken & Trono, 1995, Miller y cols., 1994, Spina y cols., 1994). Con el objeto de explicar esta observación se han propuesto dos mecanismos totalmente independientes. El primero se debería a la acción de la regulación negativa de los receptores CD4 (Lama y cols., 1999, Ross y cols., 1999), mientras que el segundo se debería a un aumento de la retrotranscripción mediado por *nef* (Aiken & Trono, 1995, Schwartz y cols., 1995)

■ Interferencia en la transmisión de señales intracelulares mediada por estimulación del receptor de células T.

La proteína Nef interactúa con serintreoninas y tirosinquinazas celulares (Baur y cols., 1997, Sawai y cols., 1995) las cuales son componentes importantes en las cascadas de señalización celulares en linfocitos T y macrófagos. Esto permite un aumento en la actividad transcripcional de factores que regulan positivamente la expresión del VIH-1 como el NFAT, NF- κ B y AP1.

Además de estas interacciones, se ha descrito recientemente que Nef induce en los macrófagos infectados la expresión y secreción de la quimioquina MIP-1, que induce la quimiotaxis y activación de los linfocitos T en reposo. Esta activación facilitaría la infección productiva por parte del VIH-1 de los linfocitos T activados (Swingler y cols., 1999).

vif: Este gen se encuentra en todos los lentivirus a excepción de los lentivirus equinos (VAIE). Codifica una proteína de 23 KDa que es sintetizada durante las etapas finales del ciclo biológico del virus y la expresión de su ARN mensajero es dependiente de Rev (Schwartz y cols., 1991). La proteína Vif está fosforilada en los residuos de serinas y treoninas (Yang y cols., 1996) y juega un papel muy importante en la replicación del virus.

Experimentos realizados en monos, con mutantes en el gen *vif* del virus de la inmunodeficiencia de simio (SIV), demostraron que éstos estaban muy atenuados en sus propiedades replicativas y patógenas (Desrosiers y cols., 1998). Esto indicaría que probablemente Vif es fundamental en el establecimiento de la infección y/o en el mantenimiento de la replicación *in vivo*.

Por otra parte, se han realizado numerosos estudios de la función de Vif *in vitro*. Los problemas replicativos observados *in vitro*, que caracterizan a los mutantes defectivos en Vif, dependen de la línea celular que se realice el estudio (Gabuzda y cols., 1992, Sakai y cols., 1993) (Sova & Volsky, 1993) y suelen producir éstos virus un menor efecto citopático. Atendiendo a este criterio se puede hablar de tres fenotipos diferentes cuando las líneas celulares son infectadas con un virus delecionado en el gen *vif*:

- **Células permisivas.** Son aquellas líneas celulares como HeLa, MT-4, C8166 y Jurkat, en las que aparentemente no se observa ningún tipo de defecto en la replicación del virus cuando son infectadas con virus defectivos en Vif
- **Células no permisivas.** Serían aquellas líneas celulares como CMSP, H9 y macrófagos en las que los virus defectivos en Vif son incapaces de replicar.
- **Células semipermissivas.** Son aquellas líneas como SupT1 y CEM en las que los virus defectivos en Vif son capaces de replicar a niveles intermedios.

vpr: Este gen codifica una proteína de 15 KDa la cual se expresa durante las etapas finales del ciclo biológico del virus. Está presente durante las primeras etapas de la infección, puesto que, mediante la interacción con la región C-terminal de la proteína p6, se empaqueta dentro de los viriones que están saliendo de las células infectadas.

Se han descrito dos funciones fundamentales de Vpr *in vitro*:

- Translocación nuclear del complejo de preintegración viral (PIC) que contiene el ADN proviral.

Se cree que la proteína Vpr participa en este proceso conjuntamente con las proteínas virales de la matriz e integrasa y es necesaria para la infección de células que no se dividen (Bukrinsky & Haffar, 1999)

- Induce y para a la célula hospedadora en la fase G2 de su ciclo celular (Poon y cols., 1998).

Vpr inhibe la transición de la fase G2 a la mitosis al impedir la activación de la quinasa celular ciclina B/ p34^{cdc2} que tiene un papel fundamental en este proceso (Emerman, 1996).

vpu: este gen codifica una proteína transmembrana de 16 KDa que tiene dos funciones importantes en el ciclo biológico del VIH-1 :

- Incrementa la salida de las partículas virales de la superficie de las células infectadas.

Esta función depende de la integridad de su dominio transmembrana y de su habilidad para formar oligómeros en el retículo endoplásmico (Paul y cols., 1998, Schubert y cols., 1996).

- Contribuye a la regulación negativa de las moléculas de CD4 de la superficie de las células infectadas.

En un principio se pensó que Vpu contribuía a la degradación de las moléculas de CD4 cuando formaban parte de complejos con el precursor de la proteína Env retenidos en el retículo endoplasmático (Willey y cols., 1992b). Sin embargo, estudios posteriores mostraron una eficiente degradación de las moléculas CD4 en ausencia de Env aunque quedaban retenidas en el retículo endoplasmático. Esto demostraba que la proteína Env no estaba directamente involucrada en la degradación de CD4, pero si actuaba indirectamente al retener estas moléculas en el retículo endoplasmático (Willey y cols., 1992a).

1.1.3.4. Regiones Terminales Repetidas (“Long Terminal Repeat o LTRs”)

Las LTRs del provirus integrado tienen la misma secuencia nucleotídica y se localizan en los extremos 5´y 3´ del VIH-1. En el extremo 3´, parte de la RTR comparte la misma secuencia nucleotídica que aproximadamente la mitad 3´ del gen *nef*. La RTR se divide en tres regiones : U3, R y U5, y contiene cuatro regiones funcionales relacionadas con la

regulación de la transcripción del virus. Dentro de U3 se localizan tres de estas regiones: la región promotora o “enhancer” (- 105 a -79), la región basal o “core promoter”(-78 a -1) y la región moduladora (-454 a -104. Importante : colocar numeración del HXB2), que contiene un elemento regulador negativo (NRE)(-340 a -184) denominado en un principio así porque deleciones en esta región aumentaban la replicación del virus (Rosen y cols., 1985). Estudios posteriores han demostrado que el NRE confiere una regulación positiva y contiene sitios específicos de unión para factores de transcripción celulares tales como USF, AP1, NFAT, hLEF, Ets o RBF-1 y 2 entre otros (Pereira y cols., 2000). Por otra parte, dentro de la región R se encuentra la secuencia de la región de transactivación denominada TAR (+1 a +60). El elemento TAR se une a la proteína vírica Tat y conjuntamente regulan la transcripción de la elongación. La interacción TAR-Tat da lugar a un complejo de estructura ternaria junto con las proteínas celulares.

La región promotora o “enhancer” contiene dos sitios cuyas secuencias están muy conservadas y que unen NFκB, tres sitios para Sp1, un TATA box y un elemento iniciador. La unión del factor de transcripción celular NFκB a los sitios conservados de la LTR, activa la transcripción del VIH-1 en respuesta a un gran número de estímulos extracelulares como proteínas víricas, ésteres de forbol y citoquinas (Gaynor, 1992).

Dentro de la región moduladora, como ya habíamos visto, se encuentra la región NRE, donde se localiza el motivo CAC(A/G)TG denominado E-box. Se ha demostrado que éste motivo es un sitio de unión para el factor de transcripción celular USF (Gregor y cols., 1990, Sawadogo & Roeder, 1985). El USF pertenece a la familia de factores de transcripción “basic helix-loop-helix/leucine zipper” y se encuentra constitutivamente expresado en un gran número de tejidos (Chodosh y cols., 1986). El USF, al unirse al E-box del virus y en cooperación con otro factor de transcripción denominado Ets-1, activa la transcripción del VIH-1 en la mayoría de células T (di Fagagna y cols., 1995, Naghavi y cols., 2001, Sieweke y cols., 1998) y en los macrófagos (Moses y cols., 1994).

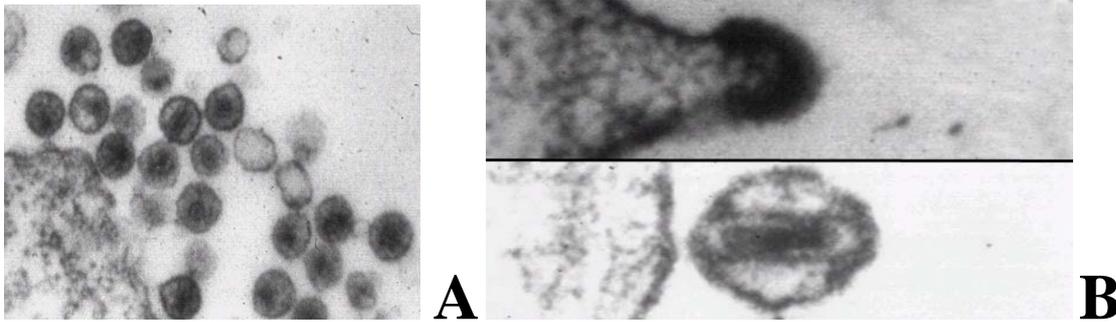
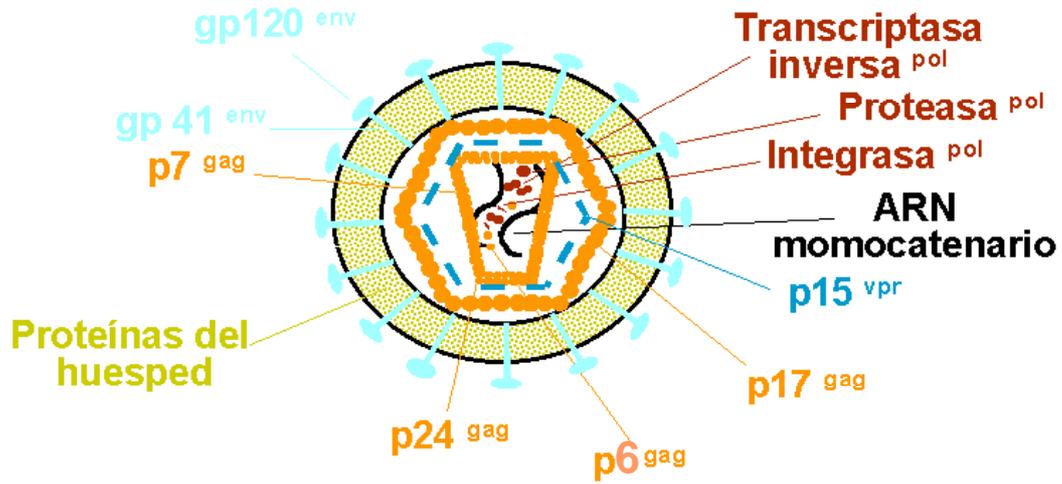


Figura 1.1.A. Microfotografía obtenida con microscopio electrónico de las partículas víricas donde se aprecia la estructura de las mismas.

Figura 1.1.B. Microfotografía obtenida con microscopio electrónico donde se puede apreciar el momento de la salida de una partícula vírica desde la superficie de una célula infectada. En la parte inferior se observa el virión inmaduro.

(Ambas fotografías se han tomado de <http://virología.ua.es/mirror/frame.htm>)

VIH-1



C

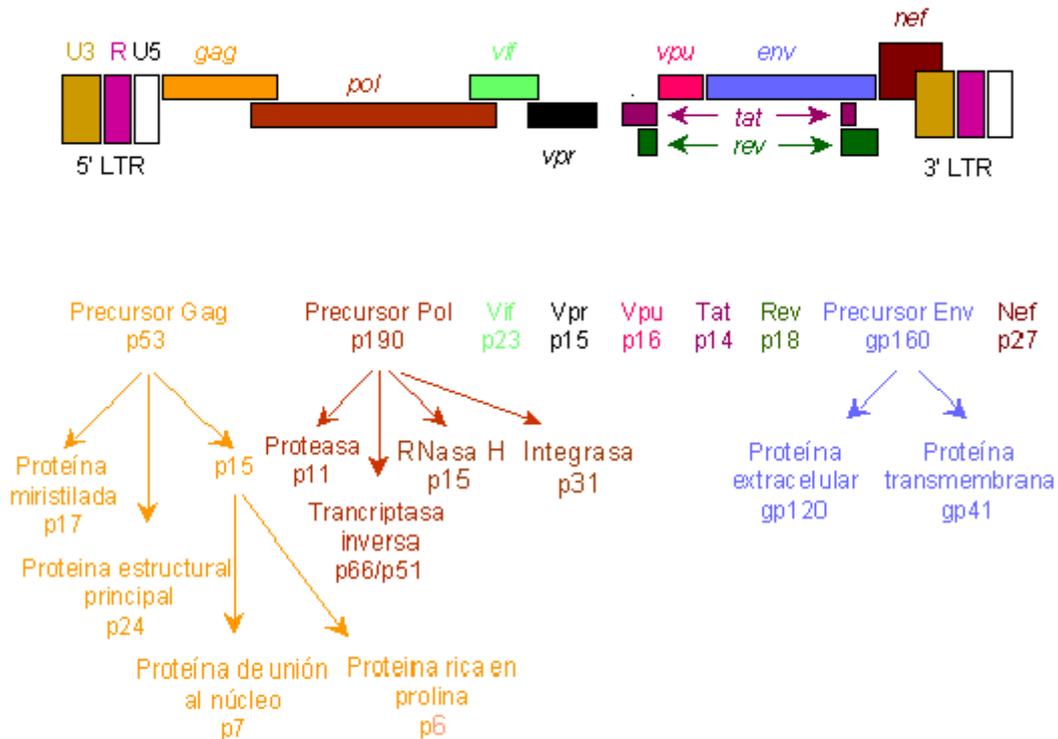


Figura 1.1.C. Estructura del virión, organización genómica y procesamiento de las proteínas del VIH-1. Nájera (1996)

1.2. Ciclo biológico.

El virus se une a la superficie celular mediante la interacción de la glicoproteína gp120 con un receptor y un correceptor. Los principales receptores y correceptores son :

- El receptor principal es la molécula CD4, localizada en la superficie de los linfocitos T colaboradores y en las células de la estirpe monocito-macrófago.

- Los correceptores serían los receptores de quimoquinas X4 (receptor de fusina), cuyo ligando natural es la quimoquina SDF-1 y R5 que tiene como ligandos naturales a las quimoquinas Rantes, MIP-1 α y MIP-1 β (Fauci, 1996).

Tras la interacción entre la gp120 y sus receptores, se produce la fusión entre las membranas celular y viral. Este proceso permite la internalización de la nucleocápside en el citoplasma celular. Posteriormente tiene lugar la desencapsidación del genoma viral y la retrotranscripción por parte de la RT de una de sus hebras de ARN a ADN de doble hebra (Fields BN, 1990). Esta síntesis del ADN , que contiene una RTR en cada extremo, se lleva a cabo en el complejo de preintegración (PIC) y posteriormente el ADN con las LTRs es transportado al núcleo. En el núcleo, el ADN no integrado se encuentra en forma lineal o circular (con una o dos LTRs). La forma lineal es la que se integra en el genoma celular mediante la acción de la integrasa viral, constituyendo el provirus integrado. El genoma del virus integrado se convierte en una estructura estable, que puede permanecer indefinidamente en el genoma de la célula huésped de forma latente, o bien replicándose a niveles bajos o bien producirse una replicación masiva con la consiguiente muerte de la célula infectada.

La iniciación de la transcripción supone la síntesis de un único precursor de ARN que es utilizado como ARNm o como genoma. El ARNm sufre un procesamiento en el extremo 5' de "capping" y en el extremo 3' es poliadenilado. Los transcritos son exportados entonces al citoplasma donde tiene lugar la traducción. Los ARNm procesados codifican las proteínas reguladoras y los ARNm no procesados codifican las proteínas precursoras estructurales.

Posteriormente, tiene lugar el ensamblaje de los viriones, que implica la interacción de los tres dominios del precursor *gag* : en su extremo C-terminal o nucleocápside (NC) con el genoma, en la porción intermedia o cápside (CA) con la región CA de otro precursor y en el

extremo N-terminal o proteína matriz (MA) con la membrana celular. Finalmente adquiere la envuelta y posteriormente se produce la salida de los viriones por gemación. En los viriones inmaduros se produce el procesamiento proteolítico del precursor *gag-pol*, mediado por la proteasa viral, produciéndose de este modo el virión maduro (ver Figura 1.2).

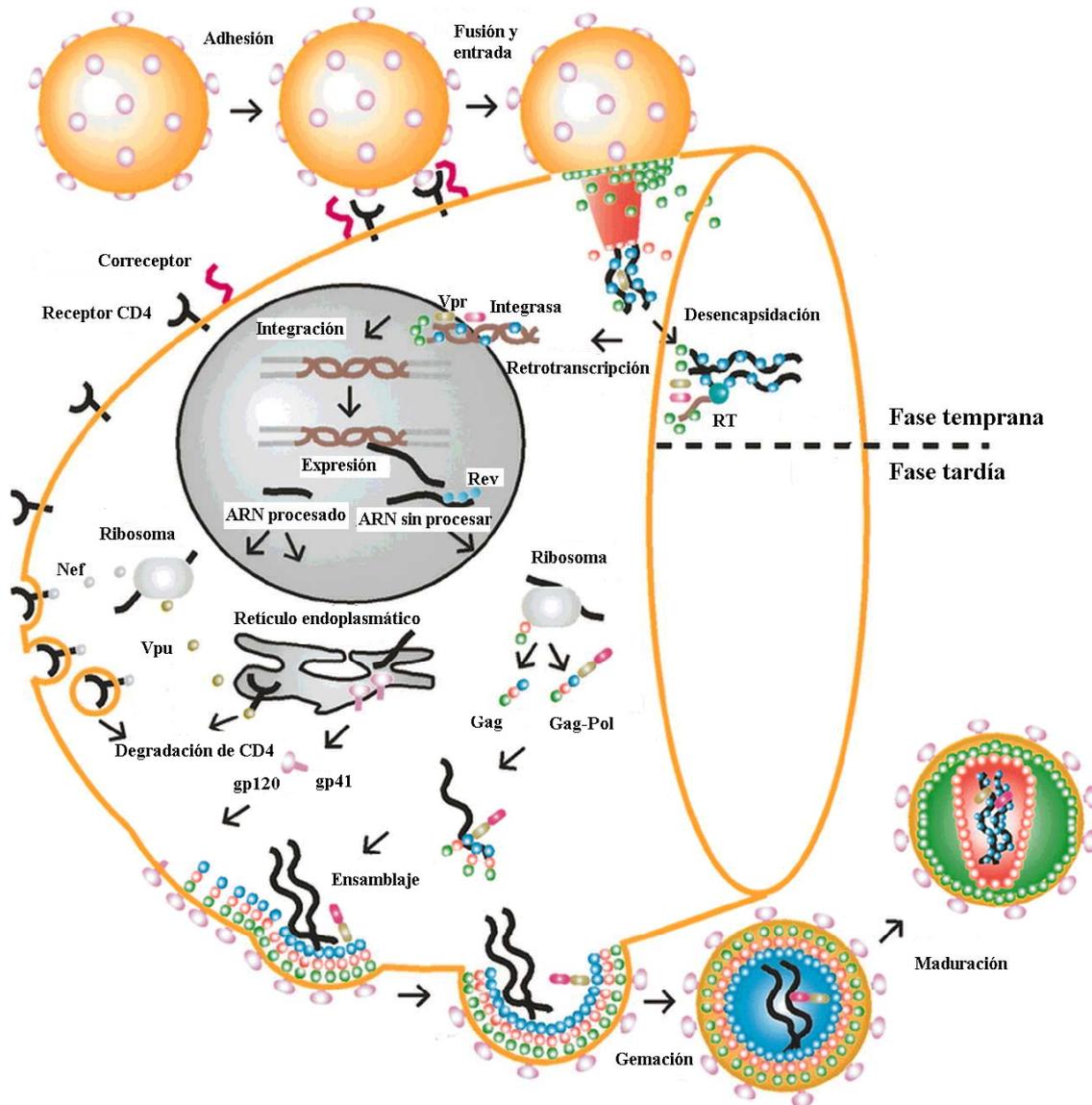


Figura 1. 2. Ciclo replicativo del VIH-1. Modificado de (Turner & Summers, 1999).

1.3. Patogénesis de la infección por el VIH-1.

La patogénesis de la enfermedad que produce el VIH-1 es compleja y heterogénea dependiendo de numerosos factores. El virus, tras su entrada, produce una infección primaria y se expande por los órganos linfoides y otros tejidos del organismo del paciente infectado. Tras dos o tres semanas post-infección, el sistema inmunológico del hospedador produce una gran respuesta mediada por CTLs y anticuerpos que lleva a una bajada en la carga viral.

Sin embargo, la gran variabilidad que posee el virus, le permite continuamente escapar del control inmune y se establece una lucha constante entre la replicación del virus el control del sistema inmune el virus y el sistema inmunológico, produciéndose una infección crónica. Esta fase suele prolongarse durante un período medio de 10 años hasta que se desarrolla la enfermedad clínica o síndrome de inmunodeficiencia adquirida (SIDA) si no se aplica un tratamiento antirretrovírico (Fauci, 1996). Algunos pacientes, sin embargo, son capaces de controlar la infección sin tratamiento y no progresar a SIDA durante más de diez años (no progresores).

Este control de la enfermedad parece estar regulado por un equilibrio entre tres factores:

a) Susceptibilidad genética a la infección por parte del huésped.

Aproximadamente entre un 25-30 % de los no progresores que se han mantenido asintomáticos durante más de 15 años, y no han sido tratados con fármacos antiretrovíricos, presentan formas mutadas de los receptores celulares CCR5 (R5) y CCR2 (Dean y cols., 1996, Smith y cols., 1997).

b) Capacidad de la respuesta inmune por parte del huésped.

A grandes rasgos, parece ser que los no progresores tienen una mejor respuesta inmune específica que los progresores frente al VIH-1. Esta respuesta vendría dada por una mayor actividad tanto de los linfocitos T cooperadores (Rosenberg y cols., 1997) como de los linfocitos T citotóxicos (Greenough y cols., 1997, Harrer y cols., 1996).

c) Presencia de virus con capacidad limitada de replicación.

La infección con formas atenuadas de virus está claramente asociada en algunos casos con una progresión lenta o de la no progresión de la enfermedad. El ejemplo más evidente se ha observado en la denominada cohorte de Australia. A partir de un donante de sangre infectado con un VIH-1 que presentaba una delección importante en el gen *nef* y que era asintomático, se infectaron nueve personas (Deacon y cols., 1995). Ocho de estos pacientes han continuado con una progresión lenta o sin ninguna progresión en el transcurso del tiempo, y todos ellos se caracterizan por tener formas defectivas del gen *nef* del virus (Learmont y cols., 1999, Oelrichs y cols., 1998). También se han descrito en la literatura otros casos de no progresores con deficiencias o alteraciones en el gen *nef* (Kirchhoff y cols., 1995, Rhodes y cols., 2000, Salvi y cols., 1998) o en otros genes tales como *rev* (Iversen y cols., 1995), *gag* (Binley y cols., 1998) y *Vpr* (Wang y cols., 1996). Finalmente, dos trabajos recientemente publicados encuentran una asociación entre no progresores o progresores lentos con la acumulación de mutaciones en los genes accesorios (Yamada & Iwamoto, 2000) y polimorfismos no habituales en diferentes regiones del genoma vírico (Alexander y cols., 2000) en comparación con pacientes que progresan a SIDA.

En el estudio de la patogénesis inducida por la infección con el VIH-1 ha sido fundamental la cuantificación del número de copias de ARN vírico que están presentes en sangre (viremia plasmática) y en tejidos de los pacientes infectados. De esta forma, se han podido cuantificar y proponer modelos sobre algunos de los mecanismos que conducen a la disminución de células T CD4⁺ y al fallo de la terapia. También se ha determinado la existencia de reservorios fundamentalmente en los tejidos linfoides (Chun y cols., 1997) y la posibilidad de erradicar el virus mediante una terapia altamente agresiva (terapia HAART). Sin embargo, por el momento parece ser que esta erradicación no es posible ya que en los pacientes que están recibiendo esta terapia altamente agresiva, aunque la viremia plasmática llega a ser indetectable, puede existir un reservorio de células T CD4⁺ no activadas que se encuentran latentemente infectadas durante un período largo de tiempo y pueden volver a producir partículas virales en respuesta a un estímulo (Finzi y cols., 1997, Wong y cols., 1997). Este reservorio tendría una vida media muy alta (43.9 meses), estimándose que se necesitarían al menos 60 años para la total erradicación del virus de esta población de células (Finzi y cols., 1999).

Mediante la cuantificación del ARN en plasma se ha podido determinar que la replicación del virus es continua a lo largo de toda la infección, incluso durante la fase asintomática que se desarrolla entre la infección primaria y el desarrollo del SIDA (Coombs y cols., 1989, Piatak y cols., 1993). Durante la infección primaria se produce una disminución del número de células T CD4⁺ y se alcanzan niveles muy elevados de virus en plasma del orden de $\geq 10^6$ copias de ARN / ml. Tras la aparición de la respuesta inmune por parte del paciente, se reducen drásticamente los niveles de ARN vírico en plasma y se restablece el número de células T CD4⁺. La principal responsable de esta drástica disminución de virus en sangre es la producción de una fuerte respuesta celular T citotóxica (Koup y cols., 1994). Los niveles de ARN vírico que se alcanzan después de la bajada, permanecen relativamente estables durante toda la fase asintomática de la enfermedad y se ha denominado el “set point”. Este valor procede de la primoinfección y predice en cada individuo la tasa de progresión de la enfermedad. En este período se produce un importante control de la replicación del virus mediante la producción de anticuerpos neutralizantes, linfocitos T citotóxicos, producción de citoquinas y citotoxicidad celular dependiente de anticuerpos mediada por macrófagos y células Natural Killer. Se obtiene por tanto un equilibrio virus-hospedador en el que la producción y eliminación de los virus en el paciente alcanzan valores semejantes [Ho, 1995 #73](Perelson y cols., 1996). En la última fase, cuando empiezan los síntomas de la enfermedad se produce un desequilibrio en favor del virus y esto lleva a una disminución de los linfocitos T CD4⁺ por debajo de las 300 células / μ l, a una disminución de la respuesta T citotóxica específica, elevación de los niveles de ARN en plasma y por consiguiente, la aparición de síntomas de la enfermedad (SIDA)(Figura 1.3).

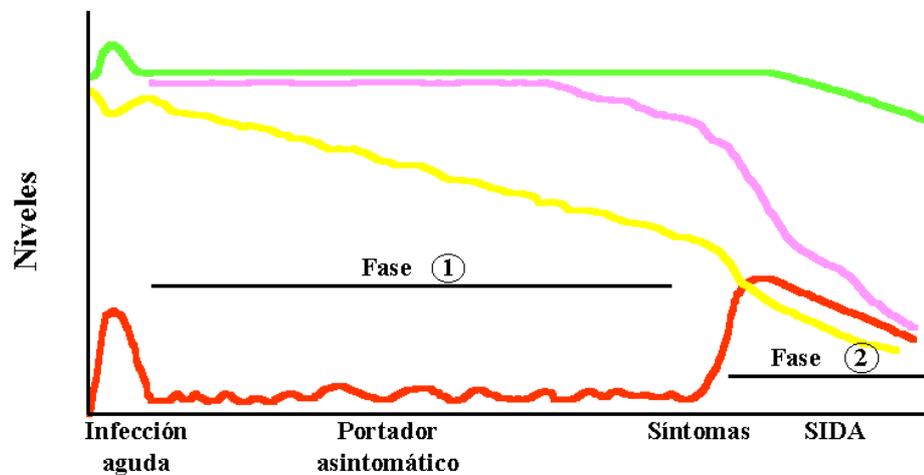


Figura 1.3. Curso clínico de la infección por el VIH-1. La línea roja representa los niveles de virus en sangre (viremia plasmática). La línea amarilla refleja el número de linfocitos CD4⁺. La línea rosa representa el número de linfocitos T citotóxicos (CTL) y la línea verde representa el número de linfocitos CD8⁺. Modificado de (Levy, 1998).

Se ha determinado que la infección por el VIH-1 es un proceso extremadamente dinámico en el que se producen continuas rondas de infección vírica en las células susceptibles del paciente al mismo tiempo que se generan nuevas células T. Se estima que cada día se generan del orden de 10^9 a 10^{10} partículas víricas a la vez que se destruyen entre 10^7 y 10^8 células T CD4⁺. A partir de estos datos se ha estimado que la vida media del virus en sangre es de aproximadamente 6 horas y de 1 a 2 días para las células infectadas productivamente [Ho, 1995 #73](Wei y cols., 1995)

1.4. Persistencia en virología.

1.4.1. Persistencia en virus ARN.

El objetivo primordial de cualquier virus es el de infectar las células del huésped con el objeto de replicar su genoma virus que vayan. Sin embargo, si ésta replicación conlleva la

muerte de la célula huésped, éste mecanismo no es la mejor solución evolutiva. La mejor solución evolutiva es aquella en la que el virus y la célula co-existen durante largo tiempo. La persistencia del virus a nivel general se definiría entonces como la infección en la cual, el virus es capaz de sobrevivir durante un período largo de tiempo en el huésped. Es un fenómeno complejo, en el que se ven involucrados un elevado número de mecanismos que llevan a un equilibrio entre la replicación vírica y la multiplicación de las células del organismo infectado. Por lo tanto, la persistencia es siempre el resultado de interacciones entre factores víricos y del hospedador. Estas interacciones ocurren en cualquier etapa del ciclo replicativo. Entre los virus ARN capaces de persistir después de una infección aguda se encuentran el del sarampión, el de la fiebre aftosa (VFA) y el VIH.

A nivel celular, los virus ARN han desarrollado diferentes estrategias, unos producen infecciones crónicas que no llevan a la muerte celular (persistencia) mientras que otros replican muy rápido, produciendo un gran número de ciclos líticos. La mayoría de los mecanismos estudiados de persistencia en virus ARN en cultivos celulares primarios o en líneas celulares establecidas, conllevan variaciones genéticas del virus, de la célula o de ambos. La muerte celular podría evitarse, y por tanto establecerse la persistencia, mediante la acumulación de genomas defectivos, con la infección de células con baja permisividad al virus o mediante la selección de variantes víricas no citolíticas (Domingo y cols., 1998).

1.4.2. Persistencia en el VIH-1

Dependiendo de la perspectiva de cada autor, se ha definido la persistencia y la latencia de una manera diferente en los artículos publicados. Con el objeto de unificar conceptos, las definiciones de latencia, persistencia y replicación activa del VIH-1 que vamos a utilizar en éste trabajo serán los seguidos por M. Stevenson (Stevenson, 1997)serían:

a) Infección latente. Se puede hablar de:

a.1.) Infección latente proviral. Se definiría como la infección en la cual el ADN (provirus) se encuentra integrado de una forma estable en el genoma celular. La expresión de los genes virales, así como la producción de las proteínas del virus, es mínima o no es suficiente como para producir nuevas partículas infecciosas. Sin embargo, el provirus integrado es capaz de producir sin ningún problema nuevos virus si la célula es activada. Es importante destacar que las condiciones iniciales de la infección necesitan ser las adecuadas para sustentar etapas de la replicación vírica que lleven al establecimiento de un provirus integrado (fusión, desencapsidación, transcripción inversa, transporte al núcleo, integración)

a.2.) Infección latente extracromosómica. Sucede cuando las condiciones en las células diana no son suficientes para sustentar etapas de la replicación vírica completa que lleven al establecimiento de un provirus integrado. Un ejemplo serían los linfocitos T no activados que permitirían la entrada del virus, pero no se producirían de una manera eficaz el resto de las etapas del ciclo replicativo (Bukrinsky y cols., 1991, Gao y cols., 1993, Stevenson y cols., 1990).

La diferencia fundamental entre ambas latencias radica en la estabilidad de el intermedio de replicación. Mientras que en las células no activadas el intermedio de replicación es relativamente lábil, en la latencia proviral el intermedio tiene la misma tasa de recambio que la célula huésped que lleva el provirus.

b) Infección restringida. Se refiere a una infección donde la célula huésped lleva un provirus integrado defectivo y por lo tanto es incapaz de producir virus infeccioso. Se pueden distinguir dos tipos de provirus defectivos:

b.1.) Provirus defectivos silenciosos. Serían aquellos que poseen defectos en las regiones que regulan la transcripción y por lo tanto replican a un nivel muy bajo.

b.2.) Provirus defectivos activos. Aquellos que tienen defectos en genes estructurales y enzimáticos, y aunque la producción de proteínas víricas puede estar mínimamente afectada la progenie de viriones no es infecciosa (por ejemplo defectos en la envuelta que impiden la entrada del virus y defectos en la retrotranscriptasa que restringen la síntesis de cADN en la célula).

c) Infección productiva.

Serían aquellas infecciones en las cuales la célula huésped es completamente permisiva para la replicación y producción de una progenie de viriones. La infección productiva de virus puede catalogarse de infección persistente cuando no existen efectos citopáticos inmediatos en la célula huésped, mientras que en la infección aguda la célula muere tras un período corto de tiempo después de la producción de la progenie de viriones.

Con el objeto de intentar caracterizar y entender mejor los procesos y mecanismos que están implicados en la persistencia del VIH-1, se han realizado diversos estudios *in vitro* que han puesto de manifiesto que la infección de células T CD4⁺ con VIH-1 recombinantes, los cuales tenían mutaciones simples o múltiples en los genes accesorios *vif*, *vpr* y *vpu*, permiten la persistencia vírica y pérdida de citopatogenicidad (Kishi y cols., 1992, Kishi y cols., 1995, Nishino y cols., 1991). Además se ha visto que se acumulan mutaciones y deleciones en los genes accesorios de los virus obtenidos de células persistentemente infectadas tras pases seriados en cultivo (Nakaya y cols., 1996, Nishino y cols., 1994).

1.5. Variabilidad en virus ARN

La variabilidad genética es una característica intrínseca a todos los seres vivos y constituye la base sobre la que actúa la selección natural como mecanismo de evolución. Los virus cuyo material genético está constituido por ARN se caracterizan por presentar una elevada tasa de mutación. Esta tasa de mutación se define como el número de bases incorrectamente incorporadas por nucleótido y ronda de copia de un ácido nucleico. Por término medio, la tasa de mutación de los virus ARN es de 10^{-3} - 10^{-5} sustituciones por nucleótido y ciclo de replicación (Drake, 1993) y como consecuencia se produce una elevada

heterogeneidad en las poblaciones de virus ARN. Esta gran heterogeneidad se describió por primera vez en poblaciones del bacteriófago Q β (Domingo y cols., 1976, Domingo y cols., 1978). Tras un análisis sistemático de clones individuales de poblaciones de virus obtenidas a partir de una sola partícula vírica infecciosa, se observó que éste bacteriófago estaba constituido por una mezcla compleja de distintos variantes que presentaban como media de 1 a 2 mutaciones entre cualquier genoma viable del fago y la secuencia promedio de la población. Estudios posteriores con otros virus ARN, tales como, el virus de la fiebre aftosa (VFA) (Domingo y cols., 1985) el virus de la estomatitis vesicular (VSV) (Steinhauer & Holland, 1987) o el VIH (Groenink y cols., 1991, Meyerhans y cols., 1989) han demostrado que los virus ARN no tienen una estructura genómica fija, sino que están constituidos por una mezcla más o menos compleja de mutantes relacionados entre sí, diferenciándose entre ellos en uno o más nucleótidos. Esta distribución de mutantes recibe el nombre de cuasiespecie por analogía con el modelo teórico descrito por Eigen (Eigen, 1971). Una cuasiespecie se caracteriza por una secuencia maestra y una secuencia consenso. La secuencia maestra es aquella que se encuentra más representada dentro de la cuasiespecie, mientras que la secuencia consenso es un promedio entre todas las secuencias que componen la cuasiespecie.

1.5.1. Variabilidad en el VIH-1

1.5.1.1. Variabilidad genética.

El VIH-1 se caracteriza, como el resto de virus ARN, por su alta variabilidad genética. Este hecho hace que la forma más adecuada de representar una población viral del VIH-1 es como una colección de genomas relacionados pero que presentan mutaciones y que se denomina, como anteriormente habíamos visto, estructura de cuasiespecie. Los estudios realizados sobre la población vírica presente en un individuo infectado, han demostrado que la población sigue una distribución variable, tanto en el tiempo (Delwart y cols., 1997, Delwart y cols., 1994, Wolinsky y cols., 1996) como en el espacio (Delassus y cols., 1992, Korber y cols., 1994).

La variabilidad del VIH-1 no es uniforme a lo largo del genoma ya que existen unas regiones más conservadas que otras. La región más conservada es el gen *pol*, mientras que la

presenta un mayor grado de variabilidad genética sería el gen *env* (Hahn y cols., 1985, Quinones-Mateu y cols., 1996). Dentro del gen de la envuelta, y concretamente en la gp120, se definieron 5 regiones hipervariables (V1-V5) y 5 regiones constantes (C1-C5) que separaban las regiones hipervariables entre sí y cuyo grado de variabilidad era mucho menor. Posteriormente se ha comprobado que el grado de variación en las regiones constantes C2 y C3 es mayor que el observado en la región V3 (Engelstad, 1996).

1.5.1.2. Variabilidad fenotípica.

La variabilidad genética que presenta el VIH-1 tiene su repercusión en las propiedades fenotípicas de los aislados de virus obtenidos de pacientes infectados. Una de las más importantes es su tropismo celular, es decir, la capacidad que tiene el virus de infectar distintos tipos celulares. Todos los aislados tienen la capacidad de replicar en CMSP, aunque algunos replican mejor en unas líneas celulares que en otras. En general, se definirían como virus T-trópicos a aquellos aislados que replican mejor en líneas T transformadas y virus M-trópicos a los que replican mejor en células monocito-macrófago. Los virus T-trópicos, replican más rápidamente y a niveles altos en CMSP (aislados “rapid-high”) e inducen la formación de sincitios en la línea celular MT-2. Por éste motivo se les suele denominar aislados inductores de sincitios o IS. Los virus M-trópicos, replican más tardíamente y a niveles más bajos en CMSP (aislados “slow-low”) y no son capaces de inducir la formación de sincitios en la línea celular MT-2, por lo que se les denomina aislados no inductores de sincitios o NIS (Cheng-Mayer y cols., 1988, Fenyó y cols., 1988, Schwartz y cols., 1989, Tersmette y cols., 1988).

La variabilidad fenotípica juega un papel importante en la patogénesis de la infección. En la primoinfección, así como en el período asintomático, los virus M-trópicos NIS son, en general, mayoritarios. Alrededor del 50 % de los pacientes mantienen este fenotipo a lo largo de la infección, mientras que en el otro 50 % de los individuos infectados se produce un cambio de fenotipo que lleva a la aparición de virus T- trópicos IS más virulentos. Este cambio de fenotipo desde virus NIS a virus IS se asocia a una rápida disminución del número de linfocitos T CD4⁺ y a una progresión de la enfermedad, junto con un menor tiempo de supervivencia tras el diagnóstico del SIDA (Lukashov & Goudsmit, 1998, Tersmette y cols., 1988)

Estudios sobre la capacidad replicativa de clones moleculares quiméricos obtenidos a partir de aislados T-trópicos y M-trópicos indican que el tropismo celular reside fundamentalmente en la región V3 de la gp120 del gen *env*, aunque parece que también son importantes las regiones V1 y V2 de la proteína gp120 (Carrillo & Ratner, 1996, Olivares y cols., 1997, Shioda y cols., 1991). La diferencia de tropismo observada entre los aislados del VIH-1 refleja una utilización diferencial de los correceptores. Mientras que los virus M-trópicos utilizan fundamentalmente para su entrada el correceptor R5, los virus T-trópicos pueden utilizar tanto el correceptor R5 como el X4 (Berger, 1997, Hoffman & Doms, 1998). Por otra parte, las cepas que utilizan R5 pueden adaptarse y utilizar el X4 tras pases seriados en células T CD4⁺ que no expresan habitualmente R5 (Harrowe & Cheng-Mayer, 1995, Klasse y cols., 1996) (Dejudicq y cols., 2000). Este cambio *in vivo* es relativamente raro, ya que como habíamos visto, solamente el 50 % de los pacientes pasa de tener mayoritariamente virus NSI a tener virus IS. Parece ser que *in vitro*, algunas cepas de virus, necesitan pocos cambios de aminoácidos para pasar de utilizar el correceptor R5 al X4 ya que se adaptan rápidamente, mientras que otras necesitan más pases en cultivo y posiblemente requieran más cambios (Dejudicq y cols., 2000)

1.6. Epidemiología molecular de la infección por el VIH-1.

La elevada diversidad genética que presenta el VIH-1 es el resultado de un conjunto complejo de factores. El primer factor que contribuyó a esta elevada diversidad genética fueron las múltiples introducciones de virus genéticamente muy distintos de primates a humanos (Hahn y cols., 2000). Estas transmisiones zoonóticas tuvieron dos orígenes diferentes. Parece ser que el VIH-1 procedería de la transmisión de un VIS del chimpanzé de la subespecie *Pan Troglodytes Troglodytes* al hombre (Gao y cols., 1999), mientras que el VIH-2 tendría su origen en un del mangabey tiznado (*Cercocebus atys*) (Hirsch y cols., 1989) al hombre respectivamente. Desde esa primera introducción en los humanos, la diversidad ha ido aumentando como consecuencia de la acumulación de un gran número de mutaciones, debido a la alta tasa de error que comete la transcriptasa inversa (Preston y cols., 1988) al recambio rápido de viriones que se produce en los individuos infectados (Ho y cols., 1995) a la recombinación (Robertson y cols., 1995) y a la presión selectiva que sufren los virus como

consecuencia de la respuesta del sistema inmunitario a la infección por el VIH-1 (Fauci, 1996).

La recombinación entre genomas víricos es uno de los factores que influyen en la variabilidad del VIH-1. La transcriptasa inversa es altamente recombinogénica (Hu & Temin, 1990) y puede recombinar genomas genéticamente muy diferentes entre sí para dar lugar a nuevas variantes víricas.

1.6.1. Clasificación y distribución de subtipos.

La gran variabilidad genética que presenta el VIH-1, queda reflejada en los análisis filogenéticos de las secuencias nucleotídicas que se han realizado con numerosos aislados de distintas procedencias geográficas. Esto ha llevado a clasificar el VIH-1 en tres grandes grupos: M (“main”), O (“outlier”) (Gurtler y cols., 1994) y N (“new”) (Simon y cols., 1998). El grupo M o principal incluye la mayor parte de los VIH-1 aislados en el mundo y, dentro de él, se pueden distinguir al menos once subtipos diferentes (A-K) (Janssens y cols., 1997, Triques y cols., 2000) genéticamente equidistantes entre sí. El subtipo B circula mayoritariamente en Europa Occidental, Estados Unidos y Australia (Casado y cols., 2000, Chaix y cols., 1993, Hu y cols., 1996). Los subtipos A, C, D, F1, F2, G, H, J y K se encuentran fundamentalmente distribuidos en África, mientras que el subtipo E (descrito recientemente como una forma circulante recombinante o CRF01 AE) está localizado principalmente en Tailandia (Ichimura y cols., 1994, Ou y cols., 1993) (ver figura 1.4). De todos los subtipos descritos hasta ahora, el subtipo C es el más prevalente, siendo responsable de aproximadamente el 50 % de las infecciones por el VIH-1 en el mundo (Anonymous, 2000). Los virus pertenecientes a éste subtipo C son mayoritariamente NSI y además tienen preferencia por el correceptor R5. Otra característica que distingue a éste subtipo de los otros es la presencia tres o más motivos NF_kB en la LTR en lugar de los dos que poseen el resto de los subtipos. Éste sitio extra parece ser que está asociado con un aumento en la actividad promotora (Montano y cols., 2000, Naghavi y cols., 1999). Por otra parte, el grupo O incluye un pequeño, aunque creciente, grupo de aislados que se caracterizan por tener una divergencia muy alta con respecto al grupo M. Ambos grupos sólo comparten aproximadamente el 50 % de homología en la secuencia de aminácidos en el gen *env* (Gurtler y cols., 1994, Vanden

Haesevelde y cols., 1994). El grupo O parece ser endémico en Camerún y en los países vecinos del África Occidental no observándose agrupamientos filogenéticos similares a los que establecen los diferentes subtipos dentro del grupo M (Loussert-Ajaka y cols., 1995). Finalmente, el grupo N está actualmente representado solamente por tres secuencias nucleotídicas obtenidas de pacientes originarios de Camerún cuyos sueros no reaccionaban frente a péptidos sintéticos específicos de los subtipos M y O, pero que si lo hacían frente a péptidos específicos del virus de la inmunodeficiencia simia de chimpancés (VIScpz).

Con respecto al VIH-2, existen muy pocas secuencias nucleotídicas en las bases de datos. Se ha descrito mediante análisis filogenéticos, que al igual que en el VIH-1, existe una gran diversidad y se han propuesto 6 subtipos equidistantes filogenéticamente unos de otros (A-F) (Chen y cols., 1997, Gao y cols., 1994). Recientemente, se ha propuesto un nuevo subtipo, aislado de un paciente asintomático de Costa de Marfil, denominado G (Yamaguchi y cols., 2000).

La variabilidad genética entre aislados del VIH presenta un rango muy amplio. Los aislados del VIH-1 y del VIH-2 difieren un 50 % en su secuencia nucleotídica, mientras que la divergencia entre aislados de los grupos M y O se reduce al 35 %. Los valores de divergencia que se obtienen si comparamos secuencias nucleotídicas entre subtipos dentro del grupo M oscilan entre el 25-30 %, y dentro de un mismo subtipo, la divergencia alcanza valores entre el 5 % y el 20 %. Por último, dentro de un mismo individuo, se han descrito valores de divergencia entre los miembros de la cuasiespecie que alcanzan valores entre el 2 % y el 10 % (Myers y cols., 1996). A la vista de estos datos, y aunque exista una elevada variabilidad intrasubtipo, así como una continua evolución de los virus, los datos actuales indicarían que cada subtipo ha sido capaz de mantener su propia identidad a lo largo de la epidemia, siendo improbable la desaparición de estas entidades en el futuro (Lukashov & Goudsmit, 1997).

La clasificación en subtipos se ha venido realizando mediante el estudio de las relaciones filogenéticas entre los aislados en los genes *gag* y *env*. En algunos aislados se da la discordancia entre las relaciones filogenéticas inferidas en un gen con respecto al otro, es decir, que un mismo aislado se agrupa en un gen con un subtipo diferente al que se agrupa en el otro. Al haber aumentado notablemente el número de secuencias de las que se disponen, y

al tener un mayor número de genomas completos, se ha puesto de manifiesto la existencia de virus “mosaico”, entendiendo por tales, aquellos virus en los que las relaciones filogenéticas establecidas en diferentes partes del genoma son discordantes. Estos virus son el resultado de distintos eventos de recombinación que se han producido entre dos virus que han infectado al mismo individuo (Korber y cols., 1997). A estos virus “mosaico” se los denomina formas circulantes recombinantes (CRFs) y hasta la fecha han sido descritas 9 CRFs.

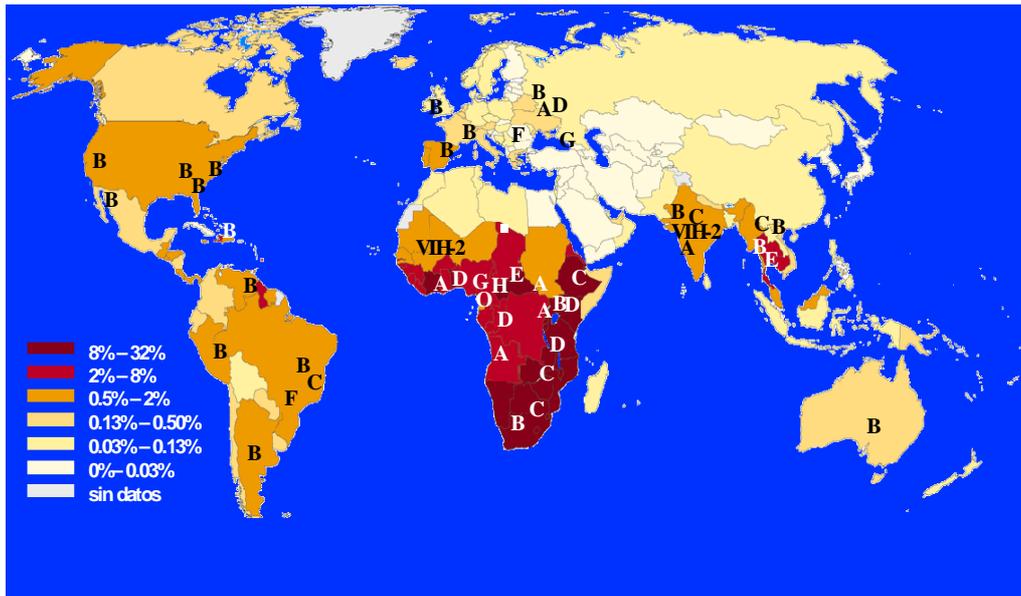


Figura 1.4. Distribución de casos y subtipos circulantes del VIH en el mundo. Los diferentes colores indican el porcentaje estimado de adultos infectados en cada región. Las letras mayúsculas reflejan los subtipos que circulan mayoritariamente en cada región (Anonymous, 1998).

1.6.2. Situación actual a nivel mundial de la epidemia de VIH/SIDA.

La pandemia causada por el VIH es uno de los problemas más graves relacionado con las enfermedades infecciosas que tiene el mundo. La pandemia sigue expandiéndose por los cinco continentes (ver Figura 1.4), y el número de personas infectadas en el año 2000 ascendió a 5,3 millones de los que 3,8 millones se habían producido en el África Subsahariana, según estimaciones del Programa Conjunto de las Naciones Unidas sobre el VIH/SIDA (ONUSIDA)(Anonymous, 2000). De acuerdo con este Programa, se ha estimado que a finales del año 2000, más de 36 millones de personas vivían con el VIH, habiendo fallecido en ese mismo año 3 millones como consecuencia de la enfermedad. El número total de defunciones desde el principio de la epidemia hasta finales de 2000 ha sido de 21,8 millones.

Sin lugar a dudas, el continente más afectado por la epidemia es África, y concretamente el África Subsahariana. En países como Sudáfrica, el VIH/SIDA se ha convertido en un problema muy grave ya que la mayor parte de las nuevas infecciones que se producen en el África Subsahariana afectan a este país. En otros países del entorno la situación tampoco es mejor, ya que en países como Zinbabwe, Bostwana, Namibia o Swazilandia al infección llega a afectar entre el 20 % - 26 % de su población adulta urbana.

La aparición del VIH en Asia fue tardía en comparación con otros continentes, pero se está propagando rápidamente en Asia meridional y sudoriental entre los usuarios de drogas por vía parenteral (UDVP) y entre los profesionales del sexo. En amplias zonas de la India y de China se agrava el tema por el gran flujo migratorio que están sufriendo, sobre todo China, de un región a otra. En contraposición a esta situación, Asia Oriental y el Pacífico conservan unas tasas bajas a pesar de tener un volumen muy elevado de población.

En América Latina, el Caribe y Europa Oriental el patrón de propagación es muy similar al de los países desarrollados afectando especialmente al grupo de homosexuales y al de UDVP. Hay que destacar el gran aumento de la Federación Rusa ya que durante el año 2000 se han producido un mayor número de nuevas infecciones por el VIH que en el conjunto de todos los años previos de la epidemia.

Finalmente, en Europa Occidental, Oceanía y en América del Norte las tasas de infección por el VIH se van reduciendo, centrándose las nuevas infecciones en pacientes heterosexuales y en UDVP. La mortalidad asociada a la enfermedad está disminuyendo por la utilización de las nuevas combinaciones de los fármacos antirretrovíricos.

1.6.3. Situación de la infección VIH/SIDA en España.

España se sitúa a la cabeza de Europa en el número de casos de SIDA notificados y el segundo, tras Portugal, en tasa de incidencia. Desde 1981, año en que se detectó el primer caso, se han acumulado un total de 59.466 casos de SIDA, de los que un 53,3 % han fallecido. Durante el año 2000 se diagnosticaron en España 2.511 casos nuevos de SIDA, disminuyendo en un 13 % el número de nuevos diagnósticos con respecto al año anterior. El número de casos nuevos ha disminuido un 64 % en los últimos 5 años debido tanto a la prevención y asistencia sanitaria, como a la aplicación de los nuevos tratamientos antirretrovíricos. Se estima que los infectados en España ascienden hasta los 120.000.

La epidemia de SIDA en nuestro país parece ser el resultado de la suma de varias subepidemias determinadas por los distintos mecanismos de transmisión, originándose a partir de una introducción principal y produciéndose una posterior diversificación del virus en pacientes de distinto origen geográfico y distintos grupos de riesgo (Casado y cols., 2000, Castilla & Gutierrez, 1997). La vía de transmisión más frecuente de la infección ha sido el compartir material de inyección para la administración parenteral de drogas en un 65 % de los casos. En segundo lugar estaría la transmisión heterosexual con un 15, y en el tercer lugar tendríamos la transmisión homosexual masculina con el 15 %. Nuestro país también es el país de Europa con más casos de SIDA notificados por transmisión madre-hijo, aunque con la aplicación de la terapia antirretrovírica a las embarazadas, la incidencia ha disminuido significativamente.

2.OBJETIVOS

El objetivo de este trabajo es el estudio de la variación de los genes accesorios y su implicación en la biología del VIH-1.

Para ello hemos llevado a cabo dos tipos de estudios:

2.1. Caracterización de muestras de ONUSIDA cocultivo negativo mediante un análisis filogenético en las regiones p17 (*gag*) y C2-V3-C3 (*env*).

2.1.1. Análisis del gen accesorio *nef* como posible responsable del fenotipo cocultivo negativo.

2.2. Estudio de la persistencia del VIH-1 *in vitro* a través del análisis del genoma de los provirus obtenidos de tres líneas celulares persistentemente infectadas con el mismo aislado s61.

2.3. Estudio de la implicación de los genes accesorios en el establecimiento de la persistencia en las tres líneas celulares.

3. MATERIALES Y MÉTODOS

3.1. Cultivos celulares.

3.1.1. Células.

Las líneas celulares utilizadas a lo largo del estudio han sido:

MT-4. Línea celular establecida de la serie linfoide T obtenida por cocultivo de linfocitos de sangre periférica de un paciente infectados con el virus de la leucemia de células T humana tipo I (HTLV- I) con linfocitos de cordón umbilical de un individuo sano (Miyoshi y cols., 1981). Esta línea fue cedida por el Dr. D.D. Richman.

H9. Línea celular establecida de la serie linfoide T derivada de las HUT78 (Popovic y cols., 1984). Cedida por Dr. R Gallo, Baltimore

Cos-1. Línea celular establecida procedente de células de riñón de mono.

U-937 clon 2. Es una línea celular promonocítica derivada de un linfoma histiocítico (Sundstrom & Nilsson, 1976).

3.1.2. Medios de cultivo.

Todas las líneas celulares utilizadas se cultivaron en medio RPMI 1640 (Bio-Whittaker) suplementado con suero bovino fetal (SBF) al 10 % (Gibco), L-glutamina 2 mM, penicilina 100 unidades/ml, estreptomycin 100 mg/ml (Bio-Whittaker) y tilosina al 0.5 % (ICN Biomedicals Inc.).

3.1.3. Cepas de virus utilizadas.

S61. Semilla obtenida por cocultivo de linfocitos de sangre periférica de un paciente español (61/89) con linfocitos de donante sano activados con interleuquina 2 (IL-2) (Sanchez-Palomino y cols., 1993). Esta semilla fue posteriormente subcultivada de nuevo con linfocitos de sangre periférica. Esta segunda semilla ha sido la utilizada en el proceso del establecimiento de la persistencia.

VM61, VH61 y VU61. Virus obtenidos de tres líneas celulares (MT-4, H9 y U-937 clon 2) persistentemente infectadas con la semilla S61.

3.2. Muestras del VIH-1.

3.2.1. Aislados naturales

En este estudio se han analizado 35 muestras, 27 cocultivo negativo y 8 cocultivo positivo como controles obtenidas a través del “Who Network for Isolation and Characterization” (ONUSIDA). Estas muestras procedían de países de diferentes continentes donde el subtipo C es prevalente. 8 muestras procedían de Brasil (BR), 7 de China (CN), 12 de India (IN), 5 de Sudáfrica (ZA) y 3 procedían de Israel (IS). La mayor parte de ellas fueron obtenidas de pacientes con baja carga viral y que no habían recibido tratamiento antirretrovírico. Estas muestras se obtuvieron entre 1997 y 1998 y se recibieron 48 después de su extracción en el laboratorio Georg Speyer Haus del Dr. Hagen von Briesen (Frankfurt, Alemania). Posteriormente, se sometieron a procedimientos estandarizados de aislamiento. Tras 4 semanas de cocultivo en CMSP, de 27 de éstas muestras no pudieron aislarse ningún virus (cocultivo negativo).

3.2.2. Provirus obtenidos de células persistentemente infectadas por el aislado s61.

Se han analizado los provirus obtenidos de tres líneas celulares persistentemente infectadas (H9, MT-4 y U-937 clon 2) con el mismo aislado s61. Se ha realizado la secuenciación del genoma completo de los provirus VH61, VM61, VU61 y del aislado inicial s61.

3.3. Obtención de ácidos nucleicos.

3.3.1. Obtención de ADN.

3.3.1.1. Obtención de ADN de las muestras de ONUSIDA.

El ADN procedente de las muestras del Programa de las Naciones Unidas para el SIDA (ONUSIDA) fue extraído con el equipo comercial QIAmp Blood kit (Qiagen), directamente de las CMSP de los pacientes o de cocultivo, en el laboratorio del Dr. Hagen von Briesen (Georg-Speyer-Haus, Frankfurt, Alemania) y remitidas a nuestro laboratorio posteriormente.

3.3.1.2. Obtención de ADN de las células persistentemente infectadas.

Para la obtención del ácido desoxirribonucleico (ADN) de las células persistentemente infectadas se partió de aproximadamente 5×10^6 células de cada línea y se siguió un método de extracción estándar con fenol (Perucho y cols., 1981). Para ello las células se resuspendieron en 1 volumen de solución (Tris 10 mM pH 7,5, EDTA 10mM, NaCl 0.15 M) a la que se adicionó 200 mg/ml de proteinasa K y SDS al 0.4%. Tras 1 hora de incubación a 65°C y toda la noche a 37°C se añadió 1 volumen de solución Tris 10mM pH 8, EDTA 10mM, NaCl 0.65 M y se realizaron tres extracciones con fenol:cloroformo:alcohol isoamílico (25:24:1) y una con cloroformo:alcohol isoamílico (24:1). El ADN contenido en la fase acuosa se precipitó con 2 volúmenes de etanol absoluto y se resuspendió en agua destilada estéril, dejándose a una concentración aproximada de 1mg/ml. La lectura de la concentración se realizó en espectrofotómetro a una longitud de onda de 260 nm, estimando que una unidad de densidad óptica equivale a una concentración de 50 mg/ml de ADN de doble banda.

3.3.2. Obtención de ARN

Para la obtención del ácido ribonucleico (ARN) se empleó el método descrito por Boom (Boom y cols., 1990). El método consiste en la incubación a temperatura ambiente durante 10 minutos de 20 μ l de sobrenadante de cultivo con una mezcla de 980 μ l de solución de lisis (120 g de isotiocianato de guanidina (Fluka), 100 ml de Tris-HCl 0.1 M pH 6.4, 22 ml de EDTA 0.2 M pH 8 y 2.6 g de Triton X-100) y 40 μ l de sílica. Por centrifugación a 14.000 r.p.m. durante 15 segundos se obtuvo un sedimento de partículas de sílica que se lavó dos veces con 1 ml de solución de lavado (120 g de isocianato de guanidina y 100 ml de Tris-HCl 0.1 M pH 6.4), dos veces con 1ml de etanol al 70% y una vez con 1 ml de acetona. Tras el último lavado se incubó la preparación para secar la sílica a 56°C durante 10 minutos. Finalmente se eluyó el ARN extraído en 125 μ l de agua destilada estéril incubando a 56°C durante 15 minutos y posterior centrifugación de la mezcla a 14.000 r.p.m. durante 2 minutos.

3.4. Reacción en cadena de la polimerasa.

3.4.1. A partir de ADN.

Se llevaron a cabo tres perfiles de amplificación:

a) Todas las reacciones de amplificación por PCR de los genomas completos se llevaron a cabo en tampón Tris-HCl 10 mM pH 8.3, KCl 50 mM, gelatina 0.01% y MgCl₂ de 1,5 mM. La concentración de desoxinucleótidos trifosfato (dNTPs) (Pharmacia) fue 0.125 mM, empleándose 100 ng de cada iniciador y 1 unidad de enzima Ampli Taq polimerasa (Perkin - Elmer) por ensayo. El volumen de reacción fue de 50 μ l y los perfiles de amplificación y los iniciadores usados dependieron de la región genómica a amplificar (ver tabla iniciadores). Los genomas completos se amplificaron en varios fragmentos grandes con una primera PCR. Para estos fragmentos, en su mayoría mayores a 1500 pares de bases, se realizaron 35 ciclos de amplificación con una desnaturalización a 94°C durante 1 minuto, hibridación durante 1 minuto entre 50°C y 60°C (dependiendo la temperatura de los iniciadores empleados) y elongación de las cadenas durante 2 minutos a 72°C. Posteriormente

se realizaron sucesivas amplificaciones internas de estos fragmentos grandes. Para estos fragmentos menores a 1500 pares de bases, se realizaron 35 ciclos de amplificación con una desnaturalización a 94°C durante 30 segundos, hibridación de los iniciadores al ADN molde durante 1 minuto entre 50°C y 60°C (dependiendo esta temperatura de hibridación de los iniciadores empleados) y elongación de las cadenas durante 1 minuto a 72°C. En ambos casos, las amplificaciones se iniciaron con una desnaturalización de 5 minutos a 94°C y finalizaron con una elongación de 10 minutos a 72°C.

b) La amplificación de las regiones genómicas p17 del gen *gag* y C2V3C3 del gen *env*, de la mayor parte de las muestras de ONUSIDA, se llevaron a cabo en dos amplificaciones consecutivas, externa e interna puesto que el ADN a amplificar provenía de CMSP y eran cocultivo negativo. Las reacciones de amplificación se llevaron a cabo en tampón 10 mM Tris pH 8.3, 25 mM NaCl, 25 µM EDTA pH 8.0 y MgCl₂ de 20mM. La concentración de desoxinucleótidos trifosfato (dNTPs) (Pharmacia) fue 2 µM, empleándose 30 ng de cada iniciador y 0.5 unidades de enzima Ampli Taq polimerasa (Perkin - Elmer) por ensayo. El volumen de reacción fue de 15 µl en la primera ronda de amplificación y de 19 µl en la amplificación interna. Las condiciones tanto para la primera PCR (externa) como para la segunda (interna) fueron 40 ciclos de amplificación con una desnaturalización a 96°C durante 1 minuto y 65°C como temperatura de hibridación y elongación durante 5 minutos.

c) Para la amplificación del gen *nef*, así como para la amplificación de algunas muestras de ONUSIDA en las regiones p17 y C2V3C3, se llevaron a cabo dos amplificaciones consecutivas, externa e interna. Se realizaron 35 ciclos de amplificación con una desnaturalización a 94°C durante 1 minuto, hibridación durante 1 minuto a 55°C y elongación de las cadenas durante 2 minutos a 72°C para la amplificación externa, y 35 ciclos de amplificación con una desnaturalización a 94°C durante 30 segundos, hibridación durante 1 minuto a 55°C y elongación de las cadenas durante 1 minuto a 72°C para la amplificación interna. En ambos casos, las amplificaciones se iniciaron con una desnaturalización de 5 minutos a 94°C y finalizaron con una elongación de 10 minutos a 72°C.

Todas las reacciones de amplificación se realizaron en los termocicladores Perkin - Elmer Cetus modelo 480 y MJ Research modelos PTC-100 y PTC-150. Los productos obtenidos se analizaron por electroforesis en geles de acrilamida al 6% (acrilamida

/bisacrilamida 30:0.8 polimerizada con persulfato amónico y Temed) en TBE (Tris- borato 0.09 M, EDTA 0.002 M) en condiciones no desnaturizantes a 200-220 V durante 1.5 a 2 horas y se visualizaron por transiluminación con luz ultravioleta tras tinción con bromuro de etidio 10 mg / ml.

3.4.2. A partir de ARN (RT- PCR).

En los casos en los que el material de partida fue ARN en lugar de ADN, se realizó un transcripción inversa para la obtención de cADN mediante la utilización del equipo comercial Titan RT-PCR system (Boehringer Mannheim). Para ello se realizó en un único paso la retrotranscripción y la primera amplificación. Se añadieron 10-20 µl del ARN, obtenido según el apartado 3.3.2., a una solución que contenía una mezcla de las enzimas Tth ADN polimerasa y Expand High Fidelity, tampón (Boehringer Mannheim), 20 unidades de Rnasina (Promega), 8.3 mM de dithiotreitol (DTT), 200 ng de cada uno de los iniciadores y 133 mM de dNTPs. Esta mezcla se llevó a un volumen final de 50 µl con agua destilada estéril. La reacción de RT-PCR se llevó a cabo mediante una incubación previa a 50°C durante 30 minutos y una posterior amplificación, siguiendo el protocolo anteriormente descrito (apartado 3.4.1.c). El producto obtenido en esta reacción se utilizó para realizar una segunda amplificación en las mismas condiciones del anterior.

3.4.3. Iniciadores empleados.

Los iniciadores utilizados en todo el estudio tanto para las reacciones de amplificación (PCR) y secuenciación vienen resumidos en la tabla 3.1.

Iniciador ^a	Secuencia 5'→ 3'	Polaridad ^b	Posición ^c
20	TCAGTCCAGCTGTCTTTTTCTGGC	-	3287-3309
27	ATAAGCTTGCAGTCTAGCAGAAGAAGA	+	7004-7030
28	ATGAATTCTGGGTCCCCTCCTGAGGA	-	7314-7339
33	TCTACTGGCTCCATTTCTTG	-	5825-5844
38	GAACCATTAGGAGTAGC	+	7698-7714
45	AGCTCCTCTGGAAAGGTGAA	+	4984-5004
46	GCGGAATTCGGGTTTATTACAGGGACA	+	4898-4918
47	GTATTAGTAGGACCTACACCT	+	2475-2495
48	AGAAAGAGCAGAAGACAGTGGCAATGA	+	6202-6228
50	TAGCCCTTCCAGTCCCCCTTTTCTTTTA	-	9068-9096
52	TAATGTATGGGAATTGGCTCAA	-	6853-6874
53	GGCGAATTCAGTCCATTGCTCTCCA	-	4284-4310
58	GCCAGAAAAGACAGCTGGACTGT	+	3287-3310
60	GCCACACATGCCTGTGTACCCACA	+	6432-6455
63	GCGCTTAAGATTCCATCTAAAAATAGTA	-	4225-4253
100	ATCTCTAGGTTTCGAACAATACAGG	+	5452-5475
101	ACCTTCGGACGTCCTTCAGTCGG	+	5862-5884
104	CAGGAGGGGATCCAGAAATTGTA	+	7318-7340
109	CCCCAGACCTGAAGCTCTCTT	-	2168-2188
113	CGCCACCTGCAGGCGGCCTTAA	-	4606-4628
115	GTTTTATTCTGCATGGGAGTGTGAT	-	9004-9025
135	TTACAGTAGAAAAATTCCCCTC	-	7360-7381
147	TCTTCTGTCAATGGCCATTGTTTAAC	-	2610-2635
163	CTGAGGGATCTCTAGTTACCAGAG	-	579-602
166	TGGAGCCAGTAGATCCTAGACTAGAGCCCT	+	5832-5861
167	TTCTCCAATTGTCCCTCATATCTCCTCCTCCA	-	7635-7664
168	AATGGTGAGTATCCCTGCCTAACTCTATT	-	8340-8368
169	AATGTCAGCACAGTACAATGTACAC	+	6945-6969
170	TCTCTAGCAGTGGCGCCCGAACAGGGAC	+	626-653
171	TTGACTAGCGGAGGCTAGA	+	761-780

172	AGACAGGATCAGAAGAA	+	995-1011
173	GGTGATATGGCCTGATGTACCATTTGCCCTG	-	1204-1235
174	TCTATCCATTCTGCAGCTTCCTCATTGAT	-	1402-1431
180	TGGAAGGGCTAATTTGGTCCCAGA	+	1-24
182	AGTAACATATCCTGCTTTTCCT	-	3905-3926
185	GAGTATTGTATGGATTTTCAGG	-	2703-2724
187	CACACACAAGGCTACTTCCCT	+	57-77
188	GCCACTCCCCTGTCCCGCCC	-	389-409
189	ATCTCGAGACCTAGAAAAACATG	+	8895-8917
205	GCAGTAGCTGAGGGGACAGATAG	+	8685-8707
214	AATAGAGTTAGGCAGGGATACTCACCATT	+	8340-8368
215	TGAAAAAGAGTGCTATTAATC	+	8644-8669
216	GAGCTCCCAGGCTCAGATCTGGTCT	-	468-492
218	AGAACCGGTCTACATAGTCTCT	-	1669-1690
246	TTCTTCTGATCCTGTCT	-	995-1011
251	TGGGTACCAGCACACAAAGGAATAGGAGGAAA	+	4152-4183
253	TAAAAGGAGAAGCCATGCATGGACAAGTAGA	+	4363-4393
254	TCACCTTTCCAGAGGAGCTTTGCTGGTCCTTTCC	-	4278-4308
278	TGTATTACTACTGCCCTTCACCTTTCCA	-	4956-4984
307	AGGGCCACAAAGGGAACCATAC	+	5582-5603
308	AAAACCACCCGGGGCACAATAA	-	6872-6893
319	GTGAATAGAGTTAGGCAGGG	+	8337-8356
320	TCCAGGTCTCGAGATGCTGCTCCCACC	-	8883-8909
321	CCCATAATAGACTGTGACCC	-	6328-6347
322	CTCTCAGAAGCAGGAGCCG	+	2205-2223
323	CCAGGCCAGATGAGAGAACCAAGGG	+	1462-1486
324	TTTCCAACAGCCCTTTTTCCTAGGGG	-	2009-2034
325	TACTAATCTAGCCTCCCCTAGTGGGATGTG	-	5206-5235
326	CTTTCCCCTGCACTGTACCC	-	4806-4825

Tabla 3.1. Iniciadores empleados en la amplificación y secuenciación de los genomas completos. ^a Número de código de nuestro laboratorio de los iniciadores. ^b (+) sentido 5'→3'. (-) Sentido 3'→5'. ^c Posición según la cepa HXB2

3.5. Purificación de los productos de PCR

En la purificación de los fragmentos amplificados se utilizó el equipo comercial “High Pure PCR Product Purification” (Boehringer Mannheim). La purificación consiste en la eliminación de los restos de iniciadores, sales, aceite, nucleótidos no incorporados y restos de la Taq polimerasa que podrían afectar a las reacciones enzimáticas posteriores. El método consiste en la adhesión del ADN por puentes de hidrógeno a una columna de sílica. Posteriormente, éste ADN retenido en la columna, es eluido por la adición de un tampón con baja concentración de sales.

3.6. Clonaje molecular.

Para realizar el clonaje molecular se utilizó el equipo comercial TA Cloning® (Invitrogen), cuyo principio se basa en el hecho de que, durante el proceso de amplificación por PCR, la enzima Taq polimerasa incorpora desoxiadenosinas adicionales en el extremo 3'. Estos residuos adicionales son utilizados para insertar el producto de PCR en un vector linealizado que contiene residuos de desoxitimidina en los extremos 3'.

3.6.1. Preparación del inserto y ligación.

Los fragmentos de ADN que se utilizaron para las inserciones en el plásmido, proceden de una amplificación por doble PCR obtenida con los iniciadores 27 y 22 para la región C2-V5 del gen de la envuelta. La ligación con el vector pCR®2.1, de 3.9 Kb, tuvo lugar en una proporción molar vector: inserto 1:2, para lo que se emplearon 50 ng de vector y 20 ng de inserto, en un volumen final de 10 µl. La reacción de ligación se llevó a cabo con 4 unidades de enzima ligasa de T4 durante toda la noche a una temperatura de 14°C.

3.6.2. Transformación de bacterias competentes.

Una vez realizada la reacción de ligación, se pasó a la transformación de las bacterias competentes con el plásmido que llevaba el inserto. Para ello se añadieron 2 μ l de β -mercaptoetanol a cada vial de 50 μ l de células E. Coli (INaF'). Tras una ligera agitación, se añadió 1 μ l de mezcla de ligación y se incubó durante 30 minutos en hielo.

Posteriormente, las bacterias se sometieron a un choque térmico durante 45 segundos a una temperatura de 42°C y 2 minutos en hielo, tras lo cual, se añadieron 250 μ l de medio SOC (Bactotripton 2%, extracto de levadura 0.5%, NaCl 10 mM, KCl 2.5 mM, MgCl₂ 10 mM, MgSO₄ 10 mM y glucosa 20 mM) a las bacterias y se incubaron durante una hora a 37°C con agitación a 225 r.p.m. Del producto de la transformación se sembraron 50 μ l y 200 μ l en placas de medio LB-agar (extracto de levadura 2%, NaCl 1%, agar 1%, pH 7.5) con ampicilina a una concentración de 50 mg/ml, para la selección de colonias resistentes a dicho antibiótico. Para la selección de colonias recombinantes por color se añadió 1.6 mg de 5-bromo-4-cloro-3-indolil-b-D-galactopiranosido (X-gal) (Boehringer Mannheim) (20 mg/ml en dimetilformamida). Las placas se incubaron a 37°C durante toda la noche. Las colonias que crecieron de color blanco eran las que llevaban el inserto, y fueron seleccionadas para su posterior análisis.

3.7. Obtención de ADN de plásmidos.

3.7.1. A pequeña escala.

Cada colonia blanca a analizar (apartado 3.6.2) se transfirió a 5 ml de medio LB (extracto de levadura 2%, NaCl 1%, pH 7.5) con ampicilina (50 ng /ml), y se incubó a 37°C durante toda la noche a 225 r.p.m. Tras centrifugación a 2500 r.p.m. durante 10 minutos, la purificación del ADN plasmídico obtenido desde el sedimento de células se realizó con el equipo comercial Wizard™ Minipreps DNA purification Systems (Promega), siguiendo las instrucciones del fabricante. El ADN obtenido se analizó en geles de agarosa al 1 % en tampón Tris-acetato 40 mM, EDTA 10 mM (TAE).

3.7.2. A gran escala.

El método es similar al anterior. Cada colonia se transfirió a 100 ml de medio LB (extracto de levadura 2%, NaCl 1%, pH 7.5) con ampicilina (50 mg/ml), y se incubó a 37°C durante toda la noche a 225 r.p.m. Tras centrifugación a 2500 r.p.m. durante 10 minutos y a 4°C, la purificación del ADN plasmídico obtenido desde el sedimento de células se realizó con el equipo comercial Wizard™ Maxipreps DNA purification Systems (Promega), siguiendo las instrucciones del fabricante. El ADN obtenido se analizó en geles de agarosa al 1% en tampón Tris-acetato 40mM, EDTA 10 mM (TAE).

3.8. Secuenciación cíclica de los productos de PCR.

3.8.1. Secuenciación manual.

La secuenciación directa de los productos de PCR es una adaptación del método de los terminadores de cadena de Sanger (Sanger y cols., 1977) basado en la capacidad de la ADN polimerasa de elongar un molde hibridado a un iniciador mediante la realización de una PCR asimétrica en presencia de didesoxinucleótidos (ddNTPs) y un desoxinucleótido (dNTP) marcado radiactivamente.

La secuenciación se realizó mediante la utilización del equipo comercial *fmol™ DNA Cycle Sequencing System* (Promega), a partir aproximadamente de 0.5 ng de ADN, previamente amplificado y purificado (según apartado 3.4.1.), 50 ng de un único iniciador por reacción, 10 mCi de $\alpha^{35}\text{S}$ -dATP (>600Ci / mmol) (Amersham) y 5 unidades de Taq polimerasa.

El perfil de ciclos utilizado fue una primera etapa de desnaturalización a 95°C durante 2 minutos seguida de una etapa de amplificación de 35 ciclos, consistentes en una desnaturalización a 95°C durante 30 segundos, una hibridación durante 30 segundos a una temperatura de 55°C y una elongación a 72°C durante 1 minuto. Los termocicladores empleados fueron los mismos que los utilizados en la PCR. Posteriormente, las reacciones de secuenciación se pararon por adición de 3 μl de solución de parada (Promega).

Los productos obtenidos en las reacciones de secuenciación se desnaturalizaron a 70°C durante 2 minutos y se analizaron por electroforesis en TBE a 45 W en geles de poliacrilamida al 6 % y Urea 8 M en TBE. Finalmente, los geles se secaron y fueron expuestos sobre una película para autorradiografía (Kodak o Agfa) a temperatura ambiente.

3.8.2. Secuenciación automática.

La secuenciación automática se lleva a cabo mediante el uso de terminadores marcados con un colorante fluorescente. La secuenciación se realiza mediante la amplificación por PCR asimétrica en presencia de los cuatro ddNTPs marcados cada uno con un colorante fluorescente distinto. Cada colorante emite luz a diferente longitud de onda cuando es excitado por un rayo láser, por lo que los cuatro colorantes, y por tanto, las cuatro reacciones, pueden ser identificadas en una sola calle del gel.

La reacción de secuenciación se realizó empleando el equipo comercial “Big Dye™ Terminator Cycle Sequencing v2.0” (Applied Biosystems, Perkin Elmer). El perfil empleado consistió en un primer ciclo de desnaturalización a 94°C durante 3 minutos seguido de 25 ciclos de desnaturalización a 96°C durante 30 segundos y renaturalización a 55°C durante 15 segundos y a 60°C durante 4 minutos. Los ddNTPs no incorporados durante la reacción fueron eliminados mediante la precipitación con MgCl₂ / etanol.

Los fragmentos obtenidos se analizaron en el secuenciador automático ABI PRISM™ modelo 377 (Applied Biosystems, Perkin Elmer).

3.9. Análisis de secuencias.

Las secuencias nucleotídicas fueron editadas, compiladas y traducidas utilizando el paquete de programas Lasergene for Windows (Dnastar Inc, Madison, Wis). Los alineamientos múltiples de secuencias nucleotídicas y de aminoácidos se realizaron con el programa CLUSTALW versión 1.7 (Thompson y cols., 1994), siendo posteriormente editados y corregidos a mano para asegurar que la lectura de fase fuese correcta. Toda aquella posición en la que no existía información completa para el conjunto de secuencias que formaban el

alineamiento (“gaps”) se excluyeron de todos los análisis filogenéticos posteriores. Las secuencias nucleotídicas de los aislados de referencia de distintos subtipos del VIH-1 se obtuvieron en la base de datos de Los Álamos (<http://hiv-web.lanl.gov>).

3.10. Análisis filogenéticos.

Los análisis filogenéticos se realizaron utilizando el programa MEGA (Kumar y cols., 1993). Las matrices de distancias se calcularon mediante el modelo de dos parámetros de Kimura. Los árboles filogenéticos se construyeron mediante el método del vecino más próximo (“neighbor-joining”; NJ) y la robustez estadística de los árboles generados se probó mediante un remuestreo de 1000 réplicas (“bootstrap”). Finalmente, los árboles filogenéticos generados fueron editados con el programa TREEVIEW versión 1.5.2. (Page, 1996).

El cálculo del número de mutaciones sinónimas y no sinónimas observadas, así como el total de posiciones sinónimas y no sinónimas posibles de los tres provirus obtenidos de las células persistentemente infectadas, se realizó con la ayuda del programa SNAP (<http://hiv-web.lanl.gov>), basado en el método descrito por Nei y Gojobori (Nei & Gojobori, 1986).

4.RESULTADOS

4.1. Análisis de muestras de ONUSIDA.

4.1.1. Análisis filogenético de muestras ONUSIDA cocultivo negativo en los genes *gag*, *env* y *nef*

En éste estudio hemos llevado a cabo un análisis de un grupo de muestras del subtipo C procedentes de distintos países del mundo. Este proyecto fue lanzado y coordinado por ONUSIDA, teniendo como objetivo la caracterización de cepas de los subtipos más prevalentes en el mundo como paso necesario en el desarrollo de un programa de vacunación. En concreto, el estudio se centró en el análisis de cepas pertenecientes al subtipo C, realizando nuestro laboratorio el estudio de un grupo de cepas que no crecían por cocultivo en CMSPs. Este proyecto está realizado en colaboración con otros laboratorios que han llevado a cabo el análisis tanto de cepas que crecían en cultivo como, de otros subtipos que han creado brotes epidémicos. Se han recibido cepas de Sudáfrica, China, India, Israel y Brasil en donde el subtipo C tiene una alta prevalencia. El análisis consistió en un estudio genético del ADN proviral de las muestras en tres regiones: p17 en *gag*, C2-V3-C3 en *env* y el gen *nef*. Se analizaron 27 muestras cocultivo negativo y 8 muestras cocultivo positivo como controles. Se amplificaron 17 muestras (12 cocultivo negativo y 5 controles positivos) en las tres regiones, mientras que de las 18 restantes (15 cocultivo negativo y 3 controles positivos), se amplificaron 9 en dos de las regiones (4 cocultivo negativo y 2 controles positivos en *gag* y *env*, y 2 muestras cocultivo negativo y 1 control positivo en *gag* y en *nef*), 4 cocultivo negativo en una región (3 en *gag* y 1 en *env*) y 5 no pudieron ser amplificadas en ninguna región. Los resultados de la amplificación así como las características de las muestras cocultivo negativo se resumen en la Tabla 4.1.

Número muestra (1)	Sexo	Edad	Residencia	Posible transmisión	Transmisión sexual	Transmisión parenteral	Estado de infección	Tratamiento	Amplificación por PCR (2)		
									p17gag	C2V3C3env	nef
BR97001 (BR1)	M	31	P. Alegre	desconocida	heterosexual	IDVP (3)	Asintomático	No	+	+	+
BR97002 (BR2)	M	43	P. Alegre	sexual	heterosexual		Enfermedad moderada	No	+	+	+
BR97003 (BR3)	M	22	P. Alegre	sexual	heterosexual		Asintomático	No	+	+	+
BR98005 (BR5)	V	44	P. Alegre	sexual	heterosexual	IDVP	Asintomático	No	+	+	-
BR98006 (BR6)	V	35	P. Alegre	sexual	heterosexual	IDVP	Asintomático	No	+	-	+
BR98007 (BR7)	V	39	P. Alegre	sexual	homosexual		Asintomático	Si	+	-	-
BR98008 (BR8)	M	35	P. Alegre	sexual	heterosexual	IDVP	Síntomas tempranos	Si	+	+	+
BR98009 (BR9)	M	42	P. Alegre	sexual	heterosexual		Asintomático	Si	+	+	+
CN98002 (CN2)	V	23	Yining			IDVP	Asintomático	No	-	+	-
CN98003 (CN3)	V	26	Yining			IDVP	Asintomático	No	+	+	-
CN98004 (CN4)	V	27	Yining			IDVP	Asintomático	No	-	-	-
CN98007 (CN7)	V	23	Yining			IDVP	Asintomático	No	-	-	-
CN98008 (CN8)	V	30	Yining			IDVP	Asintomático	No	-	-	-
IN97001 (IN1)	V	35	Chennai	sexual	heterosexual		Asintomático	No	-	-	-
IN97002 (IN2)	V	25	Chennai	sexual	heterosexual		Asintomático	No	-	-	-
IN97005 (IN5)	V	40	Gudur Ongale (Dist)	sexual	heterosexual		Síntomas tempranos	No	+	+	+
IN98006 (IN6)	M	30	Nungambakkam	sexual	heterosexual		Asintomático	No	+	+	+
IN98008 (IN8)	M	20	Chennai	sexual	heterosexual		Asintomático	No	+	+	-
IN98009 (IN9)	M	28	K. Pondichesly	sexual	heterosexual		Síntomas tempranos		+	+	+
IN98011 (IN11)	V	28	Madras	sexual	heterosexual		Asintomático	No	+	+	+
IN98014 (IN14)	V	29	Bombay	sexual	heterosexual		Asintomático	No	+	+	+
IN98019 (IN19)	V	30	Calcutta	sexual	heterosexual		Síntomas tempranos	No	+	+	-
IN98021 (IN21)	M	22	Churachand Pur		homosexual	IDVP	Asintomático	No	+	+	+
IS98001 (IS1)									+	-	+
IS98003 (IS3)									+	-	-
ZA97004 (ZA4)	M	40	Bloefoentein	sexual	heterosexual		Asintomático	No	+	+	+
ZA97007 (ZA7)	M	30	Pretoria	sexual	heterosexual		Síntomas tempranos	No	+	-	-
TOTAL =27									21	17	14

Tabla 4.1. Características epidemiológicas y virológicas de las muestras cocultivo negativo de ONUSIDA. (1) Las muestras se designan con el código de ONUSIDA seguido por el número de identificación de nuestro laboratorio. (2) Resumen de los resultados de la amplificación en los tres genes. (3) Usuarios de drogas por vía parenteral (UDVP). En la tabla no se incluyen las muestras cocultivo positivo que han sido utilizadas como controles.

Para analizar las relaciones filogenéticas entre las muestras, se alinearon conjuntamente todas las secuencias con cepas de referencia correspondientes a distintos subtipos (desde el A al H) en los genes *gag* y *env*. Las matrices de distancias obtenidas mediante el modelo de dos parámetros de Kimura se utilizaron para la construcción de los árboles filogenéticos mediante el método del vecino más próximo (“Neighbor-Joining”). Los árboles obtenidos se muestran en las figuras 4.1.A y 4.1.B

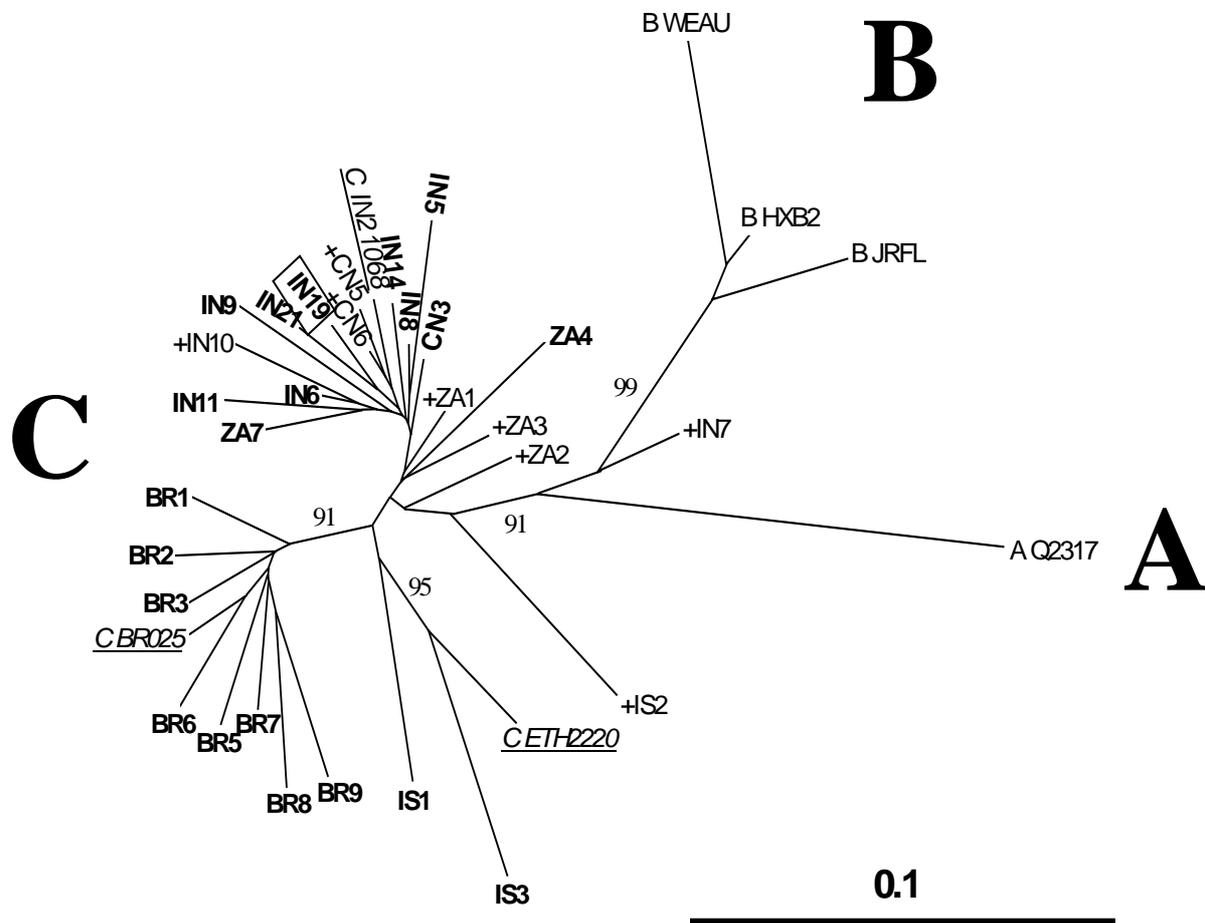


Figura 4.1.A. Análisis filogenético de las secuencias nucleotídicas de la región p17 del gen *gag* de las 21 muestras cocultivo negativo mediante el método del vecino más próximo (“Neighbor Joining”). El árbol se construyó según lo descrito en Materiales y Métodos. Las muestras de referencia se encuentran subrayadas y las muestras cocultivo positivo tienen el signo + delante del número de identificación. Las muestras cocultivo negativo se encuentran en negrita. Los números en las ramas representan los valores de remuestreo (“bootstrap”) y sólo están indicados aquellos valores por encima del 90 %. Las letras mayúsculas grandes corresponden a las muestras de referencia de los diferentes subtipos. La muestra recombinante IN19 se encuentra dentro del recuadro. La barra de la escala representa un 10 % de distancia genética entre las muestras.

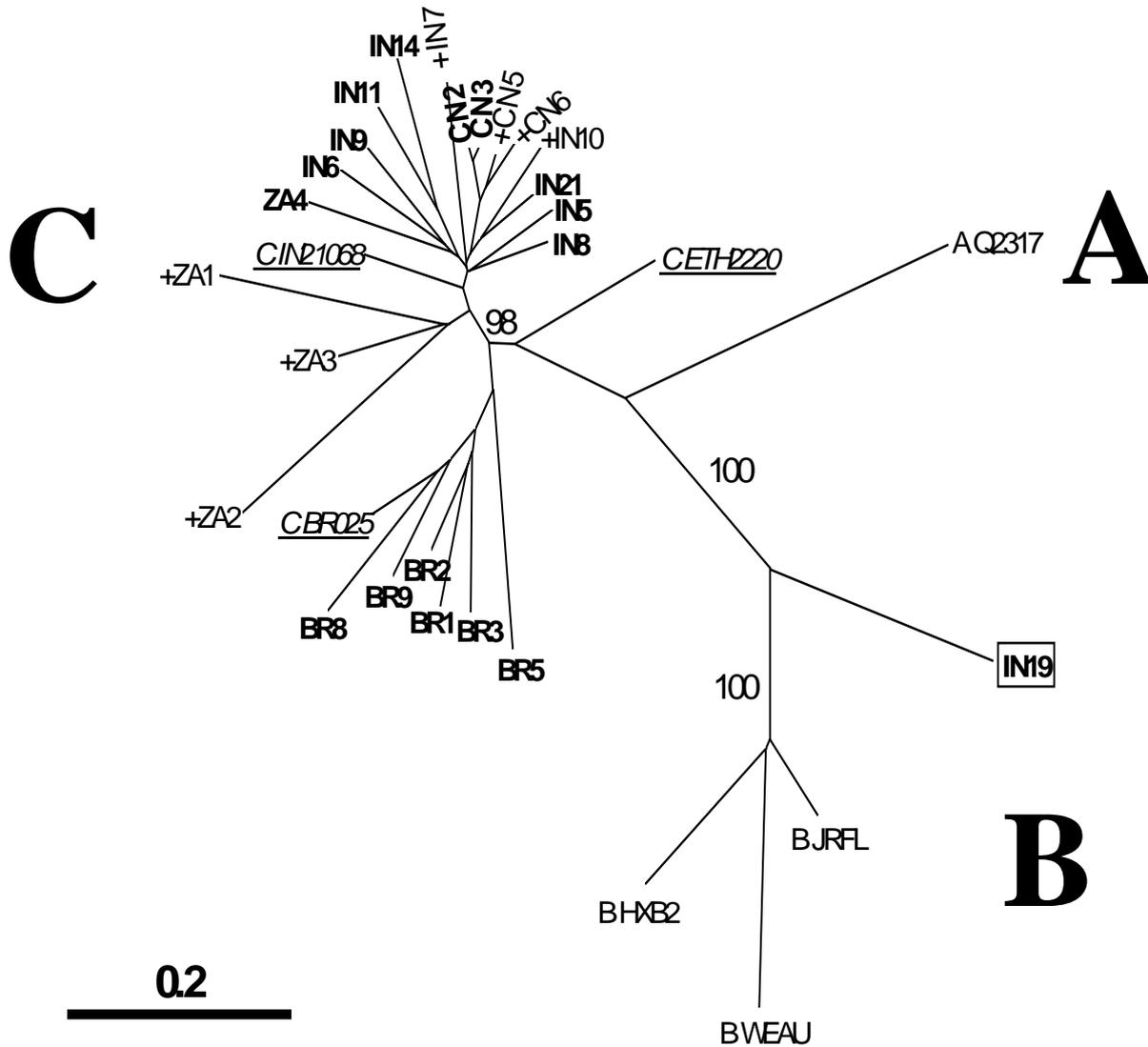


Figura 4.1.B. Análisis filogenético de las secuencias nucleotídicas de la región C2-V3-C3 del gen *env* de las 16 muestras cocultivo negativo mediante el método del vecino más próximo (“Neighbor Joining”). El árbol se construyó según lo descrito en materiales y métodos. Los símbolos son los mismos que en la Figura 4 A. La barra de la escala representa un 20 % de distancia genética entre las muestras.

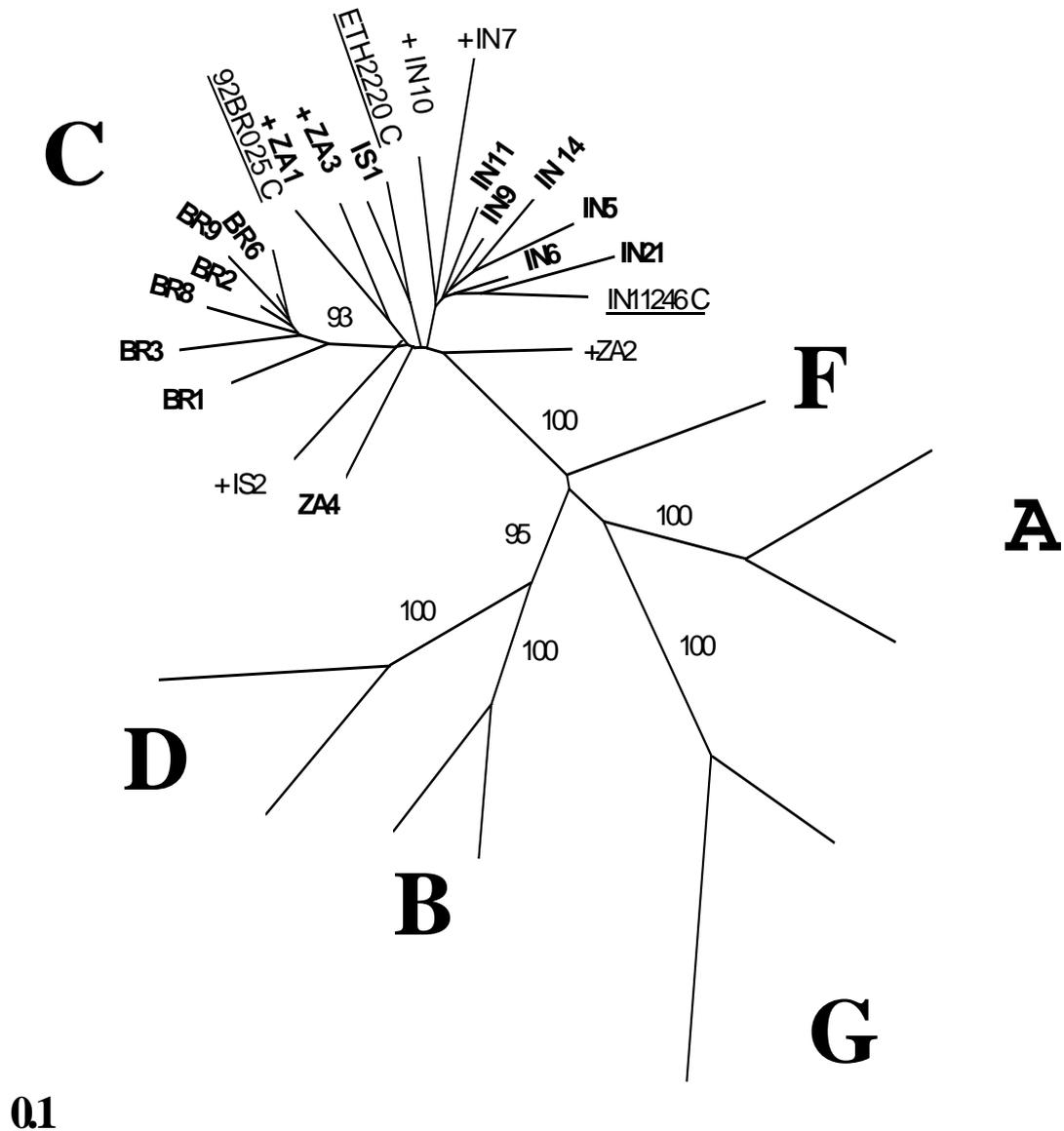


Figura 4.1.C. Análisis filogenético de las secuencias nucleotídicas del gen *nef* de las 14 muestras cocultivo negativo mediante el método del vecino más próximo (“Neighbor Joining”). El árbol se construyó según lo descrito en materiales y métodos. Los símbolos son los mismos que en las Figuras 4.1.A y 4.1.B. La barra de la escala representa un 10 % de la distancia genética entre las muestras

Como puede observarse en las figuras, todas las muestras se agrupan en la misma rama que las cepas de referencia del subtipo C obtenidas de la base de datos de Los Álamos. Este agrupamiento presenta altos valores de remuestreo (“bootstrap”) del 91 % y del 98 % en las regiones p17 y C2-V3-C3 respectivamente. Sin embargo, una de las muestras se agrupaba en la rama del subtipo C en p17, y en C2-V3-C3 se agrupaba con el subtipo B, indicando que ésta muestra (IN19) era un recombinante C / B (ver Figuras 4.1.A y 4.1.B).

En la región p17, las muestras procedentes de Brasil, así como la cepa de referencia BR025, formaban una rama independiente dentro del grupo correspondiente al subtipo C con un valor de remuestreo del 91 %. Por otra parte, 2 muestras de Israel se agrupaban con la cepa de referencia del subtipo C ETH 2220. Estas segregaciones no se observaron con un valor alto de remuestreo en el análisis filogenético de la región analizada en el gen *env* (ver figuras 4.1.A y 4.1.B).

Como también se habían realizado las secuencias nucleotídicas del gen *nef* en las muestras, se realizó el mismo análisis filogenético realizado anteriormente en p17 y en C2-V3-C3. Los resultados se muestran en la figura 4.1.C. Todas las muestras se agrupaban con las de referencia del subtipo C con unos valores de remuestreo del 100 %, confirmando que en el gen *nef* las muestras también correspondían al subtipo C como se había observado anteriormente. Se observó también la segregación de las muestras de Brasil observada anteriormente en la región p17. El recombinante C / B se agrupaba en el gen *nef* dentro del subtipo C, siendo en consecuencia un recombinante C / B / C.

4.1.2. Análisis de la variación en las proteínas p17 y en la región C2-V3-C3 de la gp120.

En la región p17 se apreciaba una zona variable en la región C-terminal de la proteína que incluía inserciones y deleciones, haciendo que 12 muestras cocultivo negativo tuviesen una longitud de la proteína de 130 aminoácidos, mientras que 5 muestras presentaban 129 aminoácidos y 4 presentaban 128. También se observó una alta variabilidad genética alrededor del aminoácido 90 (las posiciones de los aminoácidos son siempre las correspondientes al HXB2). La secuencia aminoacídica de la señal de miristilación que abarca desde la posición 1 a la 7 de la p17 se conservaba en todas las muestras excepto en la muestra BR5 que presentaba una duplicación de dos aminoácidos Arginina (R) y Alanina (A) en la posición 5. Esta muestra también tenía una deleción de cuatro aminoácidos desde la posición 122 a la 125. La muestra IN5 presentaba una proteína p17 truncada como consecuencia de dos codones de terminación (“stop codon”) en las posiciones 36 y 58. Los resultados se muestran en la tabla 4.2

En el análisis de las secuencias de la región C2-V3-C3, las secuencias de aminoácidos obtenidas de las 16 muestras cocultivo negativo tenían una longitud de 35 aminoácidos y todas presentaban la secuencia aminoacídica GPGQ en el ápice del lazo V3, característica de las secuencias del subtipo C, como se puede comprobar en la base de datos de Los Álamos (<http://hiv-web.lanl.gov>). Es importante destacar lo conservado que se encuentra el lazo V3 entre las muestras, y en concreto la secuencia GPGQ, a pesar de la diversidad geográfica de procedencia de las mismas. Por otra parte, también hay que destacar la variación de la secuencia de aminoácidos alrededor del lazo V3. Como ya había sido descrito en la literatura (Casado y cols., 2000, Chaix y cols., 1993) se observa también una mayor variación en la región adyacente a la región V3 que dentro de la misma. IN11 presentaba una deleción de dos

aminoácidos en las posiciones 336 y 337. Por otra parte, la muestra IN11 tenía en posición 306 una Glicina y la muestra IN6 presentaba en la posición 321 una Serina. Hay que destacar que estos dos residuos no han sido descritos anteriormente en el subtipo C en ninguna de las 443 secuencias existentes en la base de datos de Los Álamos (1999) (Kuiken y cols., 1999). Estas observaciones se aprecian en la figura 4.3. donde se señalan los residuos conservados y variables.

Se pueden distinguir, a nivel genético, los virus NSI de los SI en ciertas sustituciones de aminoácidos en las posiciones 307, 322 y 323 del lazo V3. La presencia de aminoácidos cargados positivamente en éstas posiciones determina un fenotipo SI. Todas las muestras cocultivo negativo de este estudio presentaban un fenotipo de NSI ya que presentaban el aminoácido Serina en la posición 307 en lugar de un aminoácido con carga positiva como la Arginina. Además tenían el aminoácido Glicina en la 322 y Aspártico en la posición 323 (aminoácidos que no presentan carga positiva). Estos resultados confirman el predominio de cepas NSI observado entre aislados del subtipo C en otros estudios (Bjorndal y cols., 1999, Van Harmelen y cols., 1999).

4.1.3. Análisis del gen *nef*

Para estudiar si existe una relación entre alteraciones en el gen *nef* y el fenotipo de cocultivo negativo que presentaban las muestras, se decidió analizar las secuencias nucleotídicas de las muestras. Se amplificaron en ésta región 14 muestras, 6 de Brasil, 6 de India, 1 de Israel y 1 una de Sudáfrica, así como 6 controles cocultivo positivo. Con el objeto de realizar un análisis de la proteína, se estudiaron las secuencias a nivel de aminoácidos. De éste análisis se pudo observar que 2 de las 14 muestras cocultivo negativo tenían claramente una proteína Nef no funcional. Una de éstas secuencias (IN14) presentaba una delección perfecta desde la primera Cisteína de la proteína hasta el comienzo de la LTR, la cual solapa con parte de la región C-terminal del gen *nef*. La otra muestra era la IN5 que tenía dos codones de terminación TAG en las posiciones 57 y 113 además de un cambio de Glicina a Lisina en la posición 2, modificando la secuencia conservada desde el aminoácido en posición 1 hasta el 7 de la señal de miristilación.

Por otra parte, la muestra IN5 presentaba otras sustituciones de aminoácidos tales como G31R, R77G, M79I, y el cambio de G132E en la segunda Glicina del motivo PGPGV. Además de éstos virus, se encontraron 7 secuencias con mutaciones que podrían tener algún efecto en la funcionalidad de la proteína Nef (ver tabla 4.4). La muestra BR8 tenía una mutación G130R en la primera Glicina del motivo conservado anteriormente descrito PGPGV. Otra muestra (IN6), presentaba el cambio de 2 prolina en el importante motivo PxxP en posiciones 69 y 75. Éste motivo se ha asociado con la unión entre quinasas celulares y la proteína Nef (Saksela y cols., 1995). La muestra IN21 presentaba una mutación (C142L) en la Cisteína que se encuentra ampliamente conservada entre los aislados de la base de datos

de Los Álamos y está implicada en la regulación negativa de los receptores CD4 (Aiken y cols., 1996) Este cambio produce la ruptura de una lámina β y podría representar un cambio estructural en la proteína Nef. Por otra parte, otras 4 muestras presentaban alteraciones que también podrían tener influencia en la función de la proteína tales como las encontradas en BR3, BR6 y BR9 e IN9. Por un lado, las muestras brasileñas presentaban deleciones en la región variable contigua a la región de la señal de miristilación. BR3 y BR6 mostraban la misma deleción de cuatro aminoácidos (del residuo 8 al 11) e IN9 tenía una deleción de un aminoácido en la posición 11. Por otra parte, BR3 tenía una sustitución de G31W y la muestra IN9 presentaba un cambio V30G. Hay que destacar que las posiciones 30 (Valina) y 31 (Glicina) se encuentran conservadas en todos los subtipos. Otras mutaciones encontradas entre las se localizaban en la región de Nef que se une al receptor CD4 en las posiciones 57 (Triptófano) y 58 (Leucina) de la proteína. Las muestras IN11 tenía el cambio W57G y IN9 presentaba el cambio W57C. Finalmente, en las posiciones 130 y 132 pertenecientes al motivo PGPGV, BR8 tenía el cambio G130R en la primera Glicina.

Por el contrario, solamente una muestra cocultivo positivo presentaba una alteración en algún motivo funcional de la proteína. La muestra IS2 presentaba una Arginina en lugar de Triptófano en la posición 57. Todas las alteraciones genéticas vistas se encuentran resumidas en la tabla 4.4. En resumen, 9 de 14 muestras incapaces de crecer en cocultivo presentaban alguna alteración en la proteína Nef, mientras que solamente 1 muestra de 6 cocultivo positivo presentaba alguna alteración.

AMINOÁCIDO ¹	1-7	8-16	21-24	30/31/47	57/58	69-80	113	129-133	142	154/155	164/165
REGION DE NEF ²	S.M. ³	P.R. ⁴	MHC I ⁵	D. VARIABLE ⁶	CD4 (WL) ⁷	(PxxP) ⁴⁸	HÉLICE b ⁹	a- Lámina β ¹⁰	Cys (b-C) ¹¹	CD4(EE) ¹²	CD4(LL) ¹³
COCULTIVO NEGATIVO											
BR1	-	-	-	-	-	-	-	-	-	-	-
BR9	-	-	-	-	-	-	-	-	-	-	-
BR2	-	-	-	-	-	-	-	-	-	-	-
BR8	-	-	-	-	-	-	-	+	-	-	-
BR3	-	+	-	+	-	-	-	-	-	-	-
BR6	-	+	-	-	-	-	-	-	-	-	-
IS1	-	-	-	-	-	-	-	-	-	-	-
IN11	-	-	-	-	+	-	-	-	-	-	-
IN14	+	+	+	+	+	+	-	-	-	-	-
IN5	+	-	-	+	+	+	+	+	-	-	-
IN9	-	+	-	+	+	-	-	-	-	-	-
IN6	-	-	-	-	-	+	-	-	-	-	-
IN21	-	-	-	-	-	-	-	-	+	-	-
ZA4	-	-	-	-	-	-	-	-	-	-	-
COCULTIVO POSITIVO											
IN7	-	-	-	-	-	-	-	-	-	-	-
IN10	-	-	-	-	-	-	-	-	-	-	-
IS2	-	-	-	-	+	-	-	-	-	-	-
ZA1	-	-	-	-	-	-	-	-	-	-	-
ZA3	-	-	-	-	-	-	-	-	-	-	-
ZA2	-	-	-	-	-	-	-	-	-	-	-

Tabla 4.4. Alteraciones a nivel de aminoácidos de los motivos conservados descritos en la proteína Nef. Las muestras que presentan alguna alteración en la secuencia se encuentran marcadas con el signo +, mientras que las muestras que no presentan alteración se señalan con el signo -. (1) Posición de los aminoácidos donde se encuentran localizados los motivos funcionales. (2) Regiones con un papel específico en la función de Nef. (3) Señal de miristilación (Niederman y cols., 1993). (4) Región polimórfica (Shugars y cols., 1993). (5) Regulación negativa del MHC I y de unión a proteín-quinasa (Schwartz y cols., 1996). (6) Región de duplicación variable (Grzesiek y cols., 1996a). (7) Sitio de unión a los CD4 (Grzesiek y cols., 1996b). (8) Región de prolinas repetidas (regulación negativa del MHC y de unión a SH₃) (Saksela y cols., 1995). (9) Hélice α-β (Grzesiek y cols., 1996a). (10) Lámina β (Grzesiek y cols., 1996a). (11) Cisteína en posición 142 (Grzesiek y cols., 1996a). (12) y (13) sitios de unión al receptor CD4 (Bresnahan y cols., 1998, Hua y cols., 1997)

4.2. Estudio del proceso de persistencia in vitro a partir de un mismo aislado del VIH-1 en tres líneas celulares.

La persistencia vírica es un proceso complejo y multifactorial en el que se ven involucrados tanto factores celulares como víricos. En nuestro laboratorio, nos propusimos estudiar el proceso del establecimiento de infecciones persistentes en tres líneas celulares diferentes con la misma cepa del VIH-1 (s61). Dos de éstas líneas eran linfocitos T CD4⁺ (células MT-4 y H9) y la tercera era una línea promonocítica (células U-937 clon 2). Para ello, con el aislado s61 se infectaron a una m.d.i. de 0.1 TCID₅₀ / ml las tres líneas celulares y se subcultivaron hasta la obtención de la persistencia. El trabajo del establecimiento de la persistencia en nuestro laboratorio así como la caracterización biológica de las células persistentemente infectadas fue realizado por Dña. Lourdes Muñoz (tesis en elaboración).

4.2.1. Criterios de infección persistente.

Los 3 criterios seguidos en éste estudio para considerar que se habían establecido las infecciones persistentes fueron los siguientes:

- a) mantenimiento del cultivo con una viabilidad mayor del 85 %, sin necesidad de aporte celular.
- b) desaparición del efecto citopático (ECP)
- c) detección de virus en las células y en el sobrenadante del cultivo

En el momento en el que los cultivos celulares presentaban éstas tres características, se consideró que las infecciones realizadas con el aislado s61 en las tres líneas celulares habían alcanzado la persistencia. La figura 4.2 muestra el proceso del establecimiento de la persistencia.

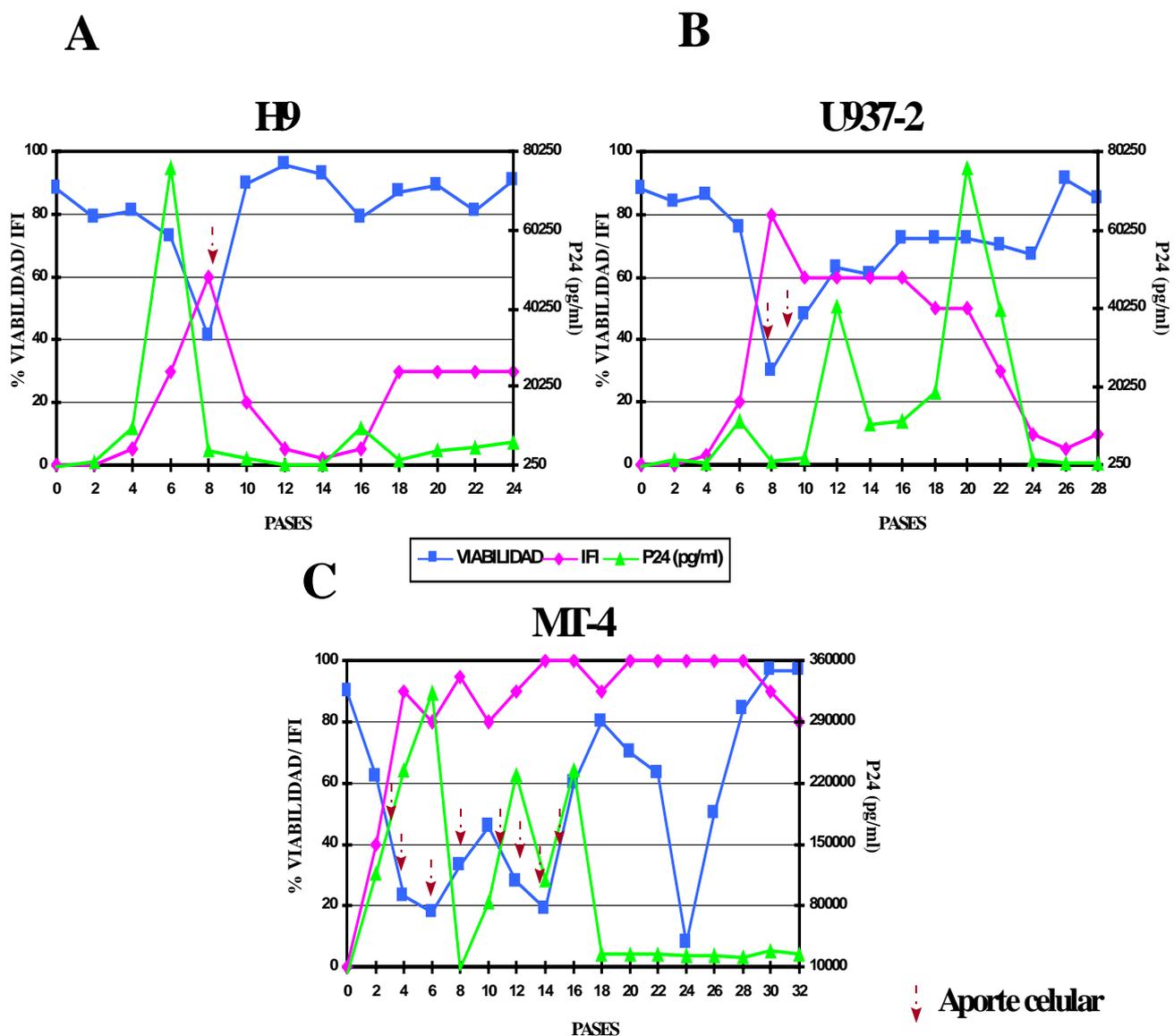


Figura 4.2. Establecimiento de las infecciones persistentes en las tres líneas celulares establecidas H9 (A), U-937 clon 2 (B) y MT-4 (C). En las gráficas se representan en el eje de abcisas el número de pases en cultivo. En los ejes de ordenadas se representan la viabilidad celular a lo largo de los pases (líneas azules) así como los valores de p24 en pg/ml (líneas verdes) y el % de células del cultivo que expresan proteínas víricas (IFI) (líneas rosas). Las flechas verticales de color marrón representan el aporte de células sin infectar a los cultivos por tener éstos una viabilidad por debajo del 40 %. En el caso de la línea celular MT-4 los valores de p24 se encuentran a distinta escala. Figura tomada de Lourdes Muñoz (Tesis Doctoral en elaboración)

4.2.2. Establecimiento de la infección persistente en la línea celular H9.

El proceso del establecimiento de la persistencia en la línea celular H9 (ver gráfico 4.2 apartado A) se siguió mediante la determinación de:

- a) **Viabilidad celular.** El cultivo se mantuvo con una viabilidad entre el 80-90 % durante los cuatro primeros pases desde el inicio de la infección, a partir de los cuales descendió hasta alcanzar en el pase 8 una viabilidad del 40 %, teniendo que aportar nuevas células sin infectar. Desde el pase 10 el cultivo se mantuvo con una viabilidad entre el 80 % y 100 %.
- b) **Efecto citopático (ECP).** A partir del pase 2 fue característico la presencia de sincitios aislados, observándose en el 50 % de las células en el pase 4. El ECP desapareció a partir del pase 18.
- c) **Inmunofluorescencia indirecta (IFI).** En el pase 3 comenzaron a expresarse las proteínas víricas en las células, alcanzándose el máximo de expresión en el pase 8 detectándose en el 60 % de las mismas. A partir de éste punto se observó un gran descenso en la expresión

alcanzando el 5 % en el pase 14. A partir de éste pase la expresión viral aumentó gradualmente, manteniéndose en un 30-40 % de las células a partir del pase 18.

d) **Detección del antígeno p24.** Comenzó a detectarse a partir del pase 2, con valores comprendidos entre 10^3 y 10^4 pg/ml en la mayor parte de las determinaciones, alcanzando el máximo de $7,6 \times 10^4$ pg/ml en el pase 6. A partir del pase 20, los valores se mantuvieron entre 2 y 3×10^3 pg/ml.

Según los criterios descritos en el apartado 4.2.1. se consideró que se había establecido la infección persistente a partir del pase 18. A las células persistentemente infectadas con el aislado s61 se las denominó H61, y al virus que se obtuvo de las mismas VH61.

4.2.3. Establecimiento de la infección persistente en la línea celular U-937 clon 2.

El proceso del establecimiento de la persistencia en la línea celular U-937 clon 2 (ver gráfico 4.2 apartado B) se siguió mediante la determinación de:

a) **Viabilidad celular.** El cultivo se mantuvo con una viabilidad entre el 85-90 % hasta el pase 6 en que comienza a descender. En los pases 8 y 9 se aportaron células y a partir de este punto la viabilidad ascendió hasta alcanzar el 60-75 %. Desde el pase 25 la viabilidad se mantuvo superior al 80 %.

b) **Efecto citopático.** A partir del pase 2 se visualizaban células gigantes aisladas junto con células de morfología irregular. Este patrón citopático desapareció a partir del pase 22.

- c) **Inmunofluorescencia indirecta.** La expresión de proteínas víricas comenzó a detectarse en el pase 4 y fue aumentando hasta alcanzar el 80 % de las células en el pase 8. A partir de éste pase disminuyó gradualmente hasta observarse en el 5-10 % de las células.
- d) **Detección de antígeno p24.** Comenzó a detectarse en el pase 2 y se podían distinguir tres fases: una primera hasta el pase 10 con valores comprendidos entre 0.2 y 1.5×10^3 pg/ml; una segunda desde el pase 12 al 22 donde los valores estaban comprendidos entre 2 y 5×10^4 pg/ml; y una tercera, a partir del pase 24 con valores comprendidos entre 2 y 4×10^3 pg/ml.

Según los criterios descritos en el apartado 4.2.1. se consideró que se había establecido la infección persistente a partir del pase 26. A las células persistentemente infectadas con el aislado s61 se las denominó U61, y al virus que se obtuvo de las mismas VU61.

4.2.4. Establecimiento de la infección persistente en la línea celular MT-4.

El proceso del establecimiento de la persistencia en la línea celular MT-4 (ver gráfico 4.3.a apartado C) se siguió mediante la determinación de:

- a) **Viabilidad celular.** Desde el inicio de la infección se observó un descenso progresivo de la viabilidad, necesitando el cultivo aporte celular desde el pase 3 al 8 y desde el 11 al 15. A partir del pase 14 la viabilidad aumentó hasta alcanzar el 80 % en el pase 18. Desde éste pase se observó un descenso lento con una caída brusca en el pase 24, manteniéndose el cultivo sin aporte celular. A partir del pase 24 ascendió gradualmente manteniéndose por encima del 80 % desde el pase 28.

- b) **Efecto citopático.** Desde el pase 4 se observó abundante lisis celular y vacuolización con algún sincitio aislado. Este efecto desaparece a partir del pase 28 comenzando a observarse el crecimiento en grupos celulares, característico de ésta línea sin infectar.
- c) **Inmunofluorescencia indirecta.** Desde el pase 1 se apreció expresión de proteínas víricas, observándose permanentemente a partir del pase 4 entre el 80 y el 100 % de las células.
- d) **Detección de antígeno p24.** Comenzó a detectarse a partir del pase 1, con valores comprendidos entre 1.5 y 2×10^4 pg/ml, ascendiendo hasta un máximo en el pase 6 de 2×10^5 pg/ml. Desde éste pase al 16, los valores descendieron a 1.5×10^5 pg/ml.

Según los criterios descritos en el apartado 4.2.1. se consideró que se había establecido la infección persistente a partir del pase 28. A las células persistentemente infectadas con el aislado s61 se las denominó M61, y al virus que se obtuvo de las mismas VM61.

4.2.5. Comparación del proceso de persistencia entre las tres líneas celulares.

- a) **Viabilidad.** Mientras que en las líneas H9 y U937 clon 2 se observó una caída en el pase 8, a partir del cual aumentó hasta mantenerse por encima del 85 %, en la línea MT-4 el descenso comenzó antes, alcanzando valores mucho más bajos que se mantuvieron durante 10 pases y no se consiguió una buena viabilidad hasta bastantes pases después.
- b) **Efecto citopático.** Desapareció en las líneas H9 y U-937 clon 2 prácticamente en el mismo momento, siendo un poco posterior en las MT-4.

c) **Inmunofluorescencia indirecta.** En la línea H9 la expresión de proteínas víricas alcanzó su mínimo en el pase 8, recuperándose posteriormente y manteniéndose en el 30-40 % de las células, a diferencia de las U-937 clon 2 que en el mismo pase comenzó a descender hasta alcanzar un 5-10 %. En las MT-4, la expresión de proteínas víricas se observó de forma constante a lo largo de la infección y se mantuvo entre el 80-90 %.

d) **Detección del antígeno p24.** Los niveles alcanzados en la línea MT-4 fueron los más altos en todo el seguimiento. Éstos fueron 10 veces superiores a los de las H9 y a los de la primera fase de las U-937 clon 2.

4.2.6. Secuenciación del genoma completo del provirus del aislado inicial s61 y de los provirus VH61, VM61, VU61 obtenidos de las células persistentemente infectadas.

Para tratar de establecer cuales son los factores víricos que pueden intervenir en la persistencia, se planteó el análisis de la secuencia completa de los provirus. Para ello, utilizamos los ADNs províricos obtenidos de las tres líneas celulares persistentemente infectadas (ver apartado 3.3.1.2). Las células seleccionadas para obtener el ADN provírico del genoma completo fueron las de los últimos pases de cada línea celular: pase 90 en las H61, y pases 100 y 92 en las MT-4 y U-937 clon 2 respectivamente. También llevamos a cabo la secuenciación del genoma completo del provirus del s61, procedente del virus con el que se habían realizado las infecciones, con el objeto de intentar ver qué cambios se habían producido en su secuencia nucleotídica durante el proceso de establecimiento de las infecciones persistentes. Asimismo, con el objeto de estudiar la evolución de las cuasiespecies clonamos y secuenciamos la región C2-V5 del gen de la envuelta y las LTRs.

4.2.6.1. Estrategia de amplificación y secuenciación.

En la gráfica (Anexo I) se indican los iniciadores que se utilizaron tanto para la amplificación como para la secuenciación de los cuatro genomas. La secuenciación se realizó por ambas bandas, tanto por la hebra de ADN positiva como por la negativa.

4.2.6.2. Descripción de las mutaciones encontradas.

Una vez amplificados y secuenciados los cuatro genomas, se procedió al alineamiento de sus secuencias nucleotídicas, tomando como secuencia de referencia la del s61 parental. De ésta manera se determinaron las mutaciones que diferenciaban al provirus del aislado inicial de los tres provirus obtenidos de las células persistentemente infectadas. La secuencia nucleotídica del genoma del s61 constaba de 9125 nucleótidos, mientras que el VH61, VM61 y VU61 tenían 9 nucleótidos de diferencia (9116 nucleótidos). El alineamiento de los genomas completos, separados por genes, se muestra en el Anexo II.

4.2.6.2.1. Frecuencias de mutación y tipos de mutaciones encontradas.

Con el objeto de estudiar la adaptación de cada virus al fenómeno de la persistencia, y describir el grado en el que afecta la presión mutacional a los genes de dichos virus, hemos analizado las frecuencias y tipos de mutaciones de los genomas completos y de cada uno de los genes por separado. Las mutaciones encontradas en los genomas completos, de los tres provirus obtenidos de las células persistentemente infectadas respecto a la secuencia del provirus del s61, fueron 75 en el caso del VH61, 70 para el VM61 y 88 para el VU61. Con estos valores calculamos las frecuencias de mutación como el cociente entre el número de mutaciones encontradas y el número total de nucleótidos secuenciados. Las frecuencias de mutación obtenidas [sustituciones observadas (s) / número de nucleótido secuenciados (nt)] eran similares, 8.2×10^{-3} s/nt en el VH61, 7.7×10^{-3} s/nt en el VM61 y de 9.6×10^{-3} s/nt para el VU61. En cuanto al tipo de mutaciones observadas, las transiciones fueron más frecuentes que las transversiones. En los tres provirus en conjunto, se observaron 156 transiciones por 77 transversiones. Centrándonos en las transiciones, los cambios más abundantes fueron el G→A

y el A→G con unos porcentajes de entre el 24 y el 30 % para cambios G→A y entre el 18 y el 21 % para cambios A→G de todas las mutaciones encontradas (ver tabla 4.5)

Para tener un análisis más detallado del tipo de mutaciones observadas, se realizó un estudio similar en cada uno de los genes que constituyen el genoma, al mismo tiempo que se estudiaron los cambios de aminoácidos asociados.

a) **Genes estructurales (*gag*, *pol*, y *env*).**

Las frecuencias de mutación observadas se encontraron en un rango entre 3.9×10^{-3} y 9.7×10^{-3} s/nt (ver tabla 4.6). Si comparamos las frecuencias de mutación entre los tres provirus en cada uno de los genes estructurales por separado, los valores observados son muy similares. Por el contrario, en la comparación entre genes, los genes *env* y *gag* presentan por término medio una mayor frecuencia de mutación que el gen *pol*, dándose una diferencia de aproximadamente 2 veces entre la frecuencia de mutación en el gen *pol* y en el gen *env*. Este valor es semejante a los valores descritos para aislados naturales de VIH-1 (Quinones-Mateu y cols., 1996). Respecto al tipo de mutaciones el número de transiciones es mayor que el número de transversiones (92 y 60 respectivamente), siendo, dentro de las transiciones, el cambio G→A el más abundante en los genes *pol* y en *env*, mientras que en el gen *gag* el cambio más frecuente es A→G. El porcentaje de mutaciones sinónimas en conjunto representa el 54.3 % del total, mientras que el de no sinónimas representa el 45.7 %.

Centrándonos en las mutaciones encontradas en cada uno de los genes estructurales, en el gen *gag*, y concretamente en la proteína p6, se observó en los tres provirus una delección

de 9 nucleótidos de las posiciones 12 a 14 (aminoácidos Ala-Pro-Pro) con respecto al s61. Como ésta zona solapa con el gen *pol*, también existe una delección de cuatro aminoácidos, en la proteína p6 pol, de los aminoácidos 28 al 31 de el motivo Ser-Pro-Thr-Arg.

El número de mutaciones observado en la proteína p24 era menor al encontrado en la proteína p17. Las mutaciones en p24 eran 10, mientras que las observadas en la p17 ascendían a 20 mutaciones.

En el gen *pol*, además de la delección descrita anteriormente, las mutaciones se distribuían de una manera más o menos homogénea a lo largo del gen, aunque el número era ligeramente menor en la retrotranscriptasa y en la integrasa. Hay que destacar que el s61 presentaba dos mutaciones naturales relacionadas con resistencia a antirretrovíricos en la retrotranscriptasa en los aminoácidos 211 (R211K) y 214 (L214F). Estos cambios de aminoácidos han sido asociados con cierta resistencia a la combinación de los antirretrovíricos AZT + 3TC (Kemp y cols., 1998). Los tres provirus obtenidos de las células persistentemente infectadas han mantenido éstas mutaciones a lo largo de los pases en cultivo. El provirus VH61, además presentaba otra mutación de resistencia, el cambio en el aminoácido 225 (P225H), que se ha descrito como una mutación que en combinación con otras, confiere resistencia al antirretrovírico Quinoxalina (Pelemans y cols., 1997, Pelemans y cols., 1998).

En cuanto al gen *env*, las mutaciones no estaban distribuídas de una forma homogénea a lo largo del gen. Las mutaciones se acumulaban mayoritariamente desde la región C1 hasta la región C2. En la región V3, en el aminoácido en posición 7 de los tres provirus, se producía el cambio de Isoleucina (Ile) a Asparragina (Asp). Este cambio se describió como el

responsable del cambio de fenotipo sincitial a no sincitial (Olivares y cols., 1997). Hay que destacar que la región V3, aparte de ésta mutación, se ha mantenido invariable a lo largo de los pases en cultivo. La conservación de esta zona con éste virus ha sido descrita anteriormente (Sanchez-Palomino y cols., 1993, Yuste y cols., 2000).

b) Genes reguladores (*tat* y *rev*).

Las frecuencias de mutación observadas en éstos genes se encuentran en un rango entre 9.8×10^{-3} y 3.2×10^{-3} s/nt (ver tabla 4.7). En el caso del gen *tat* del VU61, de las tres mutaciones observadas, 2 son transiciones G→A y una es una transversión G→C, mientras que en el gen *rev*, de las tres mutaciones observadas, 2 son transiciones A→G y la tercera también es una transversión G→C. Hay que destacar que el porcentaje de mutaciones no sinónimas representa el 92.3 % del total, mientras que el de sinónimas representa el 7.7 %. El provirus VH61 presentaba una mutación de C por T en el triplete de terminación de la proteína *rev*, con lo que la señal de parada se desplazaba hasta el siguiente TGA (tres aminoácidos).

c) Genes accesorios (*vif*, *vpu*, *vpr* y *nef*).

Las frecuencias de mutación observadas se encontraban en un rango entre 6.8×10^{-3} y 1.1×10^{-2} s/nt (ver tabla 4.8). En los genes *vpu* y *nef*, se encuentran las frecuencias de mutación más altas, junto con la LTR y la región situada entre ésta última y el gen *gag* (secuencia “leader”), encontradas en los tres provirus con valores entre 2.8×10^{-2} y 1.1×10^{-2} s/nt. Por término general las transiciones G→A son más numerosas. En conjunto, el

porcentaje de mutaciones no sinónimas representa el 60.2 % del total, mientras que las mutaciones sinónimas representan el 39.8 %, siendo mayor que en el global.

En los genes accesorios *vif*, *vpu* y *vpr* se han observado tripletes de terminación que daban lugar a proteínas no funcionales. El provirus VM61 presentaba una acumulación de transiciones G→A, dando lugar a una secuencia hipermutada, y como consecuencia a un triplete de terminación TGA en el aminoácido en posición 89 de la proteína Vif. Por otra parte, el provirus VH61 presentaba dos tripletes de terminación TAA en los aminoácidos en posición 3 de la proteína Vpr y en posición 5 de la proteína Vpu.

d) **Regiones terminales repetidas.**

Las frecuencias de mutación en ésta región se encontraban entre las más altas del genoma, con valores comprendidos en un rango entre 1.2×10^{-2} y 2.2×10^{-2} s/nt (ver tabla 4.10). El VU61 presentaba el mayor número de mutaciones, y por lo tanto, la frecuencia de mutación más alta (2.2×10^{-2} s/nt). Las mutaciones observadas suelen acumularse en zonas donde no existen secuencias nucleotídicas de unión de factores celulares con la LTR. Sin embargo, en los tres provirus, existían dos mutaciones comunes a los tres que se encontraban en las zonas de unión de los factores celulares c-Myb y USF. Las dos mutaciones se encontraban en las posiciones 155 (T→G) y 292 (T→A).

e) **Secuencia “leader”.**

En ésta región se encuentran secuencias nucleotídicas tan importantes para la replicación del virus como el sitio de unión del iniciador de la traducción del genoma proviral

(“primer binding site”) y la señal de dimerización (DIS). En los tres provirus las frecuencias de mutación observadas se encontraban entre un rango de 1.2×10^{-2} y 6.4×10^{-3} s/nt (ver tabla 4.9). Los cambios observados eran fundamentalmente las transiciones G→A o A→G.

provirus	Mutaciones totales ^a	Frecuencia de mutación ^b	Tipos de mutaciones											
			A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
VH61	75	8.2 x 10 ⁻³	2	14	0	6	1	9	18	3	1	8	7	6
VM61	70	7.7 x 10 ⁻³	4	13	0	4	0	5	24	1	0	7	5	7
VU61	88	9.6 x 10 ⁻³	7	19	2	2	3	7	27	1	1	6	8	5

Tabla 4.5. Análisis de mutaciones del genoma completo. ^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados.

Genes estructurales (*gag*, *pol* y *env*).

	Mutaciones totales ^a	Frecuencia de mutación ^b	Mutaciones sinónimas ^c	Mutaciones no sinónimas ^d	Tipos de mutaciones											
					A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
<i>gag</i>																
VH61	14	9.2 x 10 ⁻³	8	6	1	2	0	2	0	0	2	1	1	3	1	1
VM61	13	8.6 x 10 ⁻³	8	5	0	4	0	2	0	1	2	0	0	2	1	1
VU61	13	8.6 x 10 ⁻³	9	4	0	4	0	2	0	1	2	0	0	1	2	1
<i>pol</i>																
VH61	14	4.6 x 10 ⁻³	6	8	1	2	0	3	0	2	5	0	0	1	0	0
VM61	12	3.9 x 10 ⁻³	7	5	1	1	0	2	0	1	5	0	0	1	1	0
VU61	17	5.6 x 10 ⁻³	7	10	3	1	2	0	0	3	5	1	0	1	1	0
<i>env</i>																
VH61	21	8.1 x 10 ⁻³	12	9	1	5	0	1	0	2	2	1	0	3	5	1
VM61	23	8.8 x 10 ⁻³	12	11	3	3	0	0	0	3	6	1	0	2	1	4
VU61	25	9.6 x 10 ⁻³	13	12	3	7	0	0	2	2	4	1	1	2	3	0

Tabla 4.6. Análisis de mutaciones observadas en los genes estructurales. ^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados. ^c Mutaciones que no suponen cambio de aminoácido. ^d Mutaciones que suponen cambio de aminoácido.

Genes reguladores (*tat* y *rev*).

	Mutaciones totales ^a	Frecuencia de mutación ^b	Mutaciones sinónimas ^c	Mutaciones no sinónimas ^d	Tipos de mutaciones											
					A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
<i>tat</i>																
VH61	1	3.2 x 10 ⁻³	0	1	0	0	0	0	0	0	0	1	0	0	0	0
VM61	2	6.5 x 10 ⁻³	0	2	0	0	0	0	0	0	1	1	0	0	0	0
VU61	3	9.8 x 10 ⁻³	1	2	0	0	0	0	0	0	2	1	0	0	0	0
<i>rev</i>																
VH61	3	8.6 x 10 ⁻³	0	3	0	2	0	0	0	0	0	1	0	0	0	0
VM61	2	5.7 x 10 ⁻³	0	2	0	0	0	0	0	0	1	1	0	0	0	0
VU61	2	5.7 x 10 ⁻³	0	2	0	1	0	0	0	0	0	1	0	0	0	0

Tabla 4.7. Análisis de mutaciones observadas en los genes reguladores. ^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados. ^c Mutaciones que no suponen cambio de aminoácido. ^d Mutaciones que suponen cambio de aminoácido.

Genes accesorios (*vif*, *vpr*, *vpu* y *nef*).

	Mutaciones totales ^a	Frecuencia de mutación ^b	Mutaciones sinónimas ^c	Mutaciones no sinónimas ^d	Tipos de mutaciones											
					A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
<i>vif</i>																
VH61	4	6.8 x 10 ⁻³	2	2	0	0	0	0	0	2	1	1	0	0	0	0
VM61	6	1.0 x 10 ⁻²	2	4	0	0	0	0	0	0	5	0	0	0	1	0
VU61	3	5.1 x 10 ⁻³	2	1	0	1	0	0	0	1	1	0	0	0	0	0
<i>vpr</i>																
VH61	6	2.0 x 10 ⁻²	2	4	0	0	0	0	0	2	2	0	0	0	1	1
VM61	2	6.8 x 10 ⁻³	1	1	0	0	0	0	0	0	2	0	0	0	0	0
VU61	3	1.0 x 10 ⁻²	1	2	0	1	0	0	0	0	1	0	0	0	1	0
<i>vpu</i>																
VH61	7	2.8 x 10 ⁻²	3	4	0	0	0	0	1	1	3	0	0	0	1	1
VM61	3	1.2 x 10 ⁻²	0	3	0	0	0	0	0	0	2	0	0	0	0	1
VU61	5	2.0 x 10 ⁻²	2	3	0	0	0	0	1	0	3	0	0	0	0	1
<i>nef</i>																
VH61	7	1.1 x 10 ⁻²	3	4	0	2	0	0	0	0	2	0	0	1	0	2
VM61	8	1.2 x 10 ⁻²	4	4	0	2	0	0	0	1	2	0	0	1	1	1
VU61	13	2 x 10 ⁻²	5	8	1	1	0	0	0	0	6	0	0	1	1	3

Tabla 4.8. Análisis de mutaciones observadas en los genes accesorios. ^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados. ^c Mutaciones que no suponen cambio de aminoácido. ^d Mutaciones que suponen cambio de aminoácido.

Regiones terminales repetidas

LTR	Mutaciones totales ^a	Frecuencia de mutación ^b	Tipos de mutaciones											
			A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
VH61	8	1.2 x 10 ⁻²	0	3	0	0	0	1	1	0	0	2	0	1
VM61	10	1.5 x 10 ⁻²	0	3	0	0	0	1	2	0	0	3	0	1
VU61	14	2.2 x 10 ⁻²	0	3	0	0	0	1	5	0	0	2	1	2

Tabla 4.9. Análisis de mutaciones observadas en la LTR .^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados.

Región desde la LTR al principio del gen *gag*.

Sec. leader	Mutaciones totales ^a	Frecuencia de mutación ^b	Tipos de mutaciones											
			A→C	A→G	A→T	C→A	C→G	C→T	G→A	G→C	G→T	T→A	T→C	T→G
VH61	2	1.2 x 10 ⁻²	0	1	0	0	0	0	1	0	0	0	0	0
VM61	1	6.4 x 10 ⁻³	0	1	0	0	0	0	0	0	0	0	0	0
VU61	2	1.2 x 10 ⁻²	0	1	0	0	0	0	1	0	0	0	0	0

Tabla 4.10. Análisis de mutaciones observadas en la región desde la LTR al gen *gag* .^a Mutaciones de diferencia entre el provirus del aislado inicial s61 y los tres provirus obtenidos de las células persistentemente infectadas. ^b Mutaciones observadas entre el número total de nucleótidos secuenciados.

4.2.6.2.2. Relación entre mutaciones sinónimas y no sinónimas.

El estudio de la relación que existe entre las mutaciones que no suponen cambio de aminoácido (sinónimas) y las mutaciones que suponen cambio de aminoácido (no sinónimas), se utiliza generalmente para determinar si una región genómica se encuentra sometida a presión selectiva. Este método ha sido utilizado previamente para estudiar la evolución en el VIH-1 (Ganeshan y cols., 1997) (Shiino y cols., 2000). En nuestro estudio se ha realizado por grupos de genes para ver que regiones del genoma se han visto sometidas a presiones selectivas durante el proceso de la persistencia. El análisis se ha llevado a cabo con al ayuda del programa SNAP y los resultados se muestran en la tabla 4.11. y graficamente en la figura.

Genes ^a	s ^b	n ^c	S ^d	N ^e	ps ^f	pn ^g	pn / ps ^h	Selección positiva ⁱ
Estructurales	74.83	77.17	4488.84	16799.16	0.017	0.005	0.3	NO
Reguladores	1	12	447	1524	0.002	0.008	4	SI
Accesorios	23	40	1081.16	4129.84	0.02	0.01	0.5	NO

Tabla 4.11. Relación entre mutaciones sinónimas y no sinónimas. ^a Genes estructurales (*gag*, *pol* y *env*), Reguladores (*tat* y *rev*) y Accesorios (*vpr*, *vpu*, *nef* y *vif*). ^b s : Número de sustituciones sinónimas observado. ^c n: Número de sustituciones no sinónimas observado. ^d S: Número medio de sustituciones sinónimas posibles. ^e N: Número medio de sustituciones no sinónimas posibles. ^f Proporción de mutaciones sinónimas (s / S). ^g Proporción de mutaciones no sinónimas (n / N). ^h Relación entre proporción de no sinónimas y sinónimas. ⁱ Se considera que existe selección positiva cuando el cociente pn / ps es mayor de 1.

Los valores de la tabla apoyan la hipótesis de que haya existido una selección positiva en los genes reguladores, aunque éste valor puede estar sesgado debido al pequeño tamaño de *tat* (306 nucleótidos) y *rev* (351 nucleótidos para el VM61 y VU61, y 360 para el VH61).

Resultados SNAP

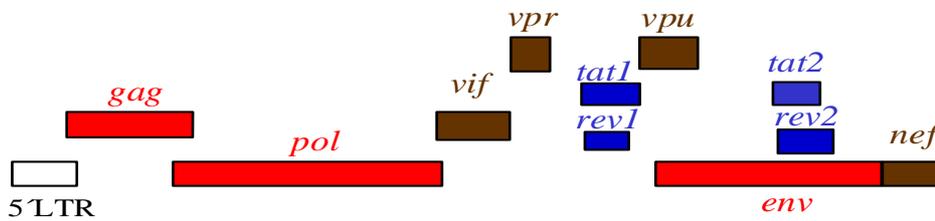
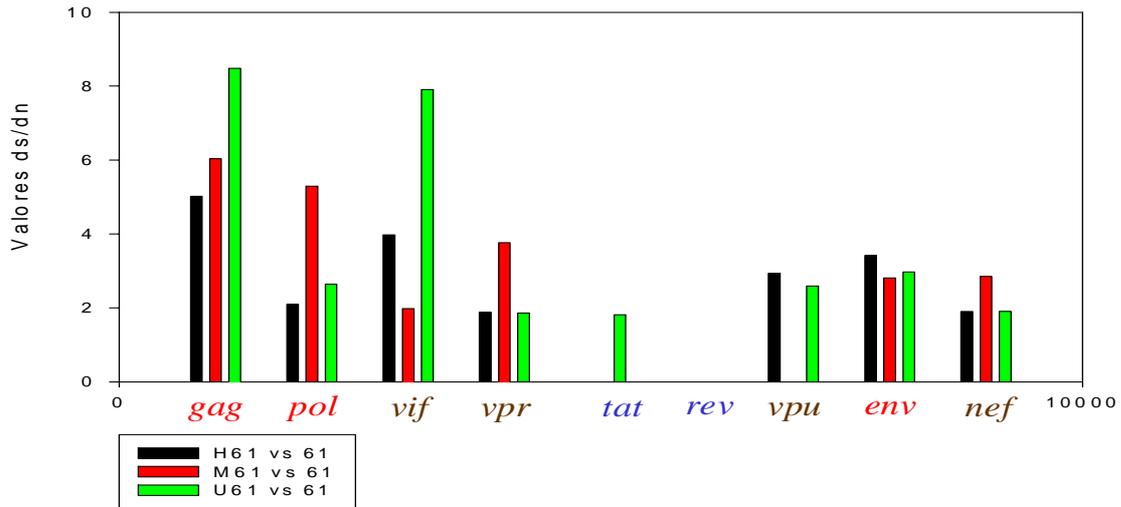


Figura 4. . Representación gráfica del análisis realizado mediante el programa SNAP. En el eje de ordenadas se representa la relación ds /dn, es decir, la relación entre mutaciones sinónimas y no sinónimas, y en el de abscisas el genoma separado por genes.

4.2.6.2.3. Descripción y localización de las mutaciones no sinónimas en los tres provirus.

Una vez estudiadas las relaciones entre las mutaciones sinónimas y no sinónimas, analizamos en detalle los cambios de aminoácido que se habían producido durante la persistencia, intentando describir las alteraciones en las proteínas víricas que pueden ser responsables de éste proceso. La descripción de los cambios de aminoácido asociados encontrados en cada uno de los genes viene descrita en las tablas 4.12, 4.13 y 4.14.

Mutaciones no sinónimas VH61.Tabla 4.12

<i>gag</i>	Localización	Frecuencia en base de datos
------------	--------------	-----------------------------

A67S	p17	Alta
G121D	p17	Alta
F6Y	p7	Muy baja
K11R	p7	Alta
T8P	p6	No observada
L30Q	p6	Alta
<i>pol</i>		
I36M	proteasa	Alta
T39A	retrotranscriptasa	Baja
D192N	retrotranscriptasa	Muy baja
P225H	retrotranscriptasa	No observada
P243L	retrotranscriptasa	No observada
D324E	retrotranscriptasa	Alta
A94T	Rnasa H	Baja
D256E	integrasa	Alta
<i>env</i>		
M4T	péptido señal	Alta
K85E	C1	Baja
P147T	V1	Muy baja
D177E	V2	No observada
E268K	C2	Alta
V283N	C2	Baja
I302N	V3	Alta
L661W	gp41	No observada
T700V	gp41	Muy baja
<i>vif</i>		
R92K		Alta
V142L		No observada
<i>vpr</i>		
Q3Stop	Dominio N-terminal	No observada
H45Y	Final de la región del lazo	Alta
D52E	Dominio helicoidal II	No observada
R86Q	Dominio rico en aminoácidos básicos	Alta
<i>vpu</i>		
Q5Stop	Dominio transmembrana	No observada
L7V	Dominio transmembrana	Alta
R35K	Dominio citoplasmático	Alta
N37K	Dominio citoplasmático	Alta
<i>tat</i>		
R73P		Alta
<i>rev</i>		
A27P	Sitio de unión al RRE	Alta
T108A	Región C-terminal	Alta
I109V	Región C-terminal	Alta
<i>nef</i>		
S8N	Sitio potencial de localización nuclear	Muy baja
I11S	Sitio potencial de localización nuclear	Baja
N46S		Alta
I101V	Dominio del core	Alta

Mutaciones no sinónimas VM61.Tabla 4.13

<i>gag</i>	Localización	Frecuencia en base de datos
D96N	p17	No observada

G121D	p17	Alta
N9S	p2	Alta
L30Q	p6	Alta
S38L	p6	Alta
<i>pol</i>		
D324E	retrotranscriptasa	Alta
H17P	Integrasa	No observada
S18N	Integrasa	Alta
I161V	Integrasa	No observada
D256E	Integrasa	Alta
<i>env</i>		
M26I	péptido señal	Alta
T49P	C1	No observada
S56A	C1	No observada
K85R	C1	Muy baja
K87E	C1	Alta
S143N	V1	Alta
V283N	C2	Alta
I302N	V3	Alta
A607T	gp41	Alta
T700V	gp41	Muy baja
V705G	gp41	No observada
<i>vif</i>		
W89Stop		No observada
R92K		Alta
G143E		No observada
R155K		Alta
<i>vpr</i>		
R86Q	Dominio rico en aminoácidos básicos	Alta
<i>vpu</i>		
L7V	Dominio transmembrana	Alta
R35K	Dominio citoplasmático	Alta
D81N	Dominio citoplasmático	Alta
<i>tat</i>		
R57K	Dominio básico	Muy baja
R73P		Alta
<i>rev</i>		
E11K	Región N-terminal	No observada
A27P	Región N-terminal	Alta
<i>nef</i>		
N46S		Alta
F68L		Muy baja
I101V	Dominio del core	Alta
A190T	Dominio del core	No observada

Mutaciones no sinónimas VU61. Tabla 4.14

<i>gag</i>	Localización	Frecuencia en base de datos
G121D	p17	Alta
F32L	p24	Muy baja
N9S	p2	Alta

K41R	p7	Baja
<i>pol</i>		
T39A	retrotranscriptasa	Baja
E173K	retrotranscriptasa	Muy alta
H208L	retrotranscriptasa	No observada
E396D	retrotranscriptasa	Baja
T33P	Rnasa H	No observada
A94T	Rnasa H	Baja
V249E	Integrasa	No observada
P261L	Integrasa	No observada
D279H	Integrasa	No observada
<i>env</i>		
D78A	C1	No observada
K85E	C1	Baja
S143N	V1	Alta
A227G	C2	Muy baja
S256L	C2	No observada
L260V	C2	No observada
V283N	C2	Alta
I302N	V3	Alta
K476R	C5	Alta
V496I	C5	Alta
T700V	gp41	Muy baja
C837F	gp41	Alta
<i>vif</i>		
I128V		Alta
<i>vpr</i>		
S41G	Dominio helicoidal I	Alta
R86Q	Dominio rico en aminoácidos básicos	Alta
<i>vpu</i>		
L7V	Dominio transmembrana	Alta
R35K	Dominio citoplasmático	Alta
N37K	Dominio citoplasmático	Alta
<i>tat</i>		
R19K		Alta
R73P		Alta
<i>rev</i>		
A27P		Alta
I109V	Región C-terminal	Alta
<i>nef</i>		
S8N		Muy baja
I11S		Baja
D28E		Alta
N46S		Alta
I100L		Alta
R105K		Alta
V114I		Alta
M173I		Alta

4.2.6.2.3. Relaciones filogenéticas de los genomas completos secuenciados.

Con el objeto de intentar ver si las mutaciones observadas en el establecimiento de la persistencia, se habían acumulado de manera similar en los provirus obtenidos de las dos líneas celulares T (H61 y M61) y si éstas eran diferentes a las observadas en el provirus obtenido de la línea celular promonocítica (U-937 clon 2), se procedió a un análisis filogenético de las secuencias de los genomas completos. Las matrices de distancias obtenidas mediante el modelo de dos parámetros de Kimura se utilizaron para la construcción del árbol filogenético mediante el método del vecino más próximo (“Neighbor-Joining”). Los genomas de los provirus obtenidos de las líneas celulares T (VH61 y VM61), se agrupaban en la misma rama, con un valor alto de remuestreo de 91 %. El genoma del provirus obtenido de la línea celular promonocítica (VU61) se agrupaba en una rama independiente de los otros 2 y separado del s61. El árbol filogenético se muestra en la Figura 4.3.

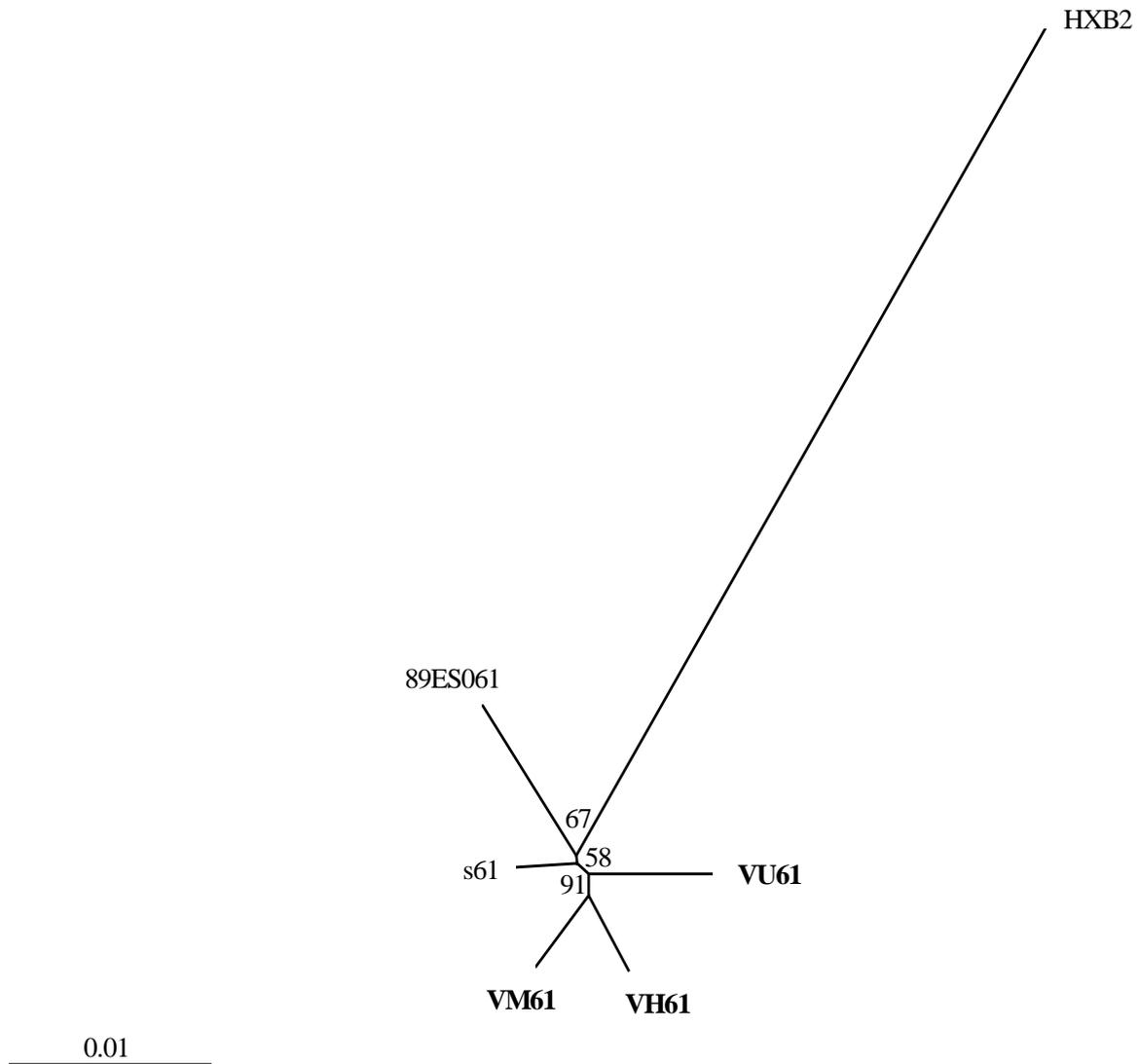


Figura 4.3. Análisis filogenético de las secuencias nucleotídicas de los genomas completos de los provirus obtenidos de la persistencia (en negrita) y del provirus s61 mediante el método del vecino más próximo (“Neighbor Joining”). Los números sobre las ramas representan los valores de remuestreo (“bootstrap”). La barra de la escala representa un 1 % de distancia genética entre las muestras.

4.2.6.2.4. Mutaciones comunes en los dos provirus VH61 y VM61 obtenidos de las líneas celulares T persistentemente infectadas.

En el análisis filogenético se ha observado que se agrupan los provirus VH61 y VM61 en la misma rama, con un alto valor de remuestreo, indicando que están filogenéticamente muy próximos. A pesar de tratarse de dos líneas independientes de obtención de persistencia, existe un gran número de mutaciones comunes. Los dos provirus compartían 34 mutaciones. Prácticamente de la mitad de todas las mutaciones observadas en los genomas del VH61 y del VM61 eran las mismas. La distribución de éstas mutaciones fue: 8 en la LTR, 1 en la secuencia leader, 7 en *gag*, 6 en *pol*, 8 en *env*, 1 en *vif*, 3 en *nef*, 1 en *rev*, 1 en *tat*, 1 en *vpr* y 2 en *vpu*. Considerando las mutaciones por genes obtenemos un número de 39 mutaciones. Esta diferencia es debida a que existen mutaciones en zonas en que varios genes solapan. Existe una mutación en la zona que solapan los genes *env*, *tat* y *rev*. Otra mutación se encuentra en la zona que solapan *gag* y *pol*. Por último, existen 2 mutaciones en la zona en la que solapan el gen *nef* y la LTR.

4.2.6.2.5. Mutaciones comunes a los tres provirus.

En el apartado anterior, hemos visto que existían mutaciones comunes en los provirus obtenidos de líneas celulares T. A continuación, nos interesaba conocer el número de mutaciones comunes que existían entre los tres provirus, independientemente de la línea celular. Éste dato nos podía indicar qué mutaciones se habían seleccionado en el proceso del establecimiento de la persistencia en los tres linajes independientes. Los tres provirus, VH61, VM61 y VU61 presentaban 24 mutaciones comunes. De éste dato se deduce que 10 mutaciones son específicas del establecimiento de la persistencia en las líneas T. La

distribución de éstas mtaciones fue: 4 en la LTR, 1 en *gag* (solapa con el gen *pol*), 4 en *pol* (1 solapa con *gag*), 1 en *vif* y 1 en *env*.

De las 24 mutaciones comunes en los tres provirus, 5 se encuentran en regiones que solapan, por lo tanto, como hemos analizado las mutaciones por genes independientes, el número de mutaciones es superior, obteniéndose un total de 28. De éstas, 23 se encontraban en las regiones codificantes, donde 14 correspondían a mutaciones sinónimas y 9 eran no sinónimas. Las otras 5 mutaciones se distribuían 2 en la LTR y 3 en la región desde la LTR al principio de *gag*. La distribución de las mutaciones no sinónimas se detalla en la Figura 4.4.

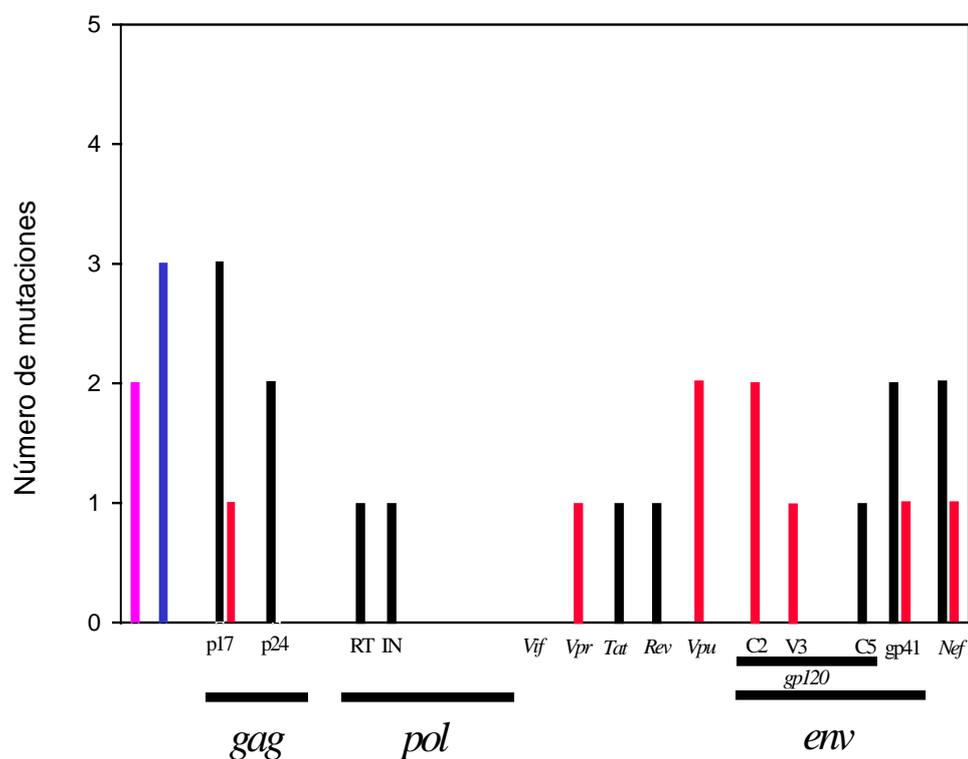


Figura 4.4. Distribución a lo largo del genoma de las mutaciones comunes de los tres provirus obtenidos de las células persistentemente infectadas. En el eje de abscisas se encuentran representadas las regiones y genes que forman el genoma del VIH-1. En el eje de ordenadas se representa el número de mutaciones observadas. Las barras de color rojo representan las mutaciones no sinónimas. Las barras de color negro representan las mutaciones sinónimas. Las barras rosa y azul representan las mutaciones encontradas en la LTR y la región de la secuencia leader.

4.2.6.2.6. Análisis de la cuasiespecie vírica en los tres provirus a un pase de persistencia reciente (pase 35) y de persistencia establecida (último pase dado en cultivo) en la región C2-V5 de la envuelta.

El propósito de éste análisis fue comparar la composición de la cuasiespecie a lo largo del proceso del establecimiento de la persistencia. Hemos considerado el pase 35 como un punto de persistencia reciente, y alrededor del pase 100 como un punto de persistencia establecida. Los alineamientos correspondientes se muestran en el Anexo III.

Para realizar el estudio de las cuasiespecies víricas de los tres provirus se realizó clonaje molecular de la región desde C2 a V5 del gen *env* (comprendida entre los primers 27 y 22). Se secuenciaron a pase 35: 10 clones del VH61, 13 clones del VM61 y 10 clones del VU61. En el punto de persistencia establecida se secuenciaron: 11 clones del VH61, 10 clones del VM61 y 12 clones del VU61. Las frecuencias de mutación observadas al comparar con la secuencia consenso obtenida en el pase correspondiente, se muestran en la tabla 4.16. En el análisis de la cuasiespecie del provirus VM61 se encontraron cuatro clones hipermutados. Éstos clones no se tuvieron en cuenta en el análisis de la frecuencia de mutación.

provirus	persistencia temprana	persistencia tardía
VH61	3.7×10^{-3}	1.7×10^{-3}
VM61	3.1×10^{-3}	2.1×10^{-3}
VU61	4.9×10^{-3}	3.1×10^{-3}

Tabla 4.16. Frecuencia de mutación expresada en s/nt observada en los tres provirus.

En el análisis de los distintivos clones que componen la cuasiespecie se encontraron en algunos casos un número elevado de clones no viables atendiendo a la secuencia de esta zona. Los clones no viables observados lo eran por la existencia de codones de terminación o, en algún caso, por la existencia de clones hipermutados. El número de clones no viables obtenidos en cada uno de los casos se encuentra resumido en la tabla 4.17.

Provirus	persistencia temprana	persistencia tardía
VH61	10/10 (100 %)	10/11 (91 %)
VM61	7/13 (54 %)	9/10 (90 %)
VU61	6/10 (60 %)	9/12 (75 %)

Tabla 4.17 Clones viables

Como se puede apreciar en las tablas anteriores, acompañando al establecimiento de la persistencia se observó una bajada global de la frecuencia de mutación y un aumento en el número de clones viables.

5. DISCUSIÓN

5. Discusión.

El trabajo realizado en ésta tesis se ha encaminado al estudio de la importancia de los genes accesorios y su implicación en la biología del VIH-1. Para ello se han realizado dos tipos de estudios:

- a) Análisis filogenético de aislados naturales, cuya principal característica era la imposibilidad de ser cocultivados y un posterior estudio de posibles alteraciones en el gen accesorio *nef* que pudiesen estar influyendo en éste fenotipo.
- b) Estudio y caracterización del establecimiento de infecciones persistentes con un mismo aislado en tres líneas celulares diferentes y participación de los genes accesorios en éste proceso.

El objeto de centrarnos en el estudio de los genes accesorios ha sido no solo por su importancia como factores de virulencia, sino porque también juegan un importante papel en las interacciones entre los virus y su hospedador (Trono, 1998).

5.1. Análisis filogenético de las muestras cocultivo negativo de ONUSIDA.

La rápida expansión del VIH en el mundo, y principalmente en el continente Africano, pone de manifiesto la urgencia del desarrollo de una vacuna. Se ha postulado que la variación genética podría representar uno de los principales obstáculos en el desarrollo de ésta vacuna, debido fundamentalmente a la falta de protección cruzada entre los subtipos.

En éste estudio, hemos analizado un grupo de muestras que eran incapaces de replicar en PBMCs, obtenidas de diferentes países y continentes, a través del programa de las Naciones Unidas para el aislamiento y caracterización de muestras. El análisis filogenético llevado a cabo ha mostrado que las muestras procedentes de Sudáfrica, India y China se agrupaban en la misma rama de los árboles filogenéticos con unos valores altos de remuestreo. Por el contrario, las muestras procedentes de Brasil y de Israel, aunque se agrupaban dentro del subtipo C, se encontraban estadísticamente segregadas. Sin embargo,

datos de un estudio paralelo con muestras cocultivo positivo, no mostraban ésta segregación y todas las muestras se agrupaban dentro del subtipo C (Dr. M. Kalish, comunicación personal).

Las muestras procedentes de Brasil se encontraban separadas estadísticamente en dos de las tres regiones genómicas analizadas (*p17-gag* y gen *nef*), aunque la distancia genética no era muy alta. Ésta segregación podía ser explicada por la estrecha relación entre las muestras brasileñas, puesto que todas ellas procedían de la ciudad de Porto Alegre y además, la muestra que hemos empleado como referencia (BR025) del subtipo C pertenece a la misma ciudad, y por ese motivo se agrupa en la misma rama que las muestras cocultivo negativo de nuestro estudio. Otra posible explicación de ésta segregación podría estar relacionada con el número limitado de muestras utilizados en nuestro estudio, así como del bajo número de países de los cuales disponíamos de muestras.

Fenotípicamente, todas las muestras analizadas fueron NSI de acuerdo a los marcadores observados en las posiciones 307, 319 y 320 de las secuencias en la región V3. El fenotipo NSI, según la literatura, es el predominante entre las muestras correspondientes al subtipo C (Bjorndal y cols., 1999, Van Harmelen y cols., 1999).

Además de éste análisis filogenético de las muestras, estábamos interesados en el estudio de la propiedad de cocultivo negativo que presentaban éstas cepas y la implicación que podría tener el gen accesorio *nef* en éste fenotipo. En éste sentido, hemos encontrado una acumulación de modificaciones en las secuencias aminoacídicas de regiones con funciones importantes en el gen *nef*. Se ha encontrado una proporción alta de alteraciones entre las muestras cocultivo negativo en contraposición a las muestras control cocultivo positivo, entre las que solo se ha encontrado una secuencia con algún tipo de alteración. La presencia de varias alteraciones en un gran número de motivos funcionales del gen, parecería indicar que la proteína Nef podría no ser funcional. Los cambios más sorprendentes se han encontrado en las muestras IN14 e IN5, que claramente presentan una proteína no funcional. La primera tenía una delección importante desde el primer nucleótido de la proteína hasta justamente el comienzo de la LTR, y la segunda presentaba dos tripletes de terminación. En ambos casos, las secuencias de la LTR que solapan con el gen *nef*, permanecían intactas.

La importancia de la acumulación de cambios se ve reflejada por el bajo número de alteraciones encontradas en las secuencias de las muestras control cocultivo positivo.

Nuestros datos se apoyan en observaciones anteriores descritas en la literatura, donde existen numerosos ejemplos en los que se relaciona la presencia de proteínas Nef no funcionales con pacientes no progresores, los cuales presentan generalmente baja carga viral y una limitada replicación vírica (Deacon y cols., 1995, Kirchhoff y cols., 1995, Salvi y cols., 1998).

En resumen, hemos analizado un grupo de muestras que se agrupaban mayoritariamente dentro del subtipo C. Solamente una de ellas fue un recombinante C/B/C. Las muestras procedentes de Brasil se encontraban segregadas de la otras pertenecientes al subtipo C, probablemente debido a un origen geográfico clonal. Finalmente, hemos encontrado una acumulación de alteraciones en el gen *nef* entre las muestras que podría estar asociado con el fenotipo de cocultivo negativo.

5.2. Establecimiento de las infecciones persistentes in vitro: mutaciones asociadas.

En ésta parte del trabajo, nos interesaba estudiar el proceso del establecimiento de la persistencia. Nuestro objetivo con éste estudio era el de intentar acercarnos, mediante un sistema *in vitro*, lo más posible a lo que sucede en una infección natural persistente. Para ello nos centramos en el estudio de la evolución que experimenta el virus en diferentes líneas celulares sometiéndolo para ello a pases sucesivos en cultivo. Lógicamente, el sistema *in vitro* tiene una serie de limitaciones importantes a la hora de intentar reproducir lo que está sucediendo *in vivo* en el paciente. De todas ellas, una de ellas es la ausencia de una presión inmune. Para intentar paliar lo más posible éstas diferencias, elegimos un aislado natural (s61) del VIH-1 para la realización del estudio. El motivo de elegir un aislado natural para realizar las infecciones en lugar de un clon infeccioso, fue el de partir de una cuasiespecie vírica heterogénea y ver la evolución de la misma y la selección de poblaciones de la cuasiespecie a lo largo de los pases en tres líneas celulares distintas. La decisión de elegir tres líneas celulares (dos líneas linfoides T CD4⁺) y una promonocítica, fue la de observar dicha evolución en diferentes líneas, ya que en el paciente las células que mayoritariamente resultan infectadas por el virus son linfocitos T CD4⁺ y monocito-macrófagos. En éste proceso, la

primera selección de variantes que ha sufrido el aislado s61 ha sido la adaptación a cada una de las líneas celulares, y posteriormente se han ido seleccionando aquellas variantes de la cuasiespecie que eran menos citopáticas para las células y que han permitido una infección persistente.

Para estudiar los cambios que se habían producido en las secuencia de los tres provirus, obtenidos de las células persistentemente infectadas, en comparación con la secuencia del aislado s61, los cuales han contribuido al establecimiento de la persistencia, se secuenciaron los genomas completos de los cuatro provirus.

5.2.1. Frecuencias de mutación observadas.

El alto grado de variación encontrado in vivo entre las secuencias nucleotídicas del VIH-1 es el resultado de la alta tasa de error asociada a la RT, a los altos niveles de replicación vírica que se producen en un paciente y al prolongado periodo de infección que caracteriza a la enfermedad (Coffin, 1995, Perelson y cols., 1996, Temin, 1989, Wain-Hobson, 1996). En nuestro caso a pesar de tratarse de un sistema in vitro también encontramos unas altas frecuencias de mutación. Las frecuencias de mutación observadas al comparar los genomas completos de los tres provirus, con el aislado s61, fueron muy similares (entre 7.7×10^{-3} y 9.6×10^{-3} s/nt). En este caso, se pueden atribuir principalmente a la alta tasa de error de la RT y la selección de variantes menos citopáticos en el establecimiento de la persistencia. Sin embargo, sí se apreciaron diferencias cuando se realizó el análisis detallado de cada gen (ver Tabla 4.5).

5.2.2. Distribución de las mutaciones.

Se ha descrito previamente que la variabilidad genética del VIH-1 no es uniforme a lo largo del genoma existiendo unas regiones más conservadas que otras (Quinones-Mateu y cols., 1996). En nuestro estudio hemos encontrado una diferencia en la distribución de las mutaciones que, en ciertos genes, es mucho más marcada que la descrita previamente. En el proceso del establecimiento de la persistencia se ha podido observar que las frecuencias de mutación más altas se encontraban generalmente localizadas en los genes accesorios, en la secuencia “leader” y en la LTR. Estos datos indican que existen diferencias en los mecanismos que actúan dentro de cada grupo de genes a la hora del establecimiento de la persistencia ya que se aceptarían más mutaciones en éstas regiones que los genes estructurales y por lo tanto, existiría una menor flexibilidad en los genes *gag*, *pol* y *env*.

Aunque hemos observado la baja capacidad para aceptar cambios en los genes estructurales, existen dentro de dichos genes unas zonas que están más conservadas que otras. En el gen *gag* se observó que la secuencia nucleotídica que codifica para la proteína p17 presentaba el doble de mutaciones que la región que codifica para la p24. En el gen *pol*, la distribución es más o menos homogénea, aunque el número de mutaciones es ligeramente menor en la retrotranscriptasa e integrasa. En el gen *env*, las mutaciones encontradas no se distribuían de una forma homogénea, existiendo zonas como la región V3 en las que sólo existía un cambio con respecto a la secuencia del s61.

Es especialmente relevante la observación de la baja variabilidad observada en el gen *env* que ha sido descrito como particularmente variable especialmente en las regiones denominadas variables V1-V5. El lazo V3 se ha venido considerando como la zona que acumula una mayor variabilidad a lo largo del genoma del VIH-1 al analizar distintos aislados primarios. Probablemente la ausencia de variación observada en este caso se deba a que se trata de un sistema *in vitro*, es decir en ausencia de presión inmune. En una infección natural V3 es el sitio con mayor poder inmunogénico por lo que se ve sometido a una presión selectiva continua por parte del sistema inmune. En un sistema *in vivo* existiría una continua dinámica de selección de variantes de escape a la respuesta inmune. En nuestro caso la única presión a la que se ve sometida esa zona del genoma durante los pases que llevan al

establecimiento de la persistencia, es la necesidad de un cambio de fenotipo sincitial a no sincitial que mapea en el lazo V3 (Olivares y cols., 1997). En el establecimiento de la persistencia, se han seleccionado los provirus que han adquirido el fenotipo de NSI ya que al no ser citopáticos, permitirían la supervivencia de las células infectadas. Recientemente se ha publicado otro caso en el que se observa una ausencia de variación en la envuelta, especialmente marcada en V3 con una ausencia total de variación. En este caso se trata también de un sistema *in vitro* en el que se estudia la acumulación de mutaciones tras pases seriados en cuello de botella (Yuste y cols., 2000)

Las diferencias en cuanto a la acumulación de mutaciones en los genes *gag* y *pol* son similares a las descritas estudiando la variación de distintos aislados de VIH-1. Probablemente estas diferencias sean debidas principalmente a restricciones de funcionalidad inherentes a dichas proteínas. En un apartado posterior analizaremos con mayor detalle la variación en los genes accesorios.

5.3. Relación filogenética entre los virus persistentes en líneas T y en U-937.

Estábamos interesados en conocer si el proceso de la persistencia se había producido de manera similar en las líneas celulares T y éste era diferente del producido en la línea promonocítica. Para ello, realizamos un análisis filogenético de las secuencias nucleotídicas de los genomas completos de los tres provirus obtenidos de las células persistentemente infectadas. El resultado de éste análisis (ver apartado 4.2.8.2.3), mostraba que la persistencia se había producido de manera similar en las células de las líneas T, ya que los provirus VH61 y VM61 se agrupaban en la misma rama del árbol filogenético con un valor de remuestreo del 91 % , y se separaban de la secuencia nucleotídica del provirus VU61 (ver resultados). Si estaban tan próximos filogenéticamente el VH61 y VM61, deberían de tener un elevado número de mutaciones comunes a los dos. En el estudio en detalle de las mutaciones comunes a los dos se observó que compartían 34 mutaciones. Este número era bastante alto puesto que las mutaciones totales observadas en todo el genoma eran de 75 para el VH61 y 70 para el VM61.

Por otra parte, nos interesaba conocer el número de mutaciones comunes que existían entre los tres provirus, independientemente de la línea celular en la que se había establecido la persistencia. El objeto de éste análisis era el de intentar determinar las mutaciones que se habían seleccionado en el transcurso del fenómeno de la persistencia. El análisis nos mostraba que existían 24 mutaciones que compartían los tres provirus (ver resultados). La probabilidad de que se fije una determinada mutación en tres líneas independientes es muy baja. Por lo tanto, cabría pensar que éstas 24 mutaciones se han seleccionado como consecuencia del proceso de persistencia. De éste dato establecimos que las mutaciones específicas de las líneas T eran 10. La distribución de éstas mutaciones fue: 4 en la LTR, 1 en *gag* (solapa con el gen *pol*), 4 en *pol* (1 solapa con *gag*), 1 en *vif* y 1 en *env*. El alto número de mutaciones comunes que se acumulan en la LTR y en el gen *pol* podrían tener que ver con la adaptación a línea celular.

5.4. Mutaciones en los genes accesorios posiblemente relacionadas con el fenotipo de persistencia.

Teniendo en cuenta, la alta acumulación de mutaciones observada en los genes accesorios de los tres provirus, así como la bibliografía existente que relaciona el fenotipo de persistencia con determinadas mutaciones en dichos genes, nos dispusimos a analizar con mayor detalle las mutaciones observadas en los genes accesorios en cada una de las líneas celulares.

5.4.1. Persistencia en la línea celular H9.

En la secuenciación del genoma completo del provirus denominado VH61 obtenido de las células H9 persistentemente infectadas, se han observado dos mutaciones que daban lugar a tripletes de terminación en los genes *vpr* (aminoácido en posición 3) y *vpu* (aminoácido en posición 5). Estas mutaciones originaban proteínas no funcionales. La no funcionalidad tanto de la proteína Vpr como de la proteína Vpu, se ha relacionado con un fenotipo de virus persistente en un clon de la línea celular MT-4 (M10) y en la línea celular H9 (Kishi y cols., 1992, Mustafa & Robinson, 1993, Nakaya y cols., 1996, Nishino y cols., 1994). Es significativo que en el estudio realizado con la misma línea celular (H9) las mutaciones que se

han encontrado asociadas a la persistencia, son proteínas Vpr y Vpu no funcionales, al igual que lo observado en nuestro estudio.

En la caracterización fenotípica de las células H9 persistentemente infectadas realizada por Lourdes Muñoz se observó que en el punto de persistencia establecida, solamente el 30 % de las células del cultivo expresaban proteínas víricas (determinado por IFI) y no se observaba expresión del receptor CD4. Estos datos podrían explicarse teniendo en cuenta la funcionalidad de la proteína Nef. Dicha proteína tiene un efecto de bajada en los niveles de expresión de CD4. El hecho de que no aumente el porcentaje de células infectadas puede deberse a que la ausencia de expresión del receptor hace que no existan células susceptibles de ser infectadas por los virus que se liberan. Este dato está apoyado por el hecho de que éstas células (H61) no son susceptibles de sobreinfectarse por las cepas de virus IIIB y s61. Además no se activó la replicación con TNF- α y PMA. Esta observación también apoya la hipótesis de la no replicación por la ausencia de células susceptibles.

Por otra parte el virus VH61 es capaz de infectar H9 y MT-4 pero no U-937 clon 2. Este dato puede estar relacionado con alguna mutación de las descritas anteriormente probablemente asociadas con la especificidad para líneas T.

5.4.2. Persistencia en la línea celular MT-4.

En la secuenciación del genoma completo del provirus denominado VM61 obtenido de las células MT-4 persistentemente infectadas, se ha observado una mutación G→A que originaba un triplete de terminación en el gen *vif* (aminoácido en posición 89). Además de ésta mutación, se observa una hipermutación G→A. La tendencia de la RT a producir éste tipo de mutaciones ha sido ampliamente descrita (Pathak & Temin, 1990, Vartanian y cols., 1991). La acumulación de éstas mutaciones en el gen *vif*, probablemente se debe al hecho de que al no dar una proteína funcional no es objeto de selección y se acumulan todas las mutaciones que se producen. La no funcionalidad de la proteína Vif, se ha relacionado con un fenotipo de virus persistente en un clon de la línea celular MT-4 (M10) (Kishi y cols., 1992, Nakaya y cols., 1996).

En la caracterización fenotípica de las células MT-4 persistentemente infectadas se observó que en el punto de persistencia establecida, el 100 % de las células del cultivo expresaban proteínas víricas (determinado por IFI). Se observaba que el 7 % de las células expresaban el receptor CD4. Al igual que las H61, no se activaba la replicación con TNF- α y PMA.

Al intentar infectar con el virus VM61 las células H9, MT-4 y U-937 clon 2, solamente se infectaron las células MT-4. Este dato era esperable teniendo en cuenta que las líneas H9 y U-937 clon 2 son no permisivas a la infección por virus defectivos en la proteína Vif (Gabuzda y cols., 1992).

En el caso de la persistencia en U-937 clon 2 no hemos encontrado mutaciones en los genes accesorios que hayamos podido asociar al fenotipo de persistencia.

5.5. Persistencia en la línea celular U-937 clon 2.

En la caracterización fenotípica de las células U937-2 persistentemente infectadas se observó que en el punto de persistencia establecida, el 1% de las células del cultivo expresaban proteínas víricas (determinado por IFI). Se observaba que el 38 % de las células expresaban el receptor CD4. A diferencia de lo observado en la persistencia en células H9 y MT4, se produce una activación en la replicación viral al activar con TNF- α y PMA. En este caso parece que la ausencia de efecto citopático es debido en gran medida a la existencia de algún defecto en la replicación del virus VU61. Este efecto estaría reflejado por la pequeña cantidad de virus infeccioso que existe en el sobrenadante ($5,6 \times 10^2$ DI₅₀/ml), la existencia de un bajo porcentaje de células infectadas (1% en IFI) a pesar de existir un gran número de células susceptibles que expresan el coreceptor CD4 (38%). Además, en experimentos de infección de células H9, MT4 y U937-2 se observa la incapacidad del virus VU61 para infectar a las tres líneas celulares. Sin embargo, la capacidad replicativa del virus se recupera al activar el cultivo con TNF- α y PMA.

Este resultado nos lleva a la hipótesis de que en el caso de la ausencia de efecto citopático a pesar de la existencia de virus en el cultivo en células U937-2 se debe más a un fenómeno de latencia que de persistencia propiamente dicha.

5.6. Otras mutaciones posiblemente relacionadas con el fenómeno de la persistencia.

Ciertas datos previamente descritos en la bibliografía así como algunas de las observaciones obtenidas de este trabajo sugieren que las mutaciones en los genes accesorios descritas anteriormente como implicadas en la persistencia no son suficientes para explicar este fenómeno.

En trabajos anteriores, al introducir por mutagénesis dirigida en un clon infeccioso deleciones en *vpr*, *vpu* y *vif* el fenotipo persistente no se observa de forma inmediata. En esos casos el cultivo sufre una crisis de viabilidad de la que sobreviven un pequeño número de células que acaban por recuperar el cultivo. Esta dinámica de obtención de persistencia parece coincidir con la necesidad de adquirir una o varias mutaciones adicionales que, junto con las mutaciones de los genes accesorios introducidas sean necesarias en el fenotipo persistente.

Por otro lado en el caso de la persistencia-latencia obtenida en la línea U937-2 no se ha observado ninguna mutación significativa en los genes accesorios.

En la secuenciación completa de los virus VH61, VM61 y VU61 se han descrito 11 mutaciones no sinónimas que se han acumulado de manera independiente en las tres líneas así como 2 mutaciones en la LTR. Entre estas mutaciones se encontrarían las que podrían ayudar a la persistencia. Las mutaciones que son candidatos más firmes a tener un papel en la persistencia son las dos mutaciones localizadas en la LTR dado que se encuentran situadas en sitios de unión de factores celulares que actuarían activando la replicación. Son sitios de unión de c-myb y de unión de USF (E-box) respectivamente. Las mutaciones en estos sitios de unión podrían inhibir la unión de dichos factores celulares manteniendo la replicación del virus en niveles bajos.

5.7. Proceso de establecimiento de persistencia reflejado en la cuasiespecie vírica.

En el proceso del establecimiento de la persistencia con un aislado s61, el cual era una cuasiespecie compleja, se han seleccionado aquellas variantes que eran menos citopáticas. Con el objeto de analizar cómo ha sido éste proceso, decidimos estudiar el espectro de variantes que componían la cuasiespecie vírica en cada uno de los provirus en dos puntos; un punto de persistencia temprana (en el pase 35) y uno de persistencia tardía (alrededor del pase

100). Para ello se clonó la región C2-V5 de los tres provirus VH61, VM61 y VU61. Del análisis de los clones pudimos deducir que en el punto de la persistencia temprana la cuasispecie no se encuentra del todo consolidada en la persistencia, mientras que en el punto de la persistencia tardía la cuasispecie se encuentra completamente adaptada a la persistencia. Esto se apoya en el hecho de que en el primer caso se observó una mayor frecuencia de mutación y un mayor número de clones no viables que cuando la persistencia se encontraba completamente establecida (ver apartado 4.2.8.2.6).

5.8. Modelos de persistencia.

Una vez analizados todos los resultados obtenidos en éste estudio, llegamos a la conclusión de que el proceso de la persistencia se había producido de manera diferente en las tres líneas celulares estudiadas, puesto que los provirus obtenidos de las células persistentemente infectadas habían sufrido diferentes alteraciones en su genoma, sobre todo en los genes accesorios. Con las observaciones descritas anteriormente en ésta discusión, y basándonos en las definiciones de persistencia postuladas por M. Stevenson (Stevenson, 1997), elaboramos una hipótesis con tres modelos de persistencia:

a) Persistencia.

Sería el caso de la línea celular MT-4 , puesto que no existen efectos citopáticos en la célula huésped, todas las células del cultivo se encuentran infectadas a pesar de mantener el cultivo una viabilidad por encima del 80 %, y no se activa la replicación del virus ni con TNF- α ni con PMA. Además, los virus que se liberan de éstas células se caracterizan por tener una proteína Vif defectiva.

b) Persistencia restringida.

Sería el caso de la línea celular H9. Sería igual al anterior salvo que existiría una restricción a la infección de todas las células en el cultivo, ya que sólo están infectadas el 30 %. Los virus que se liberan de éstas células se caracterizan por tener las proteínas Vpr y Vpu defectivas.

c) Infección latente proviral.

Sería el caso de la línea celular U-937 clon 2. Este caso sería diferente al de los otros dos, puesto que no presentaba ninguna alteración en los genes accesorios. En éste modelo tiene que ver más con la replicación del virus. El virus replica muy poco, y es estimulada ésta replicación con TNF- α y PMA, saliendo un virus totalmente viable.

6. CONCLUSIONES

- 1) Tras el análisis filio genético de aislados naturales cocultivo negativo en los genes *gag*, *env* y *nef*, se deduce que prácticamente todas las muestras se agrupan dentro del subtipo C en los tres genes estudiados, aunque se ha observado una segregación de las muestras procedentes de Brasil, que formaría un “sub-subtipo” dentro del C. Solamente un muestra (IN19), resultó ser un recombinante C / B /C.
- 2) En éste grupo de aislados naturales hemos comprobado que existe una correlación entre alteraciones en la proteína Nef y el fenotipo de cocultivo negativo. 9 de las 14 muestras que eran incapaces de replicar en cocultivo presentaban alguna alteración en la proteína, mientras que sólo 1 muestra de 6 cocultivo positivo presentaban alguna alteración.
- 3) En el estudio sobre las infecciones persistentes en tres líneas celulares diferentes (dos líneas T y una promonocítica) con el mismo aislado s61, hemos observado que se ha producido de manera diferente. De todas formas, el establecimiento de la persistencia en las dos líneas celulares T ha sido más parecida entre ellas que en la línea promonocítica. Para ello nos hemos basado en la comparación de las secuencias de los genomas completos de los tres provirus obtenidos de las células persistentemente infectadas con la secuencia del s61. Entre los provirus VH61 y VM61 se han observado 34 mutaciones comunes. De éstas, podríamos deducir que 10 mutaciones serían las específicas de la adaptación a la perisitencia en las líneas T.
- 4) En el estudio de las mutaciones comunes a los tres provirus, las cuales nos podrían indicar que podrían estar implicadas en el establecimiento de la persistencia, ya que se han seleccionado independientemente de la línea celular, hemos encontrado 28 mutaciones. De éstas 28 mutaciones comunes a los tres provirus, existen 3 que podrían tener una función más clara en el fenómeno de la persistencia. Dos se encontrarían en la LTR y disminuirían en parte la replicación de los virus, y la tercera, se encontraba en la región V3 del gen de la envuelta, y hace que los virus sean NSI, y por lo tanto, menos citopáticos.
- 5) Hemos establecido que, a lo largo del proceso del establecimiento de la persistencia, se han ido seleccionando aquellas variantes de la cuasiespecie inicial del aislado s61 que eran menos citopáticas. Por éste motivo, en la línea celular H9 se han seleccionado las

variantes que, aparte de las mutaciones vistas en el apartado anterior, presentaban las proteínas accesorias Vpu y Vpr no funcionales, y en la línea MT-4 se han seleccionado aquellas variantes que tenían la proteína accesoria Vif no funcional. Además, hemos constatado que la pérdida de la funcionalidad de Vif ha sido como consecuencia de una hipermutación. En la línea U-937 clon 2, no se han visto afectados los genes accesorios, sino que la persistencia es el resultado de la selección de aquellos variantes menos citopáticas y con una menor replicación.

6) En el estudio de la evolución de la cuasiespecie en el gen de la envuelta a un pase temprano de la persistencia y a un pase tardío, hemos determinado que a lo largo del proceso, se ha producido una homogenización de la cuasiespecie y un aumento del número de clones viables, indicando que se han ido seleccionando aquellas mutaciones que le permitían al virus coexistir con la célula sin matarla.

7) Con todos los datos vistos, hemos postulado la existencia de tres modelos de persistencia:

- a) Persistencia para el caso de las M61.
- b) Persistencia restringida para las H61.
- c) Latencia en el caso de las U61.

8) Los resultados obtenidos del estudio del establecimiento de la persistencia en tres líneas celulares diferentes con un mismo aislado, ponen de manifiesto la importancia que tienen los genes accesorios en las líneas celulares T en éste proceso.

7. BIBLIOGRAFÍA

- Aiken, C., Krause, L., Chen, Y. L. & Trono, D. (1996). Mutational analysis of HIV-1 Nef: identification of two mutants that are temperature-sensitive for CD4 downregulation. *Virology* 217, 293-300.
- Aiken, C. & Trono, D. (1995). Nef stimulates human immunodeficiency virus type 1 proviral DNA synthesis. *J Virol* 69, 5048-56.
- Alexander, L., Weiskopf, E., Greenough, T. C., Gaddis, N. C., Auerbach, M. R., Malim, M. H., O'Brien, S. J., Walker, B. D., Sullivan, J. L. & Desrosiers, R. C. (2000). Unusual polymorphisms in human immunodeficiency virus type 1 associated with nonprogressive infection. *J Virol* 74, 4361-76.
- Anonymous (1998). AIDS epidemic update. December 1998. <http://www.unaids.org>, Londres.
- Anonymous (2000). Aids epidemic update. December 2000, UNAIDS. <http://www.unaids.org> Londres.
- Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209-11.
- Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220, 868-871.
- Baur, A. S., Sass, G., Laffert, B., Willbold, D., Cheng-Mayer, C. & Peterlin, B. M. (1997). The N-terminus of Nef from HIV-1/SIV associates with a protein complex containing Lck and a serine kinase. *Immunity* 6, 283-91.
- Berger, E. A. (1997). HIV entry and tropism: the chemokine receptor connection. *Aids* 11, S3-16.
- Binley, J. M., Jin, X., Huang, Y., Zhang, L., Cao, Y., Ho, D. D. & Moore, J. P. (1998). Persistent antibody responses but declining cytotoxic T-lymphocyte responses to multiple human immunodeficiency virus type 1 antigens in a long-term nonprogressing individual with a defective p17 proviral sequence and no detectable viral RNA expression. *J Virol* 72, 3472-4.
- Bjorndal, A., Sonnerborg, A., Tscherning, C., Albert, J. & Fenyo, E. M. (1999). Phenotypic characteristics of human immunodeficiency virus type 1 subtype C isolates of Ethiopian AIDS patients. *AIDS Res Hum Retroviruses* 15, 647-53.

- Boom, R., Sol, C. J., Salimans, M. M., Jansen, C. L., Wertheim-van Dillen, P. M. & van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *J Clin Microbiol* 28, 495-503.
- Bresnahan, P. A., Yonemoto, W., Ferrell, S., Williams-Herman, D., Geleziunas, R. & Greene, W. C. (1998). A dileucine motif in HIV-1 Nef acts as an internalization signal for CD4 downregulation and binds the AP-1 clathrin adaptor. *Curr Biol* 8, 1235-8.
- Bukrinsky, M. I. & Haffar, O. K. (1999). HIV-1 nuclear import: in search of a leader. *Front Biosci* 4, D772-81.
- Bukrinsky, M. I., Stanwick, T. L., Dempsey, M. P. & Stevenson, M. (1991). Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science* 254, 423-7.
- Carrillo, A. & Ratner, L. (1996). Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: role of the V3 loop and C4 envelope determinants. *J Virol* 70, 1301-9.
- Casado, C., Urtasun, I., Martin-Walther, M. V., Garcia, S., Rodriguez, C., del Romero, J. & Lopez-Galindez, C. (2000). Genetic analysis of HIV-1 samples from Spain. *J Acquir Immune Defic Syndr* 23, 68-74.
- Castilla, J. & Gutierrez, M. A. (1997). Epidemia en España. Manual del SIDA. Edited by V. Soriano & Gonzalez-Lahoz. Madrid.
- Chaix, M. L., Chappey, C., Couillin, I., Rozenbaum, W., Levy, J. P. & Saragosti, S. (1993). Diversity of the V3 region of HIV in Paris, France. *Aids* 7, 1199-204.
- Chen, Z., Luckay, A., Sodora, D. L., Telfer, P., Reed, P., Gettie, A., Kanu, J. M., Sadek, R. F., Yee, J., Ho, D. D., Zhang, L. & Marx, P. A. (1997). Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J Virol* 71, 3953-60.
- Cheng-Mayer, C., Seto, D., Tatenos, M. & Levy, J. A. (1988). Biologic features of HIV-1 that correlate with virulence in the host. *Science* 240, 80-2.
- Chodosh, L. A., Carthew, R. W. & Sharp, P. A. (1986). A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol Cell Biol* 6, 4723-33.

- Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. A., Barditch-Crovo, P. & Siliciano, R. F. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* 387, 183-8.
- Clavel, F., Guétard, D., Brun-Vézinet, F., Chamaret, S., Rey, M. A., Santos-Ferreira, M. O., Laurent, A. G., Gauguet, C., Katlana, C., Rouzioux, C., Klatzman, D., Champalimaud, J. L. & Montagnier, L. (1986). Isolation of a new human retrovirus from West African patients with AIDS. *Science* 233, 343-346.
- Coffin, J. M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 267, 483-9.
- Cohen, G. B., Gandhi, R. T., Davis, D. M., Mandelboim, O., Chen, B. K., Strominger, J. L. & Baltimore, D. (1999). The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity* 10, 661-71.
- Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D. & Baltimore, D. (1998). HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* 391, 397-401.
- Coombs, R. W., Collier, A. C., Allain, J.-P., Nokora, B., Leuther, M., Gjerset, G. F. & Corey, L. (1989). Plasma viremia in human immunodeficiency virus infection. *New England Journal of Medicine*, 1626-1631.
- Cullen, B. R. (1998). HIV-1 auxiliary proteins: making connections in a dying cell. *Cell* 93, 685-92.
- Deacon, N. J., Tsykin, A., Solomon, A., Smith, K., Ludford-Menting, M., Hooker, D. J., McPhee, D. A., Greenway, A. L., Ellett, A., Chatfield, C. & y cols. (1995). Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* 270, 988-91.
- Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., Goedert, J. J., Buchbinder, S. P., Vittinghoff, E., Gomperts, E., Donfield, S., Vlahov, D., Kaslow, R., Saah, A., Rinaldo, C., Detels, R. & O'Brien, S. J. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study,

- Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 273, 1856-62.
- Dejuq, N., Simmons, G. & Clapham, P. R. (2000). T-cell line adaptation of human immunodeficiency virus type 1 strain SF162: effects on envelope, vpu and macrophage-tropism. *J Gen Virol* 81 Pt 12, 2899-904.
- Delassus, S., Cheynier, R. & Wain-Hobson, S. (1992). Nonhomogeneous distribution of human immunodeficiency virus type 1 proviruses in the spleen. *J Virol* 66, 5642-5.
- Delwart, E. L., Pan, H., Sheppard, H. W., Wolpert, D., Neumann, A. U., Korber, B. & Mullins, J. I. (1997). Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. *J Virol* 71, 7498-508.
- Delwart, E. L., Sheppard, H. W., Walker, B. D., Goudsmit, J. & Mullins, J. I. (1994). Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J Virol* 68, 6672-83.
- Desrosiers, R. C., Lifson, J. D., Gibbs, J. S., Czajak, S. C., Howe, A. Y., Arthur, L. O. & Johnson, R. P. (1998). Identification of highly attenuated mutants of simian immunodeficiency virus. *J Virol* 72, 1431-7.
- di Fagagna, F. D., Marzio, G., Gutierrez, M. I., Kang, L. Y., Falaschi, A. & Giacca, M. (1995). Molecular and functional interactions of transcription factor USF with the long terminal repeat of human immunodeficiency virus type 1. *J Virol* 69, 2765-75.
- Domingo, E., Baranowski, E., Ruiz-Jarabo, C. M., Martin-Hernandez, A. M., Saiz, J. C. & Escarmis, C. (1998). Quasispecies structure and persistence of RNA viruses. *Emerg Infect Dis* 4, 521-7.
- Domingo, E., Flavell, R. A. & Weissmann, C. (1976). In vitro site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Qbeta. *Gene* 1, 3-25.
- Domingo, E., Martinez-Salas, E., Sobrino, F., de la Torre, J. C., Portela, A., Ortin, J., Lopez-Galindez, C., Perez-Brena, P., Villanueva, N., Najera, R. & y cols. (1985). The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance--a review. *Gene* 40, 1-8.
- Domingo, E., Sabo, D., Taniguchi, T. & Weissmann, C. (1978). Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13, 735-44.

- Drake, J. W. (1993). Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci USA* 90, 4171-5.
- Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465-523.
- Emerman, M. (1996). HIV-1, Vpr and the cell cycle. *Curr Biol* 6, 1096-103.
- Engelstad, M. (1996). Genetic diversity of env V3 loop sequences in HIV type 1-infected homosexuals, heterosexuals, intravenous drug users, and hemophilia patients in norway: high prevalence of subtype B and detection of subtype C. *AIDS Res Hum Retroviruses* 12, 1733-8.
- Fauci, A. S. (1996). Host factors and the pathogenesis of HIV-induced disease. *Nature* 384, 529-34.
- Fenyo, E. M., Morfeldt-Manson, L., Chiodi, F., Lind, B., von Gegerfelt, A., Albert, J., Olausson, E. & Asjo, B. (1988). Distinct replicative and cytopathic characteristics of human immunodeficiency virus isolates. *J Virol* 62, 4414-9.
- Fields BN, K. D., Howley PM, y cols. (1990). *Fields Virology*.
- Finzi, D., Blankson, J., Siliciano, J. D., Margolick, J. B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T. C., Chaisson, R. E., Rosenberg, E., Walker, B., Gange, S., Gallant, J. & Siliciano, R. F. (1999). Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med* 5, 512-7.
- Finzi, D., Hermankova, M., Pierson, T., Carruth, L. M., Buck, C., Chaisson, R. E., Quinn, T. C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., Ho, D. D., Richman, D. D. & Siliciano, R. F. (1997). Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 278, 1295-300.
- Gabuzda, D. H., Lawrence, K., Langhoff, E., Terwilliger, E., Dorfman, T., Haseltine, W. A. & Sodroski, J. (1992). Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol* 66, 6489-95.
- Ganeshan, S., Dickover, R. E., Korber, B. T., Bryson, Y. J. & Wolinsky, S. M. (1997). Human immunodeficiency virus type 1 genetic evolution in children with different rates of development of disease. *J Virol* 71, 663-77.
- Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G. M., Sharp, P. M. & Hahn, B. H. (1996). Evolutionary relationships among HIV-1 group M CRF01_AG and CRF02_AG strains. *J Virol* 70, 1005-13.

- B. H. (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* 397, 436-41.
- Gao, F., Yue, L., Robertson, D. L., Hill, S. C., Hui, H., Biggar, R. J., Neequaye, A. E., Whelan, T. M., Ho, D. D., Shaw, G. M. & y cols. (1994). Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J Virol* 68, 7433-47.
- Gao, W. Y., Cara, A., Gallo, R. C. & Lori, F. (1993). Low levels of deoxynucleotides in peripheral blood lymphocytes: a strategy to inhibit human immunodeficiency virus type 1 replication. *Proc Natl Acad Sci U S A* 90, 8925-8.
- Gaynor, R. (1992). Cellular transcription factors involved in the regulation of HIV-1 gene expression. *Aids* 6, 347-63.
- Greenough, T. C., Brettler, D. B., Somasundaran, M., Panicali, D. L. & Sullivan, J. L. (1997). Human immunodeficiency virus type 1-specific cytotoxic T lymphocytes (CTL), virus load, and CD4 T cell loss: evidence supporting a protective role for CTL in vivo. *J Infect Dis* 176, 118-25.
- Gregor, P. D., Sawadogo, M. & Roeder, R. G. (1990). The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev* 4, 1730-40.
- Groenink, M., Fouchier, R. A., de Goede, R. E., de Wolf, F., Gruters, R. A., Cuypers, H. T., Huisman, H. G. & Tersmette, M. (1991). Phenotypic heterogeneity in a panel of infectious molecular human immunodeficiency virus type 1 clones derived from a single individual. *J Virol* 65, 1968-75.
- Grzesiek, S., Bax, A., Clore, G. M., Gronenborn, A. M., Hu, J. S., Kaufman, J., Palmer, I., Stahl, S. J. & Wingfield, P. T. (1996a). The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat Struct Biol* 3, 340-5.
- Grzesiek, S., Stahl, S. J., Wingfield, P. T. & Bax, A. (1996b). The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* 35, 10256-61.
- Gurtler, L. G., Hauser, P. H., Eberle, J., von Brunn, A., Knapp, S., Zekeng, L., Tsague, J. M. & Kaptue, L. (1994). A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. *J Virol* 68, 1581-5.

- Hahn, B. H., Gonda, M. A., Shaw, G. M., Popovic, M., Hoxie, J. A., Gallo, R. C. & Wong-Staal, F. (1985). Genomic diversity of the acquired immune deficiency syndrome virus HTLV- III: different viruses exhibit greatest divergence in their envelope genes. *Proc Natl Acad Sci U S A* 82, 4813-7.
- Hahn, B. H., Shaw, G. M., De Cock, K. M. & Sharp, P. M. (2000). AIDS as a zoonosis: scientific and public health implications. *Science* 287, 607-14.
- Harrer, T., Harrer, E., Kalams, S. A., Barbosa, P., Trocha, A., Johnson, R. P., Elbeik, T., Feinberg, M. B., Buchbinder, S. P. & Walker, B. D. (1996). Cytotoxic T lymphocytes in asymptomatic long-term nonprogressing HIV-1 infection. Breadth and specificity of the response and relation to in vivo viral quasispecies in a person with prolonged infection and low viral load. *J Immunol* 156, 2616-23.
- Harrowe, G. & Cheng-Mayer, C. (1995). Amino acid substitutions in the V3 loop are responsible for adaptation to growth in transformed T-cell lines of a primary human immunodeficiency virus type 1. *Virology* 210, 490-4.
- Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H. & Johnson, P. R. (1989). An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* 339, 389-92.
- Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. & Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123-6.
- Hoffman, T. L. & Doms, R. W. (1998). Chemokines and coreceptors in HIV/SIV-host interactions. *Aids* 12, S17-26.
- Hu, D. J., Dondero, T. J., Rayfield, M. A., George, J. R., Schochetman, G., Jaffe, H. W., Luo, C. C., Kalish, M. L., Weniger, B. G., Pau, C. P., Schable, C. A. & Curran, J. W. (1996). The emerging genetic diversity of HIV. The importance of global surveillance for diagnostics, research, and prevention. *Jama* 275, 210-6.
- Hu, W. S. & Temin, H. M. (1990). Retroviral recombination and reverse transcription. *Science* 250, 1227-33.
- Hua, J., Blair, W., Truant, R. & Cullen, B. R. (1997). Identification of regions in HIV-1 Nef required for efficient downregulation of cell surface CD4. *Virology* 231, 231-8.

- Ichimura, H., Kliks, S. C., Visrutaratna, S., Ou, C. Y., Kalish, M. L. & Levy, J. A. (1994). Biological, serological, and genetic characterization of HIV-1 subtype E isolates from northern Thailand. *AIDS Res Hum Retroviruses* 10, 263-9.
- Iversen, A. K., Shpaer, E. G., Rodrigo, A. G., Hirsch, M. S., Walker, B. D., Sheppard, H. W., Merigan, T. C. & Mullins, J. I. (1995). Persistence of attenuated rev genes in a human immunodeficiency virus type 1-infected asymptomatic individual. *J Virol* 69, 5743-53.
- Janssens, W., Buve, A. & Nkengasong, J. N. (1997). The puzzle of HIV-1 subtypes in Africa. *Aids* 11, 705-12.
- Kemp, S. D., Shi, C., Bloor, S., Harrigan, P. R., Mellors, J. W. & Larder, B. A. (1998). A novel polymorphism at codon 333 of human immunodeficiency virus type 1 reverse transcriptase can facilitate dual resistance to zidovudine and L-2',3'-dideoxy-3'-thiacytidine. *J Virol* 72, 5093-8.
- Kirchhoff, F., Greenough, T. C., Brettler, D. B., Sullivan, J. L. & Desrosiers, R. C. (1995). Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. *N Engl J Med* 332, 228-32.
- Kishi, M., Nishino, Y., Sumiya, M., Ohki, K., Kimura, T., Goto, T., Nakai, M., Kakinuma, M. & Ikuta, K. (1992). Cells surviving infection by human immunodeficiency virus type 1: vif or vpu mutants produce non-infectious or markedly less cytopathic viruses. *J Gen Virol* 73, 77-87.
- Kishi, M., Zheng, Y. H., Bahmani, M. K., Tokunaga, K., Takahashi, H., Kakinuma, M., Lai, P. K., Nonoyama, M., Luftig, R. B. & Ikuta, K. (1995). Naturally occurring accessory gene mutations lead to persistent human immunodeficiency virus type 1 infection of CD4-positive T cells. *J Virol* 69, 7507-18.
- Klasse, P. J., Boyd, M. T., Weiss, R. A. & Schulz, T. F. (1996). Mutations in the vpu, env, and nef genes of a syncytium-inducing variant of HIV type 1 JR-CSF that infects a range of T cell lines. *AIDS Res Hum Retroviruses* 12, 347-50.
- Korber, B., Hahn, B. H., Foley, B., Mellors, J., Leitner, T., Myers, G., McCutchan, F. E. & Kuiken, C. (1997). Human retroviruses and AIDS 1997: A compilation and analysis of nucleic acid and amino acid sequences. Edited by L. A. N. L. Theoretical Biology and Biophysics Groups. Los Alamos, Nuevo Mexico.

- Korber, B. T., MacInnes, K., Smith, R. F. & Myers, G. (1994). Mutational trends in V3 loop protein sequences observed in different genetic lineages of human immunodeficiency virus type 1. *J Virol* 68, 6730-44.
- Koup, R. A., Safrit, J. T., Cao, Y., Andrews, C. A., McLeod, G., Borkowsky, W., Farthing, C. & Ho, D. D. (1994). Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol* 68, 4650-5.
- Kuiken, C., Foley, B., Hahn, B. H., Marx, P. A., McCutchan, F. E., Mellors, J., Mullins, J. I., Wolinsky, S. M. & Korber, B. (1999). Human retroviruses and AIDS 1999: A compilation and analysis of nucleic acid and amino acid sequences. Edited by L. A. N. L. Theoretical Biology and Biophysics Groups. Los Alamos, Nuevo Mexico.
- Kumar, S., Tamura, K. & Nei, M. (1993). MEGA: Molecular Evolutionary Genetics Analysis, 1.01 edn. University Park, PA 16802: The Pennsylvania State University.
- Lama, J., Mangasarian, A. & Trono, D. (1999). Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. *Curr Biol* 9, 622-31.
- Learmont, J. C., Geczy, A. F., Mills, J., Ashton, L. J., Raynes-Greenow, C. H., Garsia, R. J., Dyer, W. B., McIntyre, L., Oelrichs, R. B., Rhodes, D. I., Deacon, N. J. & Sullivan, J. S. (1999). Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. A report from the Sydney Blood Bank Cohort. *N Engl J Med* 340, 1715-22.
- Levy, J. A. (1998). HIV and the pathogenesis of AIDS. In *Overall features of HIV pathogenesis: prognosis for long-term survival*, Second edn, pp. 311-338. Edited by A. S. o. AIDS. Washington.
- Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. & Oshiro, L. S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 225, 840-842.
- Loussert-Ajaka, I., Chaix, M. L., Korber, B., Letourneur, F., Gomas, E., Allen, E., Ly, T. D., Brun-Vezinet, F., Simon, F. & Saragosti, S. (1995). Variability of human immunodeficiency virus type 1 group O strains isolated from Cameroonian patients living in France. *J Virol* 69, 5640-9.

- Luciw, P. (1996). Human immunodeficiency viruses and their replication. In *Fields Virology*, pp. 1881 - 1952. Edited by K. D. Fields BN, Howley PM, y cols. Philadelphia: Lippincott - Raven Publishers.
- Lukashov, V. V. & Goudsmit, J. (1997). Evolution of the human immunodeficiency virus type 1 subtype-specific V3 domain is confined to a sequence space with a fixed distance to the subtype consensus. *J Virol* 71, 6332-8.
- Lukashov, V. V. & Goudsmit, J. (1998). HIV heterogeneity and disease progression in AIDS: a model of continuous virus adaptation. *Aids* 12, S43-52.
- Meyerhans, A., Cheynier, R., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B. & Wain-Hobson, S. (1989). Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* 58, 901-10.
- Miller, M. D., Warmerdam, M. T., Gaston, I., Greene, W. C. & Feinberg, M. B. (1994). The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages. *J Exp Med* 179, 101-13.
- Miyoshi, I., Kubonishi, I., Yoshimoto, S., Akagi, T., Ohtsuki, Y., Shiraishi, Y., Nagata, K. & Hinuma, Y. (1981). Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature* 294, 770-1.
- Mizutani, S., Boettiger, D. & Temin, H. M. (1970). A DNA-depenent DNA polymerase and a DNA endonuclease in virions of Rous sarcoma virus. *Nature* 228, 424-7.
- Montano, M. A., Nixon, C. P., Ndung'u, T., Bussmann, H., Novitsky, V. A., Dickman, D. & Essex, M. (2000). Elevated tumor necrosis factor-alpha activation of human immunodeficiency virus type 1 subtype C in Southern Africa is associated with an NF-kappaB enhancer gain-of-function. *J Infect Dis* 181, 76-81.
- Moses, A. V., Ibanez, C., Gaynor, R., Ghazal, P. & Nelson, J. A. (1994). Differential role of long terminal repeat control elements for the regulation of basal and Tat-mediated transcription of the human immunodeficiency virus in stimulated and unstimulated primary human macrophages. *J Virol* 68, 298-307.
- Murphy, F., Fauquet, C. M. & Bishop, D. H. L. (1995). Virus taxonomy-the classification and nomenclature of viruses: sixth report of the International Committee of Taxonomy of viruses., Viena: Springer-Verlag.

- Mustafa, F. & Robinson, H. L. (1993). Context-dependent role of human immunodeficiency virus type 1 auxiliary genes in the establishment of chronic virus producers. *J Virol* 67, 6909-15.
- Myers, G., Korber, B., Foley, B., Jeang, K., Mellors, J. & Wain-Hobson, S. (1996). Human retroviruses and AIDS 1996: A compilation and analysis of nucleic acid and amino acid sequences. Edited by L. A. N. L. Theoretical Biology and Biophysics Groups. Los Alamos, Nuevo Mexico.
- Naghavi, M. H., Estable, M. C., Schwartz, S., Roeder, R. G. & Vahlne, A. (2001). Upstream stimulating factor affects human immunodeficiency virus type 1 (HIV-1) long terminal repeat-directed transcription in a cell-specific manner, independently of the HIV-1 subtype and the core-negative regulatory element. *J Gen Virol* 82, 547-59.
- Naghavi, M. H., Schwartz, S., Sonnerborg, A. & Vahlne, A. (1999). Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. *AIDS Res Hum Retroviruses* 15, 1293-303.
- Nakaya, T., Fujinaga, K., Doi, H., Suzuki, S., Takahashi, H., Nishino, Y., Kishi, M., Azuma, I., Luftig, R. B. & Ikuta, K. (1996). Serial passage of human immunodeficiency virus type 1 generates misalignment deletions in non-essential accessory genes. *Virus Res* 46, 139-47.
- Nei, M. & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 418-26.
- Niederman, T. M., Hastings, W. R. & Ratner, L. (1993). Myristoylation-enhanced binding of the HIV-1 Nef protein to T cell skeletal matrix. *Virology* 197, 420-5.
- Nishino, Y., Kishi, M., Sumiya, M., Ogawa, K., Adachi, A., Maotani-Imai, K., Kato, S., Hirai, K. & Ikuta, K. (1991). Human immunodeficiency virus type 1 vif, vpr, and vpu mutants can produce persistently infected cells. *Arch Virol* 120, 181-92.
- Nishino, Y., Nakaya, T., Fujinaga, K., Kishi, M., Azuma, I. & Ikuta, K. (1994). Persistent infection of MT-4 cells by human immunodeficiency virus type 1 becomes increasingly likely with in vitro serial passage of wild-type but not nef mutant virus. *J Gen Virol* 75, 2241-51.
- Oelrichs, R., Tsykin, A., Rhodes, D., Solomon, A., Ellett, A., McPhee, D. & Deacon, N. (1998). Genomic sequence of HIV type 1 from four members of the Sydney Blood

- Bank Cohort of long-term nonprogressors. *AIDS Res Hum Retroviruses* 14, 811-4.
- Olivares, I., Shaw, G. & Lopez-Galindez, C. (1997). Phenotypic switch in a Spanish HIV type 1 isolate on serial passage on MT-4 cells. *AIDS Res Hum Retroviruses* 13, 979-84.
- Ou, C. Y., Takebe, Y., Weniger, B. G., Luo, C. C., Kalish, M. L., Auwanit, W., Yamazaki, S., Gayle, H. D., Young, N. L. & Schochetman, G. (1993). Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet* 341, 1171-4.
- Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12, 357-8.
- Pathak, V. K. & Temin, H. M. (1990). Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. *Proc Natl Acad Sci USA* 87, 6024-8.
- Paul, M., Mazumder, S., Raja, N. & Jabbar, M. A. (1998). Mutational analysis of the human immunodeficiency virus type 1 Vpu transmembrane domain that promotes the enhanced release of virus-like particles from the plasma membrane of mammalian cells. *J Virol* 72, 1270-9.
- Pelemans, H., Esnouf, R., Dunkler, A., Parniak, M. A., Vandamme, A. M., Karlsson, A., De Clercq, E., Kleim, J. P. & Balzarini, J. (1997). Characteristics of the Pro225His mutation in human immunodeficiency virus type 1 (HIV-1) reverse transcriptase that appears under selective pressure of dose-escalating quinoxaline treatment of HIV-1. *J Virol* 71, 8195-203.
- Pelemans, H., Esnouf, R. M., Parniak, M. A., Vandamme, A. M., De Clercq, E. & Balzarini, J. (1998). A proline-to-histidine substitution at position 225 of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) sensitizes HIV-1 RT to BHAP U-90152. *J Gen Virol* 79, 1347-52.
- Pereira, L. A., Bentley, K., Peeters, A., Churchill, M. J. & Deacon, N. J. (2000). A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res* 28, 663-8.

- Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271, 1582-6.
- Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J. & Wigler, M. (1981). Human-tumor-derived cell lines contain common and different transforming genes. *Cell* 27, 467-76.
- Piatak, M., Jr., Saag, M. S., Yang, L. C., Clark, S. J., Kappes, J. C., Luk, K. C., Hahn, B. H., Shaw, G. M. & Lifson, J. D. (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science* 259, 1749-54.
- Poon, B., Grovit-Ferbas, K., Stewart, S. A. & Chen, I. S. (1998). Cell cycle arrest by Vpr in HIV-1 virions and insensitivity to antiretroviral agents. *Science* 281, 266-9.
- Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. (1984). Detection, isolation and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 224, 497-500.
- Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988). Fidelity of HIV-1 reverse transcriptase. *Science* 242, 1168-71.
- Quinones-Mateu, M. E., Holguin, A., Dopazo, J., Najera, I. & Domingo, E. (1996). Point mutant frequencies in the pol gene of human immunodeficiency virus type 1 are two- to threefold lower than those of env. *AIDS Res Hum Retroviruses* 12, 1117-28.
- Rhodes, D. I., Ashton, L., Solomon, A., Carr, A., Cooper, D., Kaldor, J. & Deacon, N. (2000). Characterization of three nef-defective human immunodeficiency virus type 1 strains associated with long-term nonprogression. Australian Long-Term Nonprogressor Study Group. *J Virol* 74, 10581-8.
- Robertson, D. L., Sharp, P. M., McCutchan, F. E. & Hahn, B. H. (1995). Recombination in HIV-1. *Nature* 374, 124-6.
- Rosen, C. A., Sodroski, J. G. & Haseltine, W. A. (1985). The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell* 41, 813-23.
- Rosenberg, E. S., Billingsley, J. M., Caliendo, A. M., Boswell, S. L., Sax, P. E., Kalams, S. A. & Walker, B. D. (1997). Vigorous HIV-1-specific CD4⁺ T cell responses associated with control of viremia. *Science* 278, 1447-50.

- Ross, T. M., Oran, A. E. & Cullen, B. R. (1999). Inhibition of HIV-1 progeny virion release by cell-surface CD4 is relieved by expression of the viral Nef protein. *Curr Biol* 9, 613-21.
- Sakai, H., Shibata, R., Sakuragi, J., Sakuragi, S., Kawamura, M. & Adachi, A. (1993). Cell-dependent requirement of human immunodeficiency virus type 1 Vif protein for maturation of virus particles. *J Virol* 67, 1663-6.
- Saksela, K., Cheng, G. & Baltimore, D. (1995). Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. *Embo J* 14, 484-91.
- Salvi, R., Garbuglia, A. R., Di Caro, A., Pulciani, S., Montella, F. & Benedetto, A. (1998). Grossly defective nef gene sequences in a human immunodeficiency virus type 1-seropositive long-term nonprogressor. *J Virol* 72, 3646-57.
- Sanchez-Palomino, S., Rojas, J. M., Martinez, M. A., Fenyo, E. M., Najera, R., Domingo, E. & Lopez-Galindez, C. (1993). Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. *J Virol* 67, 2938-43.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-7.
- Sawadogo, M. & Roeder, R. G. (1985). Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proc Natl Acad Sci U S A* 82, 4394-8.
- Sawai, E. T., Baur, A. S., Peterlin, B. M., Levy, J. A. & Cheng-Mayer, C. (1995). A conserved domain and membrane targeting of Nef from HIV and SIV are required for association with a cellular serine kinase activity. *J Biol Chem* 270, 15307-14.
- Schubert, U., Bour, S., Ferrer-Montiel, A. V., Montal, M., Maldarell, F. & Strebel, K. (1996). The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J Virol* 70, 809-19.
- Schwartz, O., Marechal, V., Danos, O. & Heard, J. M. (1995). Human immunodeficiency virus type 1 Nef increases the efficiency of reverse transcription in the infected cell. *J Virol* 69, 4053-9.

- Schwartz, O., Marechal, V., Le Gall, S., Lemonnier, F. & Heard, J. M. (1996). Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat Med* 2, 338-42.
- Schwartz, S., Felber, B. K., Fenyo, E. M. & Pavlakis, G. N. (1989). Rapidly and slowly replicating human immunodeficiency virus type 1 isolates can be distinguished according to target-cell tropism in T- cell and monocyte cell lines. *Proc Natl Acad Sci USA* 86, 7200-3.
- Schwartz, S., Felber, B. K. & Pavlakis, G. N. (1991). Expression of human immunodeficiency virus type 1 vif and vpr mRNAs is Rev-dependent and regulated by splicing. *Virology* 183, 677-86.
- Shiino, T., Kato, K., Kodaka, N., Miyakuni, T., Takebe, Y. & Sato, H. (2000). A group of V3 sequences from human immunodeficiency virus type 1 subtype E non-syncytium-inducing, CCR5-using variants are resistant to positive selection pressure. *J Virol* 74, 1069-78.
- Shioda, T., Levy, J. A. & Cheng-Mayer, C. (1991). Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. *Nature* 349, 167-9.
- Shugars, D. C., Smith, M. S., Glueck, D. H., Nantermet, P. V., Seillier-Moiseiwitsch, F. & Swanstrom, R. (1993). Analysis of human immunodeficiency virus type 1 nef gene sequences present in vivo. *J Virol* 67, 4639-50.
- Sieweke, M. H., Tekotte, H., Jarosch, U. & Graf, T. (1998). Cooperative interaction of ets-1 with USF-1 required for HIV-1 enhancer activity in T cells. *Embo J* 17, 1728-39.
- Simon, F., Mauclore, P., Roques, P., Loussert-Ajaka, I., Muller-Trutwin, M. C., Saragosti, S., Georges-Courbot, M. C., Barre-Sinoussi, F. & Brun-Vezinet, F. (1998). Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 4, 1032-7.
- Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., Goedert, J. J., O'Brien, T. R., Jacobson, L. P., Kaslow, R., Buchbinder, S., Vittinghoff, E., Vlahov, D., Hoots, K., Hilgartner, M. W. & O'Brien, S. J. (1997). Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS),

- Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* 277, 959-65.
- Sova, P. & Volsky, D. J. (1993). Efficiency of viral DNA synthesis during infection of permissive and nonpermissive cells with vif-negative human immunodeficiency virus type 1. *J Virol* 67, 6322-6.
- Spina, C. A., Kwoh, T. J., Chowes, M. Y., Guatelli, J. C. & Richman, D. D. (1994). The importance of nef in the induction of human immunodeficiency virus type 1 replication from primary quiescent CD4 lymphocytes. *J Exp Med* 179, 115-23.
- Steinhauer, D. A. & Holland, J. J. (1987). Rapid evolution of RNA viruses. *Annu Rev Microbiol* 41, 409-33.
- Stevenson, M. (1997). Molecular mechanisms for the regulation of HIV replication, persistence and latency. *Aids* 11, S25-33.
- Stevenson, M., Stanwick, T. L., Dempsey, M. P. & Lamonica, C. A. (1990). HIV-1 replication is controlled at the level of T cell activation and proviral integration. *Embo J* 9, 1551-60.
- Sundstrom, C. & Nilsson, K. (1976). Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int J Cancer* 17, 565-77.
- Swingler, S., Mann, A., Jacque, J., Brichacek, B., Sasseville, V. G., Williams, K., Lackner, A. A., Janoff, E. N., Wang, R., Fisher, D. & Stevenson, M. (1999). HIV-1 Nef mediates lymphocyte chemotaxis and activation by infected macrophages. *Nat Med* 5, 997-103.
- Temin, H. M. (1989). Is HIV unique or merely different? *J Acquir Immune Defic Syndr* 2, 1-9.
- Tersmette, M., de Goede, R. E., Al, B. J., Winkel, I. N., Gruters, R. A., Cuypers, H. T., Huisman, H. G. & Miedema, F. (1988). Differential syncytium-inducing capacity of human immunodeficiency virus isolates: frequent detection of syncytium-inducing isolates in patients with acquired immunodeficiency syndrome (AIDS) and AIDS-related complex. *J Virol* 62, 2026-32.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-80.

- Triques, K., Bourgeois, A., Vidal, N., Mpoudi-Ngole, E., Mulanga-Kabeya, C., Nzilambi, N., Torimiro, N., Saman, E., Delaporte, E. & Peeters, M. (2000). Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. *AIDS Res Hum Retroviruses* 16, 139-51.
- Trono, D. (1998). When accessories turn out to be essential. *Nat Med* 4, 1368-9.
- Turner, B. G. & Summers, M. F. (1999). Structural biology of HIV. *J Mol Biol* 285, 1-32.
- Van Harmelen, J. H., Van der Ryst, E., Loubser, A. S., York, D., Madurai, S., Lyons, S., Wood, R. & Williamson, C. (1999). A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. *AIDS Res Hum Retroviruses* 15, 395-8.
- Vanden Haesevelde, M., Decourt, J. L., De Leys, R. J., Vanderborght, B., van der Groen, G., van Heuverswijn, H. & Saman, E. (1994). Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. *J Virol* 68, 1586-96.
- Vartanian, J. P., Meyerhans, A., Asjo, B. & Wain-Hobson, S. (1991). Selection, recombination, and G----A hypermutation of human immunodeficiency virus type 1 genomes. *J Virol* 65, 1779-88.
- Wain-Hobson, S. (1996). Running the gamut of retroviral variation. *Trends Microbiol* 4, 135-41.
- Wang, B., Ge, Y. C., Palasanthiran, P., Xiang, S. H., Ziegler, J., Dwyer, D. E., Randle, C., Dowton, D., Cunningham, A. & Saksena, N. K. (1996). Gene defects clustered at the C-terminus of the vpr gene of HIV-1 in long-term nonprogressing mother and child pair: in vivo evolution of vpr quasispecies in blood and plasma. *Virology* 223, 224-32.
- Wei, X., Ghosh, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifso, J. D., Bonhoeffer, S., Nowak, M. A., Hahn, B. H., Saag, M. S. & Shaw, G. M. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117-122.
- Willey, R. L., Maldarelli, F., Martin, M. A. & Strebel, K. (1992a). Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J Virol* 66, 7193-200.

- Willey, R. L., Maldarelli, F., Martin, M. A. & Strebel, K. (1992b). Human immunodeficiency virus type 1 Vpu protein regulates the formation of intracellular gp160-CD4 complexes. *J Virol* 66, 226-34.
- Wolinsky, S. M., Korber, B. T., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D. & Safrit, J. T. (1996). Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* 272, 537-42.
- Wong, J. K., Hezareh, M., Gunthard, H. F., Havlir, D. V., Ignacio, C. C., Spina, C. A. & Richman, D. D. (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* 278, 1291-5.
- Wyatt, R. & Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280, 1884-1888.
- Yamada, T. & Iwamoto, A. (2000). Comparison of proviral accessory genes between long-term nonprogressors and progressors of human immunodeficiency virus type 1 infection. *Arch Virol* 145, 1021-7.
- Yamaguchi, J., Devare, S. G. & Brennan, C. A. (2000). Identification of a new HIV-2 subtype based on phylogenetic analysis of full-length genomic sequence. *AIDS Res Hum Retroviruses* 16, 925-30.
- Yang, X., Goncalves, J. & Gabuzda, D. (1996). Phosphorylation of Vif and its role in HIV-1 replication. *J Biol Chem* 271, 10121-9.
- Yuste, E., Lopez-Galindez, C. & Domingo, E. (2000). Unusual distribution of mutations associated with serial bottleneck passages of human immunodeficiency virus type 1. *J Virol* 74, 9546-52.

8. ANEXO

En el anexo I se recoge una representación gráfica de los iniciadores empleados en la amplificación y secuenciación de los genomas completos de los 4 provirus secuenciados. Las flechas (→) representan los iniciadores utilizados para la hebra positiva de ADN,, y las flechas (←) indican los iniciadores utilizados para la hebra negativa de ADN.

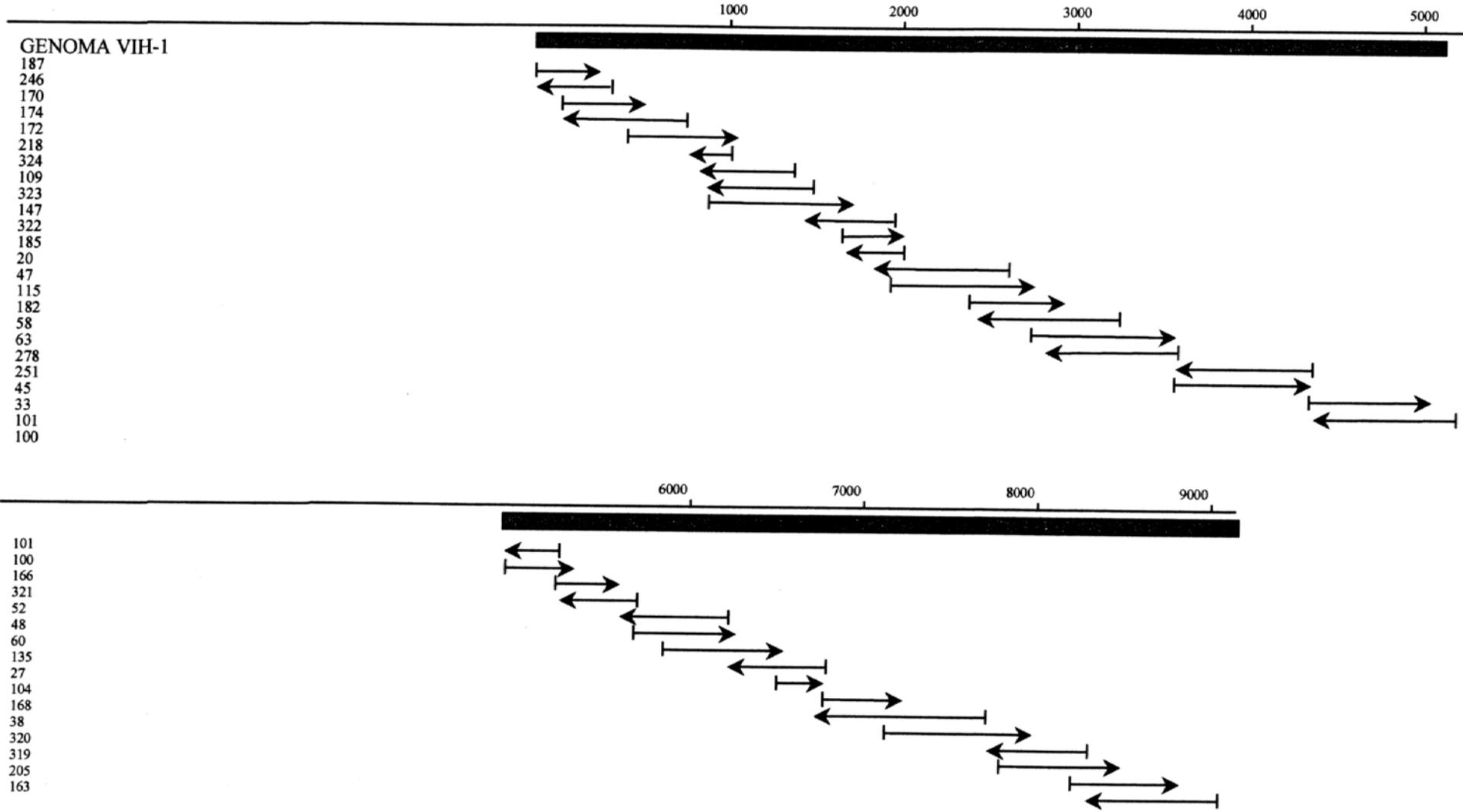
En el anexo II se muestran los alineamientos de nucleótidos de los genomas separados por proteínas. También se muestran las regiones no codificantes como la LTR y la secuencia “leader”.

En cada secuencia solo se indican aquellos aminoácidos distintos a la secuencia consenso o mayoritaria. El símbolo (.) indica identidad con la secuencia consenso. Las barras de colores encima de los alineamientos indican los sitios donde existen diferencias con la consenso. Las deleciones se indican con (-) o con una X.

Anexo I.

Estrategia de amplificación y secuenciación de los genomas completos

Estrategia de amplificación y secuenciación de los genomas completos



Anexo II
PROTEÍNA GAG

GAG.MSF.msf MSF: 503 Type: P jueves 3 de mayo de 2001 Check: 7792 ..

Name: consenso61.SEQ Len: 503 Check: 3816 Weight: 1.00

Name: consensoh61.SEQ Len: 503 Check: 5105 Weight: 1.00

Name: consensom61.SEQ Len: 503 Check: 4361 Weight: 1.00

Name: consensou61.SEQ Len: 503 Check: 4510 Weight: 1.00

//

```
1 60
consenso61.SEQ  MGARASILSGGGLDRWEKIRLRPGGKKKYRLKHLVWASRELERFAVNPGLLETSEGCRQI
consensoh61.SEQ  MGARASILSGGGLDRWEKIRLRPGGKKKYRLKHLVWASRELERFAVNPGLLETSEGCRQI
consensom61.SEQ  MGARASILSGGGLDRWEKIRLRPGGKKKYRLKHLVWASRELERFAVNPGLLETSEGCRQI
consensou61.SEQ  MGARASILSGGGLDRWEKIRLRPGGKKKYRLKHLVWASRELERFAVNPGLLETSEGCRQI
```

```
61 120
consenso61.SEQ  LGQLQPALQTGSEELRSLYNTVATLYCVHQKIDVKDTKEALDKIEEEQNKSKKKAQQAAA
consensoh61.SEQ  LGQLQPSLQTGSEELRSLYNTVATLYCVHQKIDVKDTKEALDKIEEEQNKSKKKAQQAAA
consensom61.SEQ  LGQLQPALQTGSEELRSLYNTVATLYCVHQKIDVKNTKEALDKIEEEQNKSKKKAQQAAA
consensou61.SEQ  LGQLQPALQTGSEELRSLYNTVATLYCVHQKIDVKDTKEALDKIEEEQNKSKKKAQQAAA
```

```
121 180
consenso61.SEQ  GTGNSSQVSQNYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGAT
consensoh61.SEQ  DTGNSSQVSQNYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGAT
consensom61.SEQ  DTGNSSQVSQNYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGAT
consensou61.SEQ  DTGNSSQVSQNYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKALSPEVIPMFSALSEGAT
```

```
181 240
consenso61.SEQ  PQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDR LHPVHAGPHPPGQMREPRGSDIAGTT
consensoh61.SEQ  PQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDR LHPVHAGPHPPGQMREPRGSDIAGTT
consensom61.SEQ  PQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDR LHPVHAGPHPPGQMREPRGSDIAGTT
consensou61.SEQ  PQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDR LHPVHAGPHPPGQMREPRGSDIAGTT
```

```
241 300
consenso61.SEQ  STLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSP TSILDIKQGPKEPFRDYVDRF
consensoh61.SEQ  STLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSP TSILDIKQGPKEPFRDYVDRF
consensom61.SEQ  STLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSP TSILDIKQGPKEPFRDYVDRF
consensou61.SEQ  STLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSP TSILDIKQGPKEPFRDYVDRF
```

```
301 360
consenso61.SEQ  YKTLRAEQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
```

consensoh61.SEQ YKTLRAEQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
consensom61.SEQ YKTLRAEQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
consensou61.SEQ YKTLRAEQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKA

361 420
consenso61.SEQ RVLAEAMSQVTNSATIMMQKGNFRNQRKIVKCFNCGKEGHIAKNCRAPRKRGCWKCCKEG
consensoh61.SEQ RVLAEAMSQVTNSATIMMQKGNFRNQRKIVKCFNCGKEGHIAKNCRAPRKRGCWKCCKEG
consensom61.SEQ RVLAEAMSQVTSSATIMMQKGNFRNQRKIVKCFNCGKEGHIAKNCRAPRKRGCWKCCKEG
consensou61.SEQ RVLAEAMSQVTSSATIMMQKGNFRNQRKIVKCFNCGKEGHIAKNCRAPRKRGCWKCCKEG

421 480
consenso61.SEQ HQMKDCTERQANFLGKIWPSHRGRPGNFPQNRLEPTAPPAPPEESFRFGREETTTTPSQKQE
consensoh61.SEQ HQMKDCTERQANFLGKIWPSHRGRPGNFPQNRLEPPAXXPPEESFRFGREETTTTPSQKQE
consensom61.SEQ HQMKDCTERQANFLGKIWPSHRGRPGNFPQNRLEPTAXXPPEESFRFGREETTTTPSQKQE
consensou61.SEQ HQMKDCTERQANFLGKIWPSHRGRPGNFPQNRLEPTAXXPPEESFRFGREETTTTPSQKQE

481 503
consenso61.SEQ LIDKELYPSASLKSFLGNDPLSQ
consensoh61.SEQ QIDKELYPSASLKSFLGNDPLSQ
consensom61.SEQ QIDKELYPLASLKSFLGNDPLSQ
consensou61.SEQ LIDKELYPSASLKSFLGNDPLSQ

PROTEÍNA POL

Pol.MSF.msf MSF: 1006 Type: P jueves 3 de mayo de 2001 Check: 8314 ..

Name: consenso61.SEQ Len: 1006 Check: 3981 Weight: 1.00

Name: consensoh61.SEQ Len: 1006 Check: 4388 Weight: 1.00

Name: consensom61.SEQ Len: 1006 Check: 5147 Weight: 1.00

Name: consensou61.SEQ Len: 1006 Check: 4798 Weight: 1.00

//

1 60
consenso61.SEQ FFRENLAFPQGKAREFSSEQTRANSPTSPTRELQVWGRDNNSLSEAGADRQGTVSFSFP
consensoh61.SEQ FFRENLAFPQGKAREFSSEQTRATSPTXXXXRELQVWGRDNNSLSEAGADRQGTVSFSFP
consensom61.SEQ FFRENLAFPQGKAREFSSEQTRANSPTSPTXXXXRELQVWGRDNNSLSEAGADRQGTVSFSFP
consensou61.SEQ FFRENLAFPQGKAREFSSEQTRANSPTSPTXXXXRELQVWGRDNNSLSEAGVDRQGTVSFSFP

61 120
consenso61.SEQ QITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEINLPGRWPKMIGGIGGFIKVRQYDQ
consensoh61.SEQ QITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWPKMIGGIGGFIKVRQYDQ
consensom61.SEQ QITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEINLPGRWPKMIGGIGGFIKVRQYDQ
consensou61.SEQ QITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEINLPGRWPKMIGGIGGFIKVRQYDQ

121 180
consenso61.SEQ ILVEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKKLPGMDGPKVK
consensoh61.SEQ ILVEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKKLPGMDGPKVK
consensom61.SEQ ILVEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKKLPGMDGPKVK
consensou61.SEQ ILVEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKKLPGMDGPKVK

181 240
consenso61.SEQ QWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNK
consensoh61.SEQ QWPLTEEKIKALVEICAEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNK
consensom61.SEQ QWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNK
consensou61.SEQ QWPLTEEKIKALVEICAEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNK

241 300
consenso61.SEQ KTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKEFRKYTAFTTIPSINNETPGI
consensoh61.SEQ KTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKEFRKYTAFTTIPSINNETPGI
consensom61.SEQ KTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKEFRKYTAFTTIPSINNETPGI
consensou61.SEQ KTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKEFRKYTAFTTIPSINNETPGI

301 360
consenso61.SEQ RYQYNVLPQGKWSGPAIFQSSMTKILEPFREQNPDIYIYQYMDLIVGSDLEIGQHRTKI
consensoh61.SEQ RYQYNVLPQGKWSGPAIFQSSMTKILEPFREQNPDIYIYQYMDLIVGSDLEIGQHRTKI
consensom61.SEQ RYQYNVLPQGKWSGPAIFQSSMTKILEPFREQNPDIYIYQYMDLIVGSDLEIGQHRTKI
consensou61.SEQ RYQYNVLPQGKWSGPAIFQSSMTKILEPFREQNPDIYIYQYMDLIVGSDLEIGQHRTKI

361 420
consenso61.SEQ EELRQHLLKKGFTTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVG
consensoh61.SEQ EELRQHLLKKGFTTTPDKKHQKEHPFLWMGYELHPDKWTVQLIVLPEKDSWTVNDIQKLVG
consensom61.SEQ EELRQHLLKKGFTTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVG
consensou61.SEQ EELRQLLLKKGFTTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVG

421 480
consenso61.SEQ KLNWASQIYPGIKVKQLCRLLRGTALTEVIPLTKEAELELAENREILKEPVHGVYYDPS
consensoh61.SEQ KLNWASQIYPGIKVKQLCRLLRGTALTEVIPLTKEAELELAENREILKEPVHGVYYDPS
consensom61.SEQ KLNWASQIYPGIKVKQLCRLLRGTALTEVIPLTKEAELELAENREILKEPVHGVYYDPS
consensou61.SEQ KLNWASQIYPGIKVKQLCRLLRGTALTEVIPLTKEAELELAENREILKEPVHGVYYDPS

481 540
consenso61.SEQ KDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVI
consensoh61.SEQ KELIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVI
consensom61.SEQ KELIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVI
consensou61.SEQ KDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVI

541 600
consenso61.SEQ WGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPLVVKLWYQLEQEPPIGAETFYV
consensoh61.SEQ WGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPLVVKLWYQLEQEPPIGAETFYV
consensom61.SEQ WGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPLVVKLWYQLEQEPPIGAETFYV
consensou61.SEQ WGKTPKFKLPIQKDTWETWWTEYWQATWIPEWEFVNTPLVVKLWYQLEQEPPIGAETFYV

601 660
consenso61.SEQ DGAANRETKLGKAGYVTNKGRQKVVSITD'TTNQKTELQAILLALQDSGLEVNIVTDSQYA
consensoh61.SEQ DGAANRETKLGKAGYVTNKGRQKVVSITD'TTNQKTELQAILLALQDSGLEVNIVTDSQYA
consensom61.SEQ DGAANRETKLGKAGYVTNKGRQKVVSITD'TTNQKTELQAILLALQDSGLEVNIVTDSQYA
consensou61.SEQ DGAANRETKLGKAGYVTNKGRQKVVSITDTPNQKTELQAILLALQDSGLEVNIVTDSQYA

661 720
consenso61.SEQ LGIIQAQPKSESEIVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSTGIRKVLFL
consensoh61.SEQ LGIIQAQPKSESEIVSQIIEQLIKKEKVYLTWVPAHKGIGGNEQVDKLVSTGIRKVLFL
consensom61.SEQ LGIIQAQPKSESEIVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSTGIRKVLFL
consensou61.SEQ LGIIQAQPKSESEIVSQIIEQLIKKEKVYLTWVPAHKGIGGNEQVDKLVSTGIRKVLFL

721 780
consenso61.SEQ DGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQ
consensoh61.SEQ DGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQ
consensom61.SEQ DGIDKAQEEHEKYPNNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQ

consensou61.SEQ DGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQ

781 840

consenso61.SEQ LDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFT

consensoh61.SEQ LDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFT

consensom61.SEQ LDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFT

consensou61.SEQ LDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFT

841 900

consenso61.SEQ STTVKAACWWAGIKQEFGIPYNPQSQGVVESINKELKKIIGQIRDQAEHLKTAVQMAVFI

consensoh61.SEQ STTVKAACWWAGIKQEFGIPYNPQSQGVVESINKELKKIIGQIRDQAEHLKTAVQMAVFI

consensom61.SEQ STTVKAACWWAGIKQEFGIPYNPQSQGVVESINKELKKVIGQIRDQAEHLKTAVQMAVFI

consensou61.SEQ STTVKAACWWAGIKQEFGIPYNPQSQGVVESINKELKKIIGQIRDQAEHLKTAVQMAVFI

901 960

consenso61.SEQ HNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLL

consensoh61.SEQ HNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLL

consensom61.SEQ HNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLL

consensou61.SEQ HNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLL

961 1006

consenso61.SEQ WKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED

consensoh61.SEQ WKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED

consensom61.SEQ WKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED

consensou61.SEQ WKGEGAEVIQDNSDIKVVLRKAKIIRDYGKQMAGDHCVASRQDED

PROTEÍNA ENV

Env.MSF.msf MSF: 862 Type: P jueves 3 de mayo de 2001 Check: 1900 ..
Name: consenso61.SEQ Len: 862 Check: 8514 Weight: 1.00
Name: consensoh61.seq Len: 862 Check: 7268 Weight: 1.00
Name: consensom61.SEQ Len: 862 Check: 7641 Weight: 1.00
Name: consensou61.SEQ Len: 862 Check: 8477 Weight: 1.00

//

```
1 60
consenso61.SEQ MRVMEIRKNYQHLWKWGTMLLGLMICKAAENLWVTVYYGVPVWKEATTTLFCASDAKAY
consensoh61.seq MRVTEIRKNYQHLWKWGTMLLGLMICKAAENLWVTVYYGVPVWKEATTTLFCASDAKAY
consensom61.SEQ MRVMEIRKNYQHLWKWGTMLLGLMIICKAAENLWVTVYYGVPVWKEAPTTLFCAADAKAY
consensou61.SEQ MRVMEIRKNYQHLWKWGTMLLGLMICKAAENLWVTVYYGVPVWKEATTTLFCASDAKAY

61 120
consenso61.SEQ DTEVHNVWATHACVPTDPNPQEVKLNVTENFNMWKNMVDQMHEDEIISLWDQSLKPCVK
consensoh61.seq DTEVHNVWATHACVPTDPNPQEVLEKLNVTENFNMWKNMVDQMHEDEIISLWDQSLKPCVK
consensom61.SEQ DTEVHNVWATHACVPTDPNPQEVRLNVTENFNMWKNMVDQMHEDEIISLWDQSLKPCVK
consensou61.SEQ DTEVHNVWATHACVPTAPNPQEVLEKLNVTENFNMWKNMVDQMHEDEIISLWDQSLKPCVK

121 180
consenso61.SEQ LTPLCVTLNCTDYGNDTNTNNSATNPTSSSGGMEGRGEIKNCSFNITTSIRDKVQKEYA
consensoh61.seq LTPLCVTLNCTDYGNDTNTNN..ATNTTSSSGGMEGRGEIKNCSFNITTSIRDKVQKEYA
consensom61.SEQ LTPLCVTLNCTDYGNDTNTNNSATNPTSSSGGMEGRGEIKNCSFNITTSIRDKVQKEYA
consensou61.SEQ LTPLCVTLNCTDYGNDTNTNNSATNPTSSSGGMEGRGEIKNCSFNITTSIRDKVQKEYA

181 240
consenso61.SEQ LFYSLDVIPIKDDNTSYMLRSCNTSVITQACPKVSFEPPIPIHYCAPAGFAILKCNKDFN
consensoh61.seq LFYSLDVIPIKDDNTSYMLRSCNTSVITQACPKVSFEPPIPIHYCAPAGFAILKCNKDFN
consensom61.SEQ LFYSLDVIPIKDDNTSYMLRSCNTSVITQACPKVSFEPPIPIHYCAPAGFAILKCNKDFN
consensou61.SEQ LFYSLDVIPIKDDNTSYMLRSCNTSVITQACPKVSFEPPIPIHYCAPGGFAILKCNKDFN

241 300
consenso61.SEQ GKGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEVIRSDNFSDNAKVIIVHLNESVEI
consensoh61.seq GKGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEKEVIRSDNFSDNAKNIIVHLNESVEI
consensom61.SEQ GKGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEVIRSDNFSDNAKNIIVHLNESVEI
consensou61.SEQ GKGPCTNVSTVQCTHGIRPVVLTQLVNLNGSLAEVIRSDNFSDNAKNIIVHLNESVEI

301 360
consenso61.SEQ NCTRLNNITRRSIHVGHVGPGRAIYTTGIIGKIRQAHCNISRAKWNNTLKQIVTKLREQF
```

consensoh61.seq NCTRLNNNTRRSIHVGHVGPGRAIYTTGIIGKIRQAHCNISRAKWNNTLKQIVTKLREQF
consensom61.SEQ NCTRLNNNTRRSIHVGHVGPGRAIYTTGIIGKIRQAHCNISRAKWNNTLKQIVTKLREQF
consensou61.SEQ NCTRLNNNTRRSIHVGHVGPGRAIYTTGIIGKIRQAHCNISRAKWNNTLKQIVTKLREQF

361 420
consenso61.SEQ KNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNGTAWSNNTEGNENDTITL
consensoh61.seq KNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNGTAWSNNTEGNENDTITL
consensom61.SEQ KNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNGTAWSNNTEGNENDTITL
consensou61.SEQ KNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNGTAWSNNTEGNENDTITL

421 480
consenso61.SEQ PCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLILTRDGGINQTNTEIFRPGGGDM
consensoh61.seq PCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLILTRDGGINQTNTEIFRPGGGDM
consensom61.SEQ PCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLILTRDGGINQTNTEIFRPGGGDM
consensou61.SEQ PCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLILTRDGGINQTNTEIFRPGGGDM

481 540
consenso61.SEQ KDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIIGAMLLGFLGAAGSTMGAAS
consensoh61.seq KDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIIGAMLLGFLGAAGSTMGAAS
consensom61.SEQ KDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIIGAMLLGFLGAAGSTMGAAS
consensou61.SEQ RDNWRSELYKYKVVKIEPLGIAPTKAKRRVVQREKRAVGIIGAMLLGFLGAAGSTMGAAS

541 600
consenso61.SEQ MTLTVQARQLLSGIVQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLG
consensoh61.seq MTLTVQARQLLSGIVQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLG
consensom61.SEQ MTLTVQARQLLSGIVQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLG
consensou61.SEQ MTLTVQARQLLSGIVQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLG

601 660
consenso61.SEQ IWGCSGKLICTTAVPWNTSWSNKS LNQIWDNMTWMQWEREIDNYTSLIYTLIEESQNQQE
consensoh61.seq IWGCSGKLICTTAVPWNTSWSNKS LNQIWDNMTWMQWEREIDNYTSLIYTLIEESQNQQE
consensom61.SEQ IWGCSGKLICTTAVPWNTSWSNKS LNQIWDNMTWMQWEREIDNYTSLIYTLIEESQNQQE
consensou61.SEQ IWGCSGKLICTTAVPWNTSWSNKS LNQIWDNMTWMQWEREIDNYTSLIYTLIEESQNQQE

661 720
consenso61.SEQ KNEQELLELDN WASLWNWFSITNWLWYIKIFIMIVGGLVGLRIVFTVLSIVNRVRQGYSP
consensoh61.seq KNEQELWELDN WASLWNWFSITNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVRQGYSP
consensom61.SEQ KNEQELLELDN WASLWNWFSITNWLWYIKIFIMIVGGLVGLRIVFVLSIGNRVRQGYSP

consensou61.SEQ KNEQELLELDNwasLWNWFSITNWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVRQGYSP

 721 780
 consenso61.SEQ LSFQTRLPARRGPDPEGIEEEGGGERDRDRSGQLVDGFLAIIWVDLRSLCLFSYHRLRDL
 consensoh61.seq LSFQTRLPARRGPDPEGIEEEGGGERDRDRSGQLVDGFLAIIWVDLRSLCLFSYHRLRDL
 consensom61.SEQ LSFQTRLPARRGPDPEGIEEEGGGERDRDRSGQLVDGFLAIIWVDLRSLCLFSYHRLRDL
 consensou61.SEQ LSFQTRLPARRGPDPEGIEEEGGGERDRDRSGQLVDGFLAIIWVDLRSLCLFSYHRLRDL

 781 840
 consenso61.SEQ LLIVTRIVELLGRRGWEALKYWWNLLQYWIQELKNSAVSLLNATAIAVAEGTDRVIEVLQ
 consensoh61.seq LLIVTRIVELLGRRGWEALKYWWNLLQYWIQELKNSAVSLLNATAIAVAEGTDRVIEVLQ
 consensom61.SEQ LLIVTRIVELLGRRGWEALKYWWNLLQYWIQELKNSAVSLLNATAIAVAEGTDRVIEVLQ
 consensou61.SEQ LLIVTRIVELLGRRGWEALKYWWNLLQYWIQELKNSAVSLLNATAIAVAEGTDRVIEVLQ

 841 862
 consenso61.SEQ RACRAILHIPRRIRQGLERALQ
 consensoh61.seq RACRAILHIPRRIRQGLERALQ
 consensom61.SEQ RACRAILHIPRRIRQGLERALQ
 consensou61.SEQ RAFRAILHIPRRIRQGLERALQ

PROTEÍNA VIF

Vif.MSF.msf MSF: 193 Type: P jueves 3 de mayo de 2001 Check: 2167 ..

Name: consenso61.SEQ Len: 193 Check: 3603 Weight: 1.00

Name: consensoh61.SEQ Len: 193 Check: 3078 Weight: 1.00

Name: consensom61.SEQ Len: 193 Check: 1701 Weight: 1.00

Name: consensou61.SEQ Len: 193 Check: 3785 Weight: 1.00

//

1

60

consenso61.SEQ MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSRKAKGWFYRHHYESTHPRISSEVHIPLG

consensoh61.SEQ MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSRKAKGWFYRHHYESTHPRISSEVHIPLG

consensom61.SEQ MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSRKAKGWFYRHHYESTHPRISSEVHIPLG

consensou61.SEQ MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSRKAKGWFYRHHYESTHPRISSEVHIPLG

61

120

consenso61.SEQ DAKLVITTYWGLHTGERDWHLGQGVSI EWKRKRYSTQLDPNLADQLIHLHYFDCFSES

consensoh61.SEQ DAKLVITTYWGLHTGERDWHLGQGVSI EWKRKRYSTQLDPNLADQLIHLHYFDCFSES

consensom61.SEQ DAKLVITTYWGLHTGERDWHLGQGVSI E.RKKRYSTQLDPNLADQLIHLHYFDCFSES

consensou61.SEQ DAKLVITTYWGLHTGERDWHLGQGVSI EWKRKRYSTQLDPNLADQLIHLHYFDCFSES

121

180

consenso61.SEQ RNAILGHIVSPRCEYQAGHNKVGSLQYLALTALIRPRKKIKPPLPSVTKLTEDRWNPQK

consensoh61.SEQ RNAILGHIVSPRCEYQAGHNKVGSLQYLALTALIRPRKKIKPPLPSVTKLTEDRWNPQK

consensom61.SEQ RNAILGHIVSPRCEYQAGHNKVESLQYLALTALIKPRKKIKPPLPSVTKLTEDRWNPQK

consensou61.SEQ RNAILGHVSPRCEYQAGHNKVGSLQYLALTALIRPRKKIKPPLPSVTKLTEDRWNPQK

181

193

consenso61.SEQ TKGHKSHTMNGH

consensoh61.SEQ TKGHKSHTMNGH

consensom61.SEQ TKGHKSHTMNGH

consensou61.SEQ TKGHKSHTMNGH

PROTEÍNA VPR

Vpr.MSF.msf MSF: 96 Type: P jueves 3 de mayo de 2001 Check: 7289 ..

Name: consenso61.SEQ Len: 96 Check: 6789 Weight: 1.00

Name: consensoh61.SEQ Len: 96 Check: 7472 Weight: 1.00

Name: consensom61.SEQ Len: 96 Check: 6760 Weight: 1.00

Name: consensou61.SEQ Len: 96 Check: 6268 Weight: 1.00

//

```

      1
consenso61.SEQ MEQAPEDQGPQREPYNWTTLELLEELKNEAVRHFPRILHSLGQHIYETYGDTWAGVEAI
consensoh61.SEQ ME.APEDQGPQREPYNWTTLELLEELKNEAVRHFPRILHSLGQYIYETYGETWAGVEAI
consensom61.SEQ MEQAPEDQGPQREPYNWTTLELLEELKNEAVRHFPRILHSLGQHIYETYGDTWAGVEAI
consensou61.SEQ MEQAPEDQGPQREPYNWTTLELLEELKNEAVRHFPRILHGLGQHIYETYGDTWAGVEAI
```

```

      61
consenso61.SEQ IRILQQLLFHFRIGCQHSRIGILRRRRARNGASRS
consensoh61.SEQ IRILQQLLFHFRIGCQHSRIGILRQRRARNGASRS
consensom61.SEQ IRILQQLLFHFRIGCQHSRIGILRQRRARNGASRS
consensou61.SEQ IRILQQLLFHFRIGCQHSRIGILRQRRARNGASRS
      96
```

PROTEÍNA VPU

Vpu.MSF.msf MSF: 81 Type: P jueves 3 de mayo de 2001 Check: 6748 ..

Name: consenso61.SEQ Len: 81 Check: 6860 Weight: 1.00

Name: consensoh61.SEQ Len: 81 Check: 6399 Weight: 1.00

Name: consensom61.SEQ Len: 81 Check: 6915 Weight: 1.00

Name: consensou61.SEQ Len: 81 Check: 6574 Weight: 1.00

//

1

60

consenso61.SEQ MQSLQILAVVALVVAIIIAIVVWSIVFIECRKIRRQNKIERLIDRIIERAEDSGNESDGD

consensoh61.SEQ MQSL.IVAVVALVVAIIIAIVVWSIVFIECRKIRKQKKIERLIDRIIERAEDSGNESDGD

consensom61.SEQ MQSLQIVAVVALVVAIIIAIVVWSIVFIECRKIRKQNKIERLIDRIIERAEDSGNESDGD

consensou61.SEQ MQSLQIVAVVALVVAIIIAIVVWSIVFIECRKIRKQKKIERLIDRIIERAEDSGNESDGD

61

81

consenso61.SEQ QEELSALVEMGHDPWDVNDL

consensoh61.SEQ QEELSALVEMGHDPWDVNDL

consensom61.SEQ QEELSALVEMGHDPWDVNNL

consensou61.SEQ QEELSALVEMGHDPWDVNDL

PROTEÍNA NEF

NEF.MSF.msf MSF: 206 Type: P jueves 3 de mayo de 2001 Check: 1649 ..

Name: consenso61.SEQ Len: 206 Check: 2544 Weight: 1.00

Name: consensoh61.SEQ Len: 206 Check: 3416 Weight: 1.00

Name: consensom61.SEQ Len: 206 Check: 3773 Weight: 1.00

Name: consensou61.SEQ Len: 206 Check: 1916 Weight: 1.00

//

```
1 60
consenso61.SEQ MGGKWSKSSKIGWPTVRERMRRAEPAADGVGAVSRDLEKHGAISSNNTAANNADCAWLEA
consensoh61.SEQ MGGKWSKNSKSGWPTVRERMRRAEPAADGVGAVSRDLEKHGAISSNNTAANNADCAWLEA
consensom61.SEQ MGGKWSKSSKIGWPTVRERMRRAEPAADGVGAVSRDLEKHGAISSNNTAANNADCAWLEA
consensou61.SEQ MGGKWSKNSKSGWPTVRERMRRAEPAAEVAVSRDLEKHGAISSNNTAANNADCAWLEA
```

```
61 120
consenso61.SEQ QEEEEVGFVVRPQVPLRPMTYKAAFDSLHFLKEKGGLEGIIWSQRRQDILDLDLWVHHTQGY
consensoh61.SEQ QEEEEVGFVVRPQVPLRPMTYKAAFDSLHFLKEKGGLEGIVWSQRRQDILDLDLWVHHTQGY
consensom61.SEQ QEEEEVGLPVRPQVPLRPMTYKAAFDSLHFLKEKGGLEGIVWSQRRQDILDLDLWVHHTQGY
consensou61.SEQ QEEEEVGFVVRPQVPLRPMTYKAAFDSLHFLKEKGGLEGLIWSQRRQDILDLDLWIHHTQGY
```

```
121 180
consenso61.SEQ FPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEANEENCLLHPMSQHGMEDPEKEV
consensoh61.SEQ FPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEANEENCLLHPMSQHGMEDPEKEV
consensom61.SEQ FPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEANEENCLLHPMSQHGMEDPEKEV
consensou61.SEQ FPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEANEENCLLHPMSQHGIEDPEKEV
```

```
181 206
consenso61.SEQ LVWRFDSRLAFHHKARELHPEYYKNC
consensoh61.SEQ LVWRFDSRLAFHHKARELHPEYYKNC
consensom61.SEQ LVWRFDSRLTFHHKARELHPEYYKNC
consensou61.SEQ LVWRFDSRLAFHHKARELHPEYYKNC
```

PROTEÍNA REV

Rev.MSF.msf MSF: 119 Type: P jueves 3 de mayo de 2001 Check: 4358 ..

Name: consenso61.SEQ Len: 119 Check: 8100 Weight: 1.00

Name: consensoh61.SEQ Len: 119 Check: 8506 Weight: 1.00

Name: consensom61.SEQ Len: 119 Check: 8571 Weight: 1.00

Name: consensou61.SEQ Len: 119 Check: 9181 Weight: 1.00

//

1

60

consenso61.SEQ MAGRSGDSDEELLKTVRLIKFLYQSNAPPSQEGTRQARRNRRRRWRERQRQIRSISGWIL

consensoh61.SEQ MAGRSGDSDEELLKTVRLIKFLYQSNPPPSQEGTRQARRNRRRRWRERQRQIRSISGWIL

consensom61.SEQ MAGRSGDSDEKLLKTVRLIKFLYQSNPPPSQEGTRQARRNRRRRWRERQRQIRSISGWIL

consensou61.SEQ MAGRSGDSDEELLKTVRLIKFLYQSNPPPSQEGTRQARRNRRRRWRERQRQIRSISGWIL

61

119

consenso61.SEQ SNYLGRPAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQILVESPTILDSGTKE...

consensoh61.SEQ SNYLGRPAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQILVESPAVLDSGTKEQCC

consensom61.SEQ SNYLGRPAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQILVESPTILDSGTKE...

consensou61.SEQ SNYLGRPAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQILVESPTVLDSTKE...

PROTEÍNA TAT

Tat.MSF.msf MSF: 101 Type: P jueves 3 de mayo de 2001 Check: 3232 ..

Name: consenso61.SEQ Len: 101 Check: 3465 Weight: 1.00

Name: consensoh61.SEQ Len: 101 Check: 3433 Weight: 1.00

Name: consensom61.SEQ Len: 101 Check: 3034 Weight: 1.00

Name: consensou61.SEQ Len: 101 Check: 3300 Weight: 1.00

//

1

60

consenso61.SEQ MEPVDPRLEPWKHPGSQPRTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRQRRRAPQ

consensoh61.SEQ MEPVDPRLEPWKHPGSQPRTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRQRRRAPQ

consensom61.SEQ MEPVDPRLEPWKHPGSQPRTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRQRRKAPQ

consensou61.SEQ MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRQRRRAPQ

61

101

consenso61.SEQ DSQTHQVSLSKQRASQPGGDPTGPKESKKKVERETETDPVN

consensoh61.SEQ DSQTHQVSLSKQPASQPGGDPTGPKESKKKVERETETDPVN

consensom61.SEQ DSQTHQVSLSKQPASQPGGDPTGPKESKKKVERETETDPVN

consensou61.SEQ DSQTHQVSLSKQPASQPGGDPTGPKESKKKVERETETDPVN

LTR

ltr.MSF.msf MSF: 455 Type: D jueves 3 de mayo de 2001 Check: 936 ..

Name: consenso61.SEQ Len: 455 Check: 1906 Weight: 1.00

Name: consensoh61.SEQ Len: 455 Check: 9496 Weight: 1.00

Name: consensom61.SEQ Len: 455 Check: 9190 Weight: 1.00

Name: consensou61.SEQ Len: 455 Check: 344 Weight: 1.00

//

```
1 60
consenso61.SEQ TGGAAGGGATAATTTGGTCTCAGAGAAGACAAGATATCCTTGATCTGTGGGTCCACCACA
consensoh61.SEQ TGGAAGGGATAGTTTGGTCTCAAAGAAGACAAGATATCCTTGATCTGTGGGTCCACCACA
consensom61.SEQ TGGAAGGGATAGTTTGGTCTCAAAGAAGACAAGATATCCTTGATCTGTGGGTCCACCACA
consensou61.SEQ TGGAAGGGCTAATTTGGTCCCAGAAAAGACAAGATATCCTTGATCTGTGGATCCACCACA
```

```
61 120
consenso61.SEQ CACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCAC
consensoh61.SEQ CACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCAC
consensom61.SEQ CACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCAC
consensou61.SEQ CACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCAC
```

```
121 180
consenso61.SEQ TGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCA
consensoh61.SEQ TGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCA
consensom61.SEQ TGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCA
consensou61.SEQ TGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGAGCCAGAGAAGGTAGAAGAGGCCA
```

```
181 240
consenso61.SEQ ATGAAGGAGAGAACAACACTGCTTGTTACACCCTATGAGCCNGCATGGGATGGAGGACCCGG
consensoh61.SEQ ATGAAGGAGAGAACAACACTGCTTGTTACACCCTATGAGCCAGCATGGGATGGAGGACCCGG
consensom61.SEQ ATGAAGGAGAGAACAACACTGCTTGTTACACCCTATGAGCCAGCATGGGATGGAGGACCCGG
consensou61.SEQ ATGAAGGAGAGAACAACACTGCTTGTTACACCCTATGAGCCAGCATGGGATAGAGGACCCGG
```

```
241 300
consenso61.SEQ AGAAAGAAGTGTTAGTGTGGAGGTTTTGACAGCCGCTTAGCCTTCCATCACANGGCCCGAG
consensoh61.SEQ AGAAAGAAGTGTTAGTGTGGAGGTTTTGACAGCCGCTTAGCCTTCCATCACANGGCCCGAG
consensom61.SEQ AGAAAGAAGTGTTAGTGTGGAGGTTTTGACAGCCGCTAACCTTCCATCACANGGCCCGAG
consensou61.SEQ AGAAAGAAGTGTTAGTATGGAGGTTTTGACAGCCGCTTAGCCTTCCATCACANGGCCCGAG
```

```
301 360
consenso61.SEQ AGCTACATCCGGAGTACTACAAGAAGTCTGACATCGAGCTTTCTACAAGGGACTTTCCG
consensoh61.SEQ AGCTACATCCGGAGTACTACAAGAAGTCTGACATCGAGCTATCTACAAGGGACTTTCCG
consensom61.SEQ AGCTACATCCGGAGTACTACAAGAAGTCTGACATCGAGCTATCTACAAGGGACTTTCCG
consensou61.SEQ AGCTACATCCGGAGTACTACAAGAAGTCTGACATCGAGCTTTCTACAAGGGACTTTCCG
```

361 420
consenso61.SEQ CTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT
consensoh61.SEQ CTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT
consensom61.SEQ CTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT
consensou61.SEQ CTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT

421 455
consenso61.SEQ GCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGG
consensoh61.SEQ GCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGG
consensom61.SEQ GCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGG
consensou61.SEQ GCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGG

SECUENCIA LEADER

Secuencia leader.MSF.msf MSF: 155 Type: D jueves 3 de mayo de 2001 Check: 3352 ..

Name: consenso61.SEQ Len: 155 Check: 757 Weight: 1.00

Name: consensoh61.SEQ Len: 155 Check: 799 Weight: 1.00

Name: consensom61.SEQ Len: 155 Check: 997 Weight: 1.00

Name: consensou61.SEQ Len: 155 Check: 799 Weight: 1.00

//

```
consenso61.SEQ      1                               60
consenso61.SEQ      GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGAAGCTCTCTCGACGCAG
consensoh61.SEQ     GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGAGAAAACCAGAGAAGCTCTCTCGACGCAG
consensom61.SEQ     GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGAAGCTCTCTCGACGCAG
consensou61.SEQ     GTGGCGCCCGAACAGGGACTTGAAAGCGAAAGAGAAAACCAGAGAAGCTCTCTCGACGCAG
```

```
consenso61.SEQ     61                               120
consenso61.SEQ     GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCGAAGGGACGGCGACTGGTGAGTACGC
consensoh61.SEQ     GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCGAGGGGACGGCGACTGGTGAGTACGC
consensom61.SEQ     GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCGAGGGGACGGCGACTGGTGAGTACGC
consensou61.SEQ     GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGGCGAGGGGACGGCGACTGGTGAGTACGC
```

```
consenso61.SEQ     121                               155
consenso61.SEQ     CAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAG
consensoh61.SEQ     CAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAG
consensom61.SEQ     CAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAG
consensou61.SEQ     CAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAG
```