

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE FARMACIA**

**Departamento de Bioquímica y Biología Molecular II**



**RECOMBINACIÓN DEL VIH-1 IN VITRO**

**MEMORIA PRESENTADA PARA OPTAR AL GRADO DE  
DOCTOR POR**

María José Iglesias Sánchez

Bajo la dirección del Doctor:

Cecilio López Galíndez

**Madrid, 2003**

**ISBN: 84-669-2050-1**

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

DEPARTAMENTO DE BIOQUÍMICA Y  
BIOLOGÍA MOLECULAR II

**RECOMBINACIÓN DEL  
VIH-1 *IN VITRO***

**TESIS DOCTORAL**

**MARÍA JOSÉ IGLESIAS SÁNCHEZ**

Madrid, 2003

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FARMACIA

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA  
MOLECULAR II

# **RECOMBINACIÓN DEL VIH-1 *IN VITRO***

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

**María José Iglesias Sánchez**

Director de la Tesis:

**Dr. Cecilio López Galíndez**

Centro Nacional de Microbiología

Instituto de Salud Carlos III

Madrid, Enero 2003

Quiero agradecer, en primer lugar, al Dr. Cecilio López Galindez no sólo la dirección de este trabajo, sino también la oportunidad de realizarlo y el esfuerzo y la ayuda prestados durante todo este tiempo.

Agradezco la ayuda, el interés y los consejos que en su momento me prestaron el Dr. Antonio Talavera Díaz, el Dr. Pedro Lazo y la Dra. Inmaculada Herrera.

Así mismo agradezco al Instituto de Salud Carlos III y a las personas que en él trabajan las facilidades que me han brindado. Pero sobre todo a mis compañeros del laboratorio de Virología Molecular, que no sólo me ayudaron, colaboraron y se interesaron por el trabajo, sino que me acompañaron durante mucho tiempo en los buenos y los malos ratos. A los que aún están: Isabel Olivares, Concha Casado, M<sup>a</sup> José Dueñas, Gonzalo Bello, María Pernas, Antonio Borderia, Mónica Gutiérrez, Tamara Alvaro, y Virginia Nunes; y a los que ya no están: Sonsoles Sánchez, Eloísa Yuste, Víctor Sánchez, Loles Iglesias, Guillermo Shaw, Mavi Martínez, Imelda Urtasun, Inés González y Juan Aballe.

Y a los amigos de otros laboratorios: Alicia, Susana, Paz, Fati, Javi, Marta, Elena, Beatriz, Merche y Pilar, gracias por co

### ***A mi familia***

Gracias también a mi familia y a los amigos que sin pertenecer a este "mundo de la investigación", han intentado entender en qué consistía esta, a veces, incomprensible tarea.

Quiero agradecer, en primer lugar, al Dr. Cecilio López Galíndez no sólo la dirección de este trabajo, sino también la oportunidad de realizarlo y el esfuerzo y la ayuda prestados durante todo este tiempo.

Agradezco la ayuda, el interés y los consejos que en su momento me prestaron el Dr. Antonio Talavera Díaz, el Dr. Pedro Lazo y la Dra. Inmaculada Herrera.

Así mismo agradezco al Instituto de Salud Carlos III y a las personas que en él trabajan las facilidades que me han brindado. Pero sobre todo a mis compañeros del laboratorio de Virología Molecular, que no sólo me ayudaron, colaboraron y se interesaron por el trabajo, sino que me acompañaron durante mucho tiempo en los buenos y los malos ratos. A los que aún están: Isabel Olivares, Concha Casado, M<sup>a</sup> José Dueñas, Gonzalo Bello, María Pernas, Antonio Bordería, Mónica Gutierrez, Tamara Álvaro, y Virginia Nunes; y a los que ya no están: Sonsoles Sánchez, Eloísa Yuste, Víctor Sánchez, Loles Iglesias, Guillermo Shaw, Mavi Martínez, Imelda Urtasun, Inés González y Juan Aballe.

Y a los amigos de otros laboratorios: Alicia, Susana, Paz, Fali, Javi, Marta, Elena, Beatriz, Merche y Pilar, gracias por compartir tantos cafés y buen humor.

Gracias también a mi familia y a los amigos que sin pertenecer a este “mundo de la investigación”, han intentado entender en qué consistía esta, a veces, incomprensible tarea.

## ABREVIATURAS

ADN	Ácido desoxirribonucleico
Ampli Taq	ADN polimerasa de <i>Thermophilus Aquaticus</i>
ARN	Ácido ribonucleico
ARNi	Ácido ribonucleico interferente
°C	Grados centígrados
µm	micrómetros
CMSP	Células mononucleares de sangre periférica
c.p.m.	Cuentas por minuto
DI <sub>50</sub>	Dosis infectiva 50
DNasa I	Desoxirribonucleasa I
dNTP	Desoxirribonucleótido trifosfato
ECP	Efecto citopático

Parte del trabajo de investigación discutido en esta Tesis ha sido publicado como:

María José Iglesias-Sánchez y Cecilio López-Galíndez.

“Analysis, Quantification and Evolutionary Consequences of HIV-1 *in vitro* Recombination”.

*Virology* 304, 392-402 (2002)

µm	micrómetros
ml	Mililitros.
mM	Milimolar
ng	Nanogramos.
NJ	Método del vecino más próximo (“neighbor joining”)
NSI	No inductor de sincitios
pb	Pares de bases.
PBS	Tampón salino de fosfato
PCR	Reacción en cadena de la polimerasa
pg	Picogramos.
PLL	Poli-L-lisina
r.p.m.	Revoluciones por minuto
RNasa	Ribonucleasa.
Rnasa H	Ribonucleasa H
RT	Retrotranscriptasa o transcriptasa inversa.
RT-PCR	Reacción en cadena de la polimerasa a partir de ARN
SBF	Suero bovino fetal
SDS	Dodecil sulfato sódico
SIDA	Síndrome de inmunodeficiencia adquirida
t-ARN	ARN de transferencia
TAE	Tampón Tris-acetato-EDTA
TAR	Región de respuesta transactivadora
TBE	Tampón Tris-ácido bórico-EDTA
Tris	Tris-(hidroximetil)-aminometano
u	Unidades
u. f. p.	Unidades formadoras de placas
UV	Ultravioleta
V	Voltios
VIH	Virus de la inmunodeficiencia humana

## ABREVIATURAS

ADN	Acido desoxiribonucleico
Ampli Taq	ADN polimerasa de <i>Termófilus Aquaticus</i>
ARN	Ácido ribonucleico
ARNt	Ácido ribonucleico de transferencia
°C	Grados centígrados
cm	Centímetros.:
CMSP	Células mononucleares de sangre periférica
c.p.m.	Cuentas por minuto
DI <sub>50</sub>	Dosis infectiva 50
DNasa I	Desoxirribonucleasa I.
dNTP	Desoxirribonucleótido trifosfato
ECP	Efecto citopático
EDTA	Ácido diamino tetraacético
ELISA	Ensayo Immunoenzimático
h	Horas.
IS	Inductor de sincitios
LTRs	Repeticiones largas terminales
M	Molar.
m.d.i.	Multiplicidad de infección
min	Minutos
ml	Mililitros.
mM	Milimolar
ng	Nanogramos.
NJ	Método del vecino más próximo ("neighbor joining")
NSI	No inductor de sincitios
pb	Pares de bases.
PBS	Tampón salino de fosfato
PCR	Reacción en cadena de la polimerasa
pg	Picogramos.
PLL	Poli-L-lisina
r.p.m.	Revoluciones por minuto
RNasa	Rivonucleasa.
Rnasa H	Rivonucleasa H
RT	Retrotranscriptasa o transcriptasa inversa.
RT-PCR	Reacción en cadena de la polimerasa a partir de ARN
SBF	Suero bovino fetal
SDS	Dodecil sulfato sódico
SIDA	Síndrome de inmunodeficiencia adquirida
t-ARN	ARN de transferencia
TAE	Tampón Tris-acetato-EDTA
TAR	Región de respuesta transactivadora
TBE	Tampón Tris-ácido bórico-EDTA
Tris	Tris-(hidroximetil)-aminometano
u	Unidades
u. f. p.	Unidades formadoras de placas
UV	Ultravioleta
V	Voltios
VIH	Virus de la inmunodeficiencia humana

VIH-1	Virus de la inmunodeficiencia humana tipo 1
VIH-2	Virus de la inmunodeficiencia humana tipo 2
VIS	Virus de la inmunodeficiencia simia
VSV	Virus de la estomatitis vesicular
μCi	Microcurios.
μl	Microlitros.

ÍNDICE

<b>1. INTRODUCCIÓN</b>	<b>1</b>
1.1. EL VIRUS DE LA INMUNODEFICIENCIA HUMANA	3
1.1.2. <i>Retrovirus</i>	3
1.1.3. <i>Estructura del virión y organización genómica</i>	4
1.1.4. <i>Ciclo biológico</i>	8
1.1.5. <i>Historia natural y patogénesis de la infección</i>	11
1.2. DIVERSIDAD GENÉTICA Y EPIDEMIOLOGÍA MOLECULAR DEL VIH-1	12
1.2.1. <i>Clasificación de subtipos</i>	13
1.2.2. <i>Distribución geográfica de los subtipos y formas recombinantes circulantes del VIH</i>	14
1.3. VARIACIÓN EN VIH	15
1.3.1. <i>Generación de variabilidad</i>	16
1.3.2. <i>Mecanismos de cambio en la frecuencia de los variantes de una población natural</i>	17
1.4. RECOMBINACIÓN EN VIH	18
1.4.1. <i>Transcriptasa inversa</i>	18
1.4.2. <i>Mecanismo de la transcriptasa inversa</i>	19
1.4.3. <i>Mecanismo de la recombinación retroviral</i>	22
1.4.4. <i>Modelos de recombinación retroviral</i>	23
1.4.5. <i>El proceso de transferencia de hebra</i>	23
1.4.6. <i>Las dos copias de ARN de VIH</i>	25
<b>2. OBJETIVOS</b>	<b>26</b>
<b>3. MATERIALES Y MÉTODOS</b>	<b>28</b>
3.1. CULTIVOS CELULARES	29
3.1.1. <i>Células</i>	29
3.1.2. <i>Medios de cultivo</i>	29
3.1.3. <i>Virus y semillas</i>	29
3.1.4. <i>Plaqueo en células MT4</i>	31
3.1.5. <i>Infecciones dobles</i>	31
3.1.6. <i>Cultivos secundarios de 24 horas</i>	32
3.2. OBTENCIÓN DE ÁCIDOS NUCLEICOS	32
3.2.1. <i>Obtención de ADN</i>	32
3.2.2. <i>Obtención de ARN</i>	33
3.3. REACCIÓN EN CADENA DE LA POLIMERASA (PCR)	33
3.3.1. <i>PCR utilizada para la identificación de clones biológicos recombinantes</i>	34
3.3.2. <i>PCR cuantitativa</i>	35
3.4. CONSTRUCCIÓN DE LOS PLÁSMIDOS UTILIZADOS EN LOS PATRONES DE LA PCR CUANTITATIVA	36
3.4.1. <i>Preparación del inserto y ligación</i>	37
3.4.2. <i>Transformación de bacterias competentes</i>	37
3.4.3. <i>Obtención de ADN plasmídico a gran escala</i>	38
3.5. AISLAMIENTO DE CLONES BIOLÓGICOS RECOMBINANTES	38
3.6. SECUENCIACIÓN	38
3.7. ANÁLISIS DE SECUENCIAS	39
3.8. ANÁLISIS FILOGENÉTICO	39
3.9. ANÁLISIS DE RECOMBINACIÓN	39
3.10. MUTAGÉNESIS	40
3.11. TRANSFECCIONES	41
<b>4. RESULTADOS</b>	<b>42</b>
4.1. PUESTA A PUNTO DE LAS TÉCNICAS UTILIZADAS EN EL ESTUDIO DE LA	43

## ÍNDICE

<b>1. INTRODUCCIÓN</b>	<b>1</b>
1.1. EL VIRUS DE LA INMUNODEFICIENCIA HUMANA	3
1.1.2. <i>Retrovirus</i>	3
1.1.3. <i>Estructura del virión y organización genómica</i>	4
1.1.4. <i>Ciclo biológico</i>	8
1.1.5. <i>Historia natural y patogénesis de la infección</i>	11
1.2. DIVERSIDAD GENÉTICA Y EPIDEMIOLOGÍA MOLECULAR DEL VIH-1	12
1.2.1. <i>Clasificación de subtipos</i>	13
1.2.2. <i>Distribución geográfica de los subtipos y formas recombinantes circulantes del VIH</i>	14
1.3. VARIACIÓN EN VIH	15
1.3.1. <i>Generación de variabilidad</i>	16
1.3.2. <i>Mecanismos de cambio en la frecuencia de los variantes de una población natural</i>	17
1.4. RECOMBINACIÓN EN VIH	18
1.4.1. <i>Transcriptasa inversa</i>	18
1.4.2. <i>Mecanismo de la transcriptasa inversa</i>	19
1.4.3. <i>Mecanismo de la recombinación retroviral</i>	22
1.4.4. <i>Modelos de recombinación retroviral</i>	23
1.4.5. <i>El proceso de transferencia de hebra.</i>	23
1.4.6. <i>Las dos copias de ARN de VIH</i>	25
<b>2. OBJETIVOS</b>	<b>26</b>
<b>3. MATERIALES Y MÉTODOS</b>	<b>28</b>
3.1. CULTIVOS CELULARES	29
3.1.1. <i>Células</i>	29
3.1.2. <i>Medios de cultivo</i>	29
3.1.3. <i>Virus y semillas</i>	29
3.1.4. <i>Plaqueo en células MT4</i>	31
3.1.5. <i>Infecciones dobles</i>	31
3.1.6. <i>Cultivos secundarios de 24 horas</i>	32
3.2. OBTENCIÓN DE ÁCIDOS NUCLEICOS	32
3.2.1. <i>Obtención de ADN</i>	32
3.2.2. <i>Obtención de ARN</i>	33
3.3. REACCIÓN EN CADENA DE LA POLIMERASA (PCR)	33
3.3.1. <i>PCR utilizada para la identificación de clones biológicos recombinantes</i>	34
3.3.2. <i>PCR cuantitativa</i>	35
3.4. CONSTRUCCIÓN DE LOS PLÁSMIDOS UTILIZADOS EN LOS PATRONES DE LA PCR CUANTITATIVA	36
3.4.1. <i>Preparación del inserto y ligación</i>	37
3.4.2. <i>Transformación de bacterias competentes</i>	37
3.4.3. <i>Obtención de ADN plasmídico a gran escala</i>	38
3.5. AISLAMIENTO DE CLONES BIOLÓGICOS RECOMBINANTES	38
3.6. SECUENCIACION	38
3.7. ANÁLISIS DE SECUENCIAS	39
3.8. ANÁLISIS FILOGENÉTICO	39
3.9. ANÁLISIS DE RECOMBINACIÓN	39
3.10. MUTAGÉNESIS	40
3.11. TRANSFECCIONES	41
<b>4. RESULTADOS</b>	<b>42</b>
4.1. PUESTA A PUNTO DE LAS TÉCNICAS UTILIZADAS EN EL ESTUDIO DE LA	43

RECOMBINACIÓN IN VITRO	
4.1.1. Aislamiento e identificación de los clones biológicos recombinantes	44
4.1.2. Detección y cuantificación de las diferentes especies provirales presentes en el ADN resultante de un cultivo celular en medio líquido	46
4.1.3. Cuantificación de formas recombinantes por PCR	48
4.2. ESTUDIO CUALITATIVO DE LAS VARIABLES QUE CONDICIONAN LA RECOMBINACIÓN IN VITRO	51
4.2.1. Caracterización biológica de los virus utilizados en las infecciones dobles	51
4.2.2. Influencia de la multiplicidad de infección	52
4.2.3. Variaciones del nivel de recombinación con el tiempo	53
4.2.4. La distancia genética	53
4.3. ESTUDIO CUANTITATIVO DE LA RECOMBINACIÓN EN CUATRO INFECCIONES DOBLES IN VITRO	55
4.3.1. Relación entre la distancia genética que separa los virus parentales y la proporción de virus recombinantes que aparecen en el cultivo	56
4.3.2. Evolución de la proporción de virus recombinantes con el tiempo	58
4.4. CONSECUENCIAS EVOLUTIVAS INTRODUCIDAS POR LA RECOMBINACIÓN EN UNA POBLACIÓN FORMADA INICIALMENTE POR DOS VIRUS DIFERENTES	60
4.4.1. Competición entre virus recombinantes y parentales	60
4.4.2. Selección positiva de clones recombinantes o numerosos eventos de recombinación. Contribución de estos dos fenómenos al aumento de la proporción de recombinantes con el tiempo.	61
4.4.2.1. Análisis filogenético de clones biológicos recombinantes aislados a partir del día 9 y 22	61
4.4.2.2. Análisis de recombinación de los clones biológicos	65
4.5. ESTUDIO SOBRE LA NATURALEZA DIPLOIDE O PSEUDODIPLOIDE DEL VIH-1	72
4.5.1. Análisis del proceso de formación de una placa	73
4.5.2. Ensayo de co-infección o infección doble	74
4.5.3. Análisis de la progenie de las placas que contienen dos genomas provirales diferentes	74
4.5.4. Modelo experimental para el estudio de la transmisión de información genética durante el proceso de transcripción inversa	75
4.5.5. Experimento de cotransfección	76
4.5.6. Posibles procesos responsables de la formación de las placas con tres bandas	79
4.5.7. Análisis de la progenie de las placas de tres bandas	79
<b>5. DISCUSIÓN</b>	<b>81</b>
5.1. ANÁLISIS CUANTITATIVO DE LA RECOMBINACIÓN SOBRE ADN PROVIRAL	83
5.2. EL MÉTODO DE CUANTIFICACIÓN	84
5.3. RELACIÓN ENTRE DISTANCIA GENÉTICA Y RECOMBINACIÓN	85
5.4. AUMENTO PROGRESIVO DE LA POBLACIÓN DE RECOMBINANTES EN UN CULTIVO	86
5.5. DINÁMICA DE IMPOSICIÓN DE LOS VIRUS RECOMBINANTES EN LA POBLACIÓN	87
5.6. SELECCIÓN DE PROTEINAS	87
5.7. ACERCA DE LOS VIRUS DEFECTIVOS	88
<b>6. CONCLUSIONES</b>	<b>94</b>
<b>7. BIBLIOGRAFÍA</b>	<b>97</b>
<b>8. ANEXO</b>	<b>110</b>

# *1. INTRODUCCIÓN*

En los últimos veinte años se ha dedicado un gran esfuerzo en la lucha contra el Síndrome de Inmunodeficiencia adquirida (SIDA), enfermedad causada por el virus de la inmunodeficiencia humana (VIH), que ha dado lugar a una pandemia de dimensiones inesperadas. A pesar de los enormes esfuerzos realizados en su investigación, actualmente no existe aún un tratamiento definitivo, ni una vacuna que pueda proteger contra la enfermedad.

La combinación de varios antirretrovirales contra el virus en la terapia múltiple, siendo la mejor opción contra el virus en la actualidad, no es capaz de acabar con el principal problema de todas las terapias ensayadas: la aparición de cepas resistentes.

Parece necesario un mayor esfuerzo en el estudio de los aspectos relacionados con los mecanismos que posee el virus para modificar con rapidez la información genética que transmite.

Los principales mecanismos que utiliza el virus para generar variabilidad son la mutación y la recombinación. Aunque las tasas de mutación de los retrovirus sean altas gracias a la capacidad de error de la transcriptasa inversa, la velocidad evolutiva que proporciona la recombinación empieza a ser considerada como el factor evolutivo más determinante en la pandemia (McCutchan *y cols.*, 2000). La última estimación de la frecuencia de recombinación es del orden de 3 por genoma y ciclo de replicación (Jetzt *y cols.*, 2000), de manera que supera varias veces la frecuencia de mutación, estimada en 0,25 por genoma y ciclo (Mansky y Temin, 1995).

La recombinación viral supone el intercambio de grandes fragmentos de información genética entre dos genomas diferentes, por lo que se puede considerar una forma de reproducción sexual.

Existe un gran debate sobre las causas que han provocado que el sexo, un proceso biológicamente muy costoso frente a la reproducción asexual, haya sido adoptado por la mayor parte de las especies (Arjan *y cols.*, 1999). Una de las teorías más aceptadas, explica su alta prevalencia al considerar el sexo como principal responsable de la reparación de la carga genética (Kondrashov, 1988) (Chao *y cols.*, 1992).

Aunque existen múltiples trabajos que han conseguido aislar virus recombinantes entre subtipos diferentes a partir de pacientes, los estudios *in vitro* se han enfocado más

hacia los mecanismos del proceso de recombinación que hacia el análisis de los cambios que la recombinación introduce en las poblaciones sobre las que actúa, o hacia los factores que pueden favorecer o limitar este proceso.

Este trabajo está dedicado al estudio de los cambios que introduce la recombinación en una población viral generada a partir de una infección *in vitro* con dos cepas diferentes de VIH-1.

## **1.1. EL VIRUS DE LA INMUNODEFICIENCIA HUMANA .**

### **1.1.1. El virus de la inmunodeficiencia humana.**

El VIH es el agente causal del síndrome de inmunodeficiencia adquirida (SIDA). Esta enfermedad se describió por primera vez en 1981 en varones jóvenes homosexuales que presentaban un cuadro de inmunodeficiencia caracterizado por la disminución de linfocitos T CD4+. Estudios epidemiológicos posteriores concretaron las vías de transmisión: sangre, productos derivados de ella, contacto sexual, intercambio de jeringuillas entre consumidores de drogas por vía intravenosa, y verticalmente de madre a hijo (Levy, 1993). En 1993, investigadores del Instituto Pasteur de París aislaron un retrovirus a partir de nódulos linfáticos de un paciente, y lo denominaron virus de la linfadenopatía (LAV). Este virus era capaz de replicar y causar efectos citopáticos en cultivos de células mononucleares de sangre periférica (CMSP) (Barré-Sinoussi y cols., 1983) Poco después, Gallo y cols. (Popovic y cols., 1984) y Levy (Levy y cols., 1984) aislaron un retrovirus a partir de muestras de pacientes con SIDA, al que denominaron VLTH-III. Finalmente se llegó a la conclusión de que estos tres virus eran el mismo, y se acordó llamarlo virus de la inmunodeficiencia humana (VIH).

### **1.1.2. Retrovirus**

La familia Retroviridae engloba un conjunto de virus con ARN como material genético, y una característica específica fundamental: poseen un mecanismo para integrar establemente su información genética dentro del genoma de la célula que infectan. Para lo cual llevan incorporado un conjunto de enzimas que les permite traducir su ARN genómico en ADN de doble cadena, e incorporarlo al ADN celular (Temin y Mizutani, 1970).

La información genética se organiza de una forma común para toda la familia Retroviridae: tres genes denominados gag, pol y env, codifican las proteínas estructurales, enzimáticas y de la envuelta, respectivamente. Estos tres genes están flanqueados por dos

secuencias iguales de varios cientos de nucleótidos, llamadas repeticiones largas terminales o LTRs.

El VIH-1 pertenece al género lentivirus dentro de la familia Retroviridae, que incluye otros seis géneros. Ver tabla 1.1 (Murphy y cols., 1995).

Un carácter diferenciador de este género lentivirus es que han desarrollado una estrategia que les permite transportar su genoma en forma de ADN a través de la membrana nuclear de la célula huésped. Esta característica es responsable de su capacidad única para replicar eficazmente en células que no se están dividiendo (Zennou y cols., 2000).

### 1.1.3. Estructura del virión y organización genómica

El virión consiste en una partícula esférica, de aproximadamente 100 nm de diámetro, constituida por tres capas concéntricas (Fig. 1.1). La capa interna (nucleoide o core), constituida por la proteína de la cápside (p24), tiene forma de cono truncado y contiene dos copias del ARN genómico formando un complejo ribonucleoproteico con las enzimas y las proteínas de la nucleocápside. La capa intermedia está constituida por la proteína matriz (p17). La capa más externa está constituida por una membrana lipídica derivada de la célula infectada, donde se insertan espículas constituidas por las glicoproteínas de la envuelta, gp120 y gp 41. En la envuelta existe, además, una gran concentración de proteínas celulares íntimamente asociadas como la  $\beta$ -2-microglobulina y las cadenas  $\alpha$  y  $\beta$  de los antígenos HLA DR, en una concentración mayor que la de gp120. Esta riqueza de proteínas celulares en la envuelta viral condiciona aspectos de la patogenia.

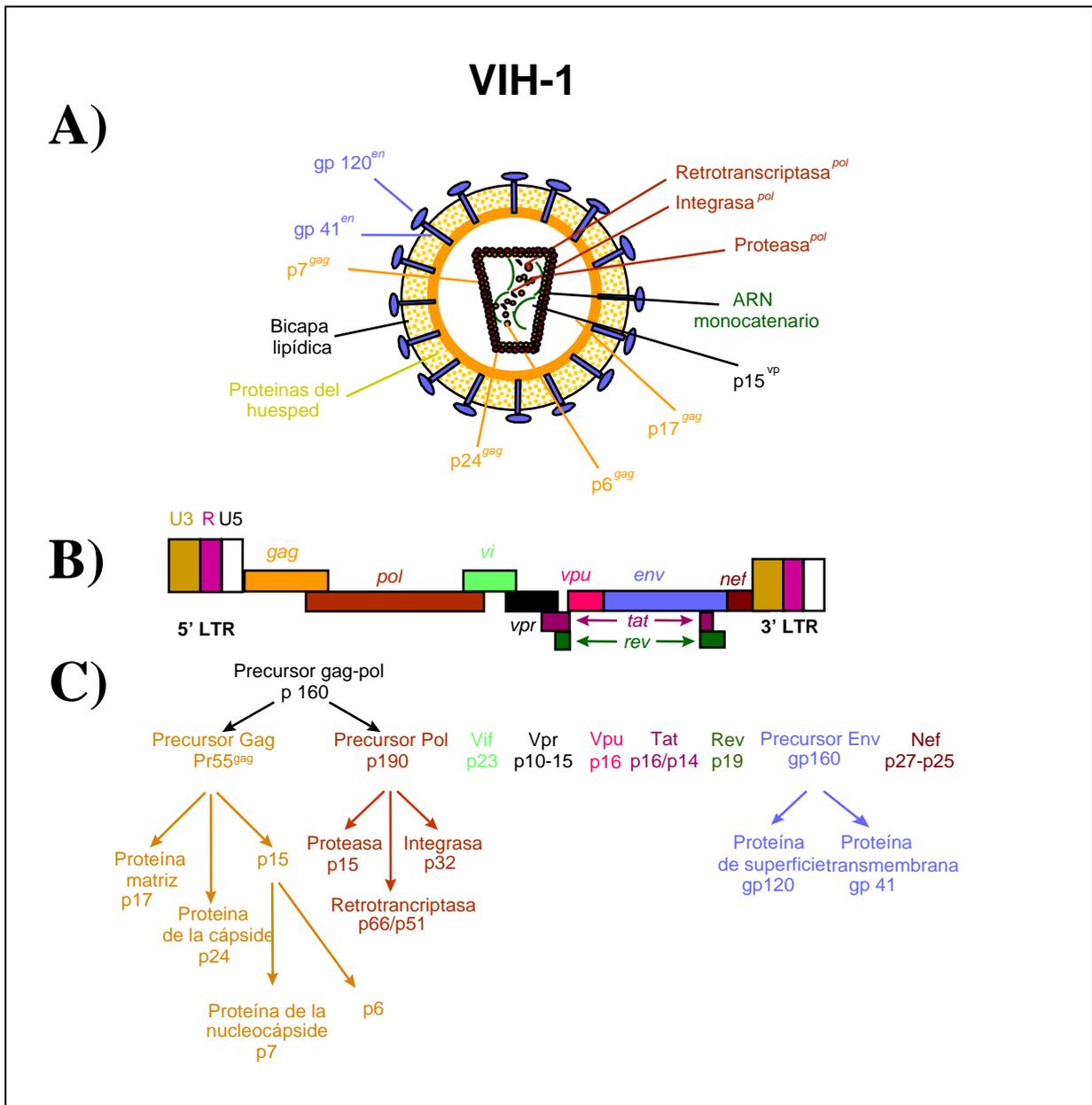
El genoma está constituido por dos copias de ARN de cadena sencilla y polaridad positiva (Hu y Temin, 1990c) de aproximadamente 9.8 Kb. Las proteínas del VIH-1 son traducidas a partir de 10 fragmentos de lectura abierta y, posteriormente, son procesadas por proteasas víricas o celulares, generándose un total de 15 proteínas víricas (Fig. 1.1).

El VIH-1 presenta tres genes estructurales: *gag*, *pol* y *env*. Las regiones *gag* y *pol* del genoma se traducen en un precursor gag-pol (p160), procesado por la proteasa viral en un precursor de gag (Pr55<sup>gag</sup>) que da origen a las cuatro proteínas que contribuyen a la formación del virión: la proteína matriz (MA, **p17**), la proteína de la cápside (CA, **p24**), y un precursor (p15) de las proteínas de la nucleocápside (NC, **p7**) y **p6** (Göttlinger y cols.,

1989). La proteína MA procede del extremo N-terminal miristilado de Pr55<sup>gag</sup>. La mayoría de las moléculas MA están asociadas a la cara interna de la bicapa lipídica, estabilizando la partícula, y un pequeño porcentaje se une a la integrasa en el complejo de preintegración, facilitando la localización nuclear del ácido nucleico viral en ausencia de mitosis. Este fenómeno le permite al VIH infectar células que no se están dividiendo, una característica inusual para un retrovirus (Lewis y cols., 1992). Mutaciones que impiden la miristilación interfieren en la asociación estable de esta proteína a la membrana, así como en la infecciosidad del virión (Bryant y Ratner, 1990; Göttlinger y cols., 1989). La proteína de la cápside forma el core de las partículas virales y es el antígeno mayoritario. Su dominio N-terminal, comprendiendo los residuos 1 al 45, contiene funciones en la maduración del virión y en la incorporación de la proteína celular cyclophilina A (CypA), la cual juega un papel importante en los primeros pasos post-infección. El dominio c-terminal (residuos 151-231) incluye la secuencia encargada de la dimerización CA-CA, e interviene en las interacciones Gag-Gag. La región NC de *gag* juega un papel fundamental en varios pasos del ciclo viral: dimerización del ARN, ensamblaje de proteínas gag, encapsidación del ARN y en la síntesis del ADN proviral por la RT. La región p6 media las interacciones entre Pr55<sup>gag</sup> y la proteína vpr, permitiendo la incorporación de vpr a los viriones. Estas proteínas son vitales para la integridad y estructura del virión (Gelderblom y cols., 1989).

El gen *pol* codifica las enzimas víricas proteasa (**p15**), retrotranscriptasa (**RT**) (**p66/p51**) e integrasa (**p32**). La proteasa cataliza el procesamiento postraduccional de los precursores a las proteínas estructurales y los enzimas. La RT presenta dos dominios, uno con actividad polimerasa ARN y ADN dependiente (Goff, 1990) y otro con actividad endonucleolítica de ribonucleasa H (RNasaH).

La integrasa presenta las actividades enzimáticas requeridas para catalizar la reacción de integración del ADN viral en el cromosoma de la célula huésped. Mutaciones que afectan a su estructura bloquean la replicación viral afectando a la integración pero no a la síntesis del ADN. Usando IN purificada a partir de viriones, o preparada por sistemas de expresión, se vió que poseía dos funciones enzimáticas: Primero reconoce y se une a ADN de doble cadena, en la secuencia ATT en el extremo de la LTR, catalizando la separación de dos bases en el extremo 3'; posteriormente une este sustrato por el grupo OH 3' libre al extremo 5' del ADN "diana", generando un provirus integrado flanqueado por dos secuencias repetidas de 5 pares de bases (Frankel y Young, 1998).



**Fig 1.1.** Estructura y organización genómica del VIH-1. A) Representación esquemática de un virión. B) Organización genómica y C) Procesamiento de las proteínas víricas.

Las proteínas codificadas por el gen *env* se originan a partir de un precursor (gp160), que es modificado posteriormente por glicosilación. Tras el procesamiento proteolítico por enzimas celulares se originan las dos proteínas de la envuelta, la proteína de superficie (SU, **gp120**) y la proteína transmembrana (TM, **gp41**). Estas dos proteínas se mantienen

asociadas y ancladas a la membrana celular por un dominio hidrofóbico cercano al extremo C-terminal. La delección de esta secuencia hidrofóbica hace que estas proteínas sean liberadas en su forma soluble. La proteína de superficie contiene los sitios de unión a los receptores celulares y los epítomos principales de neutralización por anticuerpos. La proteína transmembrana es la responsable de la fusión de la envoltura del virus con la superficie de la célula, a través de una región hidrofóbica amino terminal, una vez que se ha producido la interacción del virus con los receptores de quimioquinas.

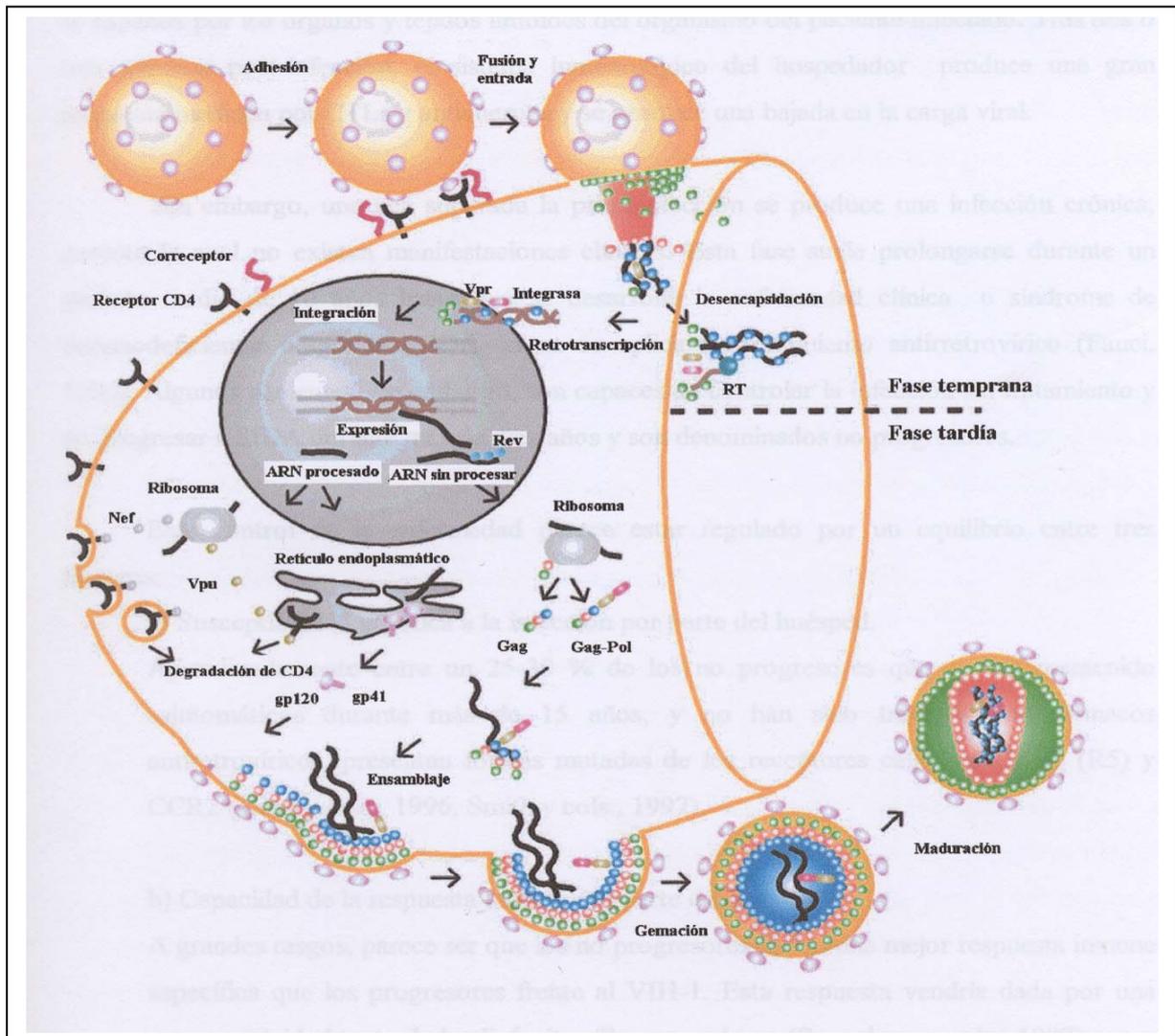
Una diferencia fundamental de los lentivirus con respecto a otros retrovirus es la existencia de varios genes reguladores. En el caso del VIH las proteínas reguladoras son *tat*, *rev*, *nef*, *vpr*, *vif* y *vpu*. La proteína **tat (p16/p14)** se une a la región TAR, localizada en la región R de la LTR del ADN viral (Feng y Holland, 1988; Roy y cols., 1990; Shibuya y cols., 1992), y, junto con otros transactivadores de origen celular (como SP1 y NFκβ), activa el inicio de la transcripción y la elongación de los transcritos formados durante la transcripción temprana (Gaynor, 1995; Sodroski y cols., 1985). **Rev (p19)** interacciona con su secuencia diana “Rev response element” (RRE), localizada en la región *env*, y regula el transporte de los ARN mensajeros (ARNm) al citoplasma, así como su procesamiento (Malim y cols., 1989). Los productos de los genes *tat* y *rev* requieren localización nuclear para ejercer su función (Ruben y cols., 1989). **Nef (p27-p25)** es una proteína miristilada. Nef tiene múltiples acciones: disminución de la expresión de CD4 en la superficie celular (Garcia y Miller, 1992), perturbación de la activación de células T y estimulación de la infectividad (Miller y cols., 1994; Spina y cols., 1994). Se la puede considerar un factor de virulencia en la patogenia de la infección por VIH (Miller y cols., 1995), ya que se han encontrado virus con delecciones en *nef* (Deacon y cols., 1995; Kirchhoff y cols., 1995) o con mutaciones puntuales (Premkumar y cols., 1996) en pacientes no progresores (LTS). La proteína **Vpr (p10-15)** es incorporada al virión por la proteína p6 y facilita la localización nuclear del complejo preintegración (Heinzinger y cols., 1994) en ausencia de mitosis (Heinzinger y cols., 1994), bloquea a las células en fase G2 del ciclo celular (Bartz y cols., 1996) y, además, es un transactivador débil. **Vif (p23)** está asociada a la infectividad y **vpu (p16)** facilita la salida de los viriones de la célula (Klimkait y cols., 1990).

El genoma viral, en su forma de provirus, se encuentra flanqueado por secuencias denominadas LTR (repeticiones terminales largas) que le permiten la replicación y la

integración en el genoma celular. Las LTRs contienen sitios de unión para proteínas celulares que activan la transcripción y están, a su vez, bajo el control de señales virales.

#### 1.1.4. Ciclo biológico.

El ciclo biológico del VIH-1 puede resumirse brevemente en los siguientes pasos, que se esquematizan en la figura 1.2.



**Fig 1.2.** Ciclo replicativo del VIH-1. Tomado de Turner y Summers, 1999, modificado por V.M. Sánchez Merino 2001.

#### Unión del virión a receptores específicos de la superficie celular.

La entrada del virus en la célula se produce a través de la interacción de la gp120 con un receptor y un correceptor celulares. El receptor es la molécula CD4, localizada en la superficie de los linfocitos T colaboradores y en células de la estirpe monocito-macrófago.

Los correceptores que utiliza el virus son receptores de quimoquinas: CXCR4, que es el receptor de la quimoquina SDF-1, y el CCR5, ò receptor fusina que es el receptor de las quimoquinas RANTES, MIP-1 $\alpha$  y MIP-1 $\beta$  (Fauci, 1996). (Zhang y cols., 1997)

El tropismo de los virus está relacionado con el uso de co-receptores: los virus M-trópicos (que replican mejor en líneas celulares monocito-macrófago) usan el receptor CCR5, mientras que los virus T-trópicos (que presentan tropismo por líneas celulares T) pueden usar tanto el CCR5 como el CXCR4 (Doranz y cols., 1996) (Livingstone y cols., 1996) (Dittmar y cols., 1997).

#### **Penetración del core en el interior de la célula.**

La interacción entre la gp120 y sus receptores, conduce a un proceso de fusión entre las membranas viral y celular, que permite la internalización de la nucleocápside, proceso mediado por la proteína gp41 (Coffin y cols., 1997).

#### **Transcripción inversa del ARN viral en ADN.**

El momento del comienzo de este proceso no es bien conocido, puede empezar antes de la penetración del virión en la célula (Zhang y Temin, 1993), concluyendo antes del transporte al núcleo tras la maduración del complejo de transcripción inversa (RTC) en forma de complejo de preintegración (PIC) (Zennou y cols., 2000). Así, mediante un complejo proceso, la RT convierte el ARN vírico en ADN de doble hebra (Fields y Knipe, 1990) y añade dos LTR que no posee el ARN del virión..

**Integración.** Una vez transportado al núcleo, el ADN proviral se integra en el genoma celular mediante la acción de la integrasa viral, constituyendo el provirus integrado. Los procesos de retrotranscripción e integración dependen también de factores celulares inducidos durante la activación celular. Así, se ha observado retrotranscripción incompleta en linfocitos CD4 en reposo y ADN proviral no integrado, pero susceptible de integración si las células se activan.

Una vez integrado, el ADN puede permanecer latente, replicarse de forma controlada o experimentar una replicación masiva, con el consiguiente efecto citopático sobre la célula infectada (Li y cols., 2000).

#### **Procesamiento de los transcritos a ARN genómico o ARN mensajero.**

La iniciación de la transcripción supone la síntesis de un único precursor ARN y depende tanto de factores celulares, como de antígenos, citoquinas o agentes mitógenos, que interaccionan con secuencias localizadas en la LTR. Además, la proteína viral Tat aumenta de 10<sup>2</sup>-10<sup>3</sup> veces la tasa de transcripción (Coffin y cols., 1997). Una fracción de

este ARN es reservada para ARN genómico y la otra es empleada como ARNm y procesada en transcritos de distintos tamaños. El ARNm sufre en 5' un proceso de "capping" y en el extremo 3' es poliadenilado seguido de un procesamiento (*splicing*), para ser transportado después al citosol. Los genes son expresados como una colección de poliproteínas precursoras.

Las proteínas y poliproteínas Gag y Gag-pol, una vez sintetizadas en el citoplasma, se localizan en la membrana celular. Las glicoproteínas de la envuelta se sintetizan en el retículo endoplásmico y se transportan hasta la membrana plasmática vía ruta secretora.)

### **Empaquetamiento y gemación (budding) de viriones.**

Las poliproteínas precursoras *gag* y *gag-pol* interactúan con dos copias de ARN y se unen formando estructuras visibles al microscopio electrónico como partículas densas en el interior de la membrana celular. Estas partículas inducen la curvatura de la membrana formando su envuelta al mismo tiempo que se incorporan las glicoproteínas de *env*, completando el proceso el desprendimiento de las partículas a partir de la membrana (proceso mediado por la proteína p6 y Ubiquitina) (Accola y cols., 2000) [Huang, 1995 #5211 (Gottlinger y cols., 1991)].

El primer paso implica contacto entre las dos cadenas de ARN y el precursor Pr55<sup>Gag</sup>. Además, la proteína Vif se une específicamente al ARN genómico (Zhang y cols., 2000a).

Numerosos estudios señalan a la proteína de la nucleocápside (NC) como fundamental en el ensamblaje de partículas y para una correcta morfogénesis.

Además, es imprescindible para este proceso la presencia en las cadenas de ARN de un conjunto de estructuras en horquilla conocidas como señal de empaquetamiento ( $\Psi$ ), que intervienen también en otras funciones virales esenciales, como activación transcripcional, y transcripción inversa.

### **Procesamiento proteolítico de las proteínas de la cápside.(Maduración).**

Inmediatamente después de la salida de los viriones por gemación, la proteasa rompe los precursores *gag* y *gag-pol* produciendo las proteínas MA, CA y NC, las cuales se reorganizan generando un virión maduro e infeccioso, capaz de iniciar un nuevo ciclo de replicación (Ganser y cols., 1999) y caracterizado por la presencia de la estructura tronco-cónica de la nucleocápside.

### 1.1.5. Historia natural y patogénesis de la infección

La cuantificación de los niveles plasmáticos de ARN viral permitió comprobar que la replicación del virus es continua durante toda la infección (Coombs y cols., 1989) y conocer mejor la patogenia de la enfermedad, cuya evolución puede dividirse en varias fases:

**Fase aguda** o primoinfección, de varias semanas de duración, durante la cual se alcanzan altos niveles de ARN en plasma ( $\geq 10^6$  copias/ml) acompañados de una disminución del número de células CD4+. Tras una fuerte respuesta citotóxica y posterior respuesta humoral, los niveles de ARN caen drásticamente y se restablece el número de células CD4. Los niveles de ARN víricos que se alcanzan en ese momento (*set point*) parece que condicionan la progresión de la enfermedad (Mellors y cols., 1995).

**Fase de replicación crónica.** Es de una duración variable de varios años, con replicación viral activa y un equilibrio dinámico entre el virus y el sistema inmune. Durante esta fase el nivel de carga viral plasmática puede ser bajo o indetectable y, desde el punto de vista clínico, se corresponde con la fase asintomática, durante la cual los niveles de CD4 se mantienen estables en general y por encima de 500 CD4/ $\mu$ l.

**Fase sintomática,** que corresponde con la fase clínica de la enfermedad, denominado el síndrome de inmunodeficiencia adquirida (SIDA), caracterizado por anomalías inmunológicas frecuentemente acompañadas de enfermedades oportunistas, desórdenes neurológicos y formas inusuales de cáncer. En esta fase se observa una elevación de la carga viral que precede al descenso en el número de linfocitos CD4.

En relación con el tropismo, existen virus que crecen en células de las líneas T y generalmente poseen fenotipo inductor de sincitios (SI) y otros con tropismo por macrófagos, que generalmente no inducen sincitios en células T (NSI). La diferente cinética de replicación ha sido también empleada para distinguir cepas de virus rápidos/altos o *rapid/high* frente a los lentos/bajos o *slow/low* y que coinciden con la clasificación anterior (Schwartz y cols., 1989).

El VIH tiene dos dianas principales: los linfocitos CD4 y los macrófagos tisulares. Se estima que se generan en torno a  $10^9$ - $10^{10}$  partículas víricas al día, a la vez que se producen y destruyen  $10^9$  células T CD4+ cada día y que la vida media del virus y de las células infectadas productivamente es aproximadamente de dos días (Ho y cols., 1995) (Wei y cols., 1995)

## 1.2. DIVERSIDAD GENÉTICA DEL VIH-1

La diversidad genética del VIH-1 se describió por primera vez en 1984, por mapeo de sitios de restricción en el ADN proviral de distintos aislados (Hahn y *cols.*). Después se comprobó que también existían diferencias entre aislados sucesivos de un mismo paciente, así como variantes con distinto tropismo a partir de un mismo aislado (Fisher y *cols.*, 1988). También se ha visto que existen diferencias entre virus que proceden de distintos órganos, y entre distintos centros germinales dentro de un bazo (Cheynier y *cols.*, 1994).

Estudios moleculares han llevado a entender el VIH-1 como una cuasiespecie, es decir, una mezcla heterogénea de variantes relacionados, pero genótipicamente distinguibles entre sí, al igual que ocurre con otros virus ARN (Domingo y Holland, 1997).

Mediante análisis genético, se han identificado cinco linajes distintos de lentivirus de primates (McCutchan y *cols.*, 1996) (Gao y *cols.*, 1994a). VIH-1 y VIH-2 pertenecen a dos de estas líneas. En el linaje del VIH-1, la gran mayoría de aislados se han caracterizado como pertenecientes a subtipos del grupo M, aunque se ha descrito otro grupo O altamente divergente. Dentro del grupo M, están reconocidos nueve subtipos, designados con las letras A-D, F-H, J y K. Los subtipos forman grupos aproximadamente equidistantes en los árboles filogenéticos, estando separados por una distancia del 25-35% en aminoácidos en *env*. Dentro de los subtipos A y F se diferencian dos sub-subtipos A1 y A2, y F1 y F2, que están más cerradamente realacionados entre sí que con el resto de subtipos.

El estudio comparativo de secuencias procedentes de África, Asia y América, puso de manifiesto que un alto porcentaje de esas cepas, eran en realidad genomas mosaico compuestos por dos o más subtipos diferentes, que a menudo reflejaban la mezcla de subtipos circulante en una población o región (Belda y *cols.*, 1998) (Bobkov y *cols.*, 1996); (Carr y *cols.*, 1996); (Choi y *cols.*, 1997); (Diaz y *cols.*, 1995); (Gao y *cols.*, 1996); (Leitner y *cols.*, 1995); (Sabino y *cols.*, 1994); (Takehisa y *cols.*, 1998) (Takehisa y *cols.*, 1997). En la actualidad se considera que más del 20% de los aislados naturales de VIH-1 son virus recombinantes (Gao y *cols.*, 1994b) ; (Kampinga y *cols.*, 1997) ; Korber y *cols.*, 1997; (Robertson y *cols.*, 1995b). La recombinación entre virus pertenecientes al mismo subtipo genético ha sido mucho menos descrita (Buttò y *cols.*, 1997) (Diaz y *cols.*, 1995); (Pedroza Martins y *cols.*, 1992). (Sala y *cols.*, 1995) (Wei y Fultz, 1998) (Zhu y *cols.*, 1995) (Casado y *cols.*, 2001), debido probablemente a la dificultad de su detección.

### 1.2.1. Clasificación de subtipos.

La acumulación de secuencias del genoma completo (“full-genome”) y la investigación de la diversidad en diferentes regiones del mundo ha generado un nuevo concepto de la pandemia (McCutchan *y cols.*, 1998) (Kuiken *y cols.*, 1999) (Kuiken *y cols.*, 1999) que reconoce a los subtipos y a los recombinantes intersubtipo la misma relevancia epidemiológica. Así, se ha propuesto una nueva nomenclatura para VIH-1 (Robertson *y cols.*, 2000) que recomienda:

- 1.- La identificación de aquellos recombinantes de VIH-1 que estén ampliamente extendidas en una población como “formas recombinantes circulantes” (CRF).
- 2.- La reclasificación de subtipos I y E como CRF.
- 3.- La validación de sub-subtipos geográficos.
- 4.- El requerimiento de al menos tres genomas completos para la identificación de nuevos subtipos o CRFs.

**Tabla 1.1.** *Distribución de subtipos y CRF.*

- **HIV-1.** - *Grupo N*

- *Grupo O*

- *Grupo M* - Subtipos: A, A2, B, C, D, F1, F2, G, H, J, K.

-*Formas recombinantes circulantes:* CRF01\_AE, CRF02\_AG, CRF03\_AB, CRF04\_cpx, CRF05\_DF, CRF06\_cpx, CRF07\_BC, CRF08\_BC, CRF09\_cpx, CRF10\_CD, CRF11\_cpx, CRF12\_BF.

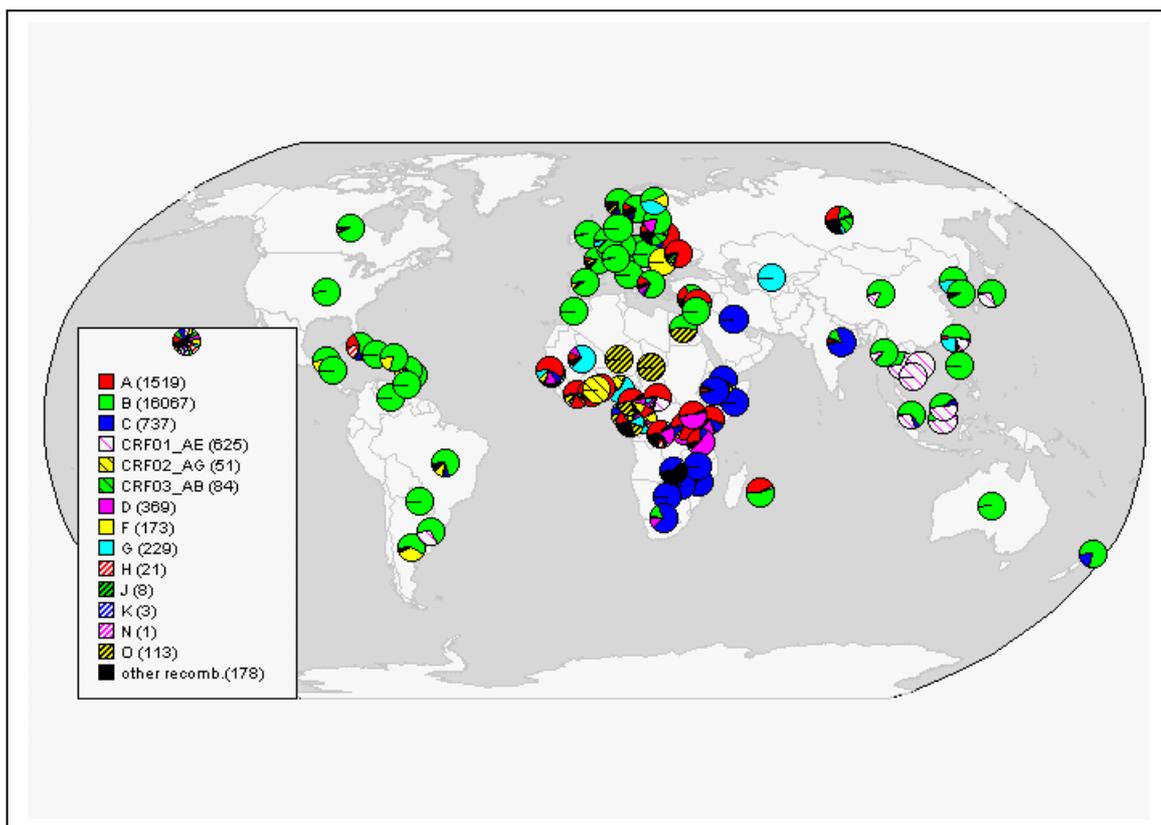
Esta elevada diversidad genética es el resultado de un conjunto complejo de factores. El primero de ellos que contribuyó a la generación de esta diversidad fue la introducción de virus genéticamente muy distintos de primates a humanos (Korber *y cols.*, 2000). Estas transmisiones zoonóticas tuvieron dos orígenes diferentes. El VIH-1 procedería de la transmisión de un VIS del chimpancé de la subespecie *Pan Troglodytes Troglodytes* al hombre (Gao *y cols.*, 1999), mientras que el VIH-2 tendría su origen en un virus del mangabey tizado (*Cercocebus atys*) (Hirsch *y cols.*, 1998). Desde esas primeras transmisiones a los humanos, la diversidad ha ido aumentando como consecuencia de la acumulación de mutaciones, (debido a la alta tasa de error de la transcriptasa inversa (Preston *y cols.*, 1988), al rápido recambio de viriones que se produce en los individuos infectados (Ho *y cols.*, 1995), a la recombinación (Robertson *y cols.*, 1995a), a la presión

selectiva que sufren los virus como consecuencia de la respuesta del sistema inmunitario a la infección (Haynes y cols., 1996), a efectos fundadores y a otros aspectos epidemiológicos.

### **1.2.2. Distribución geográfica de los subtipos y formas recombinantes circulantes del VIH.**

Las formas virales predominantes en la epidemia global pertenecen al subtipo C, seguido del B, y de las formas recombinantes CRF01-AE y CRF02-AG (McCutchan, 2000) (Hoelscher y cols., 2001). La mayor diversidad genética de HIV-1 ha sido encontrada en África, especialmente en África Central, donde los subtipos A, C y CRF02-AG son los más comunes, pero todos los tipos y subtipos están presentes en este continente, que fue la fuente de la epidemia (Vidal y cols., 2000). En el sur y el este predomina el subtipo C, mientras en el oeste es el recombinante CRF02-AG (Montavon y cols., 2000). En Europa, Norteamérica y Australia, el subtipo B es el más extendido, aunque se han encontrado en varios países europeos, otros subtipos del grupo M, incluso del grupo O, que además parecen ir en aumento (Peeters y Sharp, 2000). En Sudamérica predomina el subtipo B, aunque también se han encontrado el C y el F1 (Janini y cols., 1998), así como recombinantes B/F (Marquina y cols., 1996) (Thomson y cols., 2000). En Asia circulan varias formas, predominando los subtipos B, C y CRF01-AE.

La distribución global de diferentes formas de HIV-1, es un proceso dinámico. Debido al continuo intercambio humano entre diferentes partes del mundo, la probabilidad de emergencia de nuevas formas recombinantes aumenta, y la necesidad de determinar el papel que pueden jugar estas nuevas formas recombinantes en el futuro de la epidemia, es apremiante.



**Fig. 1.2.** Distribución geográfica de subtipos y formas recombinantes circulantes por el mundo. Tomado de HIV database LANL USA.

### 1.3. VARIACIÓN EN VIH

La familia retroviridae, en general se caracteriza por presentar unos niveles de variación genética notablemente altos (Hu y Temin, 1990a; Temin, 1993). Concretamente, el VIH-1, uno de los retrovirus más complejos ya que codifica genes adicionales implicados en regulación, además de los genes gag, pol y env de los retrovirus más simples, en su secuencia *env* es capaz de variar aproximadamente un 1% en un año dentro de una persona infectada (Temin, 1993).

#### 1.3.1. Generación de variabilidad.

La dinámica de poblaciones virales parece seguir los mismos principios básicos de la genética de poblaciones clásica para organismos superiores (Domingo, 1997).

La evolución de las poblaciones naturales puede considerarse un proceso en dos fases: una primera de generación de variabilidad, y una segunda fase en la que intervienen una serie de mecanismos que conducen a cambios en la frecuencia de los distintos variantes dentro de una población. En todas las poblaciones naturales tienen lugar una serie

de procesos responsables del mantenimiento del polimorfismo, y procesos que llevan a la imposición de unos variantes frente a otros, ya sea por un proceso selectivo o no. Los procesos que introducen variabilidad genotípica entre una generación y su descendencia pueden ser separados en dos grupos: a) Procesos de generación de mutantes durante la replicación viral, que están gobernados fundamentalmente por procesos estocásticos; y b) Procesos ajenos a la replicación, como son: expansión de la progenie viral de una célula al resto del organismo hospedador, la probabilidad de coinfección de una célula por dos virus, o la expansión desde un organismo a otros.

El destino de cada variante dependerá de una serie de factores resultantes de la interacción del virus con su entorno, como la eficacia biológica (fitness) relativa en el conjunto de la población total, así como de procesos de muestreo o restricciones drásticas de las poblaciones.

De la generación de mutantes es responsable la alta tasa de error de la transcriptasa inversa (mutación) y la capacidad de los retrovirus para intercambiar información genética entre las dos hebras de ARN que contiene el virus, por recombinación.

La tasa de mutación se define como la proporción de incorporaciones erróneas que ocurren durante la síntesis de un ácido nucleico, expresada como sustituciones por nucleótido y ciclo de replicación. La frecuencia de mutación se define como la proporción de mutantes en una población, y se expresa como sustituciones por nucleótido (Duarte y cols., 1992).

Mansky y Temin confirmaron que existían diferencias significativas en la tasa de mutación entre distintos retrovirus ( $10^{-6}$  para BLV-lacZ $\alpha$  y  $4 \times 10^{-5}$  para MLV-tk) (Mansky y Temin, 1994) (Mansky y Temin, 1995).

La tasa de mutación del VIH-1 ha sido evaluada en  $3 \times 10^{-5}$  (Mansky, 1996) (Mansky y Temin, 1995) (Mansky, 1998).

La tasa de recombinación homóloga retroviral fue determinada por Hu y Temin en 1990 en un trabajo con el virus de la necrosis esplénica (SNV) en un 2% por kb y ciclo de replicación (Hu y Temin, 1990b) (Hu y Temin, 1990a). Esta tasa de recombinación no es proporcional a la distancia entre los marcadores (Anderson y cols., 1998) .

Aunque la recombinación heteróloga es frecuente en otros virus ARN, en retrovirus parece ser bastante infrecuente comparada con la homóloga, de  $5 \times 10^{-5}$  por kilobase y ciclo

de replicación, es decir, de un 1 a un 0,1% de la tasa de recombinación homóloga (Zhang y Temin, 1993).

### **1.3.2. Mecanismos de cambio en la frecuencia de los variantes de una población natural.**

La composición de una cuasiespecie vírica va a ser el resultado de la generación de variantes creados mediante mutación y recombinación, y de una serie de fenómenos evolutivos que van a actuar sobre esa serie de variantes, favoreciendo a unos y perjudicando a otros. Estos procesos pueden ser clasificados como:

a) De selección, tanto negativa, que elimina o mantiene en baja proporción a las variantes que presentan una menor eficacia biológica, como positiva o darwiniana que permite la imposición de variantes víricas con mayor eficacia biológica.

b) De azar, por los que una variante de la población puede ser amplificada y llegar a ser dominante, independientemente de su eficacia biológica. Factores adicionales que pueden influir críticamente en la composición de las cuasiespecies in vivo son: la deriva genética, cuellos de botella genéticos, trinquete de Muller, transmisiones de grandes tamaños poblacionales, principio de exclusión competitiva, selección dependiente de frecuencia, hipótesis de la reina roja.

Un tema de gran controversia son las contribuciones relativas de cada uno de estos procesos en la evolución de una población.

El término velocidad de evolución se define como el grado de fijación o acumulación de mutaciones en función del tiempo.

## **1.4. RECOMBINACIÓN EN VIH**

La recombinación genética en el virus de la inmunodeficiencia humana tiene lugar durante la transcripción inversa del ARN viral a ADN proviral.

En 1964, Temin propuso la hipótesis del provirus, según la cual, virus ARN tumorales residían en células infectadas en forma de ADN (Temin, 1976). En 1970, demostró que preparaciones purificadas de partículas virales contenían una polimerasa que sintetizaba ADN a partir de ARN viral (Temin y Mizutani, 1970) (Baltimore, 1970).

### **1.4.1. La Transcriptasa Inversa.**

La transcriptasa inversa de VIH es un heterodímero formado por dos subunidades de 560 (p66), y 440 residuos (p51) respectivamente, que derivan de la poliproteína Pol. Cada una de las dos subunidades contiene un dominio polimerasa compuesto por cuatro subdominios llamados dedos, palma, pulgar y conexión (por analogía con una mano derecha), y la p66 contiene un dominio adicional RNAsa H.

Aunque tienen una secuencia aminoacídica idéntica, los subdominios polimerasa se organizan de manera diferente en las dos subunidades p51 y p66. P66 presenta una hendidura que contiene un sitio activo muy conservado en muchas polimerasas ( Asp 110, Asp 185 y Asp 186) mientras que p51 forma una estructura cerrada inactiva. (Frankel y Young, 1998) .

Se trata de una enzima multifuncional con tres actividades: DNA polimerasa dependiente de ARN, DNA polimerasa dependiente de ADN, y ribonucleasa H. Las actividades polimerasa y ribonucleasa H ocupan dominios separados no solapantes. La polimerasa ocupa dos tercios de la molécula en el lado amino terminal. La actividad ribonucleasa H, permite eliminar selectivamente el ARN molde a medida que se sintetiza la cadena de ADN a partir de una molécula híbrida ARN-ADN, actuando como exonucleasa y endonucleasa en sentido 5'-3', y normalmente no digiere moléculas de hebra simple.

Al igual que todas las DNAPolimerasas conocidas, la RT es incapaz de iniciar la síntesis de ADN de novo, para lo que requiere una molécula que sirva como iniciador. *In vitro* puede ser cualquier ARN o ADN, pero *in vivo* siempre es un ARNt. Además, *in vitro*, requiere la presencia de cationes divalentes, siendo preferido el Mg<sup>2+</sup>.

La molécula de ADN molde-iniciador se une a lo largo de los subdominios de la palma y conexión, en el hueco entre los dedos y el pulgar. El sitio activo para polimerización es en el centro de la palma, con la participación de tres residuos de ácido aspártico que coordinan un ión metálico implicado en la catálisis de la adición de nucleótidos. Además existe una secuencia asociada con el sitio catalítico (Tyr-X-Asp-Asp) que está altamente conservada en los enzimas de otros retrovirus.

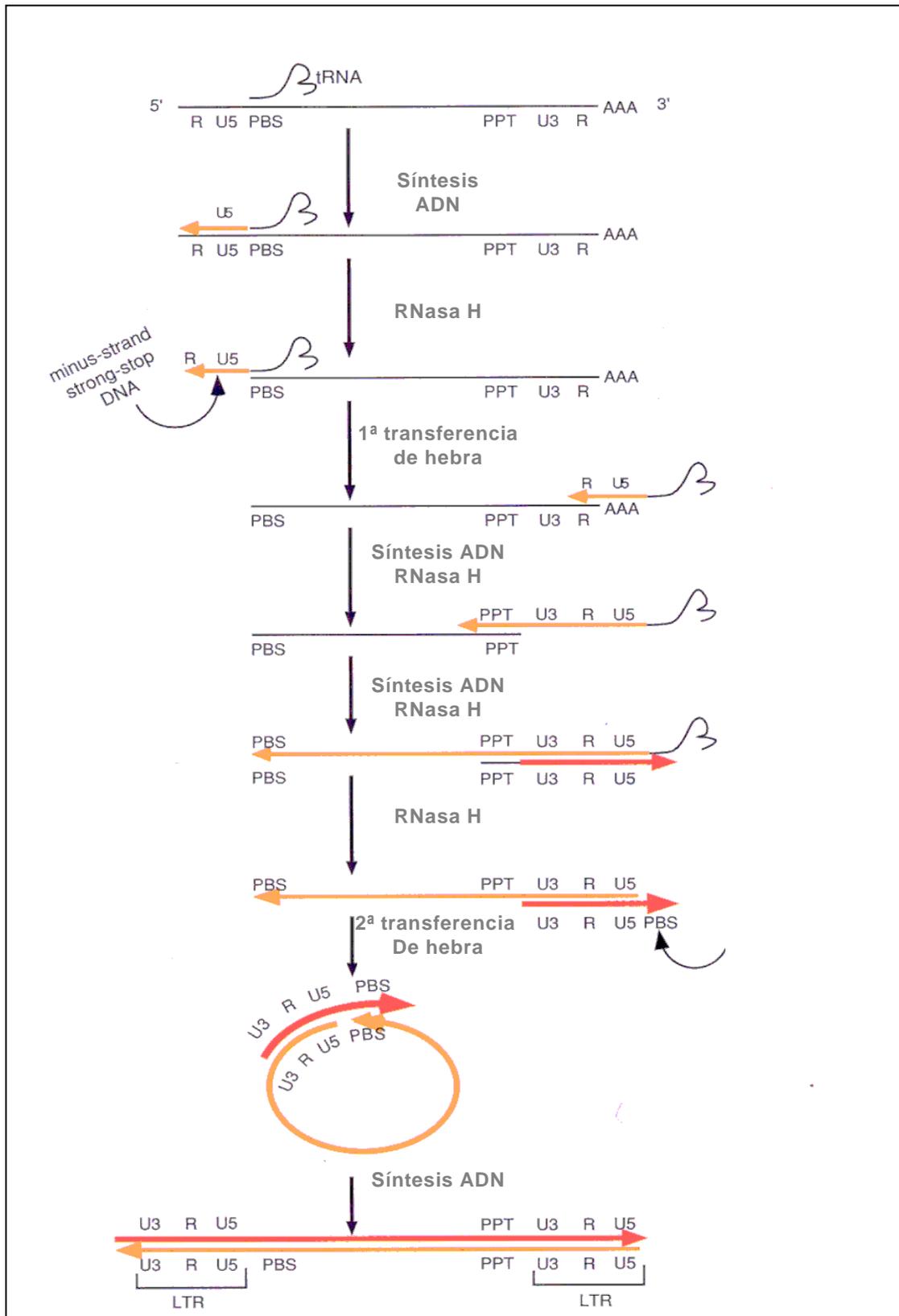
#### **1.4.2 Mecanismo de la transcripción inversa**

El modelo de transcripción inversa para VIH-1 es similar al propuesto para otros retrovirus, y fue desentrañado en sistemas *in vitro*, libres de células.

La transcripción inversa requiere dos transferencias de hebra (o saltos) para convertir el ARN viral en una molécula de ADN doble. Los requerimientos mínimos para este proceso son: dos copias de ARN simple viral, el iniciador tARN<sup>lys</sup>, y la proteína heterodimérica Transcriptasa Inversa (RT). Sin embargo, la transcripción inversa in vivo tiene lugar dentro de un complejo ribonucleoproteico que incluye también otras proteínas virales como la cápside y/o la nucleocápside, que además juega un importante papel en la dimerización del ARN y empaquetamiento de los viriones, (Farnet y Haseltine, 1991). La arquitectura de este complejo puede influir en la eficiencia del proceso. Una fracción de la RT en los viriones está asociada con la estructura nucleoproteica viral, donde parece unirse al iniciador ARNt.

La transcriptasa inversa es activada por una señal en el citoplasma celular que aún no ha sido identificada, para iniciar la síntesis a partir del iniciador tARN<sup>lys</sup> que se encuentra unido al ARN viral (Luciw y Harada, 1998) (Fields y Knipe, 1990) cerca del extremo ARN molde viral en el sitio de unión del primer (PBS). Esta interacción entre el PBS y el tARN, parece estar influenciada por la estructura secundaria de la región líder del ARN viral.

La RT extiende la primera cadena de ADN hasta el extremo 5' del genoma viral. A este fragmento de ADN de unos 100 a 150 nucleótidos se le llama “minus strand strong-stop”. Entonces la ARNasa-H media la degradación de la hebra de ARN del complejo ARN-ADN formado (Le Grice y cols., 1991) (Schatz y cols., 1990) (DeStefano y cols., 1994) (DeStefano y cols., 1991). Seguidamente, se produce la primera transferencia de hebra de la RT: este fragmento pasa a unirse al extremo 3' del ARN genómico por medio de la hibridación de las secuencias idénticas, llamadas R, repetidas en los dos extremos del genoma ARN, hibridación favorecida por la proteína NC. Además, parece que se trata de una transferencia preferentemente intramolecular (DeStefano y cols., 1991) (Hu y Temin, 1990a) (Luo y Taylor, 1990). De nuevo, esta secuencia R, sirve de iniciador de la síntesis de la cadena ADN de polaridad negativa. La actividad ARNasa H degrada la cadena molde excepto un fragmento PPT, relativamente resistente a este enzima, que ejerce de iniciador de la cadena de ADN de polaridad positiva, tomando como molde la cadena ADN negativa que se acaba de formar. Entonces se produce el segundo “salto”, del fragmento sintetizado “plus-strand-strong-stop” hasta el extremo 3' de la cadena de ADN, para completar la formación de la doble cadena ADN.



**Fig 1.4.** Esquema del mecanismo de la transcripción inversa. (Tomado de A. Telesnitsky and S. P. Goff "Retroviruses" Coffin, Hughens, Varmus 1997) ....

Los lentivirus presentan una particularidad en este paso: la síntesis de la cadena de ADN de polaridad positiva se lleva a cabo en dos fragmentos. El fragmento que se inicia en el extremo 3', acaba en la secuencia de terminación central (CTS), y el segundo comienza en una secuencia de polipurinas (cPPT) central, localizada al final de la secuencia del gen pol (justo 99 nucleótidos antes del CTS). Este particular mecanismo parece tener una importante función en la replicación del virus, ya que favorecería un pliegue en la estructura del ADN que facilita su paso al núcleo en células que no se encuentran en mitosis (Zennou y cols., 2000).

### 1.4.3. Mecanismo de la recombinación retroviral.

El requerimiento de viriones heterocigotos para detectar genomas recombinantes en las células infectadas por ellos, indica que la recombinación ocurre entre los dos ARNs empaquetados en el virión y durante el proceso de transcripción inversa (Clavel y cols., 1989) (Hu y Temin, 1990a) (Stuhlmann y Berg, 1992).

La recombinación puede potencialmente ocurrir durante la síntesis de la cadena de ADN positiva o durante la síntesis de la negativa. Existen dos modelos básicos para explicar el fenómeno de la recombinación retroviral.

El primer modelo sugiere que la hebra de ADN en crecimiento puede disociarse del ARN molde gracias a una rotura por la ribonucleasa H (RNAsa H), y entonces sería libre de unirse a una región homóloga en la otra hebra de ARN del virus. Alternativamente, una hebra de ADN en crecimiento puede ser desplazada de su unión al ARN, por otra hebra que invade el "duplex" (Diaz y cols., 2000). Ambos modos de transferencia se han demostrado *in vitro* (DeStefano y cols., 1994). La proteína de la nucleocapside parece jugar algún papel en estos procesos (Allain y cols., 1994), (DeStefano, 1995) (Peliska y cols., 1994).

El segundo modelo requiere la presencia de dos hebras de ADN de polaridad negativa sintetizadas a partir de las dos hebras de ARN del virus, para que puedan intercambiar fragmentos. Sin embargo, parece que la mayoría de las recombinaciones (98%) en un sistema *in vitro* retroviral ocurren durante la síntesis de la cadena de ADN negativa, de acuerdo con el primer modelo (Zhang y cols., 2000b).

Los requerimientos mínimos de este proceso, es decir, la transferencia de hebra entre regiones internas de dos moldes diferentes *in vitro*, son segmentos cortos de ARN de

aproximadamente 142 nucleótidos con una cierta homología de secuencia y la transcriptasa inversa viral. La depleción de la actividad RNAsa supone una disminución de la eficacia del proceso (DeStefano y cols., 1992).

Se conoce poco acerca de secuencias y estructuras que promuevan la recombinación. Estructuras en horquilla (“hairpin stem”) parecen interrumpir la síntesis por parte de la transcriptasa inversa (DeStefano y cols., 1992), sugiriendo que determinadas regiones pueden promover recombinación (Mikkelsen y cols., 2000) constituyendo regiones calientes para recombinación (“recombination hot spot”) (Zhang y Temin, 1994). En este sentido, Anders H. Lund (Lund y cols., 1999) identificó el motivo “lazo del beso” (“kissing-loop”)(que presenta una estructura secundaria implicada en la formación del dímero de ARN) como un “sitio caliente” que media recombinación en MLV.

D.P.Wooley (1998) (Wooley y cols., 1998) encontró frecuencias más altas de recombinación en las regiones inmediatas a secuencias idénticas entre dos vectores. Se han descrito sitios específicos de parada que parecen implicados también en transferencia de hebra (Kim y cols., 1997) (Wu y cols., 1996). Parece que incorporaciones de bases erróneas por parte de la transcriptasa inversa pueden promover transferencia de hebra (Diaz y cols., 2000).

Una incorporación errónea puede causar la pausa de la RT y favorecer un evento de transferencia de hebra (Perrino y cols., 1989), (Diaz y DeStefano, 1996) (Fuentes y cols., 1996).

#### **1.4.4. Modelos de recombinación retroviral**

Modelo de elección de copia o “copy-choice”. Fue primero propuesto por Weiss en 1973, y propone que la recombinación ocurre durante la síntesis de la cadena de ADN de polaridad negativa, cuando la transcripción inversa “salta” desde una cadena de ARN , para seguir copiando la información contenida en la otra (Weiss y cols., 1973). El modelo fue posteriormente modificado por Coffin (Coffin, 1979) en su formulación del modelo “forced copy-choice”, porque entiende que el salto de la RT de una cadena a la otra de ARN, es provocado por una “rotura” en la primera.(una predicción de este modelo, es que condiciones que produzcan daños en el ARN, estimularían la recombinación, pero los resultados con  $\gamma$ -irradiación fueron ambiguos (Hu y Temin, 1992). Por último, DeStefano (DeStefano, 1994) demostró que se producen “saltos” entre regiones internas de

ARN, tanto en reacciones in vitro con fragmentos de ARN, como en estudios con fragmentos de genomas virales.

El modelo de desplazamiento de banda o “strand displacement/assimilation”, surgió más tarde para explicar dos fenómenos descritos:

- a) La estructura en H encontrada por microscopía electrónica, formada por dos ADNs de doble cadena conectados por un fragmento de cadena simple (Junghans y cols., 1982).
- b) La descripción de la síntesis del ADN viral en fragmentos (Hsu y Taylor, 1982) (Boone y Skalka, 1981).

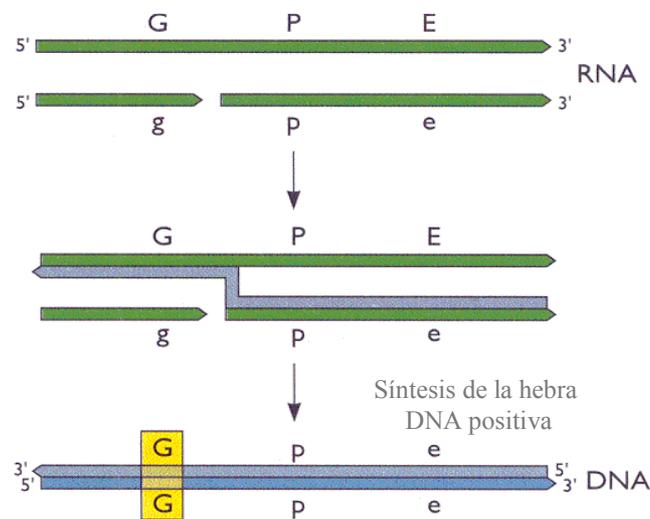
Todo ello sugería que la información genética podía ser combinada a través de fragmentos de ADN. Esto implica que la recombinación tendría lugar una vez formadas dos cadenas de polaridad positiva, cada una de ellas a partir de cada una de las cadenas de ARN contenidas en el virión.

#### **1.4.5. El proceso de transferencia de hebra.**

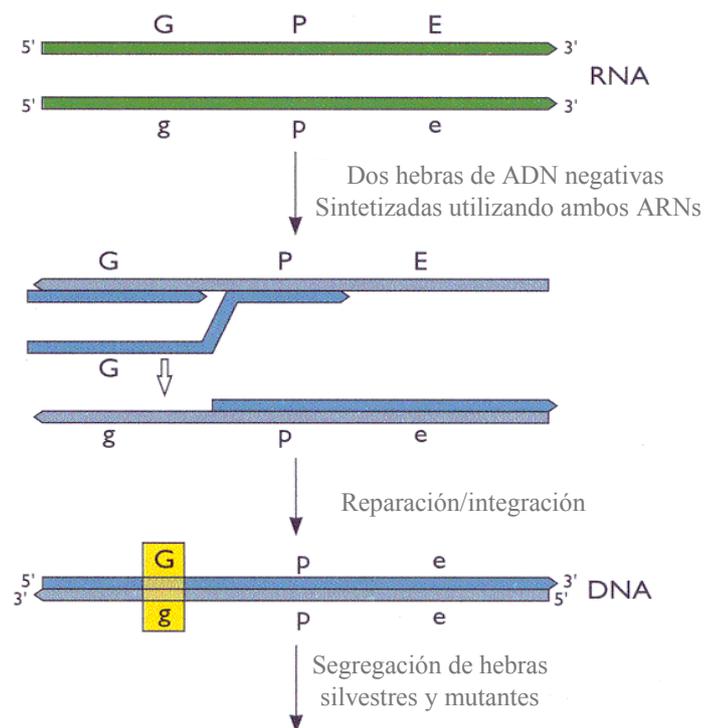
Sistemas in vitro diseñados para estudiar la recombinación, que usan ácidos nucleicos y enzimas purificados indicaron que las roturas de la RNasa H juegan un papel en la transferencia de hebra desde un ARN “donador” (Peliska y cols., 1992). Trabajos posteriores sugieren que la pausa de la transcriptasa inversa durante la síntesis de la cadena de ADN, promueve transferencia de hebra (DeStefano y cols., 1992) (DeStefano y cols., 1994). Se encontró que las posiciones de transferencia correlacionaban con secuencias y estructuras que provocan pausas. Se propuso que las paradas de la RT pueden permitir que la RNasa realice cortes en las posiciones adyacentes del ARN molde. Esta degradación del molde puede facilitar la interacción del primer con el molde aceptor (Brincat y cols., 2002).

La presencia de estructuras secundarias entre las secuencias repetidas parecen incrementar la eficiencia del proceso (Hwang y cols., 2001). La altamente estructurada horquilla TAR de HIV-1 provoca la parada de la síntesis por parte de la RT en el momento que esta alcanza la base de la horquilla (Kim y cols., 1997). Sin embargo, a pesar de los progresos en este sentido, no está clara la relación entre sitios de pausa y puntos de transferencia de hebra (Roda y cols., 2002).

### Recombinación durante la síntesis de la hebra de ADN negativa



### Recombinación durante la síntesis de la hebra de ADN positiva



**Fig 1.4.** Representación esquemática de los modelos de recombinación. En el primero se obtiene como resultado del proceso una doble hebra homoduplex. El modelo “strand displacement and assimilation”, origina una doble hebra con regiones heteroduplex. (Tomado de A. Telesnitsky and S. P. Goff “Retroviruses” Coffin, Hughens, Varmus 1997).

#### **1.4.6. Las dos copias de ARN de VIH.**

Los retrovirus se diferencian del resto de los virus en que cada virión contiene dos copias completas de ARN genómico. Se ha propuesto que la finalidad de este hecho es la de aumentar la probabilidad de éxito en la copia de ADN.: Si una de las copias tuviera un fallo, la transcripción inversa continuaría copiando la segunda hebra de ARN para completar la transcripción de un ADN completo(Coffin, 1979).

Se cree que a partir de un virión se forma una sola copia de ADN. En este sentido, los retrovirus no serían considerados diploides, ya que en cualquier posición dada del ADN de la progenie, la información genética procede de uno solo de los ARNs parentales (Coffin y cols., 1997) (Panganiban y Fiore, 1988) (Hu y Temin, 1990b) (Jones y cols., 1994). Sin embargo, la naturaleza pseudodiploide, tendría importantes implicaciones en la generación de diversidad de los retrovirus a través de la recombinación.

## *2. OBJETIVOS*

La recombinación genética es un mecanismo fundamental en la generación de variabilidad para el VIH-1. Este trabajo se ha centrado en el estudio de los factores que afectan a la capacidad de recombinación del VIH en un sistema *in vitro*, con la finalidad de entender los fenómenos evolutivos del virus *in vivo*. Para ello nos fijamos los siguientes objetivos:

1. Determinación de la influencia que ejercen sobre la recombinación del VIH-1 factores como:
  - a. el tiempo de cultivo, o el número de ciclos de recombinación.
  - b. la multiplicidad de infección,
  - c. la distancia genética entre los virus parentales.
2. Identificación de posibles zonas calientes para la recombinación mediante el análisis de las secuencias de clones recombinantes.
3. Análisis de la evolución de las poblaciones virales que resultan de una infección por dos virus diferentes.
4. Estudio sobre la naturaleza diploide o pseudodiploide del VIH. ¿Es posible la formación de dos provirus a partir de las dos copias de ARN de un virión?.

### *3. MATERIALES Y MÉTODOS*

### 3.1 CULTIVOS CELULARES

#### 3.1.1 Células.

**MT4 y MT2 :** Son líneas celulares linfoides del linaje T CD4+, obtenidas a partir de linfocitos de sangre periférica de pacientes infectados por el virus de la leucemia de las células T humana tipo 1 (HTLV-1) por cocultivo con linfocitos de sangre de cordón umbilical de un individuo sano (Miyoshi y cols., 1981). Las líneas celulares fueron cedidas por el Dr. D. D. Richman (Universidad de San Diego, California) .

**Cos-1:** Línea celular establecida procedente de riñón de mono.

#### 3.1.2 Medios de cultivo

Las líneas celulares se crecieron en medio RPMI 1640 (Bio-Whittaker) suplementado con un 10% de suero bovino fetal (SBF) (Gibco), 2 mM de L-glutamina (Flow), 100 µg/mL de estreptomicina, 100 unidades/ml de penicilina y 0.5% de tilosina (ICN Biomedicals Inc.).

#### 3.1.3 Virus e infecciones.

**3.1.3.1 Virus:** Los virus utilizados en estos estudios pertenecen a diferentes subtipos, pero todos tienen la característica de ser formadores de sincitios. Esta cualidad ha sido asociada a la utilización del receptor celular CXCR4. Todos ellos fueron obtenidos a partir del “UNAIDS/NIH/AIDS repository”.

**Tabla 3.1** *Virus utilizados en las infecciones dobles.*

<b>VIRUS</b>	<b>Subtype</b>	<b>BIOTIPO</b>
LAI	B	SI—X4
RF	B	SI—R5X4
SF2	B	SI—R5X4
92UG24	D	SI
93BR20	F	SI—R5X4
92UG29	A	SI—X4

**3.1.3.2 Semillas:** Las semillas de los virus LAI ,RF y SF2 habían sido crecidas en nuestro laboratorio en células MT2. Las semillas del resto de los virus se crecieron en células MT4.

El procedimiento seguido para su obtención consistió en la infección de 10<sup>6</sup> células MT4 con 1ml de sobrenadante del virus correspondiente, incubadas a 37°C, en estufa de CO<sub>2</sub> durante 1 hora y 35min., tras lo cual se añadió 1ml de medio fresco.

Cuando se detecta efecto citopático generalizado, y la viabilidad es menor al 50%, se pasa el cultivo directamente a  $7 \times 10^6$  células, y posteriormente a  $40 \times 10^6$ . Los sobrenadantes se recogieron y fueron congelados a  $-70^{\circ}\text{C}$  o en Nitrógeno líquido, para su conservación.

La valoración de la replicación viral en las infecciones del estudio se realizó utilizando dos parámetros:

A,- el efecto citopático , utilizado como +, ++, +++, ò +++++ según el porcentaje del cultivo afectado sea el 25, 50, 75 ò 100%.

B.- la viabilidad celular , determinada por el método de exclusión de colorante azul tripano.

La cuantificación de semillas virales se han llevado a cabo utilizando dos técnicas diferentes:

a) Plaqueo: esta técnica, que se describe en el apartado 3.1.6, permite titular las semillas mediante el conteo de placas de infección , expresándose en unidades formadoras de placa por unidad de volumen ( UPF / mL ).

b) Por dosis infectivas 50 (  $DI_{50}/\text{mL}$  ). Este método utiliza el efecto citopático producido en diluciones seriadas del virus que infectan células MT2. Para ello , se disponen  $10^5$  células MT2 en cada pocillo de una placa de 96, y se infectan con  $100\mu\text{l}$  de diluciones seriadas del virus en medio de cultivo. Las placas se incuban a  $37^{\circ}\text{C}$  en atmósfera de  $\text{CO}_2$  durante 7 días. Pasado este tiempo se lee la placa , considerando positivos todos los pocillos que presentes efecto citopático (ECP ). Las  $DI_{50}$  se calculan según el método de Spearman-Kärber (1931).

**Tabla 3.2** Resultados de la titulación de los virus utilizados en las infecciones dobles

<b>VIRUS</b>	<b>Subtype</b>	<b>TCID50/ml</b>	<b>p.f.u./ml</b>
LAI	B	$1,8 \times 10^5$	$1,8 \times 10^5$
RF	B	$1,8 \times 10^6$	
SF2	B	$5,6 \times 10^5$	
92UG24	D	$4,2 \times 10^5$	$1,75 \times 10^5$
93BR20	F	$5,6 \times 10^5$	$8 \times 10^4$
92UG29	A	$7,4 \times 10^4$	$10^5$

### **3.1.4 Plaqueo en células MT4.**

Esencialmente se siguió el protocolo de Harada y cols (1985) modificado por Sánchez-Palomino y cols. (1993) que se describe a continuación.

Para adherir las células que crecen en suspensión al plástico, se trataron placas p35 con poli-L-lisina (Sigma) a una concentración de 50 µg/mL durante 60 minutos a temperatura ambiente. A continuación se lavan las placas tres veces con PBS, y se añade 1,5 mL/pocillo de una suspensión de células MT4 a una concentración de  $3 \times 10^6$  células/mL previamente lavadas dos veces con PBS. Las placas con la suspensión celular se mantienen sobre una mesa de equilibrado durante una hora a temperatura ambiente para que se forme una monocapa homogénea. Posteriormente se retira con cuidado el PBS, y se infecta la monocapa con 500-1000 µl de una dilución de virus en medio de cultivo. Se incuban las placas en estufa durante 1 hora y 30 minutos para permitir la adsorción viral, después de la cual, se retira el inóculo, y se cubre la monocapa con 2ml de medio de cultivo con un 0,7% de agarosa (Sea-plaque Agarose, Marine Colloid Corp.) caliente que se solidificará a 4°C antes de pasar a la estufa. Las placas pueden apreciarse a simple vista después de una incubación de 4 a 7 días.

Esta técnica permite además de la titulación de virus, el aislamiento de clones biológicos. Pinchando con una micropipeta las placas de infección, se resuspenden en 100 µl de medio de cultivo, que después se utilizarán para la extracción de ADN.

### **3.1.5 Infecciones dobles.**

Las infecciones dobles se llevaron a cabo sometiendo células MT4 frescas a una infección con una mezcla de dos virus distintos. Se usaron dos multiplicidades de infección: 1 y 0.1. Para el cálculo de la multiplicidad de infección se usaron los títulos en p.f.u.s, aunque se recurrió a los títulos en TCID50 para los casos en que no disponíamos de los datos obtenidos por plaqueo. Tras una incubación de 1h y 45 minutos para permitir la adsorción del virus, se retiró el sobrenadante y se lavaron las células con PBS, para volver a resuspenderlas en medio fresco. Se recogieron sobrenadantes varias veces a lo largo de 22 días, añadiendo células cuando el efecto citopático lo requería.

En el estudio sobre recombinación, se estudiaron cuatro combinaciones de virus:

- 1.- LAI ( subtipo B) + UG24 (subtipo D)
- 2.- LAI (subtipo B) + 93BR20 (subtipo F)
- 3.- UG24 (subtipoD) + 93BR20 (subtipo F)
- 4.- LAI (subtipo B) + RF (subtipo B)

Estas infecciones dobles se realizaron en  $2 \times 10^6$  MT-4 células a un multiplicidad de infección de 0.1 p.f.u.(unidades formadoras de placas) por célula. Se añadieron  $2 \times 10^6$  cells el día 5, y  $5 \times 10^6$  los días 9, 13, 16, y 19. Los cultivos se siguieron hasta el día 22. Un esquema del procedimiento general puede verse en la fig.1.

Excepcionalmente, la infección intrasubtipo B, realizada con los virus LAI y RF, se siguió hasta el día 21, y los sobrenadantes fueron recogidos los días 8, 13, 17, debido a que el efecto citopático en esta infección era mayor que en las anteriores.

### **3.1.6 Cultivos secundarios de 24 horas.**

1 ml de cada sobrenadante recogido a partir de los cultivos de las infecciones dobles, fueron utilizados para realizar una infección sobre  $5 \times 10^6$  células, que se recogió a las 20-24 horas. La finalidad de este cultivo es conseguir que los viriones presentes en el sobrenadante infecten las células frescas, dar tiempo suficiente para que ocurra la transcripción inversa y obtener la copia proviral de los viriones, pero para no completar un ciclo entero de replicación, se recogen los cultivos antes de que se produzca un segundo ciclo de infección.

## **3.2 OBTENCIÓN DE ÁCIDOS NUCLEICOS**

### **3.2.1 Obtención de ADN**

En función del número de células disponibles, la técnica empleada para la obtención de ADN proviral, fue distinta. En infecciones con  $5 \times 10^6$  células o más, se siguió un método de extracción estandar con fenol (Perucho y cols., 1981), a partir de un sedimento de células lavado 2 veces con Na Cl 137mM, KCl 2.7mM,  $\text{Na}_2\text{HPO}_4$  8,1mM,  $\text{KH}_2\text{PO}_4$  1.47mM pH 7.4 ( PBS ). Este sedimento se resuspende en una solución Tris-HCl 10mM pH 7.5, EDTA 10mM, NaCl 0.15M, proteinasa K 200 $\mu\text{g}/\text{ml}$  y SDS 0.4%. Tras una incubación de una hora a 65°C y toda la noche a 37°C, se añadió un volumen de Tris-HCl 10mM pH 8, EDTA 10mM, NaCl 0.65M. La mezcla se extrajo 3 veces con un volumen de cloroformo-alcohol isoamílico (24:1). El ADN se precipitó con etanol y se resuspendió en agua a una concentración aproximada de 1 $\mu\text{g}/\mu\text{l}$ , determinada por medida espectrofotométrica a una longitud de onda de 260nm.

En el caso de las placas obtenidas por la técnica de plaqueo en MT-4, la obtención de ADN se realizó utilizando el equipo comercial Biorad “Instagene purification matrix”. Para ello se lleva a cabo una primera incubación a 56°C durante 30 minutos, de 200µl de la suspensión “Instagene” con 20µl de una suspensión de la placa pinchada, seguida de otra incubación de 8 minutos a 100°C. Una centrifugación a 15000 r.p.m., durante 5 minutos, permite recuperar el ADN en el sobrenadante. La solución resultante se conserva a -20°C. Excepcionalmente se utilizó este equipo comercial para la extracción de ADN a partir del residuo celular del D5 de la doble infección LAI/RF, ya que no había suficiente cantidad de células para utilizar la extracción con fenol.

### **3.2.2 Obtención de ARN**

La obtención de ácido ribonucleico (ARN) se realizó mediante el método descrito por R. Boom (Boom y cols., 1990). Consiste en una incubación a temperatura ambiente durante 10 minutos de 20 µl de sobrenadante de cultivo con 980 µl de solución de lisis (120 g de isotiocianato de guanidina, 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 µl de sílica. El sedimento obtenido de centrifugar a 15000 rpm durante 25 segundos, se lava dos veces con 1ml de solución de lavado (120 g de isotiocianato de guanidina, 100ml de Tris-HCl 0.1 M pH 6.4), dos veces con 1 ml de etanol al 70% y una vez con 1 ml de acetona. Tras este lavado se deja secar la sílica a 56°C durante 10 minutos y se eluye por resuspensión en 50µl de agua esteril e incubación a 56°C 10 minutos. El ARN se recoge en el sobrenadante tras centrifugación a 15000 r.p.m. 2 minutos y se conserva a -70°C.

### **3.3 REACCIÓN EN CADENA DE LA POLIMERASA (PCR)**

A lo largo del estudio se amplificaron diferentes fragmentos de distintas regiones genómicas del VIH-1, fundamentalmente los genes *gag pol* y *env*.

Las reacciones se llevaron a cabo en tampón Tris-HCl 10 mM pH 8.3, KCl 50 mM, gelatina 0.01% y MgCl<sub>2</sub> 1.5 mM. La concentración de desoxinucleótidos trifosfato (dNTPs) (Pharmacia) fue 0.125 mM, y la de iniciadores fue 2 ng/µl, y 0.5 unidades de enzima Ampli-Taq polimerasa (Perkin-Elmer) por ensayo, en un volumen total de 25 µl.

Los ciclos de desnaturalización del ADN se realizaron a 94-95°C durante 30 segundos, exceptuando el primero que tuvo una duración de 2 minutos 30 segundos.

La temperatura de hibridación varió desde 52°C hasta 65°C, según los iniciadores utilizados.

Dependiendo de la longitud del fragmento amplificado, el tiempo de elongación fue de 30 segundos a 6 minutos ).

### **3.3.1 PCR utilizada para la identificación de clones biológicos recombinantes.**

Una vez obtenidos los clones mediante la técnica de plaqueo descrita en el apartado 3.1.5, se pincharon individualmente, se resuspendieron en 100µl de medio de cultivo, y se almacenaron a -70°C. 20 µl de esta suspensión se utilizan para extraer DNA, con el equipo “Instagene”, según instrucciones del fabricante. El ADN de cada una de las placas fue sometido a una reacción de PCR con los cuatro iniciadores que se describen en la tabla 3.3, en las condiciones antes descritas y sometidas a 35 ciclos de temperatura con el siguiente patrón: 1 ciclo de desnaturalización, a 95°C durante 35 s, seguido de una hibridación a 52°C durante 45 s, y una elongación a 72°C de 45 s de duración. Estos 35 ciclos son precedidos por una desnaturalización de 2 min 30 s a 95°C, y finalizan con una elongación de 10 min a 72°C.

Los dos pares de iniciadores originan dos amplicones de diferente tamaño para cada virus, que pueden resolverse en geles de agarosa (TAE) al 0,9%, y que nos han permitido diferenciar todos los virus utilizados en las infecciones dobles, así como los virus recombinantes que aparecen en la progenie (ver esquema 4.2 de resultados).

Así, por ejemplo, en la doble infección B/D, los iniciadores 117 y 118 amplifican un fragmento en la región de Gag p7-p1-p6, de 152 nucleótidos con la cepa LAI, y de 117 con la cepa UG24. Los iniciadores de env, 114 y 115, producen un fragmento de 176 nucleótidos en C3-V4, en el caso de LAI, y de 167 para UG24. De esta manera, podemos distinguir cada clon aislado, si se trata de un virus parental o recombinante, según puede verse en la fig.4.2.

**Tabla 3.3.** *Iniciadores usados en la identificación de clones biológicos.*

<b>Iniciadores</b>	<b>Secuencia (5'-3')</b>	<b>Posición</b>
114E	5'-TCCTCAGGAGGGGACCCAGAAA-3',	6890 a 6912
115E	5'-TTTTATTCTGCATGGGAGGGTGAT-3'	7064 a 7040
117GD	5'-CCCCAGACCTGAAGCTCTCTT-3'	1748 a 1768
118GU	5'-AGACAGGTTAATTTTTTAGGGAA-3'	1618 a 1645

### **3.3.2 PCR cuantitativa**

La detección de formas recombinantes provirales se llevó a cabo sobre el ADN proviral procedente de los sedimentos de células MT4 resultantes del seguimiento de las infecciones dobles. Se utiliza una PCR específica que es capaz de diferenciar dos formas parentales y dos formas recombinantes distintas.

Para ello se usaron una serie de iniciadores que hibridan en dos zonas del gen env, distantes entre si unos 2500 nucleótidos aproximadamente. Cada par de iniciadores es específico de cada virus parental, y la combinación cruzada de éstos, es específica de cada especie recombinante. Esto se representa de una forma esquemática en la fig.4.3.

La secuencia y localización de estos iniciadores específicos, es la siguiente:

137B (5'-GTGAAGGAGAAATATCAGCA-3', nucleotides 5811 to 5831 in LAI)

139B (5'-CTGCTATGGCTGTGGCAT-3', nt 8265 to 8247 in LAI)

155D (5'-CTGTGTCACCTTAAACTGCATTGAAT-3', nt 6175 to 6201 in LAI)

154D (5'-TCCTGAATCCAATACTGGAG-3', nt 8218 to 8198 in LAI)

143B (5'-TCTACAAGCTCCTTGACTA-3', nt 8314 to 8294)

141F (5'-TGAGGGGGATGCAGAGGAA-3', nt 5810 to 5829)

142F (5'-TCTACCAGCTCTTTGCAAAG-3', nt 8314 to 8294)

146F (5'-GGGGAAGTGGGGCCTTTTAT-3', nt 5840 to 5860)

273R (5'-TGATGGAGATGAGGAAGAA-3', nt 5811 to 5830)

274R (5'-GTGAAGAAAAGCTCTAAGAA-3', nt 8308 to 8327)

La denominación de los iniciadores consiste en un número seguido de una letra que indica el subtipo del virus con el que hibrida, excepto en el caso de RF, cuya letra es R para diferenciarlos de los iniciadores específicos de LAI, que llevan la letra B.

Los ciclos de temperatura utilizados para las infecciones dobles B/D, B/F y D/F fueron: 94°C 2 minutos, 10 ciclos de 95°C durante 45 s, 60°C 1 min, y 72°C 4 min; seguidos de 25 ciclos a 95°C 45 s, 55°C 1 min, 72°C 4 min, acabando en una incubación de 72°C durante 10 min. La amplificación RF+LAI fue similar, cambiando únicamente la temperatura de hibridación, que fue de 63°C en los primeros 10 ciclos, y 57°C en los 25 restantes.

La cuantificación de cuatro especies moleculares distintas a partir de una mezcla, se abordó en cuatro determinaciones independientes. Cada una de ellas consta de 6 diluciones de un patrón externo y cuatro de la muestra problema, como puede verse en la fig.4.4. Se

consideró la PCR como una reacción lineal ya que se realizó siempre a concentraciones de sustrato próximas a la concentración límite.

Los patrones externos se construyeron con plásmidos, en los que se inserta una copia de la secuencia problema de la muestra, según el método expuesto en el apartado 3.4.

Para cada experimento de cuantificación, se prepararon 42 reacciones: dos de ellas son controles negativos, y 10 reacciones para cada una de las cuatro especies que pretendemos determinar: los dos provirus parentales y los dos tipos diferentes de provirus recombinantes. La fig. 4.4 muestra el resultado de uno de estos experimentos.

Los productos de la PCR se corren en un gel de agarosa al 0,9% en T.A.E., a 30-35 Voltios durante 1 hora y 30 minutos y el gel se tiñe con Bromuro de Etidio al 5%. La fotografía del gel se leyó en un scanner HP50 de Hewlett Packard y se densitometraron las bandas utilizando el programa PCBAS versión 2.08 (Isotopenmeßgeräte GMBH, 1993). La densidad óptica de las diluciones de plásmidos nos permiten elaborar una recta patrón de densidad óptica frente a número de copias, que se utilizó para extrapolar el número de copias provirales de la muestra. Para lo cual usamos el programa estadístico Graph Pad PRISM versión 2.01 (Graph Pad Software Incorporated). Las cuatro determinaciones se realizan en paralelo, sometidas a la misma reacción de PCR, para evitar que la variabilidad atribuida a esta técnica afecte a las cantidades relativas que pretendemos determinar. El número de copias de cada especie que obtenemos por este método, es traducido a porcentajes dentro de la progenie de cada infección doble. El valor final para cada muestra es la media aritmética de tres experimentos de cuantificación.

### **3.4 PLÁSMIDOS UTILIZADOS EN LOS PATRONES DE LA PCR CUANTITATIVA**

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 extremos son utilizados para insertar el producto de PCR en un vector linearizado que contiene residuos de desoxitimidina en los extremos 3'.

Se prepararon 12 insertos diferentes, uno para cada especie a cuantificar. Los nombramos con dos letras que indican la especificidad de hibridación de sus extremos con los iniciadores específicos. Por ejemplo, el inserto B-D, consiste en una secuencia

recombinante, cuyo extremo 5' hibrida con el iniciador específico de B, y el extremo 3' hibrida con el iniciador D.

Se construyeron plásmidos con cada uno de los siguientes insertos: B-B, D-D, F-F, R-R, B-D, D-B, B-F, F-B, D-F, F-D, L-R, R-L.

#### **3.4.1 Preparación del inserto y ligación.**

El fragmento de ADN que se utiliza para la inserción en el plásmido, procede de una amplificación de una muestra, con los iniciadores específicos descritos en el apartado 3.3.2. La ligación con el vector pCR®2.1, de 3,9 Kb, tuvo lugar en una proporción molar 1:2 vector :inserto. La reacción de ligación se llevó a cabo con 4 unidades de enzima T4 ligasa a 14°C durante toda la noche.

#### **3.4.2 Transformación de bacterias competentes.**

Para la transformación en bacterias competentes se añadieron 2µl de β-mercaptoetanol a cada vial de células E.coli (INαF'). Tras mezclar cuidadosamente (sin pipetear) la mezcla de reacción, se añadió 1µl de mezcla de ligación y se incubó durante 30 minutos en hielo. Posteriormente las bacterias se sometieron a un choque térmico de 45 segundos a 42°C, y 2 minutos en hielo, tras lo cual se añaden 250 µl de medio SOC (Bactotripton 2%, extracto de levadura 0.5%, NaCl 10 mM, KCl 2.5 mM, MgCl<sub>2</sub> 10 mM, MgSO<sub>4</sub> 10 mM, glucosa 20 mM) a las bacterias y se incuban durante una hora a 37°C, con agitación a 225 r.p.m. Del producto de la transformación se extendieron 50 µl y 200 µl en placas de medio LB-agar (extracto de levadura 2%, NaCl 1%, agar 1%, pH7.5) con ampicilina a una concentración de 50 µg/ml, para la selección de colonias resistentes a dicho antibiótico. Para la selección de colonias recombinantes por color, se añadió 1.6mg de 5-bromo-4-cloro-3-i-olil-β-D-galactopiranosido (X-gal) (Boehringer Mannheim) (20 mg/ml en dimetilformamida). Las placas se incubaron a 37°C durante toda la noche. Se pincharon las placas seleccionadas (blancas) y se probó la presencia de inserto mediante una PCR específica y una electroforesis en agarosa al 1%.

#### **3.4.3 Obtención de ADN plasmídico a gran escala.**

Las colonias cuyo análisis fue positivo en la detección del inserto, se cultivaron en medio LB (extracto de levadura 2%, NaCl 1%, pH 7.5) con ampicilina (50 µl/ml), y se incubaron

a 37°C durante unas 7-8 horas, al cabo de las cuales se pasa un inóculo de 1 ml a 100 o 200 ml de medio LB con ampicilina en la concentración anteriormente descrita y se incuba toda la noche.

La extracción del ADN plasmídico de este cultivo se realizó con el equipo comercial “Wizard™ Maxipreps DNA purificación Systems” (Promega) según instrucciones del fabricante. Y se resuspendió en TE pH 8. El ADN obtenido se cuantificó por espectrofotometría, a 260nm.

### **3.5 AISLAMIENTO DE CLONES BIOLÓGICOS RECOMBINANTES**

Para aislar y caracterizar virus recombinantes, se obtuvieron clones biológicos por el ensayo de plaqueo en MT4, descrito en el apartado 3.1.6. La monocapa de células se infectó con los sobrenadantes de las infecciones realizadas en medio líquido. Las placas se pincharon y se resuspendieron en 100µl de medio de cultivo. A partir del cual se extrae el ADN y se somete a una PCR descrita en el apartado 3.3.1 para identificar los que son recombinantes.

### **3.6 SECUENCIACIÓN**

La secuenciación fue realizada con el equipo comercial “Big Dye™ Terminator Cycle Sequencing v2.0” (Applied Biosystems, Perkin Elmer), que utiliza terminadores marcados con 4 colorantes fluorescentes distintos. La reacción de PCR comienza con un ciclo de desnaturalización a 96°C durante 3 minutos, seguido de 25 ciclos de 96°C 30 segundos, 55°C 15 segundos, y 60°C durante 4 minutos. Los ddNTPs no incorporados a las cadenas formadas durante la reacción fueron eliminados mediante una precipitación con MgCl<sub>2</sub> /etanol. Los fragmentos obtenidos se analizaron en el secuenciador automático ABI PRISM™ 377 DNA Sequencer (Applied Biosystems, Perkin-Elmer).

#### **3.6.1 Secuencias utilizadas para el estudio de los clones biológicos recombinantes.**

En este estudio se secuenciaron dos fragmentos del genoma viral:

1. - Desde el nt 2102 al 2714 (de LAI), que comprende la primera mitad de la RT incluyendo un pequeño fragmento del final de la proteasa.
2. - Desde el nt5850 hasta el 6648 que incluye el final del péptido señal, C1, V1, V2 y C2.

Estos fragmentos, se amplificaron con los siguientes iniciadores:

47RU (5'-AGTAGGACCTACACCT-3' nt 2060 to 2075)

20RD (5'-ACAGTCCAGCTGTCTTTTTCTGGC-3' nt 2866 to 2892)

48EU (5'-TAGAAAGAGCAGAAGACAGTGGC-3' nt 5778 to 5800)

115ED (5'-GTTTTATTCTGCATGGGAGGGTGAT-3' nt 7040 to 7064)

Para la secuenciación, se utilizó el primer 99EU (5'-AGAGCAGAAGACAGTGGC-3', nt 5782 to 5800) en env, y el 47RU en pol.

### **3.7 ANÁLISIS DE SECUENCIAS**

Las secuencias nucleotídicas fueron editadas, compiladas y traducidas utilizando el programa Lasergene for Windows (Dnastar Inc, Madison, Wis).

Los alineamientos múltiples de secuencias nucleotídicas se realizaron con el programa CLUSTALW versión 1.7 (Thompson y cols., 1994) siendo posteriormente editados y corregidos manualmente para asegurar que la fase de lectura fuera correcta. Toda aquella posición en la que no existe información completa para el conjunto de secuencias que forman el alineamiento (“gaps”) se excluyó de todos los análisis filogenéticos posteriores. Las secuencias nucleotídicas de los aislados de referencia fueron obtenidos en la base de datos de Los Álamos (<http://hiv-web.lanl.gov>).

### **3.8 ANÁLISIS FILOGENÉTICO**

Los análisis filogenéticos se realizaron utilizando el paquete de programas MEGA-2 (Kumar y cols., 1994) (Kumar y cols., 2001). Las matrices de distancias se calcularon mediante el modelo de dos parámetros de Kimura. Estas matrices sirvieron de base para la construcción de los árboles filogenéticos mediante el método del vecino más próximo (“neighbor-joining”, NJ) (programa NEIGHBOR).

La robusted 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).

### **3.9 ANÁLISIS DE RECOMBINACIÓN**

Los análisis de recombinación genética entre las secuencias nucleotídicas de los clones biológicos, se llevaron a cabo con el programa SIMPLOT versión 2.4 (Ray, 1999). Brevemente, a partir del alineamiento de secuencias, el programa calcula el porcentaje de

identidad en cada posición entre la secuencia problema y el resto de las secuencias (o grupos de secuencias) con las que deseamos realizar la comparación. Para ello, una ventana deslizante se desplaza a lo largo del alineamiento en distintos pasos, calculando el grado de similitud entre las secuencias nucleotídicas en cada paso. En todos los análisis se utilizó una ventana de 200 nucleótidos y un paso de ventana de 20 nucleótidos. Para cada fragmento (ventana) se analizan las relaciones filogenéticas entre las secuencias estudiadas, por remuestreo de 1000 réplicas (“bootstrap”). En el caso de existir evidencias de recombinaciones múltiples (secuencias mosaico), el programa nos permite identificar los sitios filogenéticamente informativos (Robertson y cols., 1995b) (Robertson y cols., 1995a).

### **3.10 MUTAGÉNESIS**

Para la construcción de los plásmidos utilizados en el ensayo de complementación se partió de dos construcciones realizadas en nuestro laboratorio (Olivares y cols., 1997), denominadas pFc, y pFAM4, que contienen todo el genoma de VIH, dividido en dos partes: pFAM4 contiene la secuencia comprendida entre la LTR 5' y un sitio RI en vpr. El segundo, pFc contiene el fragmento entre este sitio RI y la LTR3'.

Utilizamos el pFAM4 para realizar dos deleciones diferentes, por mutagénesis dirigida, dando lugar a dos plásmidos que denominamos  $\Delta p6$  y  $\Delta p24$ . Para ello, se usó el equipo comercial “QuickChange site-directed mutagenesis kit”, según instrucciones, y los iniciadores que se citan a continuación:

*$\Delta p6$ -314GU*: 5'-CCTCAGAACAGACTAGAGCCACCTTCAGCTTCCCTCAAATCACTC-3'

*$\Delta p6$ -315GD*: 5'-GAGTGATTTGAGGGAAGCTGAAGGTGGCTCTAGTCTGTTCTGAGG-3'

*$\Delta p24$ -287G*: 5'-CCGGTTCTATAAGACTCTGCATGCCGAGCAAGCTTCACAGGAGG-3'

*$\Delta p24$ -288GU*: 5'-CCTCCTGTGAAGCTTGCTCGGCATGCAGAGTCTTATAGAACCGG-3'

Los iniciadores 314 y 315 se usaron para producir una deleción directa de 99 nucleótidos en la proteína p6 de VIH en el pFAM4.

La deleción en la proteína p24, se consiguió con los iniciadores 287 y 288, que introducen 4 mutaciones puntuales en los aminoácidos 49 a 52 de p24, para constituir un sitio SphI. Tras la digestión con SphI, obtuvimos una deleción de 258 nucleótidos que incluye todos los sitios implicados en la funcionalidad de esta proteína.

Todos estos cambios se confirmaron secuenciando y analizando los productos resultantes.

### **3.11. Transfecciones**

Se llevaron a cabo co-transfecciones conjuntas de los plásmidos pFAM4 (y sus variantes delecionados) y pFc digeridos con EcoRI, los cuales son capaces de unirse entre ellos por medio de una ligación in vivo, constituyendo un genoma completo del virus dentro de la célula.

Se transfectaron 10 µl de ADN purificado de cada plásmido en  $4 \times 10^6$  células Cos-1, por electroporación a 975 µF y 250 V en un equipo “gene pulser” Bio-Rad.

Se hicieron las siguientes transfecciones en paralelo:

- 1.- 10 µl pFAM4 + 10 µl pFc
- 2.- 10 µl Δp24- pFAM4 + 10 µl pFc
- 3.- 10 µl Δp6-pFAM4 + 10 µl pFc
- 4.- 10 µl Δp24- pFAM4 + 10 µl Δp6-pFAM4 + 10 µl pFc
- 5.- 10 µl Δp24- pFAM4 + 10 µl Δp6-pFAM4 + 10 µl pFc

Control negativo: sin ADN

Los días 2, 4, 6, 8, 10, 12, 14 se recogió el sobrenadante, se centrifugó a 2500 r.p.m. para retirar restos celulares, y se ultracentrifugó a 35000 r.p.m. durante 2 horas y 20 minutos.

El sedimento se trató con DNAsa RQ1, libre de RNAsas durante 30 min a 37°C para eliminar restos de ADN plasmídico. Después se resuspendió en 1mL de medio de cultivo, que fue usado para infectar una monocapa de células MT4.

Las placas se pincharon con micropipeta y se resuspendieron en 100 µl de medio. Se realizó una extracción de ADN a partir de esta suspensión, el cual fue sometido a un análisis por PCR, con los siguientes iniciadores:

302GU 5'-CAGCATTATCAGAAGGAGCCA

14RD 5'-CTATAGATTAGGGACCACAGAGT

303GU 5'--GGGACATCAAGCAGCCATGCAA

147RD 5'-AGAAGACAGTTACCGGTAACAAATTG

## *4. RESULTADOS*

#### **4.1 PUESTA A PUNTO DE LAS TÉCNICAS UTILIZADAS EN EL ESTUDIO DE LA RECOMBINACIÓN *IN VITRO*.**

Para abordar el estudio de la recombinación *in vitro*, fue necesario poner a punto un sistema experimental que nos permitiera obtener, detectar y diferenciar los virus recombinantes dentro de una población viral.

Las poblaciones virales analizadas se obtuvieron a partir de una serie de cultivos celulares infectados simultáneamente con dos virus diferentes. El análisis se realizó sobre los sobrenadantes recogidos a distintos tiempos a lo largo del cultivo. Los procedimientos utilizados se detallan en el apartado 3.1.3.4 de Materiales y Métodos, y se esquematizan en las figuras 4.1 y 4.5.

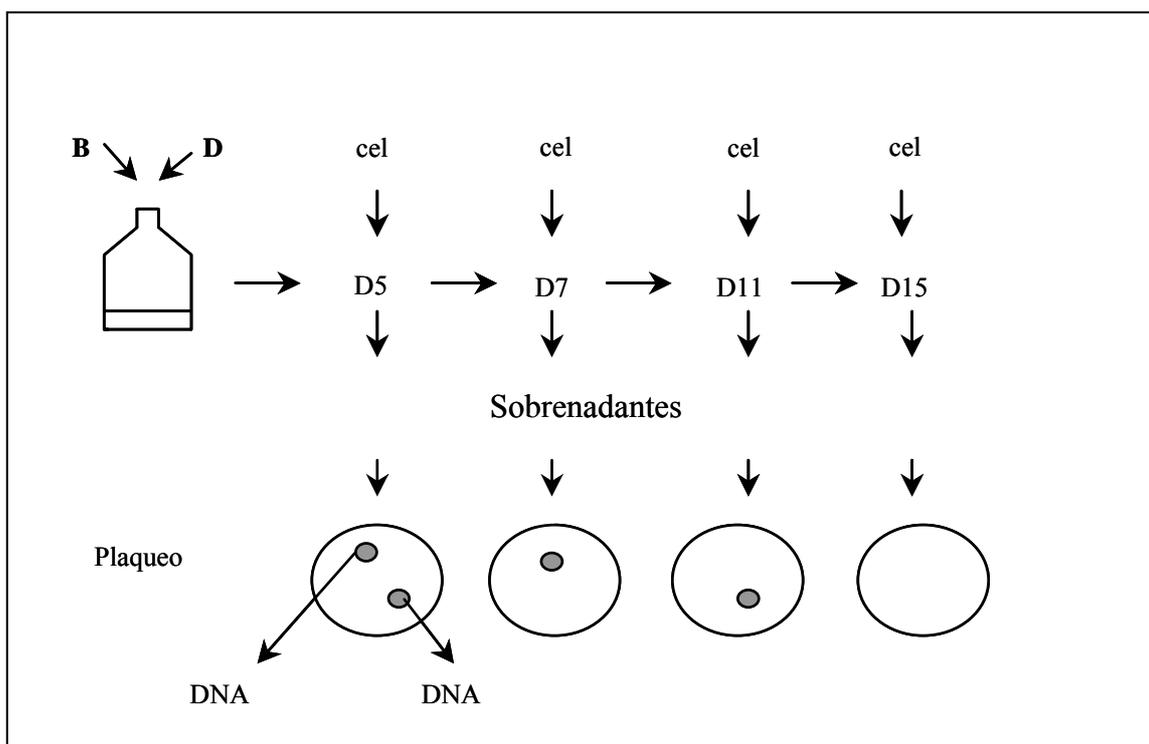
La detección de los virus recombinantes se llevó a cabo mediante la utilización de iniciadores específicos para cada virus, en reacciones de PCR realizadas sobre el ADN proviral procedente de las células infectadas. Este método, descrito anteriormente por Kuwata y cols., es la base de la técnica de cuantificación utilizada en este estudio.

Sin embargo, para la identificación de clones biológicos, se ideó un sistema de iniciadores inespecífico que podía identificar los clones recombinantes en una sola reacción de PCR.

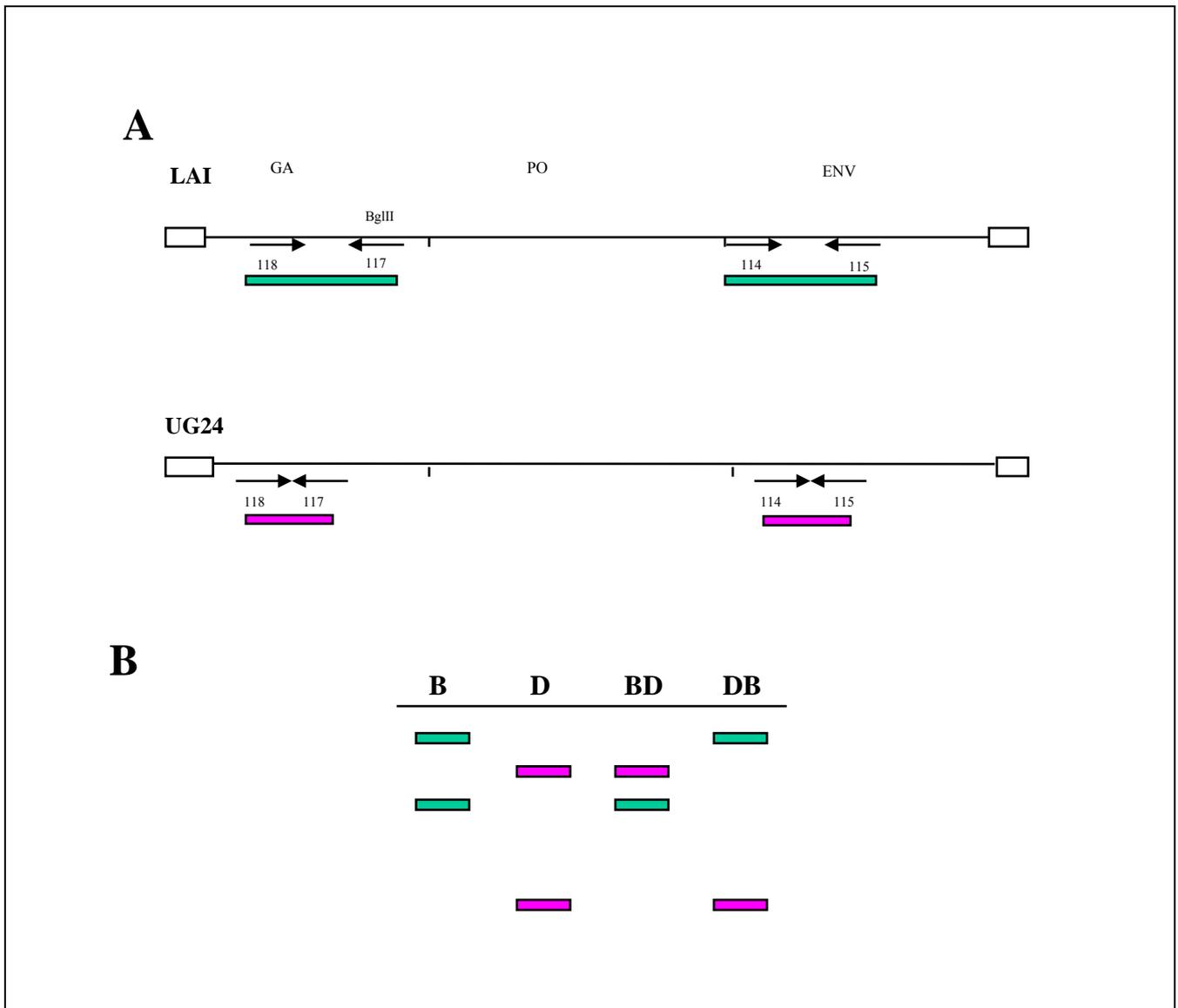
#### 4.1.1 Aislamiento e identificación de clones biológicos recombinantes.

Una vez realizada la infección doble de las células con dos virus parentales, se recogieron los sobrenadantes, que fueron utilizados para infectar una monocapa de células MT4 según la técnica de plaqueo descrita en Materiales y Métodos. Las placas de infección que se generaron en la monocapa, fueron pinchadas, se extrajo su ADN y se identificaron los clones biológicos (Fig. 4.1). Para ello se utilizan dos pares de iniciadores que dan lugar a dos fragmentos de diferente tamaño para cada virus, los cuales pueden resolverse en geles de agarosa (TAE) al 0,9%.

Por ejemplo, en el caso de LAI y UG24, los iniciadores 117 y 118 amplifican un fragmento en la región *gag* de 152 ò 117 nucleótidos según proceda del virus LAI o del UG24. Los iniciadores 114 y 115 de la región *env*, producen un fragmento de 176 nucleótidos en el caso de LAI, y de 167 para UG24. De esta manera, podemos distinguir la naturaleza parental o recombinante de cada clon aislado, según aparece en el esquema de la figura 4.2.



**Fig. 4.1.** Esquema del protocolo utilizado en el seguimiento de las infecciones dobles realizadas para el estudio de la recombinación in vitro. Los virus, que se representan con las letras B y D, infectan células MT4 en medio líquido. Se añaden células frescas cuando el efecto citopático lo requiere, y se recogen sobrenadantes varios días a lo largo del cultivo. Estos sobrenadantes se utilizan para infectar una monocapa de células MT4. Las placas que aparecen se pinchan, se extrae el ADN y se someten a la prueba de identificación de clones biológicos.



**Fig. 4.2.** Esquema de los iniciadores empleados para identificar clones biológicos recombinantes.

**A)** Las líneas horizontales representan el ADN proviral de los virus LAI (subtipo **B**), y UG24(subtipo **D**). Los rectángulos verdes representan los fragmentos amplificados pertenecientes al virus LAI, y los rectángulos rosa corresponden al virus del subtipo D.

**B)** Representación gráfica de la posición que ocupan en un gel los distintos fragmentos amplificados con estos primers, según su procedencia: parental del subtipo B ó D, o recombinantes: BD o DB. Los recombinantes BD son aquellos cuyo fragmento gag pertenece al subtipo B y el fragmento env al subtipo D, y viceversa para los recombinantes DB.

#### **4.1.2 Detección y cuantificación de las diferentes especies genómicas provirales presentes en la progenie resultante de una doble infección.**

Con la finalidad de detectar la presencia de formas recombinantes provirales directamente sobre el sedimento celular resultante de una doble infección, sin la necesidad de analizar un gran número de clones, se diseñaron una serie de iniciadores específicos de cada virus parental.

Estos iniciadores hibridan en dos zonas del gen env, distantes entre si unos 2500 nucleótidos aproximadamente. La combinación cruzada de los iniciadores es específica de cada una de las dos posibles formas recombinantes.

Se diseñó un conjunto de cuatro iniciadores, para analizar cada una de las infecciones dobles (ver Materiales y Métodos). A continuación se relacionan los que se usaron en cada una de ellas:

1.-Coinfección LAI (subtipo B) + 92UG24 (subtipo D):

BB:137-139, DD:155-154, BD:137-154, DB:155-139

2.-Coinfección LAI (subtipo B) + 93BR20 (subtipo F)

BB:137-143, FF:141-142, BF:137-142, FB:141-143

3.- Coinfección 92UG24 (subtipoD) + 93BR20 (subtipo F)

DD:155-154, FF:146-142, DF:155-142, FD:146-154

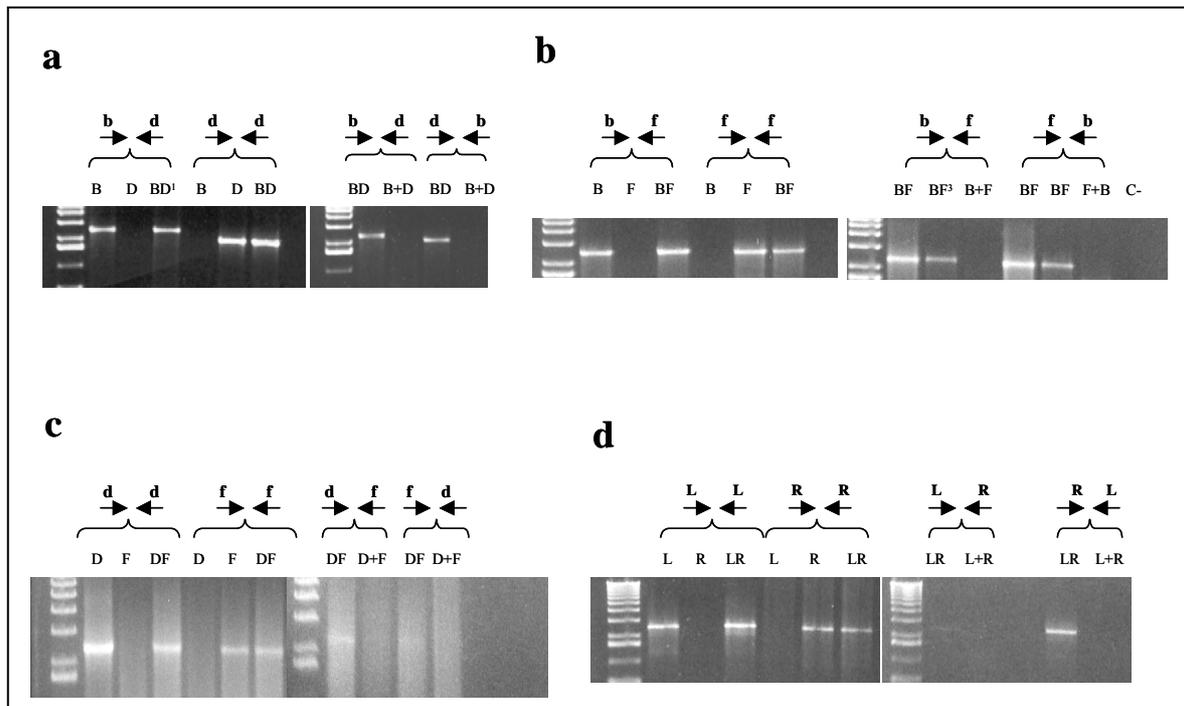
4.- Coinfección LAI (subtipo B) + RF (subtipo B)

LL:137-143, RR:273-274, L-R:137-274, R-L:273-143

Estos iniciadores, y las condiciones en que se utilizan se describen en Materiales y métodos. La prueba de su especificidad para la detección de las diferentes especies genómicas aparece en la figura 4.3.

En todos los casos, además de comprobar su especificidad, amplificando exclusivamente el virus correspondiente; se prueba también que la reacción de PCR no introduce artefactos recombinantes. Para ello, se somete una mezcla de dos ADNs parentales (R+F, por ejemplo) a una amplificación con dos iniciadores de distinta especificidad (específico de RF el iniciador de polaridad positiva, y específico de LAI el de polaridad negativa, por ejemplo). Esta amplificación puede ser positiva sobre el ADN procedente de las infecciones dobles (B/D, B/F, D/F y R/L) si en ellos hay algún provirus recombinante, pero siempre es negativa en las mezclas de ADNs parentales.

La aplicación de esta técnica al análisis del ADN del sedimento celular del día 22 de cada cultivo, permitió detectar la presencia de las cuatro especies genómicas (dos formas parentales y dos formas recombinantes) en la progenie de las cuatro infecciones, aunque la proporción en algún caso es muy baja.



**Fig. 4.3.** Especificidad de los iniciadores empleados en las cuatro infecciones dobles realizadas: en a) la de los virus pertenecientes a los subtipos B y D (B/D), en b) la infección intersubtipo B/F, en c) la infección intersubtipo D/F y en d) la infección intrasubtipo R/L. Los tres primeros canales de cada gel son el producto de una amplificación con los iniciadores específicos del primer parental, y las tres siguientes con los iniciadores del segundo parental. Los cuatro canales de la parte derecha del gel contienen amplificaciones con los iniciadores específicos cruzados, destinadas a detectar la presencia de las especies recombinantes.

Las letras localizadas sobre cada calle del gel, representan el contenido proviral del ADN: así una letra única significa un virus parental; y utilizamos dos letras consecutivas, como BD para referirnos al ADN resultante de una doble infección con B y D. Las reacciones de las calles B+D llevan una mezcla de los ADNs de los virus B y D al 50%. La ausencia de amplificación en estas dos reacciones confirma la no existencia de artefactos recombinantes procedentes de la reacción de PCR. Las flechas identifican los iniciadores utilizados.

#### **4.1.3 Cuantificación de formas recombinantes por PCR.**

Para cuantificar las formas parentales y recombinantes que éramos capaces de detectar y diferenciar gracias a la PCR específica antes descrita, se transformó la misma en una técnica cuantitativa introduciendo un patrón externo para cada una de las cuatro formas diferenciadas. Estos patrones consisten en plásmidos que contienen insertado el fragmento de secuencia que en cada caso se va a amplificar. Así, en el caso de la coinfección BF, se construyeron cuatro plásmidos. El plásmido BB (pBB) lleva insertado un fragmento de 2500 pb de longitud, correspondiente al virus parental B. El pFF lleva insertada la secuencia del virus F. Los plásmidos pBF y pFB, llevan una secuencia recombinante clonada entre todas las que pueden existir en la progenie proviral. Todas las secuencias recombinantes, tienen en común la hibridación con un iniciador B y otro F, independientemente del punto o puntos de recombinación que contenga en su interior.

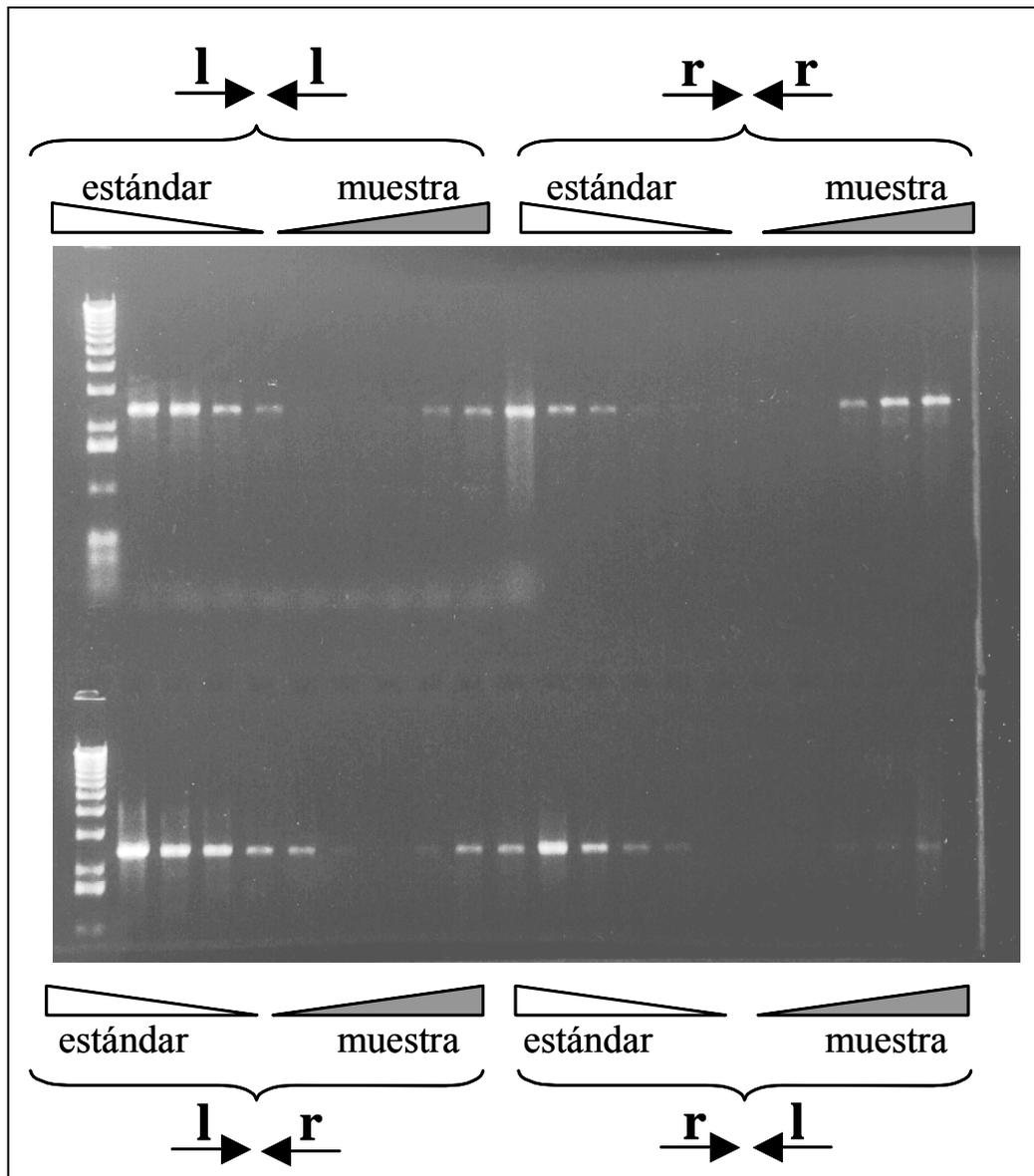
En cada experimento de cuantificación, se realizó la amplificación de las cuatro especies genómicas al mismo tiempo para evitar variaciones inherentes a la técnica de PCR. Además, se utilizó la misma dilución de ADN para las cuatro determinaciones. Cada patrón consiste en seis diluciones seriadas del plásmido correspondiente, y precede a cuatro diluciones de la muestra problema. Estas 10 reacciones acompañadas de un control negativo sin ADN, llevan la misma mezcla de reacción, con los iniciadores que corresponda. Un gel de agarosa con el resultado de la amplificación de uno de estos experimentos puede verse en la figura 4.4.

Con la densidad óptica de las bandas que corresponden a las diluciones de los plásmidos, realizamos una regresión lineal que nos permite extrapolar la cantidad de copias provirales presentes en el ADN original de la muestra a partir de la densidad óptica de las bandas obtenidas tras la amplificación. En la figura 4.5 puede verse un ejemplo del cálculo del número de copias del virus parental en una muestra, a partir de los valores de densidad óptica, con la ayuda del programa estadístico PRISMA.

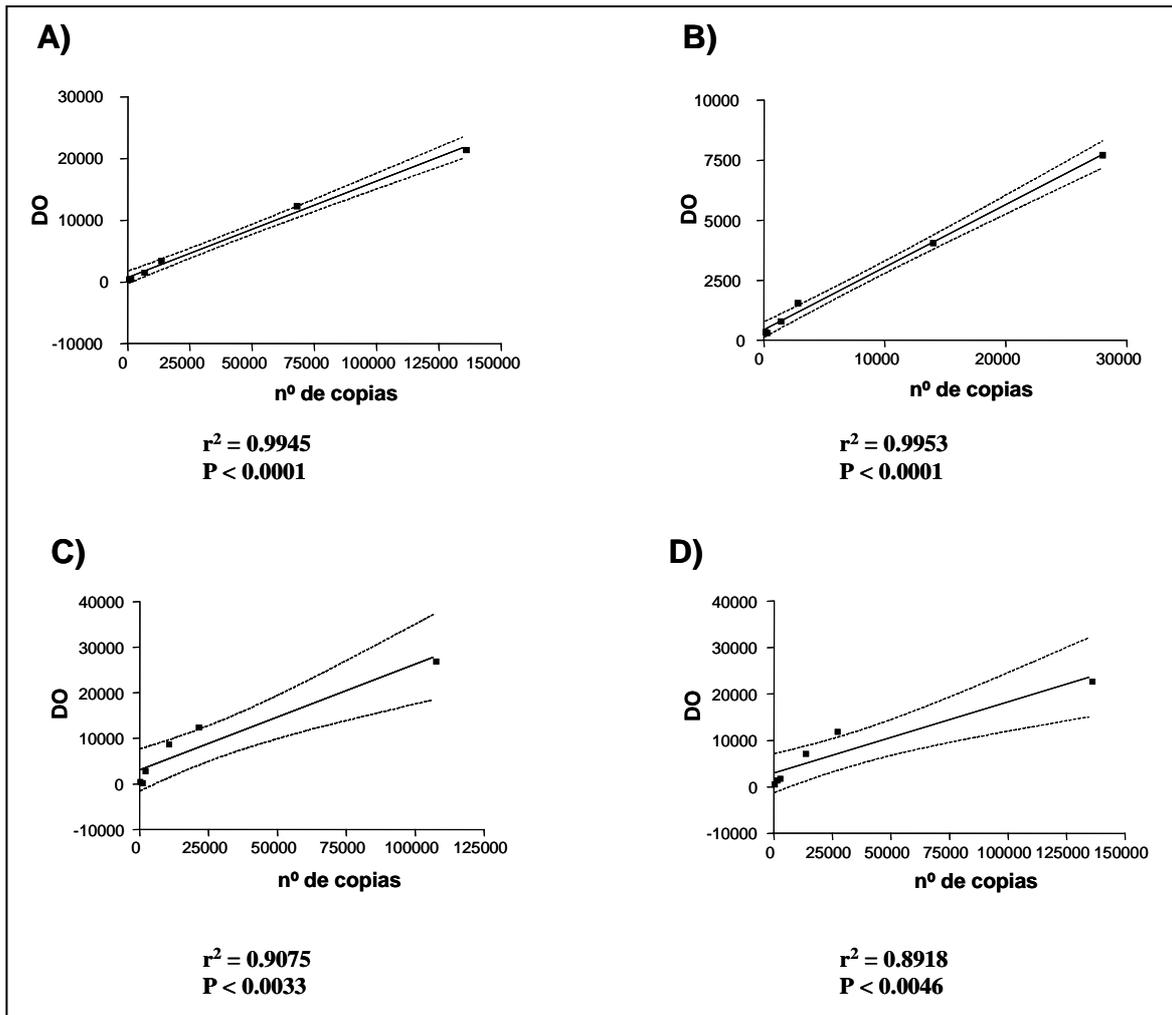
La linealidad de la relación densidad óptica frente a número de copias originales queda garantizada por la utilización de las diluciones más cercanas a la dilución límite de ADN para la amplificación. En la tabla 1 del anexo pueden verse los valores obtenidos en todos los experimentos de cuantificación realizados, y la tabla 2 recoge los valores de  $r^2$  y P correspondientes a las rectas de regresión de los patrones de amplificación. Normalmente obtenemos dos o tres valores de número de copias a partir de cada muestra, uno por cada

dilución con amplificación positiva (puede haber alguna dilución inferior a la dilución límite que no da una amplificación registrable). Se calculó siempre el promedio de los valores calculados para cada dilución.

Para cada muestra de ADN, se repitieron 3 experimentos como este. El resultado de la cuantificación se recogió como la media aritmética de los tres experimentos.



**Fig. 4.4.** Resultado de un experimento de cuantificación llevada a cabo sobre el sobrenadante del día 8 de la infección LAI/RF. Las letras representan los iniciadores utilizados: l = específico de LAI, y r = específico de RF. Cada grupo de 10 reacciones contiene 6 diluciones del plásmido empleado como patrón externo, seguido de 4 diluciones del ADN de la muestra problema.



**Fig. 4.5.** Representación gráfica de la relación densidad óptica del producto de amplificación de las distintas diluciones de plásmido que constituyen los patrones externos frente al número de copias de plásmido originales, en un experimento de cuantificación. En la parte inferior de cada representación aparecen los valores  $r^2$  y  $P$  de la regresión lineal. A) Patrón LAI. B) Patrón de RF. C) Patrón de amplificación de la especie genómica recombinante LR. D) Patrón de los recombinantes RL, es decir con su extremo 5' perteneciente a RF y su extremo 3' de LAI.

## **4.2. ESTUDIO CUALITATIVO DE LAS VARIABLES QUE CONDICIONAN LA RECOMBINACIÓN IN VITRO.**

Este trabajo se inició con una serie de experimentos que trataron de estimar la influencia que podían ejercer algunas variables sobre la recombinación.

Era necesario someter nuestros cultivos a unas condiciones que posibilitaran el proceso de la recombinación en una medida que nos permitiera detectarla. Por este motivo, se probaron unas multiplicidades de infección lo suficientemente altas para que favorecieran la coinfección de las células; además, se siguieron los cultivos durante un periodo de tiempo que nos permitiera comprobar su influencia en el proceso, y se probaron, además, varias distancias filogenéticas entre los virus que participaban en las coinfecciones.

Se realizaron una serie de coinfecciones variando la multiplicidad de infección que fue 1 y 0.1, y los virus utilizados, que fueron RF (subtipo B), LAI (subtipo B), y UG24 (subtipo D). La progenie viral resultante se analizó por plaqueo de los sobrenadantes obtenidos en cada una de las coinfecciones. Las placas resultantes, una vez pinchadas y extraído su ADN, fueron identificadas como recombinantes o parentales, a través de la prueba de identificación de clones biológicos descrita en el apartado 4.1.1.

### **4.2.1. Caracterización biológica de las cepas utilizadas en las infecciones dobles.**

Dado que la recombinación requiere la coinfección de una célula por dos virus, antes de llevar a cabo las infecciones dobles, realizamos una caracterización biológica de las cepas para seleccionar aquellas con características similares. Con ello pretendemos minimizar la posibilidad de que la eficacia en la infección por ambos virus sea muy diferente. Si la eficacia de un virus fuera muy superior a la del otro, la probabilidad de coinfección de una célula por dos virus diferentes, disminuiría considerablemente.

Así, todos los virus utilizados pertenecen al fenotipo inductor de sincitios, tienen títulos del mismo orden, y utilizan mayoritariamente el correceptor CXCR4. En la tabla 4.3, pueden verse las características de todos los virus utilizados.

En cuanto a los correceptores principales utilizados por los virus, en la tabla aparecen los que han sido descritos (los virus pueden utilizar algún correceptor minoritario aún sin describir). En este sentido, las cepas RF, SF y BR20, que también pueden usar el receptor

CCR5, no lo hacen en nuestros cultivos ya que las células MT4 no lo expresan. Consecuentemente, en nuestro sistema los virus a competir por el mismo correceptor.

**Tabla 4.1. Características biológicas de las cepas virales utilizadas en la infecciones dobles.**

VIRUS	Subtipo	Biotipo <sup>a</sup>	Correceptores <sup>b</sup>	TCID50/ml	p.f.u./ml
LAI	B	SI—X4	CXCR4	1,8x10 <sup>5</sup>	1,8x10 <sup>5</sup>
RF	B	SI—R5X4	CXCR4-CCR5	1,8x10 <sup>6</sup>	
SF2	B	SI—R5X4	CXCR4-CCR5	5,6x10 <sup>5</sup>	
92UG24	D	SI	CXCR4,CCR8*, ,V28*,CCR3	4,2x10 <sup>5</sup>	1,75x10 <sup>5</sup>
93BR20	F	SI—R5X4	CXCR4-CCR5	5,6x10 <sup>5</sup>	8x10 <sup>4</sup>
92UG29	A	SI—X4	CXCR4	7,4x10 <sup>4</sup>	10 <sup>5</sup>

<sup>a</sup> SI:inductor de sincitios, X4:que usa el correceptor CXCR4, R5:que usa el correceptor CCR5.

<sup>b</sup> La ausencia de otros correceptores en la tabla sólo significa que no han sido probados. \*indica que la utilización del correceptor se probó por ensayos de fusión , no por infección de virus.

#### 4.2.2. Influencia de la multiplicidad de infección.

Se realizaron dos infecciones dobles en paralelo con los mismos virus a dos multiplicidades de infección diferentes: por un lado los virus LAI y RF, y por otro la combinación LAI más UG24. La que se hizo a mayor m.d.i. dio origen a un mayor número de placas con genomas recombinantes que la de menor m.d.i, en los dos casos probados. Los resultados obtenidos pueden verse en la tabla 4.2.

La diferencia en nº de placas, no resultó proporcional a la diferencia en m.d.i.

Elegimos en consecuencia, la multiplicidad 0.1 para nuestros experimentos posteriores, ya que proporcionaba unos porcentajes de recombinación aceptables. Una multiplicidad de infección muy alta podría implicar una saturación del proceso, es decir llegar a un porcentaje de población recombinante máxima debido a las limitaciones del sistema de detección (ver discusión).

**Tabla 4.2 Influencia de la multiplicidad de infección en la proporción de recombinación.**

Coinfeccion	m.d.i.	Día post infecc	Nº total placas	Nº placas parentales	Nº recombinantes	% recombinación	Tipo recombinantes
LAI+RF	<b>1</b>	12	16	12RF	4	<b>25</b>	3LR, 1RL
LAI+RF	<b>0.1</b>	12	16	14RF	2	<b>12,5</b>	1LR, 1RL
LAI+UG24	<b>1</b>	15	12	9B,	2	<b>16</b>	1BD, 1DB
LAI+UG24	<b>0.1</b>	22	91	76B, 2D	10	<b>11</b>	3BD, 7DB

### 4.2.3. Variaciones del nivel de recombinación con el tiempo.

La proporción de recombinantes en la progenie de una infección doble, aumenta con el tiempo del cultivo.

Se llevaron a cabo una serie de plaques con los sobrenadantes recogidos a lo largo del seguimiento de las coinfecciones. La mayor parte de ellas requería un tiempo mínimo y la adición de células frescas, para detectar algún indicio de recombinación.

En la tabla 4.3, se reflejan los resultados de un cultivo infectado con LAI y UG24, en el que no se detectó ninguna placa recombinante hasta el día 15 post-infección, tras añadir células los días 7 y 11.

**Tabla 4.3. Influencia del tiempo en la proporción de recombinación.**

Coinfeccion	m.d.i	Día post infecc	Nº total placas	Nº placas parentales	Nº recombinantes	% recombinación	Tipo recombinantes
LAI+UG24	1	<b>3</b>	11	0B, 11D	0	<b>0</b>	
	1	<b>5</b>	11	1B, 10D	0	<b>0</b>	
	1	<b>7</b>	5	2B, 3D	0	<b>0</b>	
	1	<b>11</b>	15	15B	0	<b>0</b>	
	1	<b>15</b>	12	9B,	2	<b>16</b>	1BD, 1DB

### 4.2.4. La distancia genética y su influencia en la frecuencia de recombinación.

El “salto” o transferencia de hebra entre los dos ARNs de un virión que se requiere para que se produzca un evento de recombinación, parece estar condicionado por la homología de secuencia entre ellas. Esto implicaría frecuencias de recombinación mayores entre virus filogenéticamente más cercanos.

Para ello realizamos un experimento en el que comparamos la aparición de recombinantes en distintas combinaciones de los virus: LAI, RF, SF y UG24.

Comprobamos en una primera aproximación, que a la misma m.d.i. en las infecciones intersubtipo (LAI más UG24 que pertenecen a los subtipos B y D respectivamente) aparecieron menores proporciones de placas recombinantes que en las infecciones realizadas con dos virus pertenecientes al mismo subtipo (LAI, RF y SF pertenecen al subtipo B).

En la tabla 4.4 se muestran los resultados de estos experimentos.

Tabla 4.4. Influencia de la distancia genética en la recombinación

Coinfeccion	m.d.i.	Día post infecc	Nº total placas	Nº placas parentales	Nº recombinantes	% recombinación	Tipo recombinantes
<b>LAI+RF</b>	<b>1</b>	12	16	12RF	4	<b>25</b>	3LR, 1RL
<b>LAI+UG24</b>	<b>1</b>	15	12	9B,	2	<b>16,7</b>	1BD, 1DB

Coinfeccion	m.d.i.	Día post infecc	Nº total placas	Nº placas parentales	Nº recombinantes	% recombinación	Tipo recombinantes
<b>LAI+SF</b>	<b>0.1</b>	12	16	11LAI	3	<b>18,8</b>	3SL
<b>LAI+RF</b>	<b>0.1</b>	12	16	14RF	2	<b>12,5</b>	1LR, 1RL
<b>LAI+UG24</b>	<b>0.1</b>	22	91	76B, 2D	10	<b>11</b>	3BD, 7DB

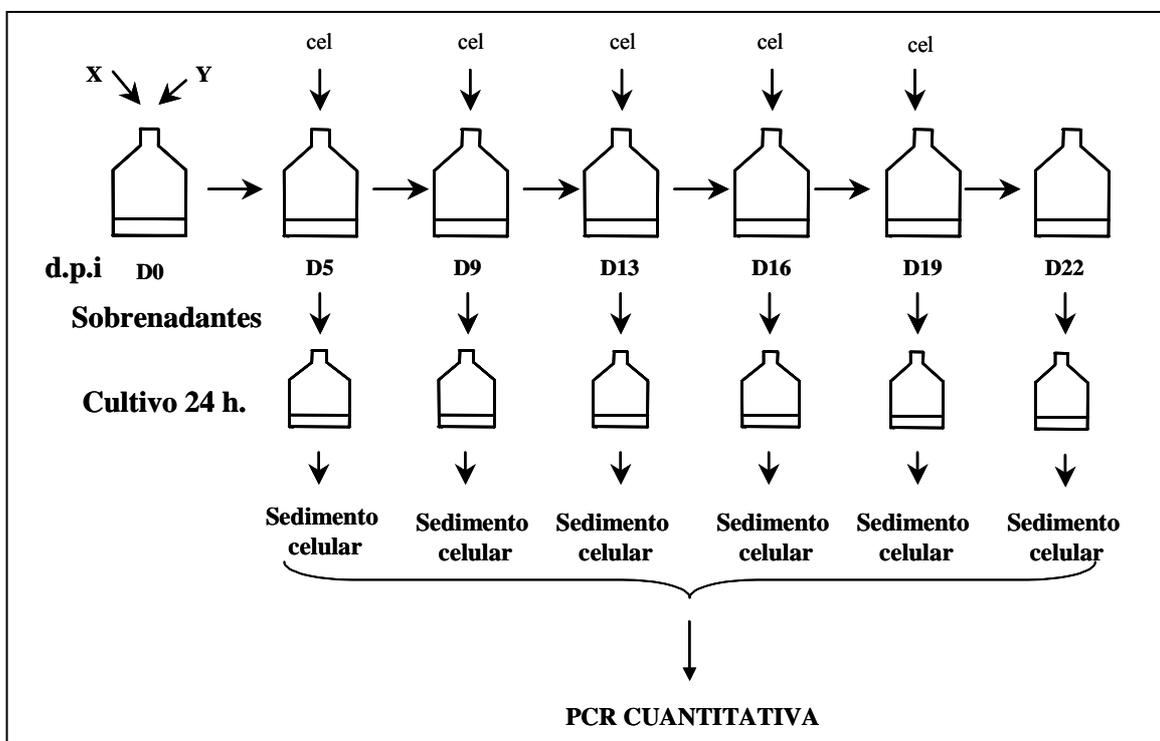
En la mayoría de los casos pudimos aislar recombinantes de dos tipos: aquellos que presentan un gen *gag* que corresponde al primer virus parental, y un gen *env* del segundo, y viceversa (ej: L-R y R-L).

Los cultivos de estos clones recombinantes, en células MT4, revelaron características fenotípicas diferentes, es decir, que no son el mismo virus recombinante sino que parecen tener una naturaleza genómica diferente que será consecuencia de diferentes eventos de recombinación. (Datos no mostrados).

### 4. 3 ESTUDIO CUANTITATIVO DE LA RECOMBINACIÓN EN EL VIRUS VIH-1.

Se realizaron siete infecciones dobles, una intrasubtipo B (LAI+RF) y seis intersubtipo: A+B, A+D, A+F, B+D, B+F y D+F.

Las infecciones se realizaron sobre células MT4, en medio líquido, mezclando previamente las semillas virales al 50%. Tras un tiempo de adsorción, se lavaron las células, y se siguieron los cultivos durante 22 días, añadiendo células cada 3 ò 4 días, según se explica en Materiales y Métodos, y como se representa en la figura IV.6.



**Fig. 4.6.** Esquema del protocolo utilizado para el seguimiento de las infecciones dobles. Una vez infectadas las células, se recoge sobrenadante los días 5, 9, 13, 16 y 19; inmediatamente después se añade a los cultivos células y medio frescos. La cuantificación de las distintas especies virales se hizo sobre el ADN proviral resultante de la infección de células frescas con los sobrenadantes del cultivo.

Para cuantificar los recombinantes presentes en los sobrenadantes recogidos, quisimos adaptar la técnica de detección descrita en el apartado 4.1.3 a una reacción RT-PCR, pero

esta técnica producía artefactos que impidieron su utilización. La enzima RT utilizada en esta reacción genera formas recombinantes a partir de una mezcla de dos virus parentales (este aspecto se comenta ampliamente en el apartado 5). Para evitar esa posibilidad la opción elegida fue llevar a cabo una una infección de células frescas con estos sobrenadantes, que se recogieron 20 a 24 horas después (ver Materiales y Métodos) para permitir que tuviese lugar la transcripción inversa de los viriones presentes en el sobrenadante, sin que llegase a producirse un ciclo de replicación nuevo, ya que a las 24 horas se estima que se ha alcanzado un máximo en la cantidad de ADN vírico en el núcleo celular (Reddy y Yin, 1999). La cuantificación se realizó sobre el ADN proviral presente en las células de estas infecciones.

#### 4.3.1 Relación entre la distancia genética que separa los virus parentales y la proporción de virus recombinantes que aparecen en el cultivo.

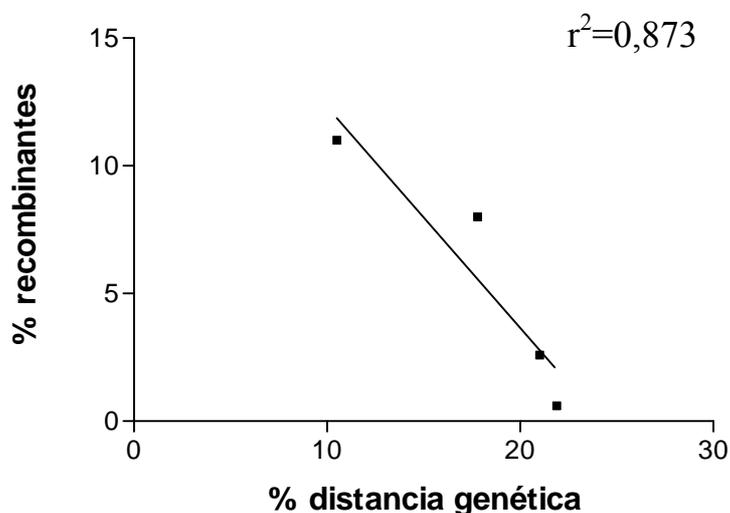
Para comprobar los resultados obtenidos en el apartado 4.2.4, donde se encontraron mayores frecuencias de recombinación entre los virus filogenéticamente más cercanos, se llevó a cabo un análisis cuantitativo de la proporción de virus recombinantes en cuatro infecciones dobles: tres de ellas entre subtipos diferentes (B+D, B+F, D+F) y una intrasubtipo (LAI+RF), descritas en el apartado 3.1.5. Se eligieron los sobrenadantes del día 5 para minimizar la influencia de los factores selectivos sobre los recombinantes que aparecen. Los resultados de las cuantificaciones realizadas se muestran en la tabla 4.5.

**Tabla 4.5. Resultados de la cuantificación de las 4 especies genómicas presentes en los sobrenadantes del día 5.**

<b>B+D</b>	%	$\delta$	<b>B+F</b>	%	$\delta$	<b>D+F</b>	%	$\delta$	<b>LAI+RF</b>	%	$\delta$
B	73	8,14	B	37	8,1	D	94	2,08	LAI	53	11,1
D	18	3,84	F	60	8,77	F	5	3,3	RF	36	12,2
Rec BD	5	3,38	Rec BF	2,4	2,54	Rec DF	0,6	0,8	Rec LR	10	2,0
Rec DB	3	3,1	Rec FB	0,2	0,21	Rec FD	0	0,0	Rec RL	1	2,3
RecBD+DB	8		Rec BF+FB	2,6		Rec DF+FD	0,6		Rec LR+RL	11	

La distancia genética entre los virus parentales se calculó mediante el alineamiento de las secuencias del gen env en el programa informático MEGA2. El resultado fue un 10,5% entre Lai y RF, un 17,8% entre UG24 (D) y Lai (B), un 21% entre Lai y BR20(F), y un 21.9% entre UG24 y BR20.

Cuando se representó la relación entre estas dos variables se obtuvo la gráfica que aparece en la figura 4.7.



**Fig. 4.7.** Representación gráfica de la relación entre el porcentaje de genomas recombinantes que aparece en la población viral resultante de una infección doble, y la distancia genética que existe entre los dos virus parentales. Los 4 valores representados son los que aparecen en la tabla 4.6, y que corresponden a 3 infecciones intersubtipo y una intrasubtipo.

Las proporciones de recombinantes obtenidas confirman los resultados de los experimentos preliminares. Los porcentajes obtenidos guardan una relación con las distancias genéticas entre los virus parentales.

En la coinfección D+F, podemos ver que el virus D constituye el 94% de la progenie, a día 5 post infección. La baja proporción del virus F en el cultivo puede explicar por sí misma la falta de recombinantes, ya que disminuye de manera considerable la probabilidad de que las células sean coinfectadas por un virus D y otro F.

Las infecciones realizadas con el subtipo A, que es el más distante filogenéticamente, no pudieron ser analizadas debido a que no pudimos poner a punto un buen sistema de iniciadores específicos. Sin embargo, pudimos utilizar los sobrenadantes del día 22 de las infecciones A+B, A+D y A+F, para realizar un plaqueo. El análisis de las placas obtenidas (23, 36 y 27 respectivamente) se realizó por el método de identificación de placas descrito en el apartado 4.1.1. y el resultado fue la ausencia de placas pertenecientes al subtipo A, así como de cualquier tipo de recombinante. Este resultado es similar al que obtenemos al analizar el sobrenadante del día 22 de la infección D+F, como puede verse en la tabla 4.6. Dado que el subtipo A es el más distante filogenéticamente de los subtipo usados en este

estudio, estos resultados concuerdan con todos los obtenidos anteriormente, aunque no puedan ser incluidos en el análisis cuantitativo.

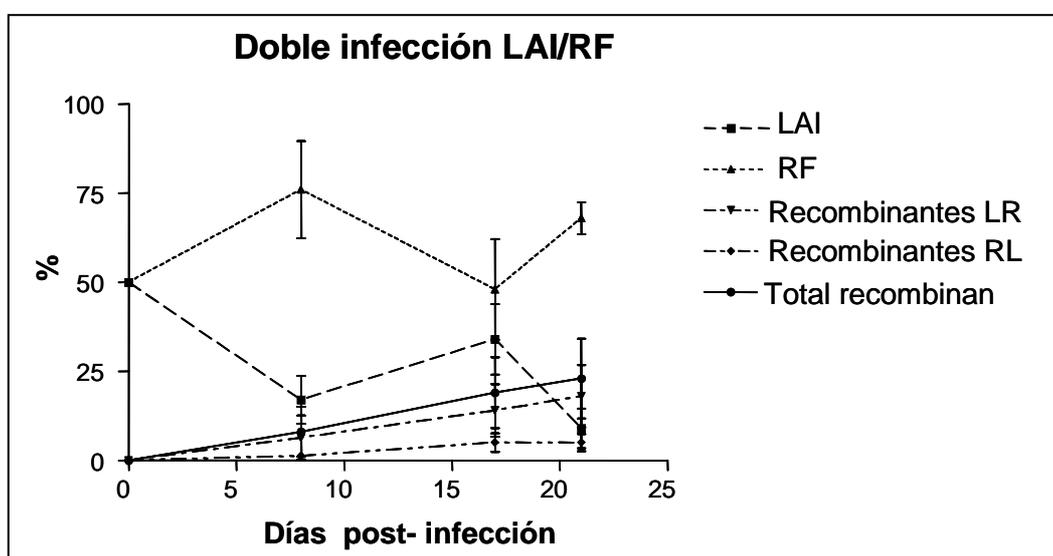
#### 4.3.2. Evolución de la proporción de los virus recombinantes con el tiempo en dos infecciones dobles.

Para estudiar la cinética de las cuatro especies genéticas que podemos cuantificar en un cultivo, elegimos dos de las infecciones dobles realizadas, una intersubtipo (B+F) y otra intrasubtipo (LAI+RF). Se realizó el análisis cuantitativo de todos los sobrenadantes recogidos a lo largo de los cultivos.

Los valores obtenidos pueden verse en las tablas 4.6 y 4.7, y las figuras 4.8 y 4.9.

**Tabla 4.6. Resultados de la cuantificación de los sobrenadantes recogidos a lo largo del cultivo de la infección LAI+RF (Intrasubtipo B)**

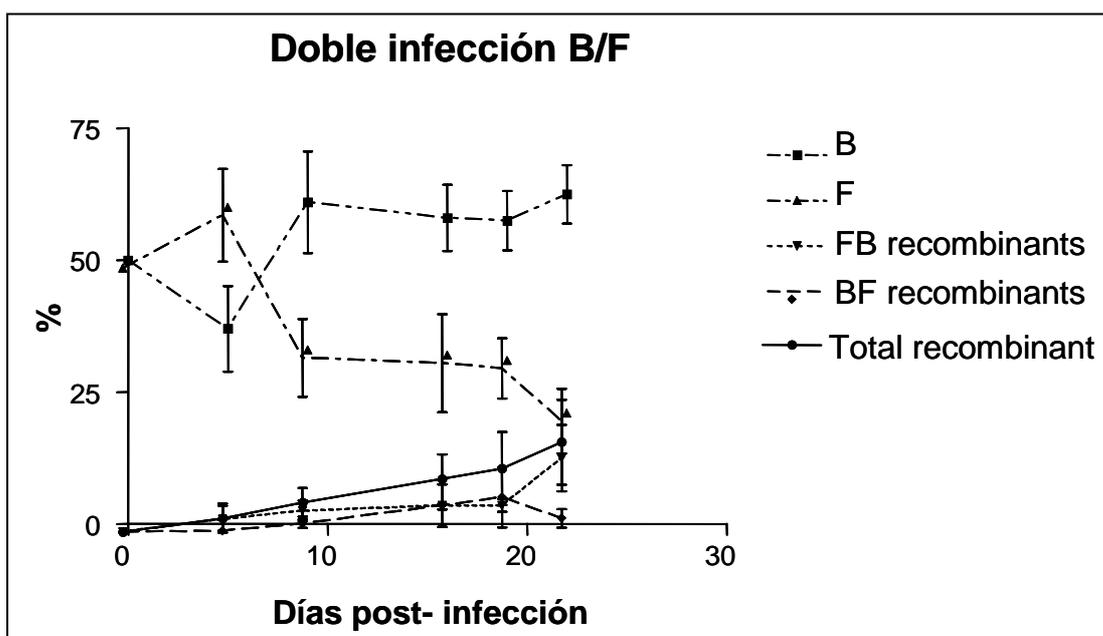
LAI+RF	D8	$\delta$	D17	$\delta$	D21	$\delta$	Pendiente
<b>L</b>	17	6,7	34	9,88	9	5,55	-1.366 $\pm$ 0.9776
<b>R</b>	76	13,6	48	14,1	68	4,46	+0.2604 $\pm$ 1.014
<b>Rec LR</b>	6,4	6,21	14	7,35	18	8,75	+0.8513 $\pm$ 0.2224
<b>Rec RL</b>	1,3	0,72	5	2,62	5	2,43	+0.2658 $\pm$ 0.04426
<b>Rec LR+RL</b>	8		19		23		+1.113 $\pm$ 0.03375



**Fig. 4.8.** Dinámica de la competición que se establece entre las cuatro especies genómicas en la población viral que resulta de la infección doble LAI/RF. Es el resultado del análisis cuantitativo llevado a cabo con los sobrenadantes de los días 8,17 y 22.

**Tabla 4.7. Resultados de la cuantificación en los sobrenadantes recogidos a lo largo del cultivo de la infección B+F(Intersubtipo)**

B+F	D5	$\delta$	D9	$\delta$	D16	$\delta$	D19	$\delta$	D22	$\delta$	Pendiente
<b>B</b>	37	8,10	61	9,66	58	6,30	57,5	5,63	62,5	5,51	+0.7369 $\pm$ 0.4129
<b>F</b>	60	8,77	33	7,37	32	9,29	31	5,69	21	6,08	-1.434 $\pm$ 0.4306
<b>Rec BF</b>	2,4	2,54	4	1,98	5	0,70	5	1,21	14	6,31	+0.4669 $\pm$ 0.1517
<b>Rec FB</b>	0,2	0,21	1,6	0,8	5	4,03	6,6	5,69	2,6	1,75	+0.2493 $\pm$ 0.0912
<b>Rec BF+FB</b>	2,6		5,6		10		12		17		+0.7161 $\pm$ 0.0750



**Fig. 4.9.** Dinámica de la competición que se establece entre las cuatro especies genómicas en la población viral que resulta de la infección doble LAI/BR20, pertenecientes a los subtipos B y F respectivamente. Es el resultado del análisis cuantitativo llevado a cabo con los sobrenadantes de los días 5, 9, 19 y 22.

En ambos casos se produjo un aumento progresivo de la proporción de genomas recombinantes, a juzgar por las pendientes continuas de las líneas que representan la suma de las dos formas recombinantes.

Los resultados obtenidos en estos dos estudios ponen de manifiesto una tendencia hacia la imposición de una de las formas recombinantes sobre una de las parentales.

La regresión lineal de los valores determinados, nos permitió el cálculo de la pendiente que corresponde a cada especie genómica, que puede ser considerada como una medida de su eficacia biológica. La pendiente correspondiente a las formas recombinantes (ver tablas 4.6 y 4.7) resultó siempre superior a la de una de las especies parentales. Incluso en el caso de

la infección R/L, la pendiente obtenida para los recombinantes RL, fue superior a la de cualquiera de los virus parentales.

Aunque en la coinfección B+D, no se ha realizado este estudio de la cinética de aparición de especies recombinantes, la comparación de los valores obtenidos a partir de los sobrenadantes de los días 5 y 22, ponen de manifiesto el mismo fenómeno: la imposición de las formas recombinantes BD sobre el parental D.

**Tabla 4.8. Variación de la proporción de recombinantes en la infección B/D entre el día 5 y 22.**

<b>B+D día 5</b>	<b>%</b>	<b>B+D día 22</b>	<b>%</b>
<b>B</b>	73	<b>B</b>	70
<b>D</b>	18	<b>D</b>	7
<b>Rec BD</b>	5	<b>Rec BD</b>	20
<b>Rec DB</b>	3	<b>Rec DB</b>	2
<b>RecBD+DB</b>	8	<b>Rec BD+DB</b>	22

#### **4.4. CONSECUENCIAS EVOLUTIVAS INTRODUCIDAS POR LA RECOMBINACIÓN EN UNA POBLACIÓN FORMADA INICIALMENTE POR DOS VIRUS DIFERENTES.**

##### **4.4.1 Competición entre virus recombinantes y parentales.**

La cuantificación de las proporciones de las cuatro especies genómicas en el sobrenadante del día 22 de las cuatro infecciones dobles, cuyos datos aparecen en la tabla 4.9, nos permitió comprobar que la proporción de recombinantes aumentó con el tiempo en tres de los cuatro casos. En estas tres infecciones, los porcentajes de recombinación aumentaron respecto al día 5, de un 11 a un 23% en la coinfección intrasubtipo, de un 8 al 22% en el caso B/D, y de un 2,6 a un 17% en el cultivo B/F. Sin embargo, en las células infectadas con el sobrenadante de la doble infección D/F, los genomas recombinantes fueron prácticamente indetectables, menos aún que a día 5. Del mismo modo que en ese sobrenadante, la proporción del virus D era muy superior a la del subtipo F (94/5), a día 22 esta relación es aún mayor (98/2), disminuyendo más la posibilidad de coinfección de una célula, y por tanto de recombinación.

**Tabla 4.9. Resultados de la cuantificación de los sobrenadantes del día 22.**

$B+D^a$	%	*	$B+F^b$	%	$\delta$	$D+F^c$	%	$\delta$	$LAI+RF^d$	%	$\delta$
B	70		B	62,5	5,5	D	98	1	LAI	9	5,6
D	7		F	21	6,1	F	2	0,6	RF	68	4,5
Rec BD	20		Rec BF	14	6,3	Rec DF	0	0,4	Rec LR	18	8,8
Rec DB	2		Rec FB	2,6	1,8	Rec FD	0	0,0	Rec RL	5	2,4
Rec BD+DB	22		Rec BF+FB	17		Rec DF+FD	0		Rec LR+RL	23	

a) Coinfección de  $Lai=B$  (subtipo B) y  $UG24=D$  (subtipo D)

b) Coinfección de  $Lai=B$  (subtipo B) y  $93BR30=F$  (subtipo F)

c) Coinfección de  $UG24=D$  (subtipo D) y  $93BR30=F$  (subtipo F)

d) Coinfección de  $Lai=L$  (subtipo B) y  $RF$  (subtipo B)

\* La desviación estándar de este resultado no pudo calcularse debido a que sólo existía un dato

A la vista de los resultados obtenidos hasta este momento, podemos decir, que en una infección in vitro de células MT4, el primer factor limitante de la recombinación es la probabilidad de coinfección de las células. Así, en el sobrenadante del día 22 del cultivo D/F no encontramos recombinación porque en los primeros ciclos, se impone fuertemente un virus parental.

#### **4.4.2. Estudio de los mecanismos responsables del aumento de la proporción de virus recombinantes con el tiempo. Selección positiva de clones recombinantes o numerosos eventos de recombinación.**

La comparación de los resultados obtenidos en el análisis cuantitativo de los sobrenadantes de los días 5 y 22 de las infecciones dobles, (tablas 4.5 y 4.9) nos llevó a preguntarnos si el aumento de formas recombinantes sería progresivo como resultado de la suma de los recombinantes creados en cada ciclo de replicación o, por el contrario, si se producirían expansiones clonales de algunos virus recombinantes con una eficacia biológica superior al resto.

##### **4.4.2.1. Análisis filogenético de clones biológicos recombinantes aislados a partir del día 9 y 22.**

Para estudiar la naturaleza del fenómeno de imposición de virus recombinantes en la progenie de las infecciones realizadas, se llevó a cabo un análisis filogenético en dos fragmentos de las zonas *env* y *pol* de genomas recombinantes clonados biológicamente. Se utilizaron un total de 22 clones recombinantes aislados a partir de los sobrenadante de los días 9 y 22 de la doble infección B/D.

La identificación de estos clones se realizó con la técnica descrita en el apartado 4.1.1. De esta manera, se seleccionaron 7 clones recombinantes aislados a partir del sobrenadante del

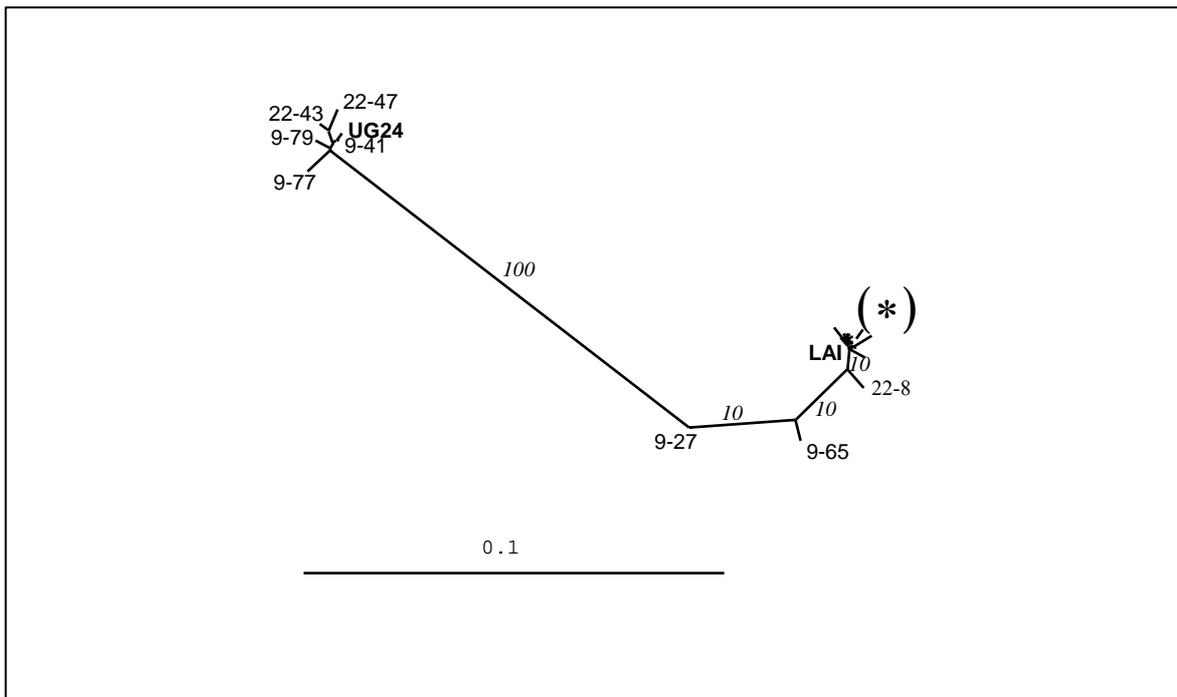
día 9 de la coinfección B/D (entre un total de 202 placas, es decir un 4%), denominados : 9-16, 9-27, 9-41, 9-48, 9-65, 9-79, 9-138. Los 12 clones aislados a partir del sobrenadante del día 22 (entre 97 placas, un 14% del total), fueron llamados : 22-2, 22-8, 22-24, 22-28, 22-34, 22-43, 22-47, 22-63, 22-67, , 22-83, 22-87, y 22-90. Entre ellos se incluyeron 3 clones identificados como parentales por este método, como control negativo de recombinación, que son: 9-77, (en el que los dos fragmentos amplificados pertenecen al subtipo D), 22-31 y 22-82 (que se identifican con el subtipo B).

Se secuenciaron dos fragmentos de 612 y 798 nt., que habían sido descritos como zonas con alta frecuencia de recombinación (Jetzt y *cols.*, 2000) con la finalidad de encontrar el máximo nº de puntos de recombinación. Los fragmentos secuenciados son los siguientes:

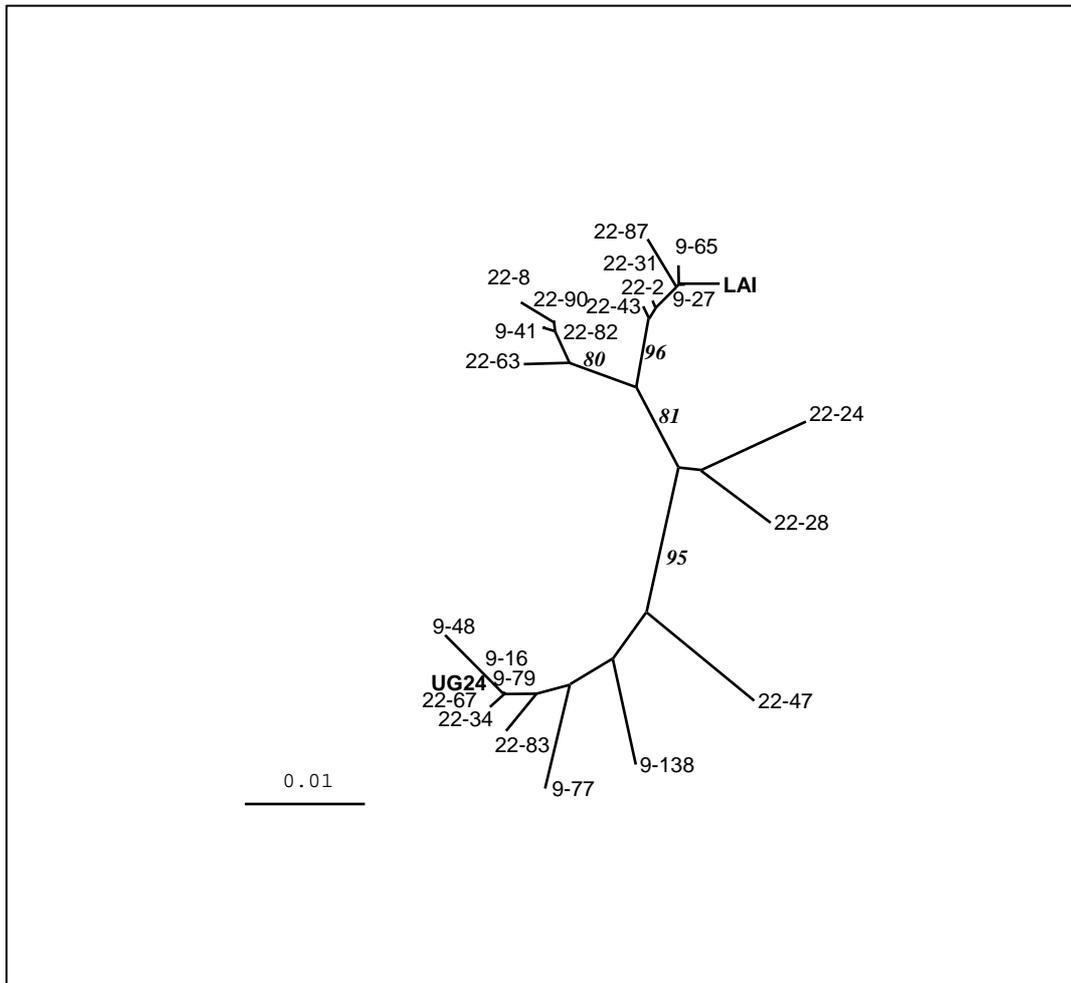
1.- Desde el nucleótido 2102 al 2714 (en la numeración de LAI), que comprende un pequeño fragmento del final de la proteasa, seguido de la primera mitad de la secuencia de la Transcriptasa Inversa.

2.- Desde el nucleótido 5850 hasta el 6648 que incluye un fragmento del final del péptido señal, seguido de varias zonas de la gp120 denominadas C1, V1, V2 y C2. Según se describe en el apartado 3.6.1.

El alineamiento de las secuencias, nos permitió obtener los árboles filogenéticos que aparecen en la figura 4.10 y 4.11.



**Fig. 4.10.** Análisis filogenético de las secuencias nucleotídicas de un fragmento de 798 nucleótidos de la región env, de 22 clones biológicos (19 recombinantes y 3 controles) hallados a partir de los sobrenadantes de los días 9 y 22 de la coinfección B/D. El análisis se llevó a cabo como se describe en Materiales y Métodos. En el alineamiento de secuencias se incluyeron las de los dos virus parentales con los que se realizó la infección, que aparecen en negrita. Los números en cursiva que aparecen en las ramas representan los valores de remuestreo ("bootstrap") y sólo están indicados aquellos valores superiores al 80%. La barra de la escala representa un 10% de distancia genética. El asterisco representa los clones 22-2, 22-24, 22-28, 22-31, 22-34, 22-63, 22-67, 22-82, 22-83, 22-87, 22-90, 9-16, 9-48, 9-138



**Fig. 4.11.** Análisis filogenético de las secuencias nucleotídicas de un fragmento de 612 nucleótidos de la región *pol*, de 22 clones biológicos recombinantes hallados a partir de los sobrenadantes de los días 9 y 22 de la coinfección B+D. El análisis se llevó a cabo como se describe en materiales y métodos. En el alineamiento de secuencias se incluyeron las de los dos virus parentales con los que se realizó la coinfección, que aparecen en negrita. Los números en cursiva que aparecen en las ramas representan los valores de remuestreo ("bootstrap") y sólo están indicados aquellos valores superiores al 80%. La barra de la escala representa un 1% de distancia genética.

En el árbol resultante del análisis filogenético de las secuencias de *env* (Fig.4.10.), pueden observarse valores de remuestreo del 100% debido a una gran distancia entre las secuencias parentales (aproximadamente un 15%), así como un claro agrupamiento de la mayor parte de estas secuencias alrededor de las dos parentales. En cada uno de los dos grupos podemos encontrar secuencias pertenecientes al día 9 y al día 22 que no parecen

agrupadas entre sí, de manera que no hay evidencia de ningún clon del día 9, que fuera el origen de clones del día 22.

Pueden verse, además, tres clones (9-27, 9-65 y 22-8) que se encuentran a diferentes distancias dentro de la rama que conecta ambos subtipos, debido a que poseen un punto de recombinación dentro de la secuencia que ha sido analizada. Su separación en el árbol indica que son el fruto de eventos de recombinación independientes.

En el árbol resultante del alineamiento de las secuencias que pertenecen al gen *pol*, se puede apreciar una menor distancia genética entre las secuencias parentales (6,5%) y menores valores de remuestreo, pero puede observarse una mayor diversidad entre la mayor parte de los clones, y mayor dispersión entre ellos. Sin embargo, también podemos ver dos grupos de secuencias: uno en torno a LAI (secuencias semejantes al parental) y un segundo grupo constituido por 5 clones: 9.41, 22.8, 22.63, 22.82 y 22.90. Este agrupamiento, podría ser el resultado de una expansión de un clon inicial, que estaría muy relacionado con el 9-41.

Por otro lado, hay 3 secuencias recombinantes independientes, con puntos de recombinación distintos, que son el 22.47, el 9.138 y el 9.77, y un dúo constituido por los clones 22-24 y 22-28.

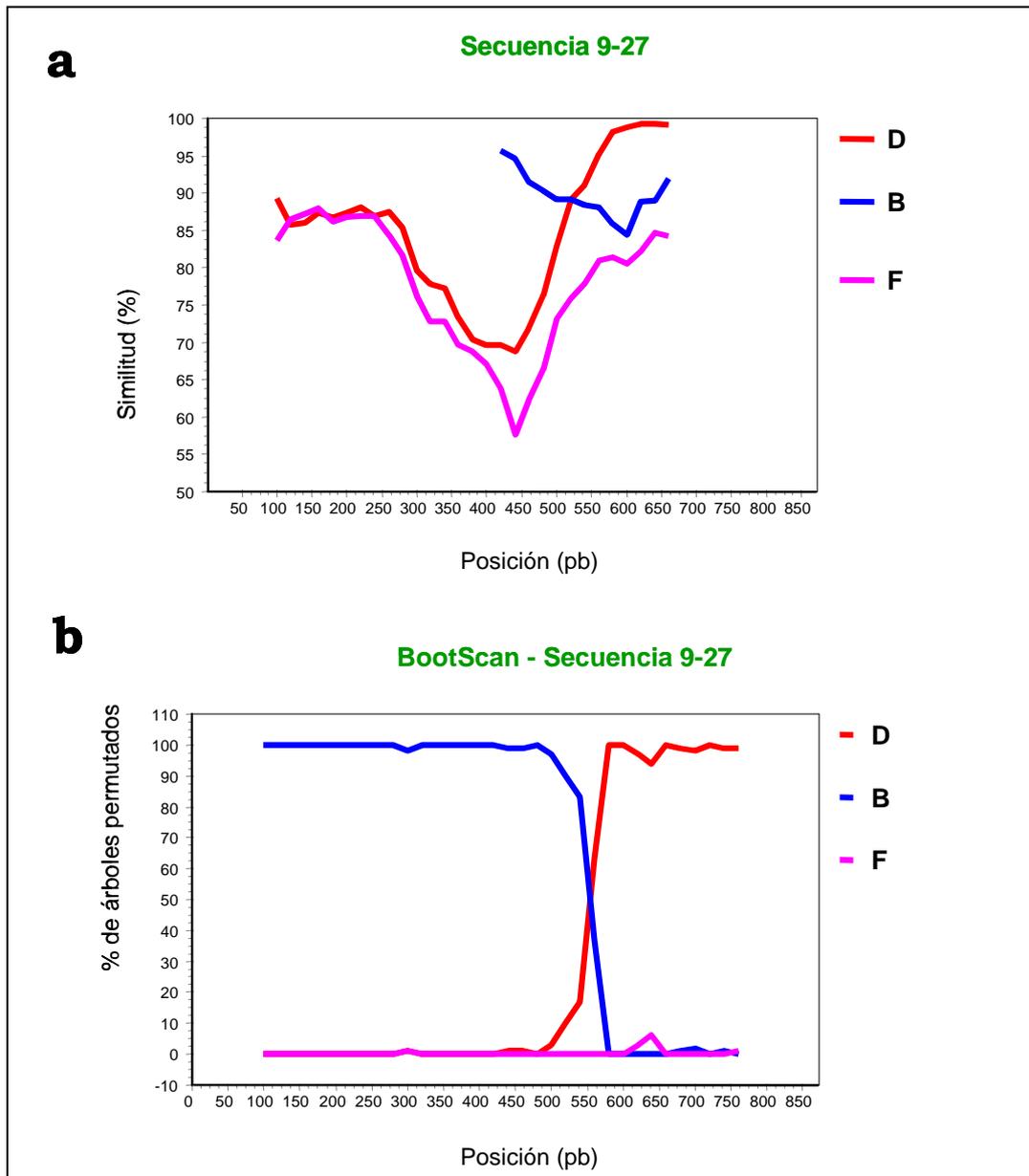
Estos datos parecen apuntar a que dos procesos han intervenido en el aumento de recombinantes en nuestros cultivos: múltiples eventos de recombinación independientes, y la expansión clonal de uno de estos productos de recombinación (9.41).

#### **4.4.2.2. Análisis de recombinación de los clones biológicos.**

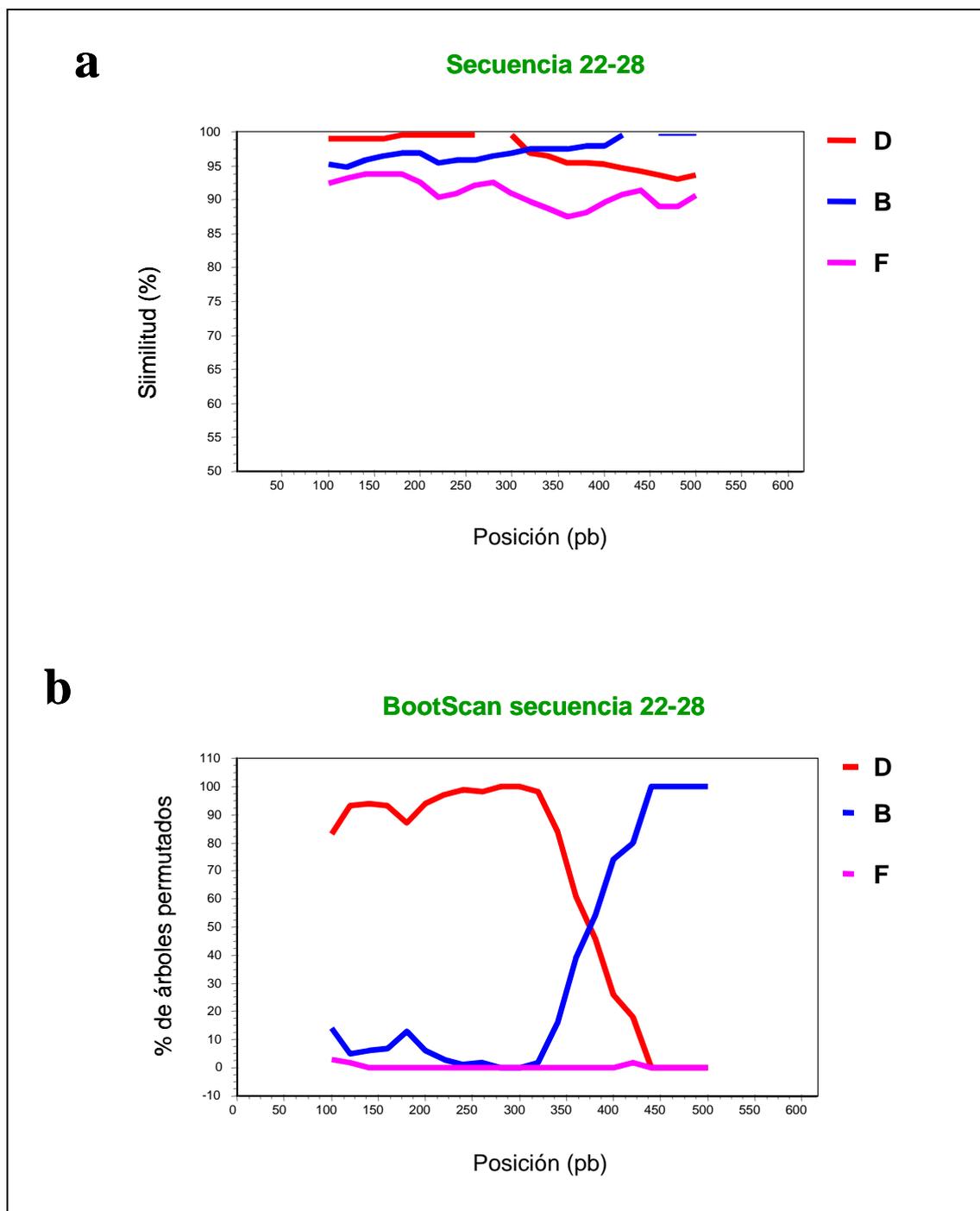
Un estudio posterior, de las secuencias gracias al programa “Simplot”, permitió identificar los puntos de recombinación en estos clones. El análisis fue realizado con una ventana de 200 nucleótidos, y un paso de ventana de 20 nt., y se utilizó como grupo externo la secuencia nucleotídica del aislado 93BR20 que pertenece al subtipo F. El programa informático Simplot (descrito en el apartado 3.9) compara la similitud de la secuencia problema con las secuencias de los virus parentales LAI y UG24. El programa genera un árbol filogenético para cada uno de los fragmentos de 200 nucleótidos en cada paso (cada 20nt). A la vez, valora el significado estadístico de las asociaciones que se establecen en cada fragmento por el remuestreo de 1000 réplicas (“bootstrap”) de las secuencias analizadas.

Se realizó el análisis de todos los clones biológicos aislados. En las figuras 4.12 a 4.14 se muestra el resultado de alguno de los clones en los que se detectó algún punto de recombinación.

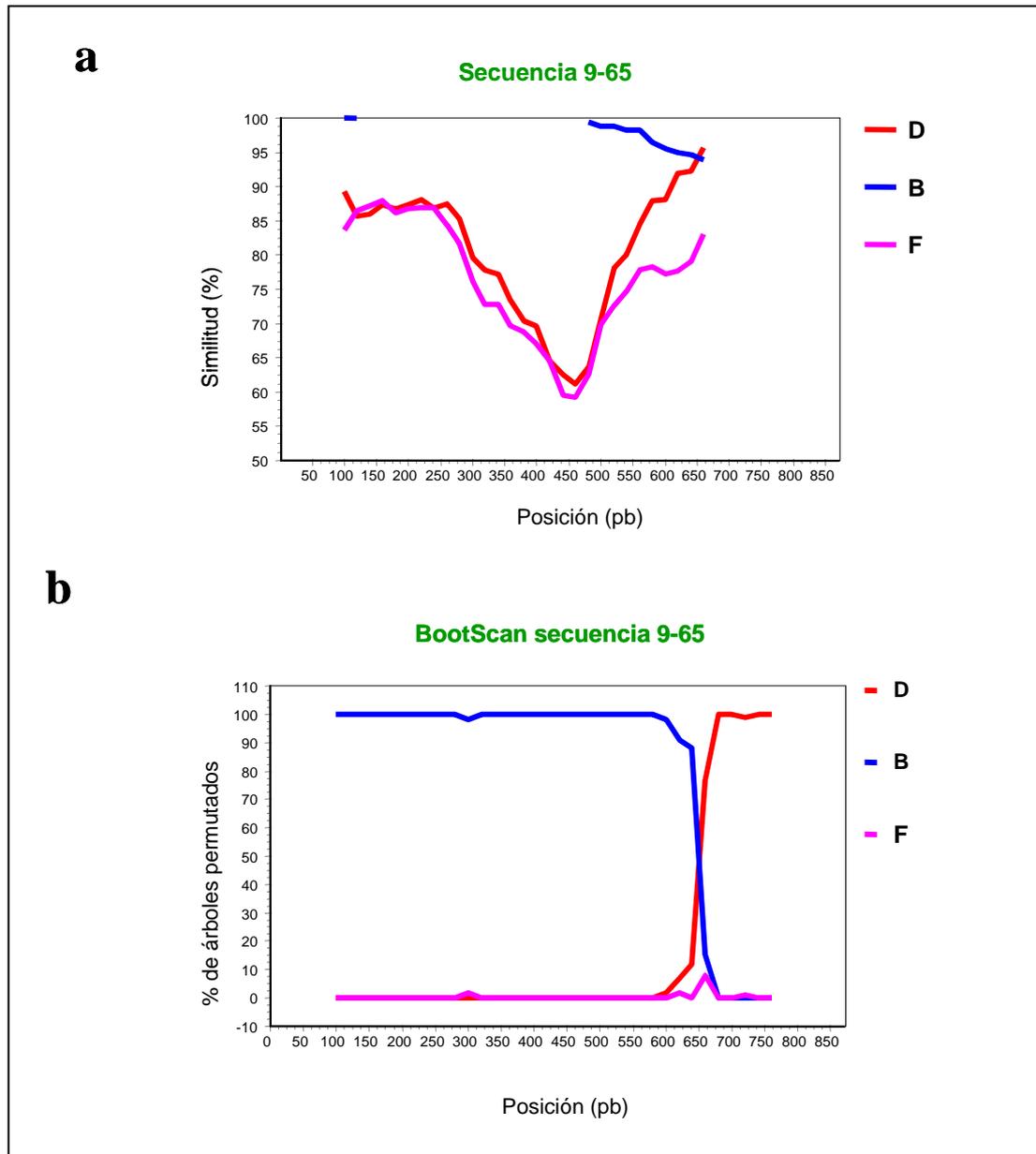
La figura 4.15 resume esquemáticamente el resultado del análisis de recombinación.



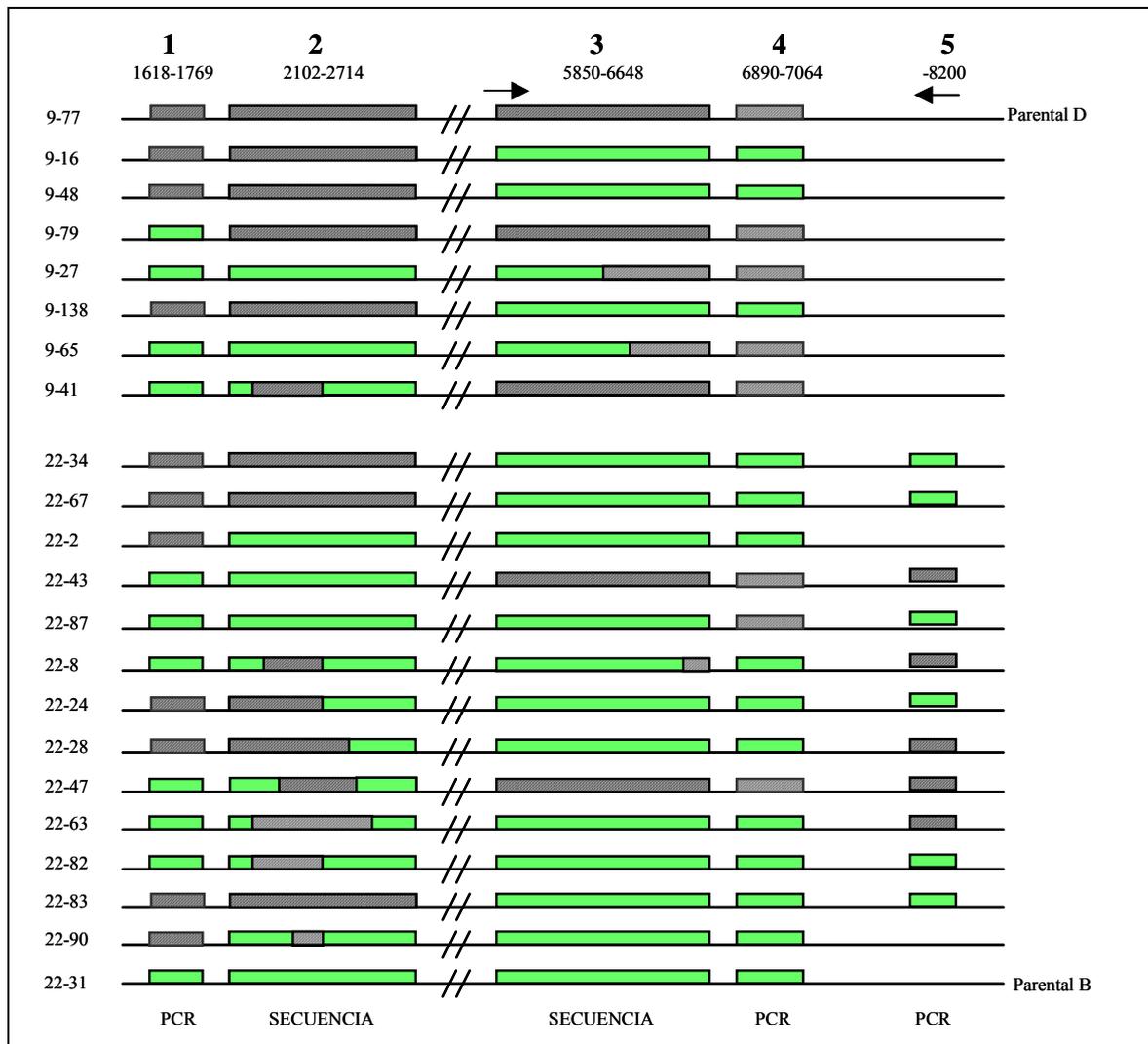
**Fig. 4.12.** Análisis de recombinación genética de la secuencia “env” del clon 9-27. (a) Similitud genética a lo largo del gen pol entre la secuencia supuestamente recombinante y las secuencias de los virus parentales. (b) Análisis por remuestreo de 1000 réplicas de las relaciones filogenéticas que se establecen entre las secuencias analizadas en los fragmentos generados en el análisis de similitud.



**Fig. 4.13.** Análisis de recombinación genética de la secuencia “pol” del clon 22-28. (a) Similitud genética a lo largo del gen pol entre la secuencia supuestamente recombinante y las secuencias de los virus parentales. (b) Análisis por remuestreo de 1000 réplicas de las relaciones filogenéticas que se establecen entre las secuencias analizadas en los fragmentos generados en el análisis de similitud.



**Fig. 4.14.** Análisis de recombinación genética de la secuencia “env” del clon 9-65. (a) Similitud genética a lo largo del gen pol entre la secuencia supuestamente recombinante y las secuencias de los virus parentales. (b) Análisis por remuestreo de 1000 réplicas de las relaciones filogenéticas que se establecen entre las secuencias analizadas en los fragmentos generados en el análisis de similitud.



**Fig. 4.15.** Análisis de clones recombinantes obtenidos de la infección B/D a día 9 y 22. El análisis fue realizado con la ayuda del programa Sim-Plot. Cada línea negra representa el genoma proviral de un clon analizado. Las secuencias analizadas se representan por barras, y sus posiciones dentro del genoma se señalan en la parte superior de la figura. El color verde representa secuencias del virus B y el gris rayado representa las secuencias del virus D. Los números que aparecen a la izquierda son la denominación de cada clon, el primero de ellos (9 ò 22) hace referencia al día del seguimiento del cultivo del que procede el clon. Las flechas indican la localización de los iniciadores utilizados en el sistema de cuantificación.

El programa detectó 17 puntos de corte dentro de las secuencias analizadas. Estos 17 puntos se encontraron dentro de 10 clones, el resto, que son 12 clones, no presentaron ninguno. De estos diez clones, 4 presentaron un solo punto de recombinación, 5 de ellos presentaron dos puntos, y uno (el 22-8) tiene 3.

**Tabla 4.10. Clasificación de los fragmentos analizados.**

	<b>Secuencias POL</b>	<b>Secuencias ENV</b>	Secuencias totales
<b>Día 9</b>	2 parental B 5 parental D 1 Recombinante	2 parental B 3 parental D 2 Recombinantes	3 Recombtes
<b>Día 22</b>	4 parental B 3 parental D 7 Recombinantes	11 parental B 2 parental D 1 Recombinante	8 Recombtes
Secuencias totales	8 Recombtes	3 Recombtes	11 Recombtes

Existen además, ocho clones en los que las dos secuencias estudiadas se identifican como procedentes de cada uno de los dos virus parentales. Esto significa que en el fragmento intermedio entre ellas, existe al menos otro punto de corte. Si este resultado lo sumamos al anterior, encontramos un total de 25 puntos de corte en 17 clones (1,47puntos por clon), siendo 11 clones los que presentan 1 punto, 4 de ellos 2, y dos tienen 3puntos.

Cuando añadimos la información genética de otros fragmentos de estos clones (que nos aporta la prueba de identificación de clones recombinantes, descritas en materiales y métodos) pudimos ver otros puntos de recombinación indirectamente localizados entre las secuencias identificadas como parentales-

**Tabla 4.11. Puntos de corte entre secuencias parentales encontrados en los clones biológicos .**

Procedencia de clones	Nº pto recombina dentro de los fragmentos secuenciados	Nº de puntos recombina totales	Promedio de puntos de recombina por clon
Día 9	4	9	1,1
Día 22	13	22	1.6
Totales	17	31	1,4

Hay una apreciable diferencia entre el número de puntos de corte encontrados en cada una de las secuencias: Los puntos de recombinación identificados, son sólo tres en el fragmento *env*, y 14 en el fragmento *pol*. Este resultado puede explicarse por la relación

entre distancia genética y frecuencia de recombinación, ya que el fragmento secuenciado en pol tiene un grado de similitud entre Lai y UG24, mucho mayor que el fragmento env. La mayor parte de los puntos de corte en las secuencias se encuentran en distinta posición, aunque existen algunos comunes. Así tenemos que los clones 9-41, 22-8, 22-63, 22-82, y 22-90, comparten un punto de recombinación idéntico, y por esta razón aparecen agrupados en el árbol filogenético. Lo cual apoya la hipótesis de una expansión del clon 9-41. Sin embargo, sólo el 9-41 y el 22-82, comparten un 2º punto de corte. El resto tienen una secuencia diferente. Es decir, aunque pueden ocurrir episodios de selección positiva de determinados clones recombinantes dentro de la progenie, nuevos eventos de recombinación se van sumando en cada ciclo de replicación (incluso dentro de los virus recombinados). Además, aunque estos cinco clones estén relacionados filogenéticamente, no es seguro que tengan un mismo origen, ya que su relación podría explicarse por la existencia de zonas “calientes” para recombinación.

Los datos indican un aumento de puntos de recombinación por clon con el tiempo de 1,1 en el día 9, a 1,6 el día 22. Como puede verse en la tabla 4.11, la proporción de clones con un solo punto de corte disminuye el día 22 a favor de los que poseen más de un corte, indicando que los eventos de recombinación son más probables en un virus recombinante que en uno parental. La explicación puede estar de nuevo en el hecho de que existe mayor similitud de secuencia entre un virus recombinante y uno parental que entre dos parentales.

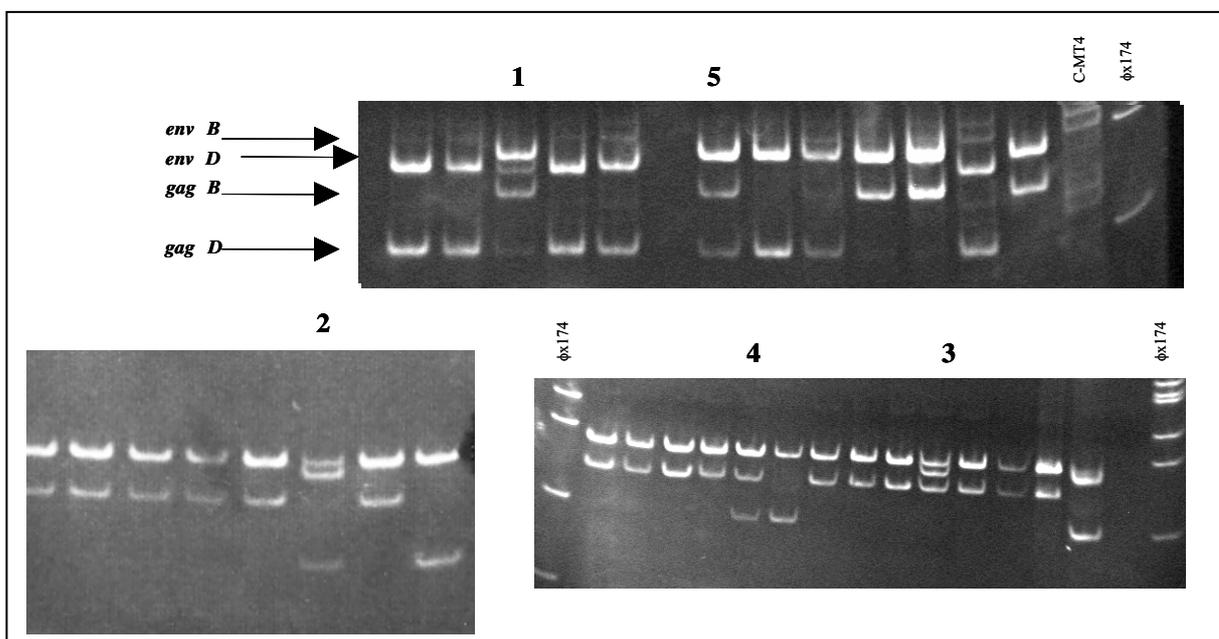
**Tabla 4.12. Proporción de clones según el nº de puntos de recombinación que presentan.**

Procedencia de clones	Clones con 1 pto de corte	clones con 2 ptos	clones con 3 puntos	Clones con 4 puntos	Clones con 5 puntos	Total clones recombinantes
Día 9	78 %	0 %	12,5 %	0 %	0 %	88 %
Día 22	57 %	14 %	14 %	7 %	0 %	93 %
Totales	73 %	9 %	13,6 %	4,5%	0 %	91 %

#### 4.5. ESTUDIO SOBRE LA NATURALEZA DIPLOIDE O PSEUDODIPLOIDE DEL VIH-1.

El VIH-1 contiene en su virión dos copias de ARN genómico completas. Se cree que sólo se origina una copia de ADN proviral a partir de ellas, por lo que se considera que el VIH-1 tiene una organización genómica pseudodiploide (Hu y Temin, 1990b). Sin embargo, se ha demostrado que cada una de estas hebras de ARN es suficiente para completar una reacción de transcripción inversa y dar origen a una copia completa de ADN (Jones y cols., 1994).

A lo largo de los múltiples experimentos de aislamiento de clones biológicos por plaqueo, llevados a cabo con los sobrenadantes de los cultivos realizados en la primera parte de este estudio, encontramos una serie de placas de infección, que contenían dos genomas provirales diferentes. En la Fig. 4.16 pueden verse algunas de estas placas que presentan tres o cuatro bandas, frente a las que presentan dos, que corresponden a un solo genoma proviral (ver también figura 4.2).



**Fig 4.16.** Caracterización por doble PCR de clones biológicos aislados a partir de sobrenadantes de las infecciones dobles B/D. La PCR amplifica una región del genoma en *gag* y otra en *env* que tienen distintos tamaños para cada uno de los virus parentales (ver figura 4.2). En estos geles aparecen incluidas placas que hemos llamado “heterozigotas” por presentar dos genomas diferentes: las que identificamos con los números 1 y 2 son placas que contienen los dos genomas parentales LAI+UG24, la número 3 contiene un genoma LAI más un recombinante BD (la secuencia de *gag* pertenece al subtipo B y la secuencia de *env* al D). Las placas 4 y 5 tienen un genoma del parental B y un recombinante DB.

Tres hipótesis pueden explicar la formación de las placas que contienen dos provirus:

**1.-** Una célula puede ser infectada simultáneamente por dos virus diferentes (**co-infección**), cada uno de ellos origina una copia de ADN, integrándose los dos dentro de la célula.

**2.-** La mayor parte de las placas se originan a partir de la infección de una célula por un virus, pero puede haber algunas formadas a partir de dos procesos de infección independientes en células muy cercanas, que denominamos **placas doblemente infectadas**.

**3.-** Una célula puede ser infectada por un **virión heterocigoto**, con dos hebras de ARN diferentes, y cada una de ellas da origen a una copia de ADN distinta.

#### **4.5.1 Análisis del proceso de formación de una placa.**

Para tratar de dilucidar cual de estas hipótesis explica mejor nuestros datos, realizamos una aproximación al proceso de formación de una placa.

Las placas de infección originadas por el VIH-1 en una monocapa de células MT4, son una acumulación de células infectadas, apreciables a simple vista como condensaciones de la monocapa, que aparecen hacia el quinto o sexto día post-infección.

Teóricamente, al igual que en otros sistemas de plaqueo, una placa se origina a partir de la infección de una célula por un virus.

En primer lugar, se hizo una estimación del número de células que contiene una placa visible, a partir del conteo de células en 9 placas pinchadas, y el promedio resultó ser del orden de 15000 células. Sin embargo, la extensión de monocapa celular que se recoge con la micropipeta es varias veces superior al tamaño de la placa de infección, de manera que estimamos que el número de células que realmente contiene la placa puede ser del orden de  $10^3$ .

Por otro lado, estimamos el número de u.f.p que libera una placa. El título medio fue del orden de  $10^4$  a  $10^5$  u.f.p. Por tanto, asumiendo que un virus da origen a un mínimo de 10 nuevos virus en un ciclo de replicación, la formación de una placa requiere al menos dos o tres ciclos de replicación viral: en el primer ciclo, una célula origina un mínimo de 10 virus que infectan otras tantas células; en el segundo ciclo se producirán 100 virus y se infectan 100 células de alrededor; y en el tercer ciclo llegan a infectarse 1000 células más haciendo

la placa visible, y siendo las responsables de originar  $10^4$  u.f.p. que es un título común a ellas.

La técnica de plaqueo se realizó en todos los casos a una multiplicidad de infección muy baja para tratar de evitar la aparición de placas doblemente infectadas. Se utilizaron aproximadamente 20 u.f.p. por cada  $3,3 \times 10^6$  de células, o lo que es igual, una m.d.i. de  $6 \times 10^{-6}$  ò  $10^{-5,2}$ .

#### **4.5.2. Ensayo de co-infección y doble infección.**

Aunque la posibilidad de co-infección o doble infección es baja debido a la m.d.i. utilizada, realizamos el siguiente ensayo para descartarla. Se mezclaron dos semillas en una proporción 1:1, LAI, con un título de  $1,8 \times 10^5$  TCID<sub>50</sub>, y UG24 ( $1,8 \times 10^5$  TCID<sub>50</sub>). Esta mezcla se diluyó y se usó para infectar una monocapa de células MT-4 a una multiplicidad de infección de  $10^{-5,2}$  usada habitualmente en nuestros plaqueos.

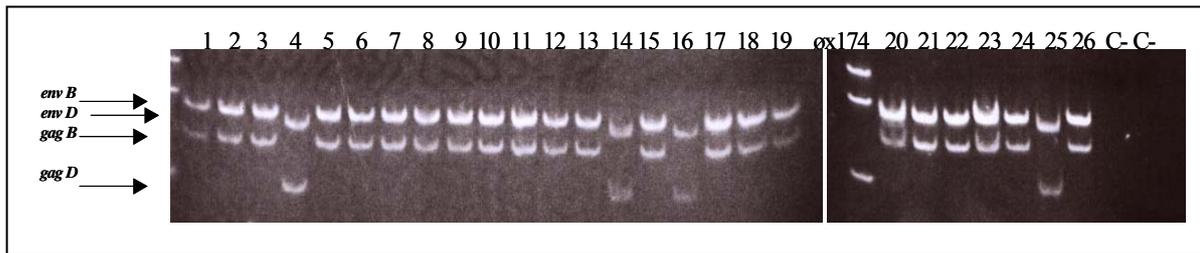
Se pincharon 130 placas, se extrajo el ADN, y se caracterizaron según se explica en el apartado 3.3.1. De las 127 placas analizadas, 126 contenían un solo genoma parental, pero en una aparecieron dos genomas diferentes. Luego, cualquiera de los dos procesos descritos de co-infección o doble infección, podía ser el origen de estas placas.

La diferencia entre una placa doblemente infectada y otra que se ha generado a partir de una célula coinfectada, es que en la segunda hay una alta probabilidad de aparición de virus recombinantes en el segundo ciclo de replicación, mientras que en la primera no hay ninguna.

#### **4.5.3. Análisis de la progenie de las placas con dos genomas provirales.**

Para tratar de identificar el origen de las placas “heterozigotas”, se realizó el siguiente análisis. Las placas que contenían dos genomas diferentes integrados, se resuspendieron y diluyeron para realizar un plaqueo. Las placas que se formaron se pincharon y se analizó su contenido proviral.

En la figura 4.17 se muestra el resultado de uno de estos experimentos. La placa “madre” contenía los genomas de LAI y de UG24, el resultado fue 25 placas que contenían un genoma parental (21 correspondían a LAI y 4 a UG24), y una placa semejante a su antecesora, con cuatro bandas, pero ningún provirus recombinante.



**Fig. 4.17.** Análisis de ADN proviral de los clones biológicos resultantes del plaqueo de un clon que presenta dos genomas provirales diferentes, correspondientes a los dos virus parentales con los que se realizó la infección doble inicial: LAI y UG24.

En otro experimento con otra placa diferente de las mismas características, se pincharon 56 placas de la progenie, y se obtuvieron 37 y 18 placas correspondientes a cada uno de los dos virus parentales y ningún recombinante.

Así mismo en una placa infectada por LAI y RF obtuvimos 13 y 9 placas de cada uno de los genomas parentales. En ningún momento obtuvimos una placa con un virus recombinante a partir de estas placas con cuatro bandas. Esto indica que el origen de estas placas es una doble infección, y no una coinfección de la misma célula, ni la infección por un virus heterocigoto.

Por tanto, el origen de estas placas “heterocigotas” parecía ser una doble infección, o dicho de otro modo, son placas dobles, formadas por la infección de dos células cercanas por dos virus diferentes.

La coinfección de una célula en las condiciones en las que hacemos el plaqueo, es por tanto muy difícil, ya que a lo largo de todos los ensayos realizados, no detectamos ninguna.

#### **4.5.4. Modelo experimental para el estudio de la transmisión de información genética durante el proceso de transcripción inversa.**

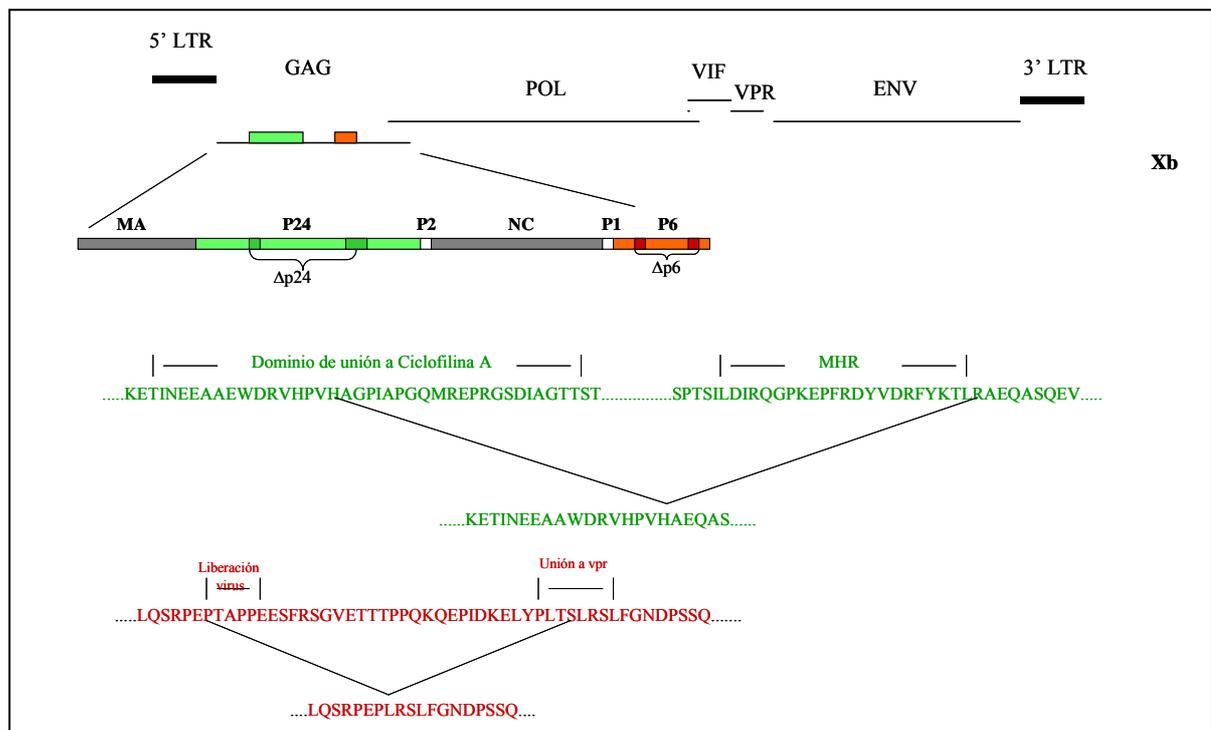
A pesar de que la coinfección quedaba casi descartada como origen de las placas “heterocigotas”, no obtuvimos ninguna evidencia de la transcripción inversa de las dos hebras del virión. No obstante, para poder confirmar que este fenómeno no ocurría nunca propusimos el siguiente sistema experimental.

Nuestro modelo se basa en un sistema que sólo permite replicar virus heterocigotos o virus recombinantes. Para ello hemos introducido dos deleciones deletéreas en el genoma del

VIH-1. Los plásmidos utilizados en este ensayo se obtuvieron a partir de dos construcciones denominadas pFc, y pFAM4, que derivan de un clon molecular obtenido en nuestro laboratorio (Olivares y cols., 1997). Este clon está dividido en dos partes: pFAM4 contiene la secuencia comprendida entre la LTR 5' y un sitio RI en vpr. El segundo, pFc contiene la secuencia comprendida entre este sitio RI y la LTR3'.

Las dos deleciones utilizadas como marcadores fueron introducidas sobre el plásmido pFAM4, por mutagénesis dirigida, dando lugar a dos construcciones que denominamos p $\Delta$ p6 y p $\Delta$ p24. El plásmido p $\Delta$ p6 contiene una deleción de 33 aminoácidos en la proteína p6, que abarca desde el residuo 7 hasta el 39 de dicha proteína, ambos incluidos, de manera que eliminamos dos dominios descritos como fundamentales para la replicación viral: el que comprende desde la prolina 7 a la prolina 10 (PTAP), que está implicado en la liberación de partículas virales, y el motivo 32 a 39 (o dominio L), de unión a Vpr.

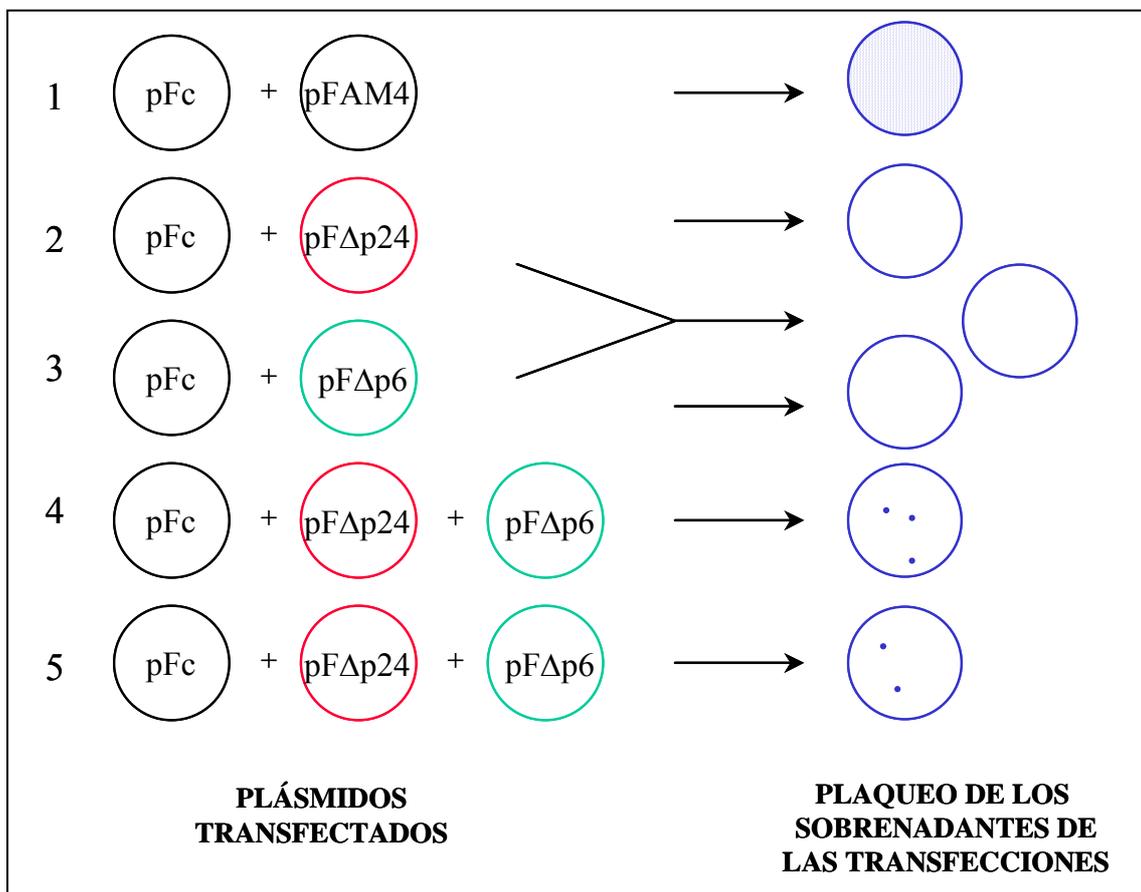
El plásmido p $\Delta$ p24 tiene una deleción en la proteína p24 de 258 nucleótidos (86 aa) que incluye todos los sitios implicados en la funcionalidad de esta proteína: dominio de unión a ciclofilina A, implicado en infectividad y maduración viral; y el dominio MHR, en el dominio C-terminal, implicado en la formación de dímeros proteicos, interacciones gag-gag y ensamblaje de proteínas virales (ver apartado 3.9 de Materiales y Métodos).



**Fig. 4.18.** Representación esquemática de las deleciones realizadas sobre el plásmido pFAM4.

#### 4.5.5. Experimento de cotransfección.

Los plásmidos descritos en el párrafo anterior, fueron transfectados en células Cos-1 siguiendo el esquema que aparece en la figura 4.19. Se recogieron los sobrenadantes de las transfecciones cada dos días. Se ultra-centrifugaron para aumentar la concentración de virus, como se describe en el apartado 3.10 de “materiales y métodos”, y se sometieron a una digestión con DNasa RQ1 para digerir los restos de plásmidos. La resuspensión del residuo viral en medio de cultivo fresco, se plaqueó sobre una monocapa de células MT4 y las placas de infección se pincharon y se caracterizaron con los iniciadores descritos en el apartado 3.10. El contenido proviral de estas placas se identificó por la longitud de los fragmentos amplificados por PCR. Una representación esquemática de este ensayo puede verse en la figura 4.19.



**Fig. 4.19.** Esquema de los plásmidos que se transfectaron en cada experimento. El número 1 corresponde al control positivo. En los tubos 2 y 3 se transfectaron cada uno de los genomas defectivos por separado. Además se realizó un plaqueo con la mezcla de los sobrenadantes 2 y 3, que se representa con una flecha bifurcada. Los tubos 4 y 5 constituyen un duplicado del experimento de co-transfección (las células transfectadas contienen dos genomas defectivos diferentes).

Tras plaquear los sobrenadantes recogidos a partir de las distintas transfecciones, comprobamos que tanto la delección en p24 como la realizada en la proteína p6, incapacitan al virus para realizar un ciclo completo de replicación, puesto que no apareció ninguna placa en los pocillos correspondientes (nº 2 y 3 en la figura). Por el contrario, en el cultivo del sobrenadante del control positivo, obtuvimos un total de  $1500 \pm 100$ .

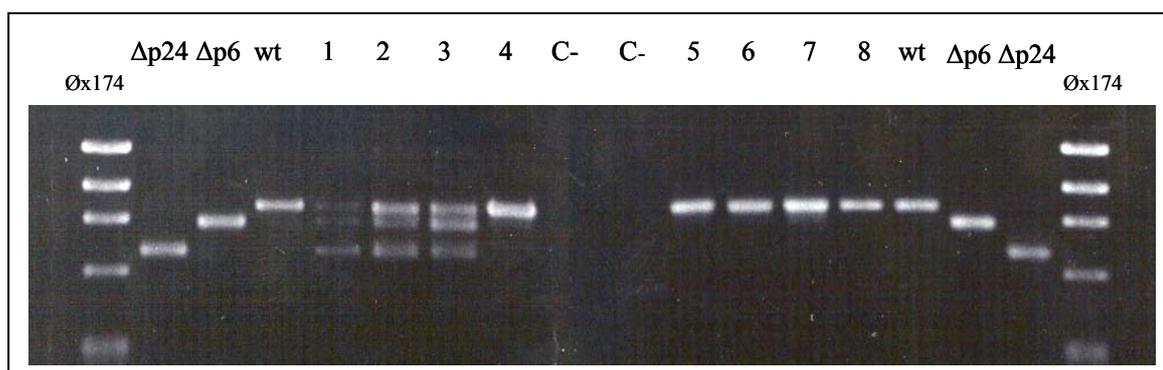
Las cotransfecciones se hicieron por duplicado, y se corresponden con los números 4 y 5 de la figura. En ellas se transfectaron los dos plásmidos p $\Delta$ p6 y p $\Delta$ p24, más el plásmido pFc. De esta forma, las células Cos-I pueden llegar a contener, tras una ligación in vivo, dos genomas de VIH-1 defectivos diferentes, que pueden complementarse.

Si esta complementación se produce, se podrían formar viriones maduros de tres tipos:

- a) Homocigotos con la delección  $\Delta$ p24.
- b) Homocigotos con la delección  $\Delta$ p6.
- c) Virus heterocigotos que contienen dos hebras de ADN con las dos delecciones diferentes.

Los viriones homocigotos, aunque sean capaces de infectar una célula, no provocan la aparición de una placa, ya que la célula infectada no podrá sintetizar la proteína delecionada, y la generación de virus se para en esa fase. Solamente los viriones heterocigotos son capaces de recombinar durante la transcripción inversa, integrar posteriormente un genoma de VIH no defectivo, y generar una placa de infección.

De la totalidad de los sobrenadantes recogidos a partir de las co-transfecciones sólo encontramos 8 placas de infección, de las cuales 5 presentaron un patrón correspondiente al virus silvestre original (placas 4, 5, 6, 7 y 8 en la Fig. 4.20.) producto de la infección por un virión heterocigoto y la recombinación durante la transcripción inversa..



**Fig. 4.20.** Análisis del ADN proviral de los virus que aparecen en el experimento de complementación entre dos genomas defectivos. Tras pinchar las placas y extraer el ADN, se realiza una PCR doble (ver apartado 3.11). La longitud de la banda en las placas 4 a 7, es la que corresponde al virus silvestre y al plásmido pFAM4. Son, por tanto, resultantes de un proceso de recombinación, que revierte las dos deleciones originales. Las placas 1 a 3 presentan tres bandas, que corresponden a la secuencia que contiene la deleción p24, la que contiene la deleción de p6, y la del virus silvestre original.

En el control positivo, que se representa en la figura 4.19 con el nº 1, donde se transfectó el plásmido sin deleción pFAM4, obtuvimos un total de  $1500 \pm 100$  placas. Suponiendo que la cantidad de virus obtenida en las transfecciones denominadas 4 y 5 es similar a este, que la proporción de virus heterocigotos es un tercio del total, y teniendo en cuenta que la distancia entre los marcadores es de 450 nucleótidos, las cinco placas que aparecieron con el genoma silvestre permiten calcular una tasa del 1.1% de recombinantes por kilobase y ciclo de replicación, para este fragmento del genoma de VIH-1.

Las placas que aparecen en el gel con los números 1, 2 y 3, presentan tres bandas con distinto peso molecular y con distinta intensidad, cuyos tamaños corresponden al virus silvestre, y a cada uno de los genomas con una deleción, luego las células que constituyen la placa llevan integrados dos genomas defectivos y uno recombinante.

#### 4.5.6. Posibles mecanismos de generación de las placas con tres bandas.

Las placas identificadas con los números 1, 2, y 3 presentan un patrón complejo, que puede ser explicado por dos mecanismos:

- 1.- La coinfección de una célula por dos virus homocigotos defectivos. Estos dos genomas se complementan y dan origen a un virus heterocigoto que dará lugar a un genoma recombinante en un ciclo posterior de formación de la placa.
- 2.- La transcripción inversa de las dos hebras de ARN de un virión heterocigoto.

La aparición del virus silvestre en estas tres placas sería el resultado de un proceso de recombinación en un segundo ciclo de replicación, dado que la coinfección de una célula por tres viriones diferentes es altamente improbable.

#### **4.5.7. Análisis de la progenie de las placas con tres bandas.**

El resultado de una infección en monocapa con la suspensión de las placas 1, 2 y 3 del experimento de complementación, fue la aparición de numerosa progenie, con un título similar al obtenido en experimentos anteriores. De esta progenie hemos analizado 58 placas: 20 de la placa 1, 26 de la número 2 y 10 de la placa 3. La PCR reveló un mismo patrón para todas ellas. Todas habían recuperado el genoma silvestre original.

Aunque nuestro ensayo no descarta la coinfección como origen de las placas 1, 2 y 3, la probabilidad teórica de que ocurra una coinfección en nuestras condiciones experimentales es muy baja. Así, la probabilidad de que una célula de la monocapa sea infectada es aproximadamente de  $10^{-4}$  (obtenemos 300 p.f.u. por pocillo en el control positivo, y el número de células en un pocillo es  $3,3 \times 10^6$ ). La probabilidad de coinfección de una célula por dos virus es  $10^{-4} \times 10^{-4} = 10^{-8}$ . Y la probabilidad de que esto ocurra tres veces es  $10^{-8} \times 10^{-8} \times 10^{-8} = 10^{-24}$ .

Podemos concluir que, aunque la transcripción inversa de las dos hebras de un virión no ha sido demostrada, y además, sería un proceso poco frecuente, es una posibilidad que debemos tener en cuenta, ya que, junto con la recombinación, constituiría un mecanismo muy poderoso de reversión de mutaciones deletéreas.

## *5. DISCUSIÓN*

La recombinación ejerce una notable influencia en la evolución de las especies. En el caso del VIH, la recombinación se ha descrito como el mecanismo más importante en la evolución de este virus (McCutchan, 2000) (Jung y cols., 2002). La organización genómica del virus parece estar diseñada para facilitar la recombinación en cada ciclo: en primer lugar, el virus transporta dos cadenas de ARN con toda la información necesaria para crear un virus completo en cada una de ellas (Jones y cols., 1994), y sin embargo, parece que sólo se origina un provirus a partir de cada virión (Hu y Temin, 1990b). ¿Cuál puede ser la función de la segunda hebra?

Lo más probable es que el virus utilice estas dos cadenas en cada ciclo de replicación, aunque la información genética que contengan sea exactamente la misma. Este intercambio genético no es necesario para su ciclo vital, y en muchos casos no aporta nada, pero puede ser ventajoso a la hora de corregir errores. Así, se ha obtenido una media de dos a tres eventos de recombinación (fuera de las LTRs) por genoma y ciclo de replicación para dos cepas de HIV-1, cuya diferencia de secuencia es aproximadamente del 5%. Esto implica que ambos ARNs son utilizados durante la transcripción inversa, y que la recombinación es intrínseca al proceso de replicación (Jetzt y cols., 2000).

Este trabajo se ha centrado en el estudio de algunos aspectos relacionados con la recombinación en el VIH-1 en un modelo *in vitro*.

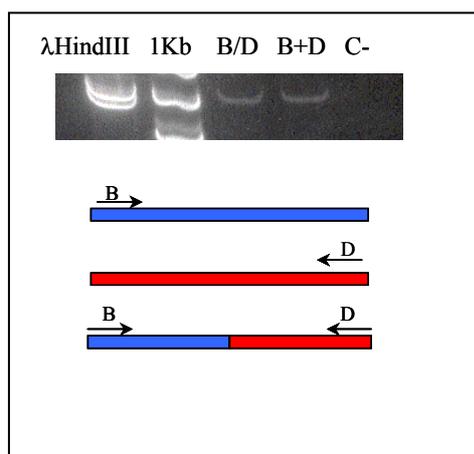
En primer lugar, hemos realizado un estudio comparativo, determinando las proporciones de recombinantes en la progenie de seis infecciones dobles con virus de diferentes subtipos y dos infecciones intrasubtipo. La finalidad de este estudio era determinar si la recombinación es un fenómeno dependiente o independiente de la distancia genética. En una primera aproximación por plaqueo de sobrenadantes, encontramos que las proporciones de virus recombinantes más altas se daban en las infecciones intrasubtipo, mientras que, en cuatro de los seis infecciones intersubtipo, no detectamos ningún recombinante (A/B, A/D, A/F y D/F), coincidiendo con las distancias genéticas más altas. En las infecciones B/D y B/F se obtuvieron proporciones intermedias. Las cuatro infecciones donde se obtuvieron porcentajes apreciables se sometieron a una cuantificación más precisa mediante la técnica de PCR descrita en el apartado 3.3.2 de Materiales y Métodos.

### 5.1. ANÁLISIS CUANTITATIVO DE LA RECOMBINACIÓN SOBRE ADN PROVIRAL.

La cuantificación de la recombinación, se llevó a cabo sobre la progenie viral resultante de cuatro infecciones dobles con virus pertenecientes a los subtipos B, D y F de VIH-1, en células MT4.

Las proporciones de las cuatro especies genéticas presentes en estos cultivos se determinaron en el ADN proviral resultante de cultivos secundarios de células MT4 infectadas con los sobrenadantes. Se adoptó este procedimiento debido a que la determinación directa sobre los sobrenadantes, mediante una RT-PCR producía falsos positivos.

En un control negativo que consistía en una mezcla de dos ARNs parentales, la reacción RT-PCR produjo una amplificación con los iniciadores correspondientes a los genomas recombinantes. Es decir, la síntesis de la cadena de ADN empieza a partir del iniciador específico de subtipo B, y copia la primera parte del genoma de este virus, pero acaba copiando una parte del virus de subtipo D. En la figura 5.1 puede observarse una amplia banda en ese control negativo. Esto demuestra que en la reacción de transcripción inversa realizada *in vitro*, ocurren también los “saltos de hebra” que se producen *in vivo* dentro del complejo de preintegración. Este fenómeno había sido descrito anteriormente (Fang y cols., 1998)



**Fig. 5.1.** Resultado obtenido en la reacción de RT-PCR. En el canal BD, aparece el resultado de la amplificación de una muestra de ARN obtenido a partir de una infección doble con los virus B y D, constituido en parte por recombinantes BD. El canal B+D corresponde a una mezcla al 50% de ARN de virus B más virus D, donde no hay genomas recombinantes. En los dos casos se obtiene amplificación con los iniciadores específicos de un recombinante.

El cultivo secundario a partir de los sobrenadantes, que se describe en el apartado 4.3, nos permite detectar los virus recombinantes infectivos presentes en los sobrenadantes, a través de su copia de ADN proviral. Es decir, mediante esta infección secundaria “traducimos” la población viral que existe en el sobrenadante en forma de ARN a ADN proviral.

Este procedimiento introduce dos alteraciones sobre la población viral total:

- 1) Sólo detectaremos los virus recombinantes que sean infectivos, y no aquellos genomas que resulten defectivos tras la recombinación.
- 2) Existirán algunos viriones heterocigotos en el sobrenadante del cultivo inicial, que tras el proceso de transcripción inversa en esta infección, pasarán a integrarse como recombinantes, aumentando su proporción.

El hecho de que la mayoría de los virus detectados sean viables nos parece un aspecto muy interesante de nuestro método. En cuanto al aumento de la proporción de recombinantes que supone este método, es asumible teniendo en cuenta que se trata de una cuantificación comparativa.

## **5.2. EL MÉTODO DE CUANTIFICACIÓN**

La cuantificación de los genomas recombinantes y parentales está basado en una PCR cuantitativa con patrones externos. La amplificación de las diluciones seriadas de los plásmidos que constituyen los patrones, fue lineal con respecto al número de copias de ADN originales, como se muestra en la Fig 4.5. Para conseguir esta linealidad en la amplificación por PCR, se utilizaron siempre diluciones seriadas inmediatas a la dilución límite, tanto para los patrones como para las muestras. Los valores de  $r^2$  de la regresión lineal fueron mayores de 0,90 en el 85% de los casos, y se muestran en la tabla 8.2 del anexo.

Teniendo en cuenta el carácter estocástico de la PCR, se ha tratado de buscar las condiciones que minimizan este efecto, por ello, las cuatro especies genómicas presentes en una muestra (los dos virus parentales y los dos tipos de recombinantes), se amplificaron en reacciones paralelas, y a partir de la misma dilución de ADN proviral. Cada muestra de ADN se sometió a tres determinaciones independientes, siendo el valor estimado, la media aritmética de estas tres réplicas. En la tabla 8.1 del anexo pueden encontrarse todos los valores de estas determinaciones, así como la medida de su dispersión.

Es conveniente subrayar en este punto que los genomas recombinantes detectados en este estudio no corresponden a una única forma recombinante, sino que es una población heterogénea que tiene en común la hibridación con dos iniciadores específicos en sus extremos. El sistema de PCR con iniciadores específicos utilizado detecta todas las formas recombinantes que tienen uno o un n° impar de puntos de recombinación, pero no detectará los que tengan dos puntos o n° par. Aunque esta falta de sensibilidad no es despreciable, la mayor parte de los recombinantes tendrán sólo un punto de cruce. Además nuestra cuantificación no es absoluta, sino comparativa entre las distintas combinaciones de cepas estudiadas.

### **5.3. RELACIÓN ENTRE DISTANCIA GENÉTICA Y RECOMBINACIÓN**

Para que tenga lugar un proceso de recombinación es necesario que antes se haya producido la coinfección de una célula por dos viriones distintos. Este paso depende de la multiplicidad de infección, así como de las características de los virus que infectan. Por ejemplo, las cepas con cinéticas de replicación similares, favorecerán la coinfección, mientras que un virus con una cinética mucho más rápida que el otro, ocupará la mayor parte de las células antes de que el segundo empiece a infectar. Esta podría ser la explicación de lo ocurrido en la doble infección D/F. A día 5 el virus D ocupa el 94% del ADN proviral del cultivo, mientras el virus F sólo representa el 5%, y el 0,6% corresponde a la población de virus recombinantes. Esta proporción de virus recombinantes disminuye con el tiempo, y a día 22 es indetectable, debido a que la probabilidad de coinfección de una célula por dos virus diferentes es muy baja.

Las cepas elegidas para este estudio se seleccionaron con características fenotípicas similares (ver tabla 4.1). Esta diferencia entre los virus D y F no estaba prevista, ya que en la caracterización biológica presentan un título similar, y usan el mismo correceptor principal CXCR4, ya que las células MT4 utilizadas en las infecciones no presentan el correceptor CCR5.

A pesar de que todos los virus utilizados en las infecciones dobles tenían características biológicas semejantes, y que las infecciones se realizaron en las mismas condiciones, los porcentajes de virus recombinantes obtenidos a día 5 fueron muy diferentes en las distintas combinaciones. Pensamos, por tanto, que podía haber otros factores limitantes de la recombinación.

Cuando representamos la proporción de virus recombinantes que aparecen en los cultivos el día 5, frente a la distancia genética entre los virus parentales usados en la infección doble, encontramos una buena correlación, con un valor de  $r^2 = 0,87$  en la regresión lineal.

La influencia de la distancia genética en la recombinación se ve reflejada en los resultados del análisis clonal de los recombinantes B/D que se describen en el apartado 4.4.2.2, donde encontramos un número mayor de eventos de recombinación en el fragmento *pol*, que en la secuencia de *env* (ver figura 7) siendo la distancia genética entre los fragmentos de *env* dos o tres veces más alta en relación con los de *pol* (ver árboles en las figuras 4.10 y 4.11) (Quiñones-Mateu y cols., 1999).

Esta correlación predice proporciones de recombinación mayores de las obtenidas para secuencias más parecidas, como en los experimentos de Kuwata y cols. (Kuwata y cols., 1997), que obtienen porcentajes por encima del 20% a los 15 días post-infección entre dos virus con una homología del 86%, o como los porcentajes de recombinación del 51% descritos entre dos secuencias iguales por (Zhang y cols., 2000b) (Li y Zhang, 2002) (Zhang y cols., 2000b).

Existen varios trabajos que justifican la influencia de la distancia genética en la proporción de recombinantes presentes en la progenie:

1.- Zhang y Temin (Zhang y Temin, 1994) describen cómo los puntos de recombinación, están corrientemente localizados en fragmentos cortos de identidad de secuencia entre las dos hebras de un virión heterocigoto. De manera que, dos hebras más homólogas tendrán más secuencias comunes.

2.- También se han descrito incompatibilidades entre diferentes virus a la hora de encapsidar juntas dos hebras diferentes (Kaye y Lever, 1998).

#### **5.4. AUMENTO PROGRESIVO DE LOS VIRUS RECOMBINANTES EN UNA POBLACIÓN**

En los primeros experimentos realizados en este trabajo, pudimos observar un aumento de la proporción de virus recombinantes en los cultivos, con el tiempo o con los ciclos de replicación. Este aumento fue confirmado con la cuantificación de las formas recombinantes en tres de las cuatro infecciones dobles estudiadas mediante la PCR cuantitativa.

Además, el estudio cinético que se llevó a cabo con las infecciones B/F y L/R aportó dos aspectos a resaltar:

- a) un aumento continuo del total de las especies recombinantes, y
- b) la imposición progresiva de estas sobre uno de los parentales.

Hay que señalar que en nuestro sistema de cuantificación, hemos analizado solamente alrededor de un 25% del genoma (2500 nucleótidos de un total de 9800). Aunque las proporciones de recombinantes no se incrementan linealmente con la distancia entre los marcadores (Anderson y cols., 1998), podemos esperar que la proporción de recombinantes, si se estudia la totalidad del genoma, sea más alta de la obtenida para este fragmento. Y dado que para la cuarta parte del genoma encontramos porcentajes de recombinantes de hasta el 23% en la infección R/L el día 22 del cultivo, se podría considerar que la totalidad de los genomas serían recombinantes.

### **5.5. DINÁMICA DE IMPOSICIÓN DE LOS VIRUS RECOMBINANTES EN LA POBLACIÓN.**

El análisis realizado sobre los clones aislados a partir de los sobrenadantes de la doble infección B/D, tenía como objetivo determinar si el aumento progresivo de los genomas recombinantes en la progenie era debido a la expansión clonal de virus con una eficacia biológica muy alta, o se debía a la suma de nuevos fenómenos de recombinación en cada ciclo de replicación. Los árboles filogenéticos de las figuras 4.10 y 4.11, muestran un predominio de secuencias independientes entre las que parecen ser recombinantes, que tendrían su origen en eventos de recombinación independientes. Pero además, existe un grupo de cinco secuencias agrupadas (9-41, 22-8, 22-90, 22-82, 22-63) en el árbol correspondiente a las secuencias “*pol*”. El programa Simplot, detectó un punto de cruce común a cuatro de estos cinco clones y con el 22-24 (figura 4.13), aunque las secuencias de todos ellos son muy diferentes. La agrupación de estas secuencias y el punto de recombinación compartido por todas ellas, indican un origen filogenético común, aunque no podemos descartar como causa de ese punto de recombinación, una estructura secundaria del ARN, que en ese punto favorezca la parada de la transcripción inversa y la transferencia de hebra (Wooley y cols., 1998).

En cualquier caso, aunque se produzcan ocasionalmente expansiones clonales de algunos virus, la mayoría de los clones recombinantes que hemos analizado, son consecuencia de eventos de recombinación independientes.

El análisis de recombinación realizado sobre los clones biológicos, nos ha permitido llevar a cabo una estimación de los fenómenos evolutivos que están operando dentro de esta población viral originada a partir de una mezcla de dos virus parentales al 50%. Así, hemos podido observar diferencias en cuanto a la naturaleza de los clones según procedan del día 9 o del día 22. Puede apreciarse un claro aumento del número de puntos de recombinación por clon con el tiempo. Mientras en las secuencias analizadas el día 9, encontramos una media de 1,1 puntos de corte por clon, en el día 22 la media fue de 1,6 (ver tabla 4.10). Este hecho refuerza la idea de que nuevos eventos de recombinación independientes se van sumando, incluso dentro de los clones recombinantes.

### 5.6. SELECCIÓN DE PROTEINAS.

Los cambios que se observan en las proporciones de las secuencias entre el día 9 y 22, son un reflejo de la presencia relativa de las dos proteínas de los dos subtipos B y D. No se observó ningún cambio con el tiempo en *gag* (fragmento 1 en la Fig 4.15); pero en el fragmento 3 que representa una fracción de *env*, la imposición con el tiempo de la proteína gp120 de subtipo B desde 3/7(43%) a día 9, a 10/13 (77%), sugiere una mayor eficacia biológica que la del subtipo D. En la región de la transcriptasa inversa, la mayoría de la población final de virus presenta una proteína recombinante entre los dos subtipos. Esto indicaría que la recombinación está favorecida en este fragmento, por oposición a lo que ocurre en el fragmento 3, donde se impone el parental B. Estos datos sugieren que a lo largo del tiempo de evolución del cultivo, se ha seleccionado la mejor combinación de proteínas, de una manera rápida gracias a la recombinación.

La región de la proteína *env* secuenciada, se eligió por estar descrita como el fragmento de máxima recombinación en VIH-1(Jetzt y cols., 2000), con la finalidad de encontrar el máximo número de puntos de recombinación en nuestros clones. Sin embargo, entre los clones del día 22 analizados, sólo encontramos una secuencia recombinante en el gen *env*. Este resultado sólo indica que el fragmento secuenciado no incluye el punto caliente descrito.

### 5.7 ACERCA DE LOS VIRUS DEFECTIVOS

No está claro en que medida los virus defectivos juegan un papel en la patogénesis del SIDA (F Leri J. Virol. 92), pero se ha descrito que una alta proporción de las cepas de HIV encontradas en pacientes son defectivas para la replicación. Algo similar ocurre en cultivos de líneas celulares que, conteniendo provirus defectivos, son capaces de liberar partículas infecciosas (Meyerhans y cols., 1989) (Boulerice y cols., 1990) (Inoue y cols., 1991). Se ha probado además que determinadas construcciones de VIS con genes reguladores deletados no revierten en cultivos de larga duración (Guan y cols., 2001b) (Guan y cols., 2001a).

La complementación en trans suministra un modelo útil para estudiar las interacciones que pueden ocurrir *in vivo*. Este fenómeno probablemente es el responsable de la alta difusión de virus defectivos y puede ser vital en la patogénesis del SIDA.

Nosotros hemos construido dos genomas defectivos, que no pueden originar viriones competentes en replicación, mediante dos deleciones, una en la proteína de la cápside (p24), y otra en p6. Esto quedó demostrado mediante el plaqueo de los sobrenadantes de las transfecciones 2 y 3, que se corresponden con cada una de las deleciones señaladas. En ninguno de los dos casos apareció una sola placa. Por el contrario, el virus silvestre originó un total de 1500.

No obstante, estos plásmidos pueden complementarse en una cotransfección, aportando cada uno de ellos la proteína que le falta al otro para producir viriones maduros, que poseen todas las proteínas que necesita un virus para poder infectar una célula.

La progenie viral resultante de la transfección de las cel Cos-1 con ambos plásmidos, estaría compuesta por:

A.- viriones homocigotos de dos tipos: aquellos con p24 deletada, y los que contienen dos ARNs con la deleción en p6; y

B.- viriones heterocigotos, que contienen una deleción diferente en cada una de las dos moléculas de ARN.

Los primeros (homocigotos) son incapaces de replicar en las células MT4, ya que, aunque pudieran infectar y retrotranscribir su ARN, serían incapaces de producir una progenie viral que diera lugar a una placa (hemos parado el proceso en un único ciclo de

replicación, por tanto). Sólo la coinfección de una célula por dos virus homocigotos diferentes puede dar origen a una placa.

Por otro lado, los viriones heterocigotos pueden infectar estas células, y dar lugar a una progenie viral por dos mecanismos:

1- Recombinación: Durante la transcripción inversa, puede producirse un salto de hebra, originando un ADN recombinante sin ninguna de las dos deleciones. Este proceso es el responsable de la formación de cinco placas (números 4, 5, 6, 7 y 8) de la figura 4.17.

2- Complementación en la célula MT4 entre los ADNs integrados procedentes de cada una de los dos ARNs diferentes de un virión heterocigoto. Para que esto ocurra sería necesaria la transcripción inversa de cada una de las dos hebras de ARN del virión. Este puede ser el caso de las placas 1, 2 y 3, cuyo ADN proviral está compuesto, como puede apreciarse en la fig2, por dos ADNs delecionados, más uno recombinante.

No podemos descartar la posibilidad de que estas placas que presentan tres genomas provirales diferentes sean el producto de una coinfección por dos viriones defectivos homocigotos o heterocigotos. No obstante, es altamente improbable:

Por un lado, hemos determinado la probabilidad de obtención de una placa doblemente infectada, que fue de 1/126. Estas placas se caracterizan porque su progenie está compuesta solamente de virus parentales, nunca por virus recombinantes, ya que no se produjo la coinfección de una célula.

La posibilidad de que ocurra una coinfección es mucho más baja que la anterior. Una prueba de ello es que no hemos encontrado ninguna a lo largo de los numerosos experimentos de plaqueo realizados en este estudio, pero además, las estimaciones realizadas sobre la probabilidad de que ocurran, son muy bajas. Una estimación aproximada es la que se desarrolla en el apartado 4.5.4. Dado que la m.d.i. del experimento es del orden de  $10^{-4}$ , la probabilidad que tiene una célula de ser infectada al mismo tiempo por dos virus es  $10^{-4} \times 10^{-4} = 10^{-8}$ . Teniendo en cuenta que el número de placas formadas en el control positivo es de 1500, las tres placas con genomas defectivos no pueden ser consecuencia de este proceso.

Hay que señalar en este punto, que la probabilidad de coinfección aumentaría considerablemente si pensamos que dos o más viriones pueden formar agregados que no se desagregarían con la dilución. Sin embargo, el hecho de que a lo largo de un gran número

de experimentos de plaqueo, nunca hayamos encontrado este fenómeno, sugiere que la frecuencia de agregación, o la infectividad de los agregados no es alta. Aunque no podemos descartarlo, recientemente, se ha descrito que la posible agregación viral no constituye un factor a tener en cuenta en el proceso de la doble infección (Wei-Shau Hu., "The nature of HIV-1 double infection", Third VIH DRP Symposium on Antiviral Drug Resistance. Washington 8-11 diciembre 2002). Cada vez es más aceptado el hecho de que la reinfección es más frecuente que la coinfección simultánea de una célula, como proceso necesario para la recombinación.

La progenie viral generada a partir de estas tres placas se caracterizó, y el resultado fue que los provirus integrados en las 46 placas analizadas correspondían al patrón del virus silvestre. Este resultado es esperable dado que el virus recombinante ya estaba presente en las tres placas "madre" de la generación anterior, es decir, se había originado en algún ciclo de la formación de las placas 1, 2 o 3; y además, el virus recombinante es mucho más eficaz biológicamente que un proceso de complementación entre dos genomas defectivos.

Si bien hay que tener en cuenta que las condiciones experimentales están seleccionando los virus recombinantes, ya que los parentales (defectivos) son inviables, no se aprecia una disminución del título (nº de unidades formadoras de placa por ml) respecto al que tiene una placa del virus silvestre (son del orden de  $10^4$  a  $10^5$ ). Esta disminución del título sería esperable en el caso de que el origen de las placas fuera la coinfección, debido a la proporción mayoritaria de virus homocigotos en la progenie.

Además, si se estuvieran produciendo una proporción apreciable de virus parentales, obtendríamos una proporción de placas de complementación, semejante a la que encontramos en el sobrenadante de las transfecciones (3/8).

Podemos concluir que, aunque no hemos podido probar definitivamente la transcripción inversa de las dos hebras de un virión, nuestros resultados sugieren que es una posibilidad a tener en cuenta, ya que, aunque se trataría de un fenómeno muy poco frecuente, constituiría, junto con la recombinación, un valioso mecanismo de reversión de mutaciones deletéreas.

Aunque el modelo de recombinación de elección de copia ("copy-choice") parece ser el más utilizado, aún está por demostrar que el segundo modelo (strand displacement) no

sea utilizado. Este modelo necesita la síntesis de las dos copias de ADN de polaridad negativa. De manera que los retrovirus podrían ser al menos, parcialmente diploides.

La ventaja evolutiva fundamental que aporta al virus este proceso, es que multiplica las posibilidades de recombinación entre dos marcadores, lo cual adquiere especial importancia en los casos de mutaciones deletéreas o mutaciones de resistencia a antivirales. Cuando estas mutaciones se encuentran muy cercanas dentro del genoma, la probabilidad de recombinación en el fragmento intermedio disminuye mucho, como es el caso de nuestro sistema experimental. La transcripción inversa de las dos hebras del virión aumenta notablemente las posibilidades de recombinación, y recuperación de material genético por parte de este virus.

Independientemente de que el proceso responsable sea la retrotranscripción de las dos hebras de un virión, o la coinfección de una célula por dos virus diferentes, los resultados ponen de manifiesto que a partir de una célula que contiene dos provirus integrados diferentes, pueden generarse viriones heterocigotos, que en la siguiente infección pueden recombinar generando un genoma no defectivo. En este segundo ciclo, la recombinación es un fenómeno mayoritario, frente a la complementación. Por tanto, podemos concluir, que la recombinación es con mucho el principal mecanismo de rescate genómico, frente a la complementación, que es mucho menos frecuente.

Las implicaciones epidemiológicas que este mecanismo tendría son notables. Por un lado, cualquier genoma defectivo integrado, puede ser recuperado si la célula que lo contiene es reinfectada. Un virión heterocigoto formado por esta célula, puede originar un recombinante durante la transcripción inversa en la infección de la siguiente célula. Si además se produce la transcripción inversa de las dos hebras, tendremos integrado el genoma defectivo en una célula más, dando una oportunidad más a la recombinación para recuperar ese genoma.

En suma, los resultados de este estudio *in vitro*, han aportado datos en cuanto a la cuantificación, dinámica, prevalencia, naturaleza de los virus recombinantes y consecuencias evolutivas de la recombinación, lo cual puede ayudar a comprender la historia natural y biología del VIH-1.

En primer lugar, hemos puesto a punto un sistema de diferenciación y cuantificación de cuatro especies genómicas presentes en una misma muestra ADN.

En segundo lugar, nuestro método de identificación de clones biológicos es capaz de diferenciar todos los virus utilizados en este estudio.

A partir del seguimiento de una serie de infecciones dobles, hemos sido capaces de estudiar la influencia que ejercen el tiempo, la multiplicidad de infección y la distancia genética sobre la proporción de recombinantes en un cultivo in vitro.

Las consecuencias que tiene esta influencia de la distancia genética sobre la frecuencia de recombinación, son muy variadas:

1- Los recombinantes se convierten en catalizadores de la recombinación, ya que disminuyen la distancia genética entre los componentes de una población, y por tanto la facilitan.

2- También hemos observado una gran diferencia en la frecuencia de recombinación entre unas regiones y otras del genoma viral, coincidiendo las más homólogas con mayores niveles de recombinación.

3.- La recombinación, facilita la selección de la mejor combinación de proteínas en la población viral que resulta de una doble infección.

4.- Uno de los resultados más importantes del estudio es la imposición de las formas recombinantes sobre una de las parentales en la progenie de tres infecciones dobles. Una de las consecuencias evolutivas de la pérdida de un parental, es una población con una distancia genética global menor, pero una mayor heterogeneidad debido a la generación de formas recombinantes diferentes. Esta situación favorecerá nuevos procesos de recombinación, conduciendo hacia una convergencia genética.

Nuestros datos indican que la recombinación realiza un importante papel generando múltiples posibilidades genómicas de aumentar la eficacia biológica, que purificadas por selección pueden jugar un importante papel en la dinámica de evolución del VIH-1.

Nuestra demostración de que un alto número de recombinantes puede aparecer en un corto periodo de tiempo, y que el conjunto tiene mayor fitness que los virus parentales, tiene importantes implicaciones en la patogénesis del SIDA y el hecho de que esta recombinación sea más fácil entre virus filogenéticamente más cercanos, tiene implicaciones epidemiológicas: la recombinación entre dos virus divergentes en una zona de coexistencia se dará antes o después, pero en el momento de aparición de recombinantes, el proceso se acelera, produciendo infinidad de variantes diferentes.

Los resultados *in vitro*, nos permiten prever que en un paciente infectado con dos virus suficientemente cercanos, se producirá la aparición de recombinantes con rapidez, y la imposición de estos sobre los parentales.

Los últimos datos epidemiológicos están confirmando estos resultados: ciertas FRCs empiezan a jugar un papel importante en la pandemia. Una proporción media de 18.1% de las muestras analizadas por la red WHO-UNAIDS corresponden a genomas recombinantes intersubtipo, La mitad de ellas está constituido por las FRC AE (3,6%), AG (5,3%), y AB (1%) (Osmanov y cols., 2002).

## *6. CONCLUSIONES*

1. En cultivos de células MT4 infectados por cuatro combinaciones diferentes de dos virus con características biológicas semejantes, detectamos la aparición de virus recombinantes desde el día 5 post-infección, con proporciones entre un 1 y un 11% de la población global
2. La proporción de virus recombinantes en el sobrenadante del quinto día de cuatro infecciones dobles, fue inversamente proporcional a la distancia genética que separa a los virus parentales.
3. La proporción de virus recombinantes aumentó progresivamente con los ciclos de replicación, en tres de las cuatro infecciones estudiadas, hasta superar la proporción de uno de los virus parentales.
4. En una infección doble realizada con los virus UG24 y 93BR20 detectamos una proporción de recombinantes inferior al 1% a día 5, que llegó a ser indetectable el día 22.
5. La eficacia biológica del conjunto de virus recombinantes detectados en las infecciones dobles ha sido superior a uno de los virus progenitores.
6. El aumento progresivo de la proporción de virus recombinantes con el tiempo, fue debido a la suma de dos procesos: fundamentalmente la aparición continua de nuevos recombinantes en cada ciclo de replicación y también a la expansión clonal de alguno de ellos.
7. Se ha detectado un aumento del número de puntos de recombinación por clon de 1,1 a 1,7, entre los clones analizados el día 9 y y el día 22 de la doble infección B/D.
8. Hemos encontrado puntos de recombinación en un clon del día 9 que se repiten en clones del día 22, y van acompañados de otros nuevos. Este hecho sugiere que los virus recombinantes sufren nuevos eventos de recombinación en ciclos sucesivos. La recombinación parece facilitar nuevos eventos de recombinación, probablemente debido a que disminuye la distancia genética media en la población.
9. La población de recombinantes en un cultivo doblemente infectado por los virus LAI y UG24 evolucionó seleccionando la secuencia de LAI en el gen de la envuelta *env*, y una proteína recombinante en RT.

10. Hemos detectado viriones heterocigotos maduros, infectivos y competentes en replicación, generados por complementación entre dos genomas defectivos con amplias deleciones en las proteínas p24 y p6.
11. La recombinación entre los dos genomas defectivos se encontró en cinco placas. Estos virus son consecuencia de eventos de recombinación independientes dentro de un fragmento del genoma de 450 pares de bases, lo que permite establecer una tasa de recombinación para este fragmento del VIH-1 del 1% por Kb y ciclo de replicación.
12. Entre los viriones detectados en estos ensayos de complementación, hemos encontrado tres casos que sugieren la transcripción inversa de las dos moléculas de ARN de un virión heterocigoto.

## *7. BIBLIOGRAFÍA*

- Accola, M. A., Strack, B., and Gottlinger, H. G. (2000). Efficient particle production by minimal Gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. *J Virol* **74**(12), 5395-402.
- Allain, B., Lapadat-Tapolsky, M., Berlioz, C., and Darlix, J. L. (1994). Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *Embo J* **13**(4), 973-81.
- Anderson, J. A., Bowman, E. H., and Hu, W.-S. (1998). Retroviral Recombination Rates Do Not Increase Linearly with Marker Distance and Are Limited by the Size of the Recombining Subpopulation. *Journal of Virology* **72**, 1195-1202.
- Arjan, J. A., Visser, M., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L., and Lenski, R. E. (1999). Diminishing returns from mutation supply rate in asexual populations. *Science* **283**(5400), 404-6.
- Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **226**, 1209-1211.
- Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871.
- Bartz, S. R., Rogel, M. E., and Emerman, M. (1996). Human immunodeficiency virus type 1 cell cycle control: Vpr is cytostatic and mediates G<sub>2</sub> accumulation by a mechanism which differs from DNA damage checkpoint control. *J Virol* **70**, 2324-2331.
- Belda, F. J., Barlow, K. L., Murphy, G., Parry, J. V., and Clewley, J. P. (1998). A dual subtype B/E HIV type 1 infection with a novel V3 loop crown motif among infections acquired in Thailand and imported into England. *AIDS Res Hum Retroviruses* **14**(10), 911-6.
- Bobkov, A., Cheingsong-Popov, R., Salminen, M., McCutchan, F., Louwagie, J., Ariyoshi, K., Whittle, H., and Weber, J. (1996). Complex Mosaic Structure of the Partial Envelope Sequence from a Gambian HIV Type 1 Isolate. *AIDS Research and Human Retroviruses* **12**, 169-171.
- Boom, R., Sol, C. J. A., Salimans, M. M. M., Jansen, C. L., Wertheim-van Dillen, P. M. E., and Van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *J Clin Microbiol* **28** (3), 495-503.
- Boone, L. R., and Skalka, A. M. (1981). Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with melittin. I. Kinetics of synthesis and size of minus- and plus-strand transcripts. *J Virol* **37**(1), 109-16.
- Boulerice, F., Bour, S., Geleziunas, R., Lvovich, A., and Wainberg, M. A. (1990). High Frequency of Isolation of Defective Human Immunodeficiency Virus Type 1 and Heterogeneity of Viral Gene Expression in Clones of Infected U-937 Cells. *Journal of Virology* **64**, 1745-1755.
- Bryant, M., and Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc Natl Acad Sci U S A* **87**(2), 523-7.
- Buttò, S., Argentini, C., Mazzella, A. M., Iannotti, M. P., Leone, P., Nicolosi, A., and Rezza, G. (1997). Dual infection with different strains of the same HIV-1 subtype. *AIDS* **11**, 694-696.
- Carr, J. K., Salminen, M. O., Koch, C., Gotte, D., Arntstein, A. W., Hegerich, P. A., St. Louis, D., Burke, D. S., and McCutchan, F. E. (1996). Full-Length Sequence and Mosaic Structure of a Human Immunodeficiency Virus Type 1 Isolate from Thailand. *Journal of Virology* **70**, 5935-5943.

- Casado, C., Garcia, S., Rodriguez, C., del Romero, J., Bello, G., and Lopez-Galindez, C. (2001). Different evolutionary patterns are found within human immunodeficiency virus type 1-infected patients. *J Gen Virol* **82**(Pt 10), 2495-508.
- Chao, L., Tran, T., and Mththews, C. (1992). Muller's ratchet and the advantage of sex in the RNA virus phi 6. *Evolution* **46**, 289-299.
- Cheyrier, R., Henrichwark, S., Hadida, F., Pelletier, E., Oksenhendler, E., Autran, B., and Wain-Hobson, S. (1994). HIV and T cell expansion in splenic white pulps is accompanied by infiltration of HIV-specific cytotoxic T lymphocytes. *Cell* **78**(3), 373-87.
- Choi, D. J., Dube, S., Spicer, T. P., Slade, H. B., Jensen, F. C., and Poiesz, B. J. (1997). HIV Type 1 Isolate Z321, the Strain Used to Make a Therapeutic HIV Type 1 Immunogen, Is Intersubtype Recombinant. *AIDS Research and Human Retroviruses* **13**, 357-361.
- Clavel, F., Hoggan, M. D., Willey, R. L., Strebel, K., Martin, M. A., and Repaske, R. (1989). Genetic Recombination of Human Immunodeficiency Virus. *Journal of Virology* **63**, 1455-1459.
- Coffin, J., S. Hughes, and Varmus, a. H. (1997). "Retroviruses." Cold Spring Harbor Laboratory Press, New York.
- Coffin, J. M. (1979). Structure, replication, and recombination of retrovirus genomes: some unifying hypothesis. *J. Gen. Virol.* **42**, 1-26.
- Coombs, R. W., Collier, A. C., Allain, J.-P., Nokora, B., Leuther, M., Gjerset, G. F., and Corey, L. (1989). Plasma viremia in human immunodeficiency virus infection. *N Engl J Med*, 1626-1631.
- 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., Lawson, V. A., Crowe, S., Maerz, A., Sonza, S., Leramont, J., Sullivan, J. S., Cunningham, A., Dwyer, D., Dowton, D., and Mills, J. (1995). Genomic Structure of an Attenuated Quasi Species of HIV-1 from a Blood Transfusion Donor and Recipients. *Science* **270**, 988-991.
- DeStefano, J. J. (1994). Kinetic analysis of the catalysis of strand transfer from internal regions of heteropolymeric RNA templates by human immunodeficiency virus reverse transcriptase. *J Mol Biol* **243**(4), 558-67.
- DeStefano, J. J. (1995). Human immunodeficiency virus nucleocapsid protein stimulates strand transfer from internal regions of heteropolymeric RNA templates. *Arch Virol* **140**(10), 1775-89.
- DeStefano, J. J., Bambara, R. A., and Fay, P. J. (1994). The mechanism of human immunodeficiency virus reverse transcriptase- catalyzed strand transfer from internal regions of heteropolymeric RNA templates. *J Biol Chem* **269**(1), 161-8.
- DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., and Fay, P. J. (1991). Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *J Biol Chem* **266**(12), 7423-31.
- DeStefano, J. J., Mallaber, L. M., Rodríguez-Rodríguez, L., Fay, P. J., and Bambara, R. A. (1992). Requirements for Strand Transfer between Internal Regions of Heteropolymer Templates by Human Immunodeficiency Virus Reverse Transcriptase. *Journal of Virology* **66**, 6370-6378.
- Diaz, L., Cristofaro, J. V., and DeStefano, J. J. (2000). Human immunodeficiency virus reverse transcriptase base misincorporations can promote strand transfer. *Arch Virol* **145**(6), 1117-31.
- Diaz, L., and DeStefano, J. J. (1996). Strand transfer is enhanced by mismatched nucleotides at the 3' primer terminus: a possible link between HIV reverse transcriptase fidelity and recombination. *Nucleic Acids Res* **24**(15), 3086-92.

- Diaz, R. S., Sabino, E. C., Mayer, A., Mosley, J. W., and Busch, M. P. (1995). Dual human immunodeficiency virus type 1 infection and recombination in a dually exposed transfusion recipient. The Transfusion Safety Study Group. *J Virol* **69**(6), 3273-81.
- Dittmar, M. T., McKnight, A., Simmons, G., Clapham, P. R., and Weiss, R. A. (1997). HIV-1 tropism and co-receptor use. *Nature* **385**, 495-496.
- Domingo, E. (1997). RNA Virus Evolution, Population Dynamics, and Nutritional Status. *Biological Trace Element Research* **56**, 23-30.
- Domingo, E., and Holland, J. J. (1997). RNA Virus Mutations and Fitness for Survival. *Annual Review of Microbiology* **51**, 151-178.
- Doranz, B. J., Rucker, J., Yi, Y., Smyth, R. J., Samson, M., Peiper, S. C., Parmentier, M., Collman, R. G., and Doms, R. W. (1996). A dual-tropic primary HIV-1 isolate that uses fusin and the beta- chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* **85**(7), 1149-58.
- Duarte, E., Clarke, D., Moya, A., Domingo, E., and Holland, J. J. (1992). Rapid fitness losses in mamalian RNA virus due Muller's ratchet. *Proc.Natl.Acad. Sci.USA* **89**, 6015-6019.
- Fang, G., Zhu, G., Burger, H., Keithly, J. S., and Weiser, B. (1998). Minimizing DNA recombination during long RT-PCR. *J Virol Methods* **76**(1-2), 139-48.
- Farnet, C. M., and Haseltine, W. A. (1991). Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol* **65**(4), 1910-5.
- Fauci, A. S. (1996). Resistance to HIV-1 infection: It's in the genes. *Nature Medicine* **2**, 966-967.
- Feng, S., and Holland, E. C. (1988). HIV-1 tat trans-activation requires the loop sequence within tar. *Nature* **334**(6178), 165-7.
- Fields, B., and Knipe, a. D. (1990). "Fields Virology." Second edition. ed. 2 vols. Raven Press.
- Fisher, A. G., Ensoli, B., Looney, D., Rose, A., Gallo, R. C., Saag, M. S., Shaw, G. M., Hahn, B. H., and Wong-Staal, F. (1988). Biologically diverse molecular variants within a single HIV-1 isolate. *Nature* **334**(6181), 444-7.
- Frankel, A. D., and Young, J. A. (1998). HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* **67**, 1-25.
- Fuentes, G. M., Palaniappan, C., Fay, P. J., and Bambara, R. A. (1996). Strand displacement synthesis in the central polypurine tract region of HIV-1 promotes DNA to DNA strand transfer recombination. *J Biol Chem* **271**(47), 29605-11.
- Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science* **283**(5398), 80-3.
- 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., and Hahn, B. H. (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes [In Process Citation]. *Nature* **397**(6718), 436-41.
- Gao, F., Morrison, S. G., Robertson, D. L., Thornton, C. L., Craig, S., Karlsson, G., Sodroski, J., Morgado, M., Galvao-Castro, B., von Briesen, H., Beddows, S., Weber, J., Sharp, P. M., Shaw, G. M., and Hahn, B. H. (1996). Molecular Cloning and Analysis of Functional Envelope Genes Human Immunodeficiency Virus Type 1 Sequence Subtypes A through G. *J.Virol.* **70**, 1651-1667.
- Gao, F., Yue, L., Craig, S., Thornton, C. L., Robertson, D. L., McCutchan, F., Bradac, J. A., Sharp, P. M., and Hahn, B. (1994a). Genetic Variation of HIV type 1 in four World Health Organization-Sponsored

Vaccine Evaluation Sites: Generation of Functional Envelope (Glycoprotein 160) Clones Representatives of Sequence Subtypes A,B,C, and E. *AIDS Research and Human Retroviruses* **10**, 1359-1368.

Gao, F., Yue, L., Robertson, D. L., Hill, S. C., Hui, H., Biggar, R. J., Neequye, A. E., Whelan, T. M., HO, D. D., Shaw, G. M., Sharp, P. M., and Hahn, B. H. (1994b). Genetic Diversity of Human Immunodeficiency Virus Type 2: Evidence for Distinct Sequence Subtypes with Differences in Virus Biology. *Journal of Virology* **68**, 7433-7447.

Garcia, J. V., and Miller, A. D. (1992). Downregulation of cell surface CD4 by nef. *Res Virol* **143**(1), 52-5.

Gaynor, R. B. (1995). Regulation of HIV-1 gene expression by the transactivator protein Tat. *Curr Top Microbiol Immunol* **193**, 51-77.

Gelderblom, H. R., M. Özel, and Pauli., a. G. (1989). Morphogenesis and morphology of HIV structure-function relations. *Arch Virol* **106**, 1-13.

Goff, S. P. (1990). Retroviral reverse transcriptase: synthesis, structure, and function. *J Acquir Immune Defic Syndr* **3**(8), 817-31.

Gottlinger, H. G., Dorfman, T., Sodroski, J. G., and Haseltine, W. A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A* **88**(8), 3195-9.

Göttlinger, H. G., Sodroski, J. G., and Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* **86**(15), 5781-5.

Guan, Y., Whitney, J. B., Detorio, M., and Wainberg, M. A. (2001a). Construction and in vitro properties of a series of attenuated simian immunodeficiency viruses with all accessory genes deleted. *J Virol* **75**(9), 4056-67.

Guan, Y., Whitney, J. B., Liang, C., and Wainberg, M. A. (2001b). Novel, live attenuated simian immunodeficiency virus constructs containing major deletions in leader RNA sequences. *J Virol* **75**(6), 2776-85.

Hahn, B. H., Robertson, D. L., McCutchan, F. E., and Sharp, P. M. Recombination and diversity of HIV: implications for vaccine development, 87-94.

Haynes, B. F., Pantaleo, G., and Fauci, A. S. (1996). Toward an Understanding of the Correlates of Protective Immunity to HIV Infection. *Science* **271**, 324-327.

Heinzinger, N. K., Bukinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M. A., Gendelman, H. E., Ratner, L., Stevenson, M., and Emerman, M. (1994). The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc Natl Acad Sci U S A* **91**(15), 7311-5.

Hirsch, V. M., Sharkey, M. E., Brown, C. R., Brichacek, B., Goldstein, S., Wakefield, J., Byrum, R., Elkins, W. R., Hahn, B. H., Lifson, J. D., and Stevenson, M. (1998). Vpx is required for dissemination and pathogenesis of SIV(SM) PBj: evidence of macrophage-dependent viral amplification [see comments]. *Nat Med* **4**(12), 1401-8.

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., and Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123-126.

Hoelscher, M., Kim, B., Maboko, L., Mhalu, F., von Sonnenburg, F., Birx, D. L., and McCutchan, F. E. (2001). High proportion of unrelated HIV-1 intersubtype recombinants in the Mbeya region of southwest Tanzania. *Aids* **15**(12), 1461-70.

- Hsu, T. W., and Taylor, J. M. (1982). Single-stranded regions on unintegrated avian retrovirus DNA. *J Virol* **44**(1), 47-53.
- Hu, W. S., and Temin, H. M. (1990a). Retroviral recombination and reverse transcription. *Science* **250**, 1227-1233.
- Hu, W. S., and Temin, H. M. (1990b). Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A* **87**(4), 1556-60.
- Hu, W. S., and Temin, H. M. (1990c). Genetic consequences of packing two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A* **87**, 1556-1560.
- Hu, W. S., and Temin, H. M. (1992). Effect of gamma radiation on retroviral recombination. *Journal of Virology* **66**, 4457-4463.
- Inoue, M., Hoxie, J. A., Reddy, M. V. R., Srinivasan, A., and Reddy, E. P. (1991). Mechanisms associated with the generation of biologically active human immunodeficiency virus type 1 particles from defective proviruses. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 2278-2282.
- Janini, L. M., Tanuri, A., Schechter, M., Peralta, J. M., Vicente, A. C., Dela Torre, N., Pieniazek, N. J., Luo, C. C., Ramos, A., Soriano, V., Schochetman, G., Rayfield, M. A., and Pieniazek, D. (1998). Horizontal and vertical transmission of human immunodeficiency virus type 1 dual infections caused by viruses of subtypes B and C. *J Infect Dis* **177**(1), 227-31.
- Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D., and Dougherty, J. P. (2000). High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol* **74**(3), 1234-40.
- Jones, J. S., Allan, R. W., and Temin, H. M. (1994). One retroviral RNA is sufficient for synthesis of viral DNA. *J Virol* **68**(1), 207-16.
- Jung, A., Maier, R., Vartanian, J. P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S., and Meyerhans, A. (2002). Recombination: Multiply infected spleen cells in HIV patients. *Nature* **418**(6894), 144.
- Junghans, R. P., Boone, L. R., and Skalka, A. M. (1982). Retroviral DNA H structures: displacement-assimilation model of recombination. *Cell* **30**(1), 53-62.
- Kampinga, G. A., Simonon, A., Perre, P. v. d., Karita, E., Msellati, P., and Goudsmit, J. (1997). Primary Infections with HIV-1 of Women and Their Offspring in Rwanda. Findings of Heterogeneity at Seroconversion, Coinfection, and Recombinations of HIV-1 Subtypes A and C. *Virology* **227**, 63-76.
- Kaye, J. F., and Lever, A. M. (1998). Nonreciprocal packaging of human immunodeficiency virus type 1 and type 2 RNA: a possible role for the p2 domain of Gag in RNA encapsidation. *J Virol* **72**(7), 5877-85.
- Kim, J. K., Palaniappan, C., Wu, W., Fay, P. J., and Bambara, R. A. (1997). Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1. *J Biol Chem* **272**(27), 16769-77.
- Kirchhoff, F., Greenough, T. C., Brettler, D. B., Sullivan, J. L., and Desrosiers, R. C. (1995). Brief Report: Absence of Intact *nef* Sequences in a Long-Term Survivor with Nonprogressive HIV-1 Infection. *New England Journal of Medicine* **332**, 228-232.
- Klimkait, T., Strebel, K., Hoggan, M., Martin, M. A., and J.M.Orenstein (1990). The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release. *J Virol* **64**, 621-629.

- Kondrashov, A. (1988). Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435-440.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B. H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains [see comments]. *Science* **288**(5472), 1789-96.
- Kuiken, C., Folley, B., Hahn, B., Marx, P., McCutchan, F., Mellors, J., Mullins, J., Wolinsky, S., Korber, B., and (Eds) (1999). "Human Retrovirus and AIDS." Theoretical Biology and Biophysics, Group T 10, Los Alamos National Laboratory, Los Alamos, New Mexico.
- Kumar, S., Tamura, K., Jakobsen, I. B., and Nei, M. (2001). MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* **17**(12), 1244-1245.
- Kumar, S., Tamura, K., and Nei, M. (1994). MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. *Comput Appl Biosci* **10**(2), 189-91.
- Kuwata, T., Miyazaki, Y., Igarashi, T., Takehisa, J., and Hayami, M. (1997). The Rapid Spread of Recombinants during a Natural In Vitro Infection with Two Human Immunodeficiency Virus Type 1 Strains. *Journal of Virology* **71**, 7088-7091.
- Le Grice, S. F., Naas, T., Wohlgensinger, B., and Schatz, O. (1991). Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase. *Embo J* **10**(12), 3905-11.
- Leitner, T., Escanilla, D., Marquina, S., Wahlberg, J., Broström, C., Hansson, H. B., Uhlén, M., and Albert, J. (1995). Biological and Molecular Characterization of Subtype D, G, and A/D Recombinant HIV-1 Transmissions in Sweden. *Virology* **209**, 136-146.
- Levy, J. A. (1993). Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* **57**(1), 183-289.
- Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J., and Oshiro, L. S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**, 840-842.
- Lewis, P., Hensel, M., and Emerman, M. (1992). Human immunodeficiency virus infection of cells arrested in the cell cycle [published erratum appears in EMBO J 1992 Nov;11(11):4249]. *Embo J* **11**(8), 3053-8.
- Li, L., Yoder, K., Hansen, M. S., Olvera, J., Miller, M. D., and Bushman, F. D. (2000). Retroviral cDNA integration: stimulation by HMG I family proteins. *J Virol* **74**(23), 10965-74.
- Li, T., and Zhang, J. (2002). Intramolecular recombinations of Moloney murine leukemia virus occur during minus-strand DNA synthesis. *J Virol* **76**(19), 9614-23.
- Livingstone, W. J., Moore, M., Innes, D., Bell, J. E., Simmonds, P., and Group, E. H. T. S. (1996). Frequent infection of peripheral blood CD8-positive T-lymphocytes with HIV-1. *Lancet* **348**, 649-631.
- Luciw, P. A., and Harada, S. (1998). AIDS 1998. Virology: overview. *Aids* **12**(Suppl A), S1-3.
- Lund, A. H., Mikkelsen, J. G., Schmidt, J., Duch, M., and Pedersen, F. S. (1999). The kissing-loop motif is a preferred site of 5' leader recombination during replication of SL3-3 murine leukemia viruses in mice. *J Virol* **73**(11), 9614-8.
- Luo, G. X., and Taylor, J. (1990). Template switching by reverse transcriptase during DNA synthesis. *J Virol* **64**(9), 4321-8.
- Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V., and Cullen, B. R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**(6212), 254-7.

- Mansky, L. M. (1996). Forward Mutation Rate of Human Immunodeficiency Virus Type 1 in a T Lymphoid Cell Line. *AIDS Research and Human Retroviruses* **12**, 307-314.
- Mansky, L. M. (1998). Retrovirus mutation rates and their role in genetic variation. *J Gen Virol* **79**(Pt 6), 1337-45.
- Mansky, L. M., and Temin, H. M. (1994). Lower Mutation Rate of Bovine Leukemia Virus Relative to That of Spleen Necrosis Virus. *Journal of Virology* **68**, 494-499.
- Mansky, L. M., and Temin, H. M. (1995). Lower In Vivo Mutation Rate of Human Immunodeficiency Virus Type 1 than That Predicted from the Fidelity of Purified Reverse Transcriptase. *Journal of Virology* **69**, 5087-5094.
- Marquina, S., Leitner, T., Rabinovich, R. D., Benetucci, J., Libonatti, O., and Albert, J. (1996). Coexistence of Subtype B,F, and B/F *env* Recombinant of HIV Type 1 in Buenos Aires, Argentina. *AIDS Research and Human Retroviruses* **12**, 1651-1654.
- McCutchan, F. E. (2000). Understanding the genetic diversity of HIV-1. *Aids* **14**(Suppl 3), S31-44.
- McCutchan, F. E., Salminen, M. O., Carr, J. K., and Burke, D. S. (1996). HIV-1 genetic diversity. *Aids* **10 Suppl 3**, S13-20.
- McCutchan, F. E., Sanders-Buell, E., Salminen, M. O., Carr, J. K., and Sheppard, W. H. (1998). Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in San Francisco Men's Health Study participants. *AIDS Res Hum Retroviruses* **14**(4), 329-37.
- McCutchan, F. E., Viputtigul, K., de Souza, M. S., Carr, J. K., Markowitz, L. E., Buapunth, P., McNeil, J. G., Robb, M. L., Nitayaphan, S., Birx, D. L., and Brown, A. E. (2000). Diversity of envelope glycoprotein from human immunodeficiency virus type 1 of recent seroconverters in Thailand. *AIDS Res Hum Retroviruses* **16**(8), 801-5.
- Mellors, J. W., Kingsley, L. A., Rinaldo, J., Ch.R., Todd, J. A., Hoo, B. S., Kokka, R. P., and Gupta, P. (1995). Quantitation of HIV-1 RNA in Plasma Predicts Outcome after Seroconversion. *Ann Intern Med* **122**, 573-579.
- Meyerhans, A., Cheynier, R., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B., and Wain-Hobson, S. (1989). Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58**(5), 901-10.
- Mikkelsen, J. G., Lund, A. H., Duch, M., and Pedersen, F. S. (2000). Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo. *J Virol* **74**(2), 600-10.
- Miller, M. D., M.T. Warmerdam, K.A. Page, M.B. Feinberg, and Greene., a. W. C. (1995). Expression of the Human Immunodeficiency Virus type 1 (HIV-1) *nef* gene during HIV-1 production increases progeny particle infectivity independently of gp160 or viral entry. *J. Virol* **69** (1), 579-84.
- Miller, M. D., Warmerdam, M. T., Gaston, I., Greene, W. C., and 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**, 101-113.
- Miyoshi, I., Kubonishi, I., Yoshimoto, S., Akagi, T., Ohtsuki, Y., Shiraishi, Y., Nagata, K., and 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 (London)* **294**(5843), 770-1.
- Montavon, C., Toure-Kane, C., Liegeois, F., Mpoudi, E., Bourgeois, A., Vergne, L., Perret, J. L., Boumah, A., Saman, E., Mboup, S., Delaporte, E., and Peeters, M. (2000). Most *env* and *gag* subtype A HIV-1 viruses

circulating in West and West Central Africa are similar to the prototype AG recombinant virus IBNG. *J Acquir Immune Defic Syndr* **23**(5), 363-74.

(1995). Virus taxonomy - the classification and nomenclature of viruses: sixth report of the International Committee of Taxonomy of viruses. Murphy, F., Fauquet, C. M., and Bishop, D. H. L.

Olivares, I., Shaw, G., and López-Galíndez, C. (1997). Phenotypic switch in a spanish HIV Type 1 isolate on serial passage on MT-4 cells. *AIDS Res Hum Retroviruses* **13** (11), 979-84.

Osmanov, S., Pattou, C., Walker, N., Schwardlander, B., and Esparza, J. (2002). Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr* **29**(2), 184-90.

Panganiban, A. T., and Fiore, D. (1988). Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* **241**(4869), 1064-9.

Pedroza Martins, L., Chenciner, N., and Wain-Hobson, S. (1992). Complex Intrapatient Sequence Variation in the V1 and V2 Hypervariable Regions of the HIV-1 gp120 Envelope Sequence. *Virology* **191**, 837-845.

Peeters, M., and Sharp, P. M. (2000). Genetic diversity of HIV-1: the moving target. *Aids* **14**(Suppl 3), S129-40.

Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J. (1994). Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. *Biochemistry* **33**(46), 13817-23.

Perrino, F. W., Preston, B. D., Sandell, L. L., and Loeb, L. A. (1989). Extension of mismatched 3' termini of DNA is a major determinant of the infidelity of human immunodeficiency virus type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* **86**(21), 8343-7.

Popovic, M., Sarngadharan, M. G., Read, E., and 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.

Premkumar, D. R. D., X.Z. Ma, R. K. Maitra, B.K. Chakrabarti, J. Salkowitz, B. Yen-Lieberman, M.S. Hirsch, and Kestler, a. H. W. (1996). The nef gene from a long-term HIV type 1 nonprogressor. *AIDS Res Hum Retroviruses* **12** (4), 337-45.

Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988). Fidelity of HIV-1 reverse transcriptase. *Science* **242**, 1168-73.

Quiñones-Mateu, M. E., Albright, J. L., Torre, V., Reinis, M., Vandasova, J., Bruckova, M., and Arts, E. J. (1999). Molecular epidemiology of HIV type 1 isolates from the Czech Republic: identification of an env E subtype case [In Process Citation]. *AIDS Res Hum Retroviruses* **15**(1), 85-9.

SimPlot for Windows 95/NT 2.4. Baltimore, MD: Distributed by author

Reddy, B., and Yin, J. (1999). Quantitative intracellular kinetics of HIV type 1. *AIDS Res Hum Retroviruses* **15**(3), 273-83.

Robertson, D. L., Anderson, J. P., Bradac, J. A., Carr, J. K., Foley, B., Funkhouser, R. K., Gao, F., Hahn, B. H., Kalish, M. L., Kuiken, C., Learn, G. H., Leitner, T., McCutchan, F., Osmanov, S., Peeters, M., Pieniazek, D., Salminen, M., Sharp, P. M., Wolinsky, S., and Korber, B. (2000). HIV-1 nomenclature proposal [letter]. *Science* **288**(5463), 55-6.

Robertson, D. L., Hahn, B. H., and Sharp, P. M. (1995a). Recombination in AIDS Viruses. *Journal of Molecular Evolution* **40**, 249-259.

- Robertson, D. L., Sharp, P. M., McCutchan, F. E., and Hahn, B. H. (1995b). Recombination in HIV-1. *Nature* **374**, 124-125.
- Roy, S., Delling, U., Chen, C. H., Rosen, C. A., and Sonenberg, N. (1990). A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev* **4**(8), 1365-73.
- Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W. A., and Rosen, C. A. (1989). Structural and functional characterization of human immunodeficiency virus tat protein. *J Virol* **63**(1), 1-8.
- Sabino, E. C., Shpaer, E. G., Morgado, M. G., Korber, B. T. M., Diaz, R. S., Bongertz, V., Cavalcante, S., Galvão-Castro, B., Mullins, J. I., and Mayer, A. (1994). Identification of Human Immunodeficiency Virus Type 1 Envelope Genes Recombinant between Subtypes B and F in Two Epidemiologically Linked Individuals from Brazil. *Journal of Virology* **68**, 6340-6346.
- Sala, M., Pelletier, E., and Wain-Hobson (1995). HIV-1 gp 120 Sequences from a Doubly Infected Drug User. *AIDS Research and Human Retroviruses* **11**, 653-655.
- Schatz, O., Mous, J., and Le Grice, S. F. (1990). HIV-1 RT-associated ribonuclease H displays both endonuclease and 3'---- 5' exonuclease activity. *Embo J* **9**(4), 1171-6.
- Schwartz, S., Felber, B. K., Feny, E., and 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 U S A* **86**, 7200-3.
- Shibuya, H., Irie, K., Ninomiya-Tsuji, J., Goebel, M., Taniguchi, T., and Matsumoto, K. (1992). New human gene encoding a positive modulator of HIV Tat-mediated transactivation. *Nature* **357**(6380), 700-2.
- Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., and Haseltine, W. (1985). Location of trans-activating region on the genome of human T-cell lymphotropic virus type III. *Science* **229**, 74-77.
- Spina, C. A., Kwok, T. J., Chowers, M. Y., Guatelli, J. C., and 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.
- Stuhlmann, H., and Berg, P. (1992). Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *J Virol* **66**(4), 2378-88.
- Takehisa, J., Osei-Kwasi, M., Ayisi, N. K., Hishida, O., Miura, T., Igarashi, T., Brandful, J., Ampofo, W., Netty, V. B. A., Mensah, M., Yamashita, M., Ido, E., and Hayami, M. (1997). Phylogenetic Analysis of HIV Type 2 in Ghana and Intrasubtype Recombination in HIV Type 2. *AIDS Research and Human Retroviruses* **13**, 621-623.
- Takehisa, J., Zekeng, L., Ido, E., Mboudjeka, I., Moriyama, H., Miura, T., Yamashita, M., Gurtler, L. G., Hayami, M., and Kaptue, L. (1998). Various types of HIV mixed infections in Cameroon. *Virology* **245**(1), 1-10.
- Temin, H. (1976). The DNA Provirus Hypothesis. *Science* **192**, 1075-1080.
- Temin, H. M. (1993). Retrovirus variation and reverse transcription: Abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA* **90**, 6900-6903.
- Temin, H. M., and Mizutani, a. S. (1970). RNA-dependent DNA polymerase in virions of Rous Sarcoma virus. *Nature* **226**, 1211-1213.
- Thompson, J. D., Higgins, D. G., and 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**(22), 4673-80.

- Thomson, M. M., Villahermosa, M. L., Vazquez-de-Parga, E., Cuevas, M. T., Delgado, E., Manjon, N., Medrano, L., Perez-Alvarez, L., Contreras, G., Carrillo, M. G., Salomon, H., and Najera, R. (2000). Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina [letter] [In Process Citation]. *Aids* **14**(7), 897-9.
- Vidal, N., Peeters, M., Mulanga-Kabeya, C., Nzilambi, N., Robertson, D., Ilunga, W., Sema, H., Tshimanga, K., Bongo, B., and Delaporte, E. (2000). Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the democratic republic of congo suggests that the HIV-1 pandemic originated in central africa [In Process Citation]. *J Virol* **74**(22), 10498-507.
- Wei, Q., and Fultz, P. N. (1998). Extensive diversification of human immunodeficiency virus type 1 subtype B strains during dual infection of a chimpanzee that progressed to AIDS. *J Virol* **72**(4), 3005-17.
- 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., and Shaw, G. M. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117-122.
- Weiss, R. A., Mason, W. S., and Vogt, P. K. (1973). Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. *Virology* **52**(2), 535-52.
- Wooley, D. P., Bircher, L. A., and Smith, R. A. (1998). Retroviral recombination is nonrandom and sequence dependent. *Virology* **243**(1), 229-34.
- Wu, H., Myszka, D. G., Tendian, S. W., Brouillette, C. G., Sweet, R. W., Chaiken, I. M., and Hendrickson, W. A. (1996). Kinetic and structural analysis of mutant CD4 receptors that are defective in HIV gp120 binding. *Proc Natl Acad Sci U S A* **93**(26), 15030-5.
- Zennou, V., Petit, C., Guetard, D., Nerhass, U., Montagnier, L., and Charneau, P. (2000). HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**(2), 173-85.
- Zhang, H., Pomerantz, R. J., Dornadula, G., and Sun, Y. (2000a). Human immunodeficiency virus type 1 Vif protein is an integral component of an mRNP complex of viral RNA and could be involved in the viral RNA folding and packaging process. *J Virol* **74**(18), 8252-61.
- Zhang, J., Tang, L. Y., Li, T., Ma, Y., and Sapp, C. M. (2000b). Most retroviral recombinations occur during minus-strand DNA synthesis. *J Virol* **74**(5), 2313-22.
- Zhang, J., and Temin, H. M. (1993). Rate and Mechanism of Nonhomologous Recombination During a Single Cycle of Retroviral Replication. *Science* **259**, 234-238.
- Zhang, J., and Temin, H. M. (1994). Retrovirus Recombination Depends on the Length of Sequence Identity and Is Not Error Prone. *Journal of Virology* **68**, 2409-2414.
- Zhang, L., Carruthers, C. D., He, T., Huang, Y., Cao, Y., Wang, G., Hahn, B., and Ho, D. D. (1997). HIV Type 1 Subtypes, Coreceptor Usage, and CCR5 Polymorphism. *AIDS Research and Human Retroviruses* **13**, 1357-1365.
- Zhu, T., Wang, N., Carr, A., Wolinsky, S., and Ho, D. D. (1995). Evidence for coinfection by multiple strains of Human Immunodeficiency Virus Type 1 Subtype B in an acute seroconverter. *J. Virol.* **69**, 1324-1327.

## *8. ANEXO*

## RESULTADOS PCR CUANTITATIVA

COINFECCIÓN	1º RESULT.	2º RESULT	3º RESULT	4º RESULT	PROMEDIO	DESVIACIÓN ESTÁNDAR
<b>111BD D22(390)</b>						
B	70					
D	7					
BD	20					
DB	2					
<b>111BD D5(390)</b>						
B	15	67	65	76	55,75	27,58
D	82,5	17	24	18	35,375	31,57
BD	1,8	9,7	6	3,8	5,325	3,38
DB	0,3	6,5	5	0,7	3,125	3,10
<b>111BD D13(390)</b>						
B	98,6	99,5			99,05	0,64
D	0,7	0,07			0,385	0,45
BD	0,3	0,4			0,35	0,07
DB	0,4	0			0,2	0,28
<b>111BD D22</b>						
B	68	59	51	34	53	14,45
D	16	30	26	26	24,5	5,97
BD	5	4	17	20	11,5	8,19
DB	10	7	6	19,6	10,65	6,20
<b>111DF D22</b>						
D	98,5	97,5	96,6		97,53	0,95
F	0,9	1,3	2,15		1,45	0,64
DF	0,6	1,3	1,3		1,07	0,40
FD	0	0	0		0,00	0,00
<b>111DF D22(440)</b>						
D	97,5	97,5	100		98,33	1,44
F	2,4	2,5	0		1,63	1,42
DF	0	0	0		0	0,00
FD	0	0	0		0	0,00
<b>111DF D5 (440)</b>						
D	96	95	92		94,33	2,08
F	1,4	4,4	8		4,60	3,30
DF	1,5	0,3	0		0,60	0,79
FD	0	0	0		0,00	0,00
<b>111DF D9 (440)</b>						
D	95	92	94		93,67	1,53
F	4,7	8	5,8		6,17	1,68
DF	0	0,6	0		0,20	0,35
FD	0	0	0		0,00	0,00
<b>452RL D21</b>						
L	23,5	28,5	27		26,33	2,57
R	76	52,5	58		62,17	12,29
LR	0,5	18,3	15		11,27	9,47
RL	0	0,8	0		0,27	0,46
<b>452RL D21(499)</b>						
L	14,4	9,1	3,3		8,93	5,55
R	73,4	65,4	66		68,27	4,46
LR	9,5	18,2	27		18,23	8,75
RL	2,6	7,3	3,9		4,60	2,43

**452RL D8(499)**

L	23,4	10	17		16,80	6,70
R	61	88	78		75,67	13,65
LR	13,5	2	3,7		6,40	6,21
RL	2,1	0,7	1,1		1,30	0,72

**452RL D13(499)**

L	99,9	100	99,9		99,93	0,06
R	0	0	0,1		0,03	0,06
LR	0	0	0		0,00	0,00
RL	0,1	0	0		0,03	0,06

**452RL D17(499)**

L	45,3	27	29,7		34,00	9,88
R	32	59,5	51,3		47,60	14,12
RL	21	6,3	14		13,77	7,35
LR	1,8	7	5		4,60	2,62

**111BF D22**

B	29	25,6	39	39	33,15	6,90
F	60	65	53	49	56,75	7,14
BF	10	7,4	7,4	12	9,20	2,23
FB	0,6	2,2	0,6	0,85	1,06	0,77

**111BF D22(390)**

B	65	69	59	57	62,50	5,51
F	23	14	28	18	20,75	6,08
BF	8,8	11,2	12	23	13,75	6,31
FB	2,2	5	0,8	2,4	2,60	1,75

**111BF D5(390)**

B	33	28	43	45	37,25	8,10
F	65	70	51	55	60,25	8,77
BF	2	1,7	6	0	2,43	2,54
FB	0,4	0,3	0	0	0,18	0,21

**111BF D9(390)**

B	65	51	73	56,5	61,38	9,66
F	30	42	25	36	33,25	7,37
BF	3,7	5,55	1,2	5,2	3,91	1,98
FB	0,75	1,8	1,2	2,6	1,59	0,80

**111BF D16(390)**

B	53	55,7		65	57,90	6,30
F	40	35		22	32,33	9,29
BF	5,7	5		4,3	5,00	0,70
FB	1,1	3,7		9	4,60	4,03

**111BF D19(390)**

B	54,5	54	64		57,50	5,63
F	37	29	26		30,67	5,69
BF	6,2	3,8	5,3		5,10	1,21
FB	2,1	13	4,7		6,60	5,69

Valores de regresión lineal de las cuantificaciones realizadas sobre la coinfección B+F

	$r^2$			<b>P</b>		
111BF,D5						
BB	0,9997	0,9651	0,9776	1	0,4	0,4
FF	0,996	0,9812	0,9979	1	1	1
BF	0,9869	0,9912	0,9888	1	1	0,3
FB	0,9179	0,7391	0,844	1	1	1
111BF,D9	r2			P		
BB	0,9852	0,9917	0,9809	1	1	1
FF	0,8267	0,8591	0,6438	1	0,4	1
BF	0,9723	0,9689	0,9934	1	0,4	1
FB	0,7123	0,9536	0,9546	1	0,4	1
111BF,D16	r2			P		
BB	0,9145	0,9567	0,9918	1	1	1
FF	0,9496	0,9688	0,9459	1	0,4	0,4
BF	0,9677	0,844	0,9699	1	0,8	1
FB	0,9091	0,9639	0,8941	1	1	1
111BF,D19	r2			P		
BB	0,9906	0,9671	0,9865	1	0,4	1
FF	0,9999	0,958	0,9848	1	1	0,9
BF	0,963	0,899	0,9653	0,667	0,8	1
FB	0,9023	0,9596	0,9814	1	0,9	0,4
111BF,D22	r2			P		
BB	0,9852	0,997	0,9999	1	1	1
FF	0,8897	0,9692	0,9869	0,9	1	1
BF	0,9949	1	0,9877	1	0,67	0,4
FB	0,9164	0,8469	0,9313	0,5	1	1

# Analysis, Quantification, and Evolutionary Consequences of HIV-1 *in Vitro* Recombination

María José Iglesias-Sánchez and Cecilio López-Galíndez<sup>1</sup>

Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, 28220 Madrid, Spain

Received April 10, 2002; returned to author for revision May 21, 2002; accepted July 9, 2002

HIV-1 recombination was studied *in vitro* by viral cocultivation of four combinations of strains of subtypes B, D, and F. Viral cocultivations were performed in MT-4 cells and maintained for 22 days. The parental and recombinant forms were quantified by a specific PCR system in an *env* fragment of 2500 nucleotides. On day 5, there was a close correlation between the proportion of recombination and the genetic distance between strains. In three of the four viral combinations studied, a steady increase in the proportion of recombinant genomes was observed over time. This rise coincided with the progressive loss of one of the parental strains, resulting in less diverse viral populations. Nucleotide sequencing of biological recombinant clones from the B/D cocultivation revealed a higher number of recombination events in *pol* than in *env* gene, and an increasing number of crossovers per clone with time. © 2002 Elsevier Science (USA)

## INTRODUCTION

Recombination is considered to be the most important evolutionary mechanism in the natural history of HIV (McCutchan, 2000), although the error-prone viral replication system produces a large number of point mutations (Domingo *et al.*, 1996; Mansky and Temin, 1995). The importance of recombination in HIV-1 and *Retroviridae* is related to the incorporation of two copies of viral RNA in the virion (Hu and Temin, 1990b) and to special characteristics of the reverse transcriptase enzyme (RT) (Coffin *et al.*, 1997). Genetic recombination allows evolution to occur faster than in its absence due to the incorporation of larger genomic fragments (Barton and Charlesworth, 1998). Apart from evolutive consequences, recombination could recover fitness from parental strains (Lai, 1992) as in biological clones subjected to Müller's ratchet (Escarmis *et al.*, 1999), or fill the phylogenetic gap between different subpopulations within patients (Casado *et al.*, 2001), or contribute to antiviral resistance (Kellam and Larder, 1995; Moutouh *et al.*, 1996).

Molecular epidemiological studies have revealed the importance of recombination in the natural history of HIV-1. The first recombinant viruses detected were of subtype E strains, and were, in fact, A–E recombinant viruses or a circulating recombinant form (CRF01), according to the approved nomenclature of isolates (Korber *et al.*, 1998). In Africa, all HIV-1 types and subtypes are present and frequently cocirculate in the same region. In some studies it has been estimated that between 10 and 20% of isolates are recombinant viruses (Kampinga *et al.*,

1997; Hoelscher *et al.*, 2001). Recombinant forms are detected with increasing frequency and currently represent the most important challenge in the molecular epidemiology of HIV (Liitsola *et al.*, 1998; McCutchan, 2000; Thomson *et al.*, 2000). Studies have been undertaken to evaluate and quantify recombination in other retroviruses using reporter genes in artificial constructs (Hu and Temin, 1990a) and also *in vitro* in HIV-1 (Kuwata *et al.*, 1997).

We have carried out a comparative study on HIV-1 recombination in an *in vitro* model by cocultivating two strains from subtypes B, D, and F in MT-4 cells. The amounts of parental and recombinant viruses in the viral cocultivations were estimated at two time points, and kinetic and clonal analyses were carried out to evaluate the nature and dynamics of the different genetic forms in the cocultivations.

## RESULTS

### Biological characterization of the strains

Before establishing the viral cocultivations, we biologically characterized the strains to select those with similar characteristics. All the viruses had the syncytium-inducing phenotype and were able to form plaques in the MT-4 plaque assay. They were titrated in tissue culture-infective doses (TCID<sub>50</sub>) and in plaque-forming units (PFU) in MT-4 cells. All viruses displayed similar titers (summarized in Table 1). MT-4 cells, which express the CXCR4 but not the CCR5 coreceptor, were used for the cocultivations. The coreceptor usage of the strains is also shown in Table 1 (Björndal *et al.*, 1997; Doms *et al.*, 1998). All strains used as main coreceptor the CXCR4 molecule, although subtype D strain 92UG24 could use other minor coreceptors such as CCR8, V28, and CCR3,

<sup>1</sup> To whom correspondence and reprint requests should be addressed. E-mail: clopez@isciii.es.

TABLE 1  
Biological Characteristics of the Strains

Strains	Subtype	Biotype <sup>a</sup>	Coreceptor <sup>b</sup>	Titer TCID50/ml	Titer PFU/ml
LAI	B	SI—X4	CXCR4	$1.8 \times 10^5$	$1.8 \times 10^5$
RF	B	SI—X4/R5	CXCR4-CCR5	$1.8 \times 10^6$	
92UG24	D	SI—X4	CXCR4, CCR8 <sup>c</sup> , V28 <sup>d</sup> , CCR3 <sup>d</sup>	$4.2 \times 10^5$	$1.75 \times 10^5$
93BR20	F	SI—X4/R5	CXCR4-CCR5	$5.6 \times 10^5$	$8 \times 10^4$

<sup>a</sup> The biotype is characterized by the SI, syncytium induction in MT-2 cell phenotype, and by the coreceptor usage X4 for CXCR4 and R5 for CCR5.

<sup>b</sup> Data from Trkola *et al.*, 1998, Rucker *et al.*, 1997, Björndal *et al.*, 1997 and Doms *et al.*, 1998.

<sup>c</sup> Indicates that the coreceptor used was determined only by fusion assays rather than by virus infection.

<sup>d</sup> The virus isolate used the CCR3 coreceptor, whereas the cloned *env* from this isolate did not (Björndal *et al.*, 1997).

and subtype F strain 93BR20 was dual tropic (CXCR4-CCR5).

To analyze the viral replication of each strain, RT activity in the culture supernatant and cell viability were monitored; the results are shown in Figs. 1A and 1B, respectively. As can be seen, the four strains displayed similar kinetics and levels of RT activity, although subtype D strain had a peak of RT activity on day 9 postinfection. The MT-4 cell viability data showed very similar profiles of cell death, indicating a similar viral replication.

#### Detection and quantification of the different HIV-1 proviral DNAs present in the viral cocultivations

The study of *in vitro* HIV-1 recombination was carried out by cocultivating different combinations of strains in MT-4 cells. Four virus combinations, three inter-subtype (B/D, B/F, and D/F) and one intra-subtype B (LAI and RF strains), were evaluated. To detect and quantify the different species present in each viral cocultivation, a set of specific primers was designed (see Materials and Methods). The two types of recombinant genome were detected using crossed primers. The specificity of each primer combination was tested, as can be seen in Fig. 2. We also checked that the PCR method performed with crossed primers amplified the two types of recombinant genomes, but the technique failed to yield amplification from a mixture of parental DNAs, as shown in lanes B+D and D+B of Fig. 2. One of the quantifications performed for the B/B (LAI/RF) cocultivation on day 5 is presented in Fig. 3. The quantification of the parental forms is shown in the top part of the gel and those of the two recombinant forms are shown below. The regression lines of each standard species of the RF, LAI, RF-Lai recombinant (RL), and Lai-RF recombinant (LR) had associated correlation coefficients ( $r^2$ ) greater than 0.89 (data not shown).

#### Analysis of the prevalence of the recombinant forms during cocultivations

To study the prevalence of the recombinant forms in the viral cocultivations, the latter were maintained and

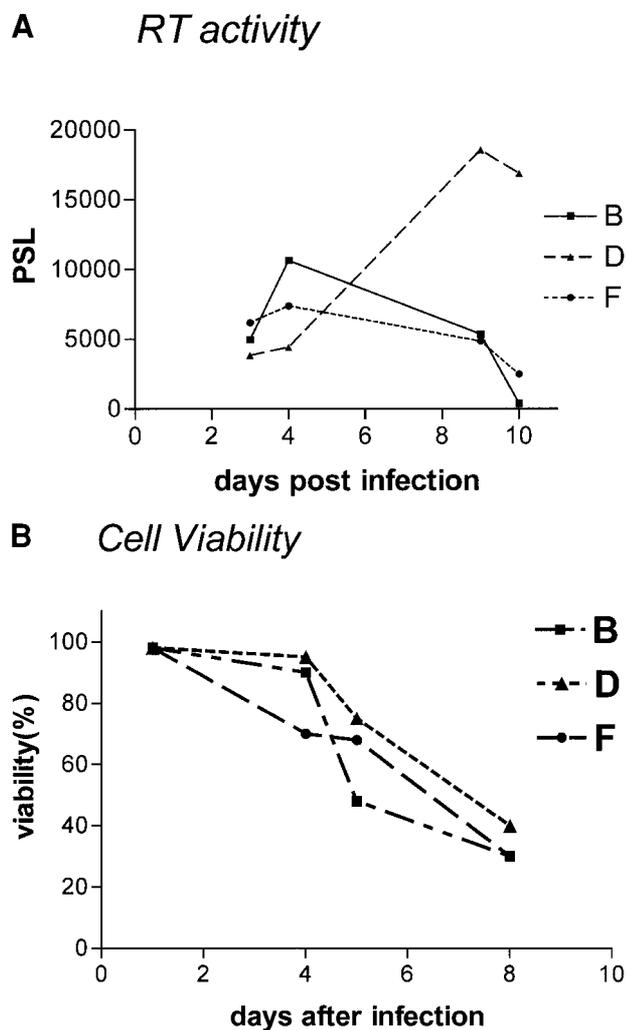
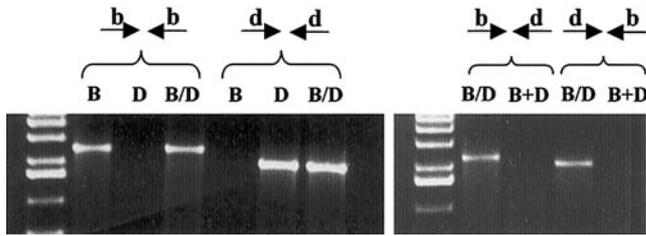


FIG. 1. Kinetics of viral replication of viruses in MT-4 cells. The kinetics of the viral replication of each strain used in the viral cocultivations were analyzed in single infections in MT-4 cells as described under Materials and Methods. The viral replication in (A) was measured by the RT activity assay, performed as indicated under Materials and Methods, and as previously described (---). The intensity of the radioactive spot was quantified with the help of the PC Bas program and the titers were expressed in arbitrary PSL units. (B) Cell viability measured with the trypan blue dye.



**FIG. 2.** Analysis of specificity of primers used to detect different genomes in the B/D cocultivation by PCR. To detect the different proviral DNAs in the cocultivation, two pairs of specific primers were synthesized to amplify by PCR the parental strains B and D. All primers are designated with an arrow and a lowercase letter, indicating the subtype or strain they amplify. In the left-hand gel, we used primers specific for subtype B, as indicated by the arrows above the gel. Lane B shows the band obtained with the DNA of a subtype B culture. Lane D shows the lack of amplification of the PCR with these primers with DNA from a subtype D culture but they could amplify DNA from B/D cocultivation, as shown in lane B/D. The gel on the right shows the result of the amplification of recombinants using crossed primers, indicated by the arrows at the top of the gel. On the left, the primers used were upstream b and downstream d, and the reciprocal combination on the right. In lane B/D, we can see the band corresponding to the amplification of proviral DNA from the B/D cocultivation. These crossed primers were unable to amplify DNA from a 50:50 mixture of B and D DNAs, as shown by the absence of a band in the B+D lane. The right-hand part of the gel shows the similar results obtained with the reciprocal primers.

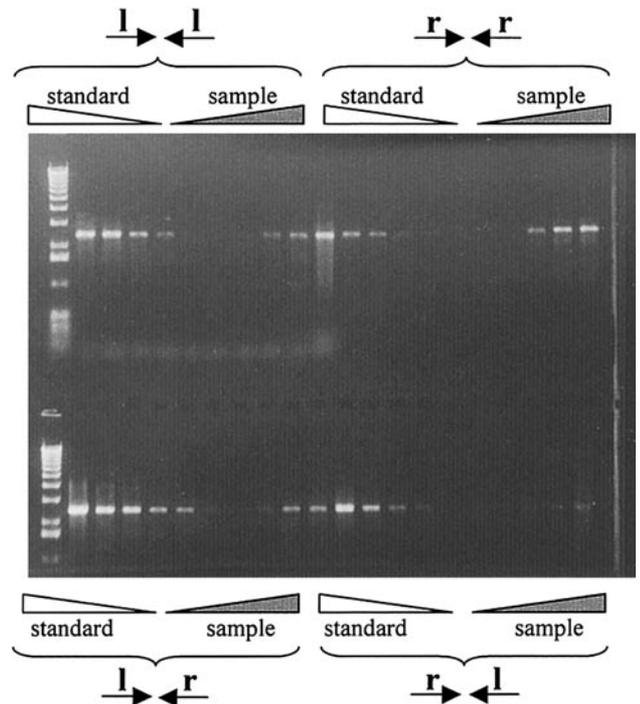
evaluated on days 5 and 22; the results are shown in Tables 2 and 3, respectively. By day 5, recombinant forms were detected in all the cocultivations performed; their proportions are summarized in Table 2. In the B/D mixture we found 8% of recombinant genomes, with 5% corresponding to the BD recombinants (i.e., the 5'-end is of subtype B and the 3'-end is of subtype D) and 3% to the DB genomes. The recombinant genomes were identified in the same way thereafter. A total of 2.6% recombinant species was estimated to be present in the B/F cocultivation. Almost all of them were of the BF type, while a very small number were of the FB type. In the D/F subtype combination, only 0.6% of recombinants were detected, all of which were of the DF type. In the study carried out on subtype B, 11% of genomes were estimated to be recombinant, 10% corresponding to the LR type and 1% corresponding to the RL genomes. The low percentage of recombinants in D/F culture can be explained by the very unbalanced proportion of parental viruses in the cocultivation (see Table 1), which reduced the possibility of cell coinfection. In the group of strains studied, we found a good inverse relation (with an  $r^2$  value of 0.87) between the proportion of recombinants and the genetic distance between strains.

The viral cocultivations were maintained for 22 days and the recombinant forms were again quantified at the end of the culture. The results from day 22 are shown in Table 3. All the recombinant forms present in significant percentages on day 5 displayed higher percentages on

day 22, except in the D/F cocultivation, in which the recombinant forms became undetectable. Total recombinant percentages rose from 8 to 22% in the B/D (2.8 $\times$ ), from 2.6 to 17% in the B/F (6.5 $\times$ ), and from 11 to 22% within subtype B (2 $\times$ ). In two viral combinations, B/D and B/F, the increase in recombinant forms yielded higher percentages than one of the parental strains (see Table 2). In the B/F combination the proportion of parental B strain on day 22 had decreased but had not yet been exceeded by the recombinant forms (see next paragraph). All quantifications were derived from the analysis of approximately 25% of the proviral genome.

### Kinetic analysis of the viral forms present in the cocultivations

To extend the study of the dynamics of the prevalence of the recombinant forms arising during viral cocultivation, we analyzed the B/F and B/B (Lai/RF) combinations at different times (Figs. 4A and 4B). We observed that the total percentage of the recombinant forms increased



**FIG. 3.** PCR Quantification of the different HIV-1 proviruses found in the B/B cocultivation. Proviral DNA obtained from supernatant of day 5 in the RF/LAI culture was used to amplify the four different species LAI, RF, and recombinant forms, LR and RL. An arrow with a lowercase letter on top indicates the primers used for the amplification of the different DNAs. For example, the lowercase l identifies the LAI strain-specific primer. In each part of the gel is shown the quantification of each genomic form grouped by a brace. The left part of each group shows the amplification obtained with decreasing dilutions of the corresponding standard plasmid of each species. The dilutions and the PCR conditions are described under Materials and Methods. The same DNA obtained from the cocultivation was used to quantify the Lai and RF strains and the recombinant forms.

TABLE 2  
Quantification of the Parental and Recombinant Species Present in the Supernatants of the Cocultivations on Day 5

B/D <sup>a</sup>	%	$\delta$	B + F <sup>b</sup>	%	$\delta$	D + F <sup>c</sup>	%	$\delta$	LAI + RF <sup>d</sup>	%	$\delta$
B	73	8.1	B	37	8.1	D	94	2.1	LAI	53	11.1
D	18	3.8	F	60	8.8	F	5	3.3	RF	36	12.2
Rec BD	5	3.4	Rec BF	2.4	2.5	Rec DF	0.6	0.8	Rec LR	10	2.0
Rec DB	3	3.1	Rec FB	0.2	0.2	Rec FD	0	0.0	Rec RL	1	2.3
$\Sigma$ BD + DB <sup>e</sup>	8		$\Sigma$ BF + FB	2.6		$\Sigma$ DF + FD	0.6		$\Sigma$ LR + RL	11	

Note. Percentages in this table are the means and standard deviations ( $\delta$ ) of three independent quantifications. Values were obtained, as described under Materials and Methods, by comparison with the corresponding standard (as shown in Figs. 2 and 3).

<sup>a</sup> Cocultivation of subtype B (strain LAI) and subtype D (strain UG24).

<sup>b</sup> Cocultivation of subtype B (strain LAI) and subtype F (strain 93BR30).

<sup>c</sup> Cocultivation subtype D (strain UG24) and subtype F (strain 93BR30).

<sup>d</sup> Cocultivation within subtype B (strain LAI) and (strain RF).

<sup>e</sup> Total percentages of the two recombinant forms.

with time in both viral cocultivations, reaching higher proportions in the B/B mixture than in the Lai parental strain. In the B/F cocultivation, the proportion of the parental subtype F strain (93BR020) diminished and by day 22 it had almost been surpassed by the recombinant forms (see Fig. 4A). The determination of the proportion of each of the four viruses in the culture allowed the regression line of each species to be calculated. The slope thus obtained may be considered to be a measure of the fitness of each virus (Holland *et al.*, 1991; Yuste *et al.*, 1999). In the B/F cocultivation, B virus had a slope of  $0.74 \pm 0.41$ ; F virus had a slope of  $-1.43 \pm 0.43$ , and the total recombinant forms had a slope of  $0.72 \pm 0.07$ . In the B/B (Lai/RF) cocultivation, Lai virus had a slope of  $-1.37 \pm 0.98$ , that of RF virus was  $0.26 \pm 1.01$ , and the slope of the total recombinant forms was  $1.11 \pm 0.03$ . The lower fitness of one of the parental strains in these viral combinations explains the loss of the parental viruses and the dominance of the recombinant forms.

### Phylogenetic analysis of recombinant viruses arising in the B/D cocultivation

To study the nature of the recombinant viruses arising in the viral cocultivations, 22 biological clones were obtained from the supernatants of the B/D combination on days 9 and 22. They were sequenced and analyzed as explained under Materials and Methods. The phylogenetic trees derived are shown in Fig. 5. In the *env* gene tree, three viruses (22-8, 9-65, and 9-27) were in branches rooted between the parental strains identifying them as possible recombinant genomes. In the tree corresponding to *pol* gene, we observed more recombinant genomes, of which some, such as 9-41, 22-8, 22-63, 22-82, and 22-90 or samples 22-24 and 22-28, formed clusters. From this phylogenetic analysis it is evident that more recombinant events occurred in *pol* gene than in *env* gene.

To map the break points and confirm the recombina-

TABLE 3  
Quantification of the Parental and Recombinant Species Present in the Supernatants of the Cocultivations on Day 22

B/D <sup>a</sup>	%	$\delta$ <sup>b</sup>	B + F <sup>c</sup>	%	$\delta$	D + F <sup>d</sup>	%	$\delta$	LAI + RF <sup>e</sup>	%	$\delta$
B	70		B	62	5.5	D	98	1	LAI	9	5.6
D	7		F	21	6.1	F	2	0.6	RF	68	4.5
Rec BD	20		Rec BF	14	6.3	Rec DF	0	0.4	Rec LR	18	8.8
Rec DB	2		Rec FB	3	1.8	Rec FD	0	0.0	Rec RL	5	2.4
$\Sigma$ BD + DB <sup>f</sup>	22		$\Sigma$ BF + FB	17		$\Sigma$ DF + FD	0		$\Sigma$ LR + RL	23	

Note. Percentages in this table are the means and standard deviations ( $\delta$ ) of three independent quantifications. Values were obtained, as described under Materials and Methods, by comparison with the corresponding standard (as shown in Figs. 2 and 3).

<sup>a</sup> Cocultivation of subtype B (strain LAI) and subtype D (strain UG24).

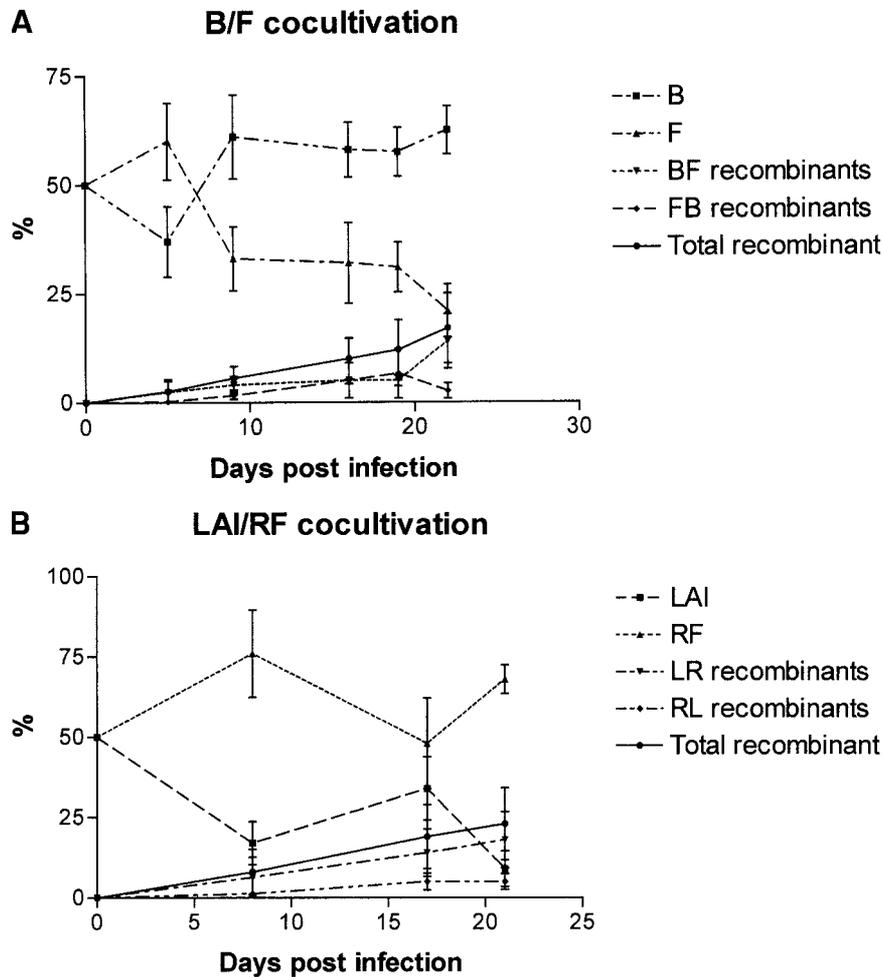
<sup>b</sup> The standard deviation could not be calculated because only one cocultivation was carried out in this combination of viruses.

<sup>c</sup> Cocultivation of subtype B (strain LAI) and subtype F (strain 93BR30).

<sup>d</sup> Cocultivation of subtype D (strain UG24) and subtype F (strain 93BR30).

<sup>e</sup> Cocultivation within subtype B (strain LAI) and (strain RF).

<sup>f</sup> Total percentages of the two recombinant forms.



**FIG. 4.** Kinetic analysis of the B/F and the B/B (LAI/RF) cocultivations. Quantifications of the B/F and B/B cocultivations, performed on different days, are shown. The previously described quantifications were carried out on days 5, 9, 19, and 22 for the B/F cocultivation (A) and on days 8, 17, and 22 for the B/B (LAI/RF) cocultivation (B). The graph illustrates the percentages of the four genetic species and the total recombinant forms over time.

tion in the clones, sequences from the different recombinant viruses were analyzed with Simplot and Mega-2 programs. We found three break points in the *env* region and 17 break points in the RT in *pol* gene. Of the break points in RT, clones 9-41, 22-8, 22-24, 22-82, and 22-90 shared one recombination point, indicating a possible common ancestor (see Fig. 6). However, all these clones differed from clone 9-41 in other crossovers in RT except virus 22-82, which was distinct in *env* gene. These results corroborate the frequent occurrence of new recombination events in recombinant clones.

#### Genetic characterization of the recombinant viruses from the BD cocultivation

The maintenance of the cocultivations for 22 days allowed the study of the prevalence of the recombinant and parental viruses competing in the same medium. In Fig. 6 we present the genetic characterization, in five genomic regions, of the biological clones obtained on days 9 and 22 in the B/D cocultivation. We observed a

random presence of the two parental sequences in the viruses of day 9 in the five different regions analyzed. It is of particular note that by day 22, the frequency of parental sequences had departed from this random distribution in *pol* and *env* genes. In *pol* gene (fragment 2 in Fig. 6), three of the viruses had B sequences (frequency of 0.23), three were subtype D (0.23), and seven were recombinant forms (0.53). In *env* gene (fragment 3), subtype B sequences were found in 10 of the clones (0.72), subtype D in 2 (0.13), and also one recombinant sequence (0.08) were detected. In region 4 at the 3'-end of *env* gene, 10 clones were B (0.72) and 2 clones were D (0.13). Thus, over only 13 days of cocultivation, we observed a dramatic change in the frequency of the parental sequences in *pol* and *env* genes in the recombinant viruses. From the selection of the sequences of the two subtypes in the clones analyzed we deduced that subtype B gp120 protein was fitter than the gp120 of subtype D. Furthermore, recombination does not seem to be promoted in this fragment of the gp120 protein. In con-

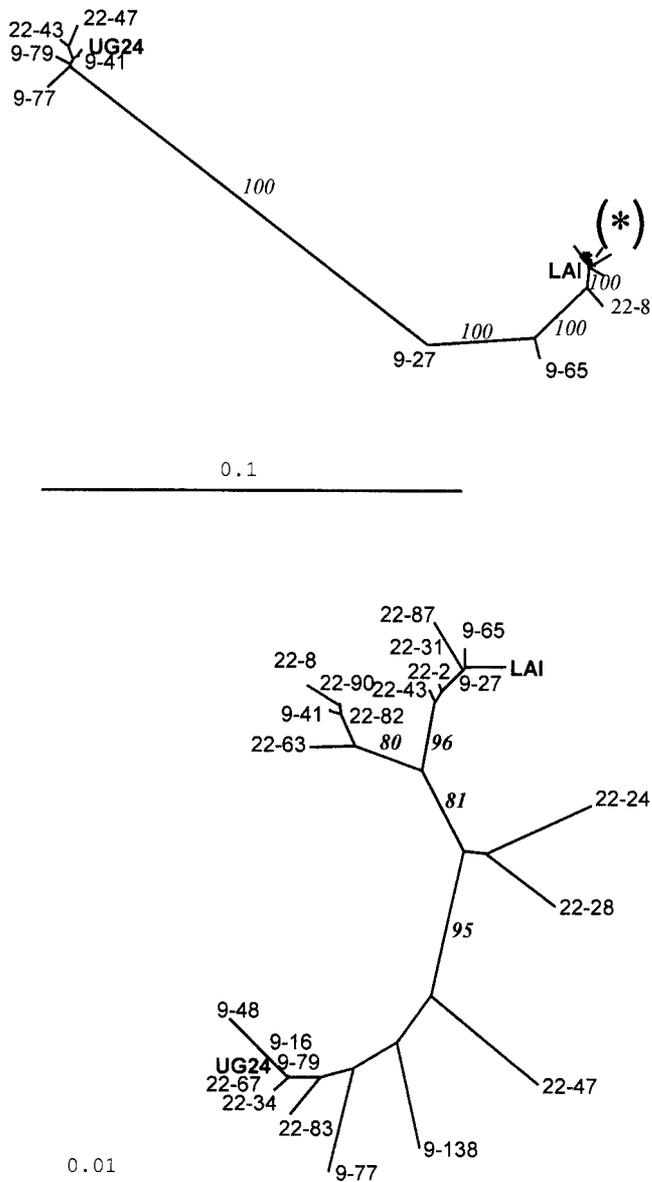


FIG. 5. Phylogenetic trees derived by the Neighbor-Joining method corresponding to the sequences from the clones obtained from the B/D cocultivation on days 9 and 22. Clones were picked in an MT-4 plaque assay as indicated under Materials and Methods and designated by two numbers. The first indicates the day and the second indicates the clone number. Clones obtained on day 9 are named 9-16, 9-27, 9-41, 9-48, 9-65, 9-77, 9-79, 9-138, and the 14 recovered on day 22 were designated 22-2, 22-8, 22-24, 22-28, 22-31, 22-34, 22-43, 22-47, 22-63, 22-67, 22-82, 22-83, 22-87, and 22-90. The upper tree corresponds to the sequences in *env* gene and the lower one to the RT sequence in *pol* gene. The start of the *env* tree corresponds to the clones 22-2, 22-24, 22-28, 22-31, 22-34, 22-63, 22-67, 22-82, 22-83, 22-87, 22-90, 9-16, 9-48, and 9-138. Numbers on the trees indicate the bootstrap value derived from 199 replicates. Bars represent 10% genetic variation in *env* tree and 1% in *pol* gene.

trast, recombinant forms and subtype D were the dominant forms in the RT protein, given that on day 9, only one recombinant was detected, but seven were present on day 22. Thus, this study has demonstrated the rapid

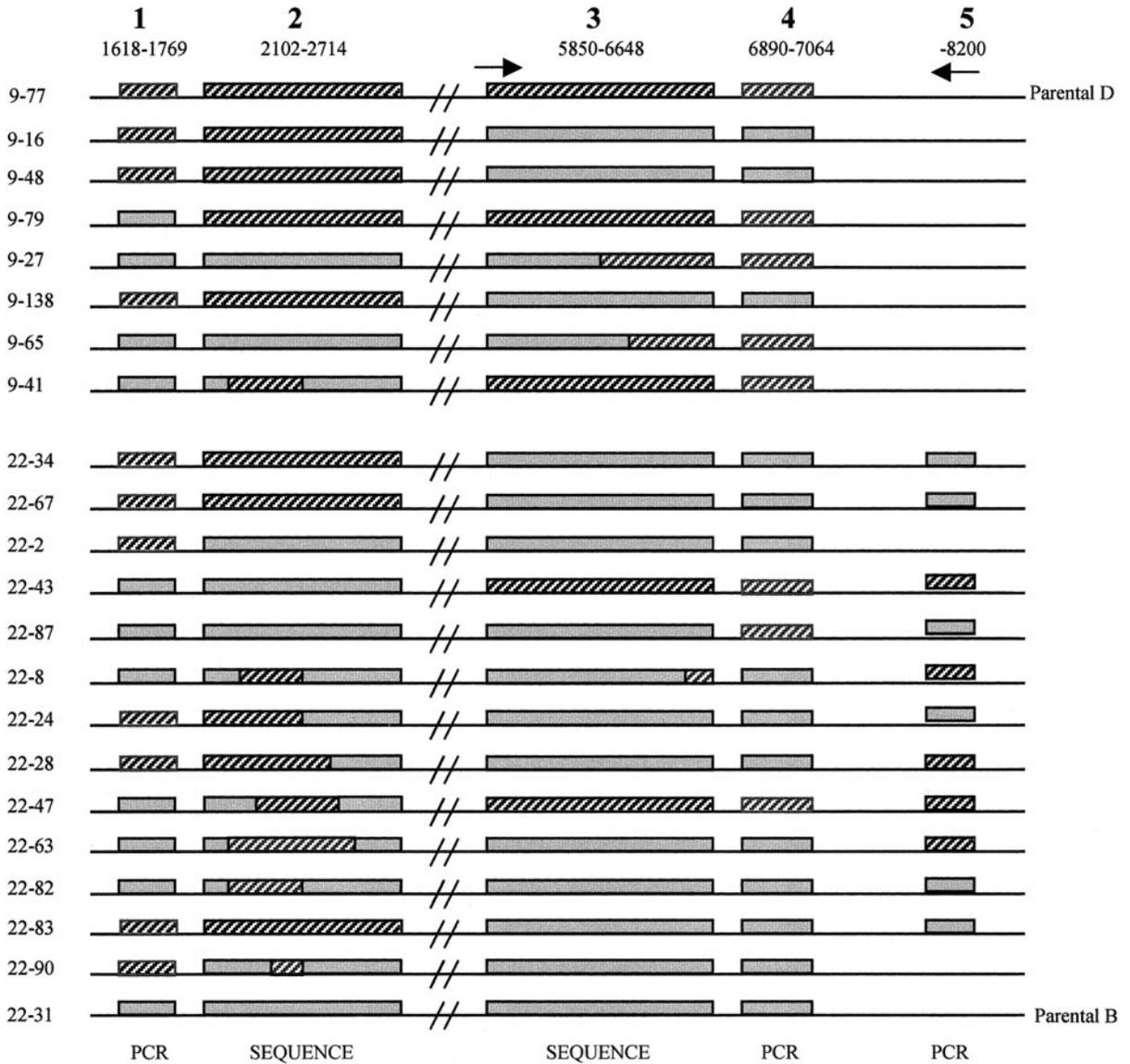
incorporation of proteins or protein fragments from the two subtypes into the viruses through genetic recombination.

DISCUSSION

In this study we have quantified and analyzed the appearance of recombinant viruses with HIV-1 strains from B, D, and F subtypes in MT-4 cocultivations. Proportions of the diverse genetic forms produced in the viral cocultivations were determined on proviral DNA obtained from secondary cell cultures infected with viral cocultivation supernatants. This procedure was adopted because RT-PCR produced unspecific recombinant genomes from the mixture of two parental RNAs (data not shown), as previously described (Fang *et al.*, 1998). The estimation of the amounts of recombinant viruses in the secondary culture has the advantage that the detected provirus comes from the infective viruses in the supernatant of the viral cocultivation. The quantity of recombinant viruses in the secondary viral culture reflects the infective viruses produced in the cocultivation.

The quantification of parental and recombinant genomes in the cultures was based on PCR methodology. The amplification of the standard plasmid serial dilutions from each proviral DNA was linear (see Figs. 3 and 4), and the  $r^2$  values were greater than 0.90 in 85% of the cases. Quantification was carried out with the dilutions close to the limiting dilution where a better linear correlation exists. In addition, to minimize the variability of the assay, all amplifications were performed on the same proviral DNA dilution in parallel amplifications, and each quantification was carried out in three independent experiments. The estimated value is the mean of the three replicates (see standard deviations in tables). The quantification performed yields a percentage with respect to the parental strains and is thus a relative estimate. The PCR system with crossed primers employed to detect the recombinant forms does not identify all possible recombinant viruses; it discriminates those in which the sequences hybridizing with the primers belong to different strains, but not recombinant genomes with an uneven number of internal crossovers. It should be stressed that the recombinant genomes we detected in the study are not a single homogeneous population but rather a collection of the different recombinant forms present in the culture.

For recombination to occur, the first and most limiting factor is the coinfection of the same cell by two different viruses. This depends on the multiplicity of infection (m.o.i.) and the characteristics of the infecting viruses. The strains selected for the study displayed similar growth and biological characteristics, including coreceptor usage, as shown in Fig. 1 and Table 1. Strains with similar replication dynamics will favor cell coinfection, whereas viruses with a very rapid replication cycle will



**FIG. 6.** Schematic representation of genetic characteristics in different genomic regions of biological clones from B/D cocultivation. In this characterization we considered the 1618–1769 region of *gag* gene corresponding to the p24 protein (labeled 1) and the region of 6890–7067 (labeled 4) in the 3′-end of *env* gene both analyzed by length variation of the amplicons. Two regions were sequenced, one in *pol* gene (2101–2714) (labeled 2) and another in *env* gene (5850–6649) (labeled 3), and finally, the amplification of the downstream primer in position 8200 in *env* gene (labeled 5) used in the PCR amplification. Each line represents the genome of each of the clones analyzed. The first line represents the genotype of the parental subtype D 92BR020 strain, which is identified by the dark hatching; in the bottom line the genotype of the LAI strain is identified by the light shading. The number on the left identifies the clones that were picked on day 9, in the upper part, and those selected on day 22, at the bottom. Numbers at the top of the graph identify the genomic region analyzed; the techniques used are listed beneath. On day 9, in the *gag* gene of the seven clones analyzed, four were subtype B (frequency of 0.57) and three (0.43) were subtype D. In *pol* gene this distribution was 2 B (0.24), 4 D (0.57), and one recombinant form (0.14). In *env* gene the distribution was again 3 B (0.43) to 3 D (0.43) and one recombinant (0.14). On day 22, in region 1 (*gag* gene) of the 13 clones analyzed, 6 were B (0.46) and 7 were D (0.54). However, in *pol* gene the proportion was 3 B (0.23), 3 D (0.23), and seven recombinant forms (0.53) were detected. In *env* gene the subtype B sequences found in the clones were 10 (0.72) and only 2 D (0.13), and 1 (0.08) recombinant sequence was detected. In region 4, 10 clones were B (0.72) and 2 D (0.13). Finally, in region 5, the distribution was six clones of subtype B (0.77) and five clones of subtype D (0.33).

dominate the culture and displace the other virus. This could have occurred in the D/F cocultivation. On day 5, virus D replicated more rapidly and made up 94% of the culture, whereas virus F constituted only 5% of the total (see Table 1). Only 0.6% of recombinant genomes were detected. The proportion of parental strains was similar

on day 22 with 98% of subtype D genomes and 2% of subtype F, but no recombinant genomes were detected. Thus, the dissimilarity of the viral replication cycle is an important factor that could account for the lack of recombination. However, we do not know the cause of the difference in the viral replication of the D/F combination

because there was no such difference in the D/F combination, in which there was, in fact, a high percentage of recombination. Moreover, there were no significant differences in the viral dynamics between any of the strains tested on day 5 (see Table 2 and Fig. 1).

Factors other than cell coinfection may be also operating to cause the appearance of recombinant viruses. The first factor to be considered is the homology of the viral sequences, as previously described in spleen necrosis virus (SNV) (Zhang and Temin, 1994). There is evidence that the similarity of the sequences is important for recombination because, on day 5, we found a significant inverse correlation between the genetic distance and the proportion of recombination. Moreover, in the clonal analysis with the B/D combination we recorded more recombination events in *pol* gene (13) than in *env* gene (2) (see Fig. 6) because in general *env* displayed a two- to three-fold greater degree of divergence than did *pol* (Quiñones-Mateu and Domingo, 1999). Among other factors that could be intervening in the generation of recombinant viruses is the physicochemical stability of the heterozygous virions, because of the more labile interaction or hindrance between proteins and nucleic acids from different subtypes. This interaction is also related to the genetic distance.

The clonal analysis performed in the B/D cocultivation on days 9 and 22 allowed us to examine whether the dominance of the recombinant genomes was due to clonal expansion of variants with greater fitness or to the generation of new variants. The phylogenetic trees in the two genes analyzed, shown in Fig. 6, displayed a widespread presence of independent sequences, but also five viruses (20%) formed a cluster in the *pol* tree. The analysis of the crossovers performed by the Simplot program confirmed these results detecting 19 break points among the clones, although a common break point was found in the viruses forming the cluster in *pol*. Thus, in the recombinant viruses present on day 22 in the cocultivation, the majority of the clones resulted from the generation of new recombinant viruses rather than to clonal expansion.

The changes observed in the proportions of the sequences found in the cocultivations are a reflection of the relative presence of the two proteins of the subtypes. No changes were observed in *gag* (fragment 1) over time. However, in the *env* fragment 3, the increasing dominance with time of the subtype B gp120 protein, from 3 of 7 (0.43) on day 9, to 10 of 13 (0.77) on day 22, implies that it is fitter than subtype D protein. In contrast, in the reverse transcriptase, the viral cocultivation resulted in a substantial presence of recombinant proteins in which subtype D was dominant. These results indicate that recombination is favored in reverse transcriptase but not in gp160, and that subtype D reverse transcriptase and subtype B gp160 displayed higher fitness in this viral combination. The result in the *env* region suggests

that the diagnostic PCR fragment, which also included one of the hot spots defined by Jetzt *et al.* (2000), with 22% recombination, lies outside the sequenced region.

In this *in vitro* study, in spite of the limitations of the culture system, we have obtained several results that could further our understanding of the natural history of HIV-1. We found that gp120 from subtype B was selected over subtype D in the B/D cultivation. In contrast, we detected more recombination in the RT protein where subtype D protein predominated over subtype B. Although these differences were derived from a small number of clones and time points, these results nevertheless illustrate the selection of some inter-subtype recombinant genomes. *In vivo*, in the circulating recombinant forms (CRFs), crossovers are found frequently at the 5'-end of the RT (see <http://www.lanl.gov>). This result could also be the consequence of secondary structures in the RNAs (Wooley *et al.*, 1998). There were very few break points in the sequenced region at the 5'-end of the gp120, as also happens occasionally *in vivo* (2 of the 14 CRFs). Our results confirm that some regions are more prone to recombination than others, as previously described (Jetzt *et al.*, 2000). Additionally, as shown in the D/F combination, some recombinant viruses are generated with greater difficulty because of the differences in the viral cycles. One of the most important results of the study is the observation of the dominance of recombinant forms over one of the parental viruses, as was seen in the three combinations with significant recombination (see Table 2). This is supported by the dynamic analysis performed in the B/B and B/F combinations and by the fitness values of the recombinant forms in relation to the parental strains. The evolutionary consequence of the loss of one parental virus is a population with smaller genetic distances between sequences but with higher heterogeneity because of the generation of recombinant forms and the production of overall genetic convergence. In this scenario of more similar sequences, new recombination events are favored.

It is important to highlight that in our quantification system, we have only screened approximately 25% of the genome (2400 nucleotides in *env* gene out of the 9400 nucleotides of the proviral genome). Although we have obtained evidence that recombination occurs preferentially in defined regions, if the high percentages obtained could be extrapolated to the complete genome, then almost all genomes would be recombinant by day 22.

In summary, the results of an *in vitro* study have provided insights into the quantification, dynamics, prevalence, and nature of recombinant viruses, and the evolutionary consequences of recombination, which could further our understanding of the natural history and biology of HIV-1.

## MATERIALS AND METHODS

### Cells and viruses

Cocultivations of two HIV-1 strains were carried out in MT4 cell cultures grown in RPMI with 10% fetal calf serum. Lai and RF strains were used for subtype B; for subtype D we employed strain 92UG24, and we employed strain 93BR020 for subtype F. All strains were obtained from the UNAIDS/NIH/AIDS repository.

### Biological characteristics of the strains

The strains were selected on the basis of them having similar biological properties such as coreceptor usage, virus titer, infectivity, and virus production. Viral strains used the CXR4 coreceptor and formed syncytia. They were grown in MT-4 cells and titered in an MT-4 plaque assay, as described in Harada *et al.*, 1985 with modifications (Sánchez-Palomino *et al.*, 1993b), and by employing the TCID<sub>50</sub> method of Spearman and Kärber. RT activity was measured in 10  $\mu$ l of the cocultivation supernatant, essentially as described in Victoria Johnson, 1990. Cell viability was calculated by the trypan blue dye method.

### Viral cocultivations

Four cocultivations of two viruses were studied, three of them between subtypes B/D, B/F, and D/F, and one within B subtype. MT-4 cells ( $2 \times 10^6$ ) were simultaneously infected with two HIV-1 strains, at an m.o.i. of 0.1 PFU/cell for each virus. Both viruses were mixed prior to the addition to the cells, adsorbed for 1 h and 45 min at 37°C, and later removed by centrifugation. The cell pellet was washed with PBS and resuspended in fresh medium. Viral infections were maintained for 22 days with the addition of cells when the cytopathic effect was evident:  $2 \times 10^6$  cells on day 5, and  $5 \times 10^6$  cells on days 9, 13, 16, and 19. Supernatants from the cultures were recovered on days 5, 9, 13, 16, 19, and 22, and the cell pellet was recovered on day 22. Intra-B subtype (Lai/RF) cocultivation was carried out under the same conditions, although cells were added on days 8, 13, and 17 and recovered on day 21, due to the different dynamics of viral replication.

### Detection of recombinant viruses

To detect the recombinant forms arising in the viral cultivations, cell cultures infected with 500  $\mu$ l of the supernatants on days 5, 9, 13, 16, 19, and 22 were used to infect fresh  $5 \times 10^6$  MT-4 cells. DNA was extracted from these cells and recovered 24 h postinfection by standard techniques. It was quantified at 260 nM and the HIV-1 proviral DNA was estimated. Exceptionally, due to low cell numbers, DNA from the Lai/RF cocultivation on day 5 was extracted with a kit (Instagene, Bio-Rad, CA).

The different proviruses in the cultures were detected

by specific PCR amplification in *env* gene in a region of approximately 2500 nucleotides from the signal peptide to 100–200 nt before the end of gp41 protein. To amplify the HIV-1 proviral DNA from the different strains, a set of different primers was designed. Every primer pair was specific for each virus and did not amplify the other virus in the cocultivation. The recombinant forms were detected using crossed pairs, one primer from each of the parental viruses (one up and another down), as previously described (Kuwata *et al.*, 1997). The B-specific primers in the B/D cocultivation were 137B, 139B and 155D, 154D for D amplification. In the B/F culture, primers 137b and 143b were used for the amplification of subtype B and 141f, 142f for F. To detect the presence of D subtype, primers 155d and 154d were used, and primers 146f, 142f for the detection of subtype F in the DF cocultivation. Within subtype B, primers 137B, 143B were specific for LAI and primers 273R, 274R for RF strain. The sequence and localization of the primers is as follows:

137B (5'-GTGAAGGAGAAATATCAGCA-3', nucleotides (nt) 5811 to 5831)

139B (5'-CTGCTATGGCTGTGGCAT-3', nt 8265 to 8247)

155D (5'-CTGTGTCACCTTTAACTGCATTGAAT-3', nt 6175 to 6201)

154D (5'-TCCTGAATCCAATACTGGAG-3', nt 8218 to 8198)

143B (5'-TCTACAAGCTCCTTGTACTA-3', nt 8314 to 8294)

141F (5'-TGAGGGGGATGCAGAGGAA-3', nt 5810 to 5829)

142F (5'-TCTACCAGCTCTTTGCAAAG-3', nt 8314 to 8294)

146F (5'-GGGGAAGTGGGGCCCTTTAT-3', nt 5840 to 5860)

273R (5'-TGATGGAGATGAGGAAGAA-3', nt 5811 to 5830)

274R (5'-GTGAAGAAAAGCTCTAAGAA-3', nt 8308 to 8327)

The nucleotide position of the primers in the genome is in relation to the HXB2 sequence (Ratner *et al.*, 1985). Amplification was carried out for the B/D, B/F, and D/F cultures with the following amplification cycles: 94°C/2 min and 10 cycles of 95°C/45 s, 60°C/1 min, and 72°C/4 min, followed by 25 cycles of 95°C/45 s, 55°C 1 min, 72°C/4 min, and a final incubation at 72°C for 10 min. The B/B (RF/LAI) amplification was similar, but hybridization temperatures were 63 and 57°C, respectively.

### Quantification of the proviral species by quantitative PCR

To quantify the parental and recombinant forms present in the different cultures, control plasmids were constructed, inserting amplicons obtained from PCR carried out with specific primers for the parental viruses and with crossed primers for the recombinant viruses (see Fig. 2). Amplicons were cloned in plasmid pCR2.1 using the TA cloning kit (Invitrogen CA). In each viral cocultivation, quantification of each of the four proviral DNAs was carried out on the same DNA sample obtained from the coculture. The two parental strains and the two recombinant forms were quantified by comparison of

serial dilutions of total DNA from the coculture with corresponding dilutions of the DNA standard (six tubes of the standard and four of the culture). Quantification was achieved by comparing the band intensity of each amplified product with the corresponding control plasmid in 0.9% agarose gels in TAE. The intensity of the bands was measured using a densitometer with the help of the PC Bas program. An example of this procedure is shown in Fig. 3. The proviral DNA copy number was determined by extrapolating from the regression line derived from the different DNA concentrations of the corresponding control plasmid (data not shown). This DNA quantification was carried out for each virus (the two parental strains and the two recombinants species) and the final copy number expressed as percentages. This number is the mean of the three replicates. A total of 278 lineal regressions was calculated, yielding  $r^2$  values greater than 0.95 in 69% of cases and greater than 0.90 in 85% of cases (data not shown).

### Isolation and characterization of biological recombinant clones

To isolate and characterize recombinant viruses, biological clones were obtained by the MT-4 plaque assay of culture supernatants, as previously described (Harada *et al.*, 1985). Individual plaques were picked at random and resuspended in 100  $\mu$ l of complete medium and kept frozen ( $-70^\circ\text{C}$ ) until analysis. DNA was extracted from each clone in 20  $\mu$ l of this medium following the In-stagene kit (Bio-Rad) instructions. Recombinant viruses were detected by PCR in the two genes, *gag* and *env*. To this end the following primers were used: 117GD (5'-CCCCAGACCTGAAGCTCTCTT-3' nt 1748 to 1768) and 118GU (5'-AGACAGGTTAATTTTTTAGGGAA-3' nt 1618 to 1645) in *gag* gene, and 114EU (5'-TCCTCAGGAGGGGACCCAGAAA-3', nt 6890 to 6912) and 115ED (5'-TTTTAT-TCTGCATGGGAGGGTGAT-3' nt 7064 to 7040) in *env* gene. The two primer pairs gave four amplicons of different sizes that could be resolved in 0.9% TAE agarose gels. *Gag* primers gave fragments of 152 nt for subtype B Lai strain and of 117 nt for subtype D 92Ug24. The *env* amplification produced fragments of 176 nt from B virus and 167 nt for D strains. Eight plaques with a recombinant genome were selected from the B/D culture supernatant on days 9 and 14 recombinant from day 22.

### Sequence analysis of the biological recombinant viruses

Biological clones were characterized by nucleotide sequencing in two regions in *pol* and *env* genes. These regions included the two fragments with the most cross-overs in the HIV-1 genome (Jetzt *et al.*, 2000). *pol* gene was analyzed from nucleotide 2102 to 2714 region covering a small fragment of protease and the first half of RT up to residue 253. The primers used were 47RU (5'-

AGTAGGACCTACACCT-3' nt 2060 to 2075) and 20RD (5'-ACAGTCCAGCTGTCTTTTTCTGGC-3' nt 2866 to 2892). In *env* gene the region analyzed was from nucleotide 5850 to nucleotide 6648, including a fragment from the signal peptide to the C2 region. The primers used were 48EU (5'-TAGAAAGAGCAGAAGACAGTGGC-3' nt 5778 to 5800) and 115ED (5'-GTTTTATTCTGCATGGGAGGGTGAT-3' nt 7040 to 7064). The sequenced fragment was of 612 nt in the *pol* region and of 798 nt in the *env* region.

Sequencing was performed in an ABI-prism 377. Sequences were aligned using the CLUSTALW program, version 1.7 (Thompson *et al.*, 1994), and subsequently edited manually. Genetic distances tables were estimated with Kimura's two-parameter algorithm and using the MEGA2 program. Phylogenetic trees were constructed using the Neighbor-Joining method and edited with the Tree View program.

Recombinant sequences and break points were analyzed with the help of the SIMPLOT program (Ray, 1999) with a window of 200 nucleotides and a sliding step of 20 nucleotides. Genetic distances between the sequences of strains were calculated with the MEGA 2 program (Kumar *et al.*, 2001) using sequences retrieved from the Los Alamos database. Genetic distances were estimated in 10.5% for the two B strains, 17% in the B/D mixture, 21% in the B/F, and 21.9% in the D/F strains.

### ACKNOWLEDGMENTS

We acknowledge Antonio Borderias for the statistical calculations, and Isabel Olivares and Concepción Casado for helpful suggestions. MT-4 cells were kindly provided by Dr. D. D. Richman (UCSD). The different strains were obtained from the Joint United Nations Program on HIV/AIDS through the NIH/AIDS Research and Reference Reagent Program.

### REFERENCES

- Barton, N. H., and Charlesworth, B. (1998). Why sex and recombination? *Science* **281**(5385), 1986–1990.
- Björndal, Å., Deng, H., Jansson, M., Fiore, J. R., Colognesi, C., Karlsson, A., Albert, J., Scarlatti, G., Littman, D. R., and Fenyö, E. M. (1997). Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *J. Virol.* **71**, 7478–7487.
- Casado, C., Garcia, S., Rodriguez, C., del Romero, J., Bello, G., and Lopez-Galindez, C. (2001). Different evolutionary patterns are found within human immunodeficiency virus type 1-infected patients. *J. Gen. Virol.* **82**(Pt. 10), 2495–2508.
- Coffin, J., Hughes, S., and Varmus, A. H. (1997). "Retroviruses." Cold Spring Harbor Laboratory Press, New York.
- Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S., and Holland, J. J. (1996). Basic concepts in RNA virus evolution. *FASEB J.* **10**, 859–863.
- Doms, R. W., Edinger, A. L., and Moore, J. P. (1998). Coreceptor use by primate lentivirus. In "Human Retrovirus and AIDS" (B. Korber, C. Kuiken, B. Folley, B. Hahn, F. McCutchan, J. Mellors, and J. Sodroski, Eds.), pp. III–36–79. Theoretical Biology and Biophysics, Group T 10, Los Alamos National Laboratory, Los Alamos, NM.
- Escarmis, C., Davila, M., and Domingo, E. (1999). Multiple molecular

- pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* **285**(2), 495–505.
- Fang, G., Zhu, G., Burger, H., Keithly, J. S., and Weiser, B. (1998). Minimizing DNA recombination during long RT-PCR. *J. Virol Methods* **76**(1-2), 139–148.
- Harada, S., Koyanagi, Y., and Yamamoto, N. (1985). Infection of HTLV-III/LAV in HTLV-I-carrying cells MT-2 and MT-4 and application in a plaque assay. *Science* **229**, 563–566.
- Hoelscher, M., Kim, B., Maboko, L., Mhalu, F., von Sonnenburg, F., Birx, D. L., and McCutchan, F. E. (2001). High proportion of unrelated HIV-1 intersubtype recombinants in the Mbeya region of southwest Tanzania. *AIDS* **15**(12), 1461–1470.
- Holland, J. J., de la Torre, J. C., Clarke, D. K., and Duarte, E. (1991). Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *J. Virol* **65**(6), 2960–2967.
- Hu, W. S., and Temin, H. M. (1990a). Genetic consequences of packaging two RNA genomes in one retroviral particle: Pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. USA* **87**(4), 1556–1560.
- Hu, W. S., and Temin, H. M. (1990b). Genetic consequences of packing two RNA genomes in one retroviral particle: Pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. USA* **87**, 1556–1560.
- Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D., and Dougherty, J. P. (2000). High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **74**(3), 1234–1240.
- Kampinga, G. A., Simonon, A., Perre, P. V. D., Karita, E., Msellati, P., and Goudsmit, J. (1997). Primary infections with HIV-1 of women and their offspring in Rwanda. Findings of heterogeneity at seroconversion, coinfection, and recombinations of HIV-1 subtypes A and C. *Virology* **227**, 63–76.
- Kellam, P., and Larder, B. A. (1995). Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. *J. Virol.* **69**(2), 669–674.
- Korber, B., Kuiken, C., Foley, B., Hahn, B., McCutchan, F., Mellors, J. W., Sodroski, J., Eds. (1998). Human retroviruses and AIDS 1998: A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.
- Kuiken, C., Folley, B., Hahn, B., Marx, P., McCutchan, F., Mellors, J., Mullins, J., Wolinsky, S., Korber, B., and (Eds) (1999). "Human Retrovirus and AIDS." Theoretical Biology and Biophysics, Group T 10, Los Alamos National Laboratory, Los Alamos, New Mexico.
- Kumar, S., Tamura, K., Jakobsen, I. B., and Nei, M. (2001). MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* **17**(12), 1244–1245.
- Kuwata, T., Miyazaki, Y., Igarashi, T., Takehisa, J., and Hayami, M. (1997). The rapid spread of recombinants during a natural in vitro infection with two human immunodeficiency virus type 1 strains. *J. Virol.* **71**, 7088–7091.
- Lai, M. M. (1992). Genetic recombination in RNA viruses. *Curr. Top. Microbiol. Immunol.* **176**, 21–32.
- Liitsola, K., Tashkinova, I., Laukkanen, T., Korovina, G., Smolskaja, T., Momot, O., Mashkilleysen, N., Chaplinskias, S., Brummer-Korvenkontio, H., Vanhatalo, J., Leinikki, P., and Salminen, M. O. (1998). HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad. *AIDS* **12**(14), 1907–1919.
- Mansky, L. M., and Temin, H. M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**, 5087–5094.
- McCutchan, F. E. (2000). Understanding the genetic diversity of HIV-1. *AIDS* **14**(Suppl. 3), S31–S44.
- Moutouh, L., Corbeil, J., and Richman, D. D. (1996). Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc. Natl. Acad. Sci. USA* **93**, 6106–6111.
- Quiñones-Mateu, M. E., and Domingo, E. (1999). Nucleotide diversity in three different genomic regions of Venezuelan HIV type 1 isolates: A subtyping update. *AIDS Res. Hum. Retroviruses* **15**(1), 73–79.
- Ray, S. C. (1999). SimPlot for Windows 95/NT 2.4. Baltimore, MD: Distributed by author.
- Rucker, J., Edinger, A. L., Sharron, M., Samson, M., Lee, B., Berson, J. F., Yi, Y., Margulies, B., Collman, R. G., Doranz, B. J., Parmentier, M., and Doms, R. W. (1997). Utilization of chemokine receptors, orphan receptors, and herpesvirus-encoded receptors by diverse human and simian immunodeficiency viruses. *J. Virol.* **71**, 8999–9007.
- Sánchez-Palomino, S., Rojas, J. M., Martínez, M. A., Fenyő, E. M., Nájera, R., Domingo, E., and López-Galíndez, C. (1993a). Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. *J. Virol.* **67**(5), 2938–2943.
- Sánchez-Palomino, S., Rojas, J. M., Martínez, M. A., Fenyő, E. M., Nájera, R., Domingo, E., and López-Galíndez, C. (1993b). Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. *J. Virol.* **67**, 2938–2943.
- Thompson, J., Higgins, D., and Gibson, T. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
- Thomson, M. M., Villahermosa, M. L., Vazquez-de-Parga, E., Cuevas, M. T., Delgado, E., Manjon, N., Medrano, L., Perez-Alvarez, L., Contreras, G., Carrillo, M. G., Salomon, H., and Najera, R. (2000). Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina [letter]. *AIDS* **14**(7), 897–899.
- Trkola, A., Ketas, T., Kewalramani, V. N., Endorf, F., Binley, J. M., Katinger, H., Robinson, J., Littman, D. R., and Moore, J. P. (1998). Neutralization sensitivity of human immunodeficiency virus type 1 primary isolates to antibodies and CD4-based reagents is independent of coreceptor usage. *J. Virol.* **72**, 1876–1885.
- Victoria Johnson, R. E. B. A. J. K. (1990). Reverse transcriptase (RT) activity assay. In "Techniques in HIV Research" (A. A. B. D. Walker, Ed.), pp. 98–102. Stockton Press, New York.
- Wooley, D. P., Bircher, L. A., and Smith, R. A. (1998). Retroviral recombination is nonrandom and sequence dependent. *Virology* **243**(1), 229–234.
- Yuste, E., Sánchez-Palomino, S., Casado, C., Domingo, E., and López-Galíndez, C. (1999). Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *J. Virol* **73**(4), 2745–2751.
- Zhang, J., and Temin, H. M. (1994). Retrovirus recombination depends on the length of sequence identity and is not error prone. *J. Virol.* **68**, 2409–2414.