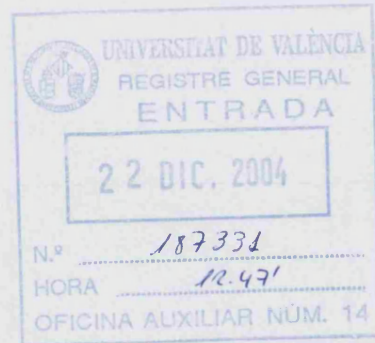




VNIVERSITAT  
Q D  
VALENCIA



INSTITUTO CAVANILLES DE BIODIVERSIDAD Y BIOLOGÍA EVOLUTIVA

ROBUSTEZ GENÉTICA, EPISTASIA Y EVOLUCIÓN DE LOS  
VIRUS DE RNA

Rafael Sanjuán Verdeguer

2005

UMI Number: U607706

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607706

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

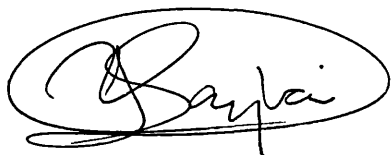
UNIVERSITAT DE VALÈNCIA  
BIBLIOTECA CÈNTRIC  
Nº Registre ..... 23.410 .....  
DATA ..... 17-02-05 .....  
SIGNATURA Biblioteca  
TD-B 758  
Nº LINDS: i18679316

30 cm.

b16612474

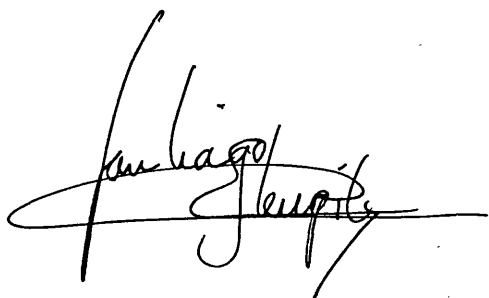


Memoria presentada por Rafael Sanjuán Verdeguer para optar al grado de Doctor en Biología por la Universidad de Valencia.

A handwritten signature in black ink, enclosed within an oval shape. The signature is cursive and appears to read 'Rafael'.

Rafael Sanjuán Verdeguer.

Este trabajo ha sido dirigido por los Doctores Santiago Francisco Elena Fito y Andrés Moya Simarro.

A handwritten signature in black ink, consisting of several overlapping loops and strokes. The name 'Santiago' is partially legible.

Santiago Francisco Elena Fito.

A handwritten signature in black ink, featuring a large, stylized 'M' and 'S'.

Andrés Moya Simarro.



# ÍNDICE

<b>PRÓLOGO</b> .....	<b>3</b>
<b>INTRODUCCIÓN</b> .....	<b>7</b>
<b>MUTACIÓN Y SELECCIÓN</b> .....	<b>9</b>
Efecto de las mutaciones .....	9
Equilibrio mutación-selección .....	11
Mutación recurrente .....	12
Epistasia, mutación y selección .....	14
<b>EVOLUCIÓN ADAPTATIVA</b> .....	<b>17</b>
El proceso de adaptación .....	17
Tasa de adaptación en asexuales .....	18
Tasa de adaptación y recombinación.....	19
Paisajes adaptativos .....	21
<b>DERIVA GENÉTICA</b> .....	<b>22</b>
Tamaño poblacional efectivo .....	22
Trinquete de Muller .....	23
Epistasia y trinquete de Muller .....	25
Deriva, epistasia y recombinación.....	27
Deriva y adaptación .....	28
<b>EL VIRUS DE LA ESTOMATITIS VESICULAR</b> .....	<b>29</b>
Características generales.....	29
Estructura del genoma y del virión .....	30
Ciclo de infección .....	32
<b>OBJETIVOS</b> .....	<b>35</b>
<b>MATERIAL Y MÉTODOS</b> .....	<b>39</b>
Estructura del clon infeccioso .....	41
Selección de las mutaciones a introducir .....	42
<b>TÉCNICAS BÁSICAS</b> .....	<b>43</b>
Cultivo celular .....	43
Infección con VSV en líquido .....	45
Titulación de VSV .....	45
Producción de anticuerpo monoclonal I <sub>1</sub> .....	46
Stock e infecciones con el virus Vaccinia vTF-7.....	46
<b>OBTENCIÓN DE LOS VIRUS MUTANTES</b> .....	<b>47</b>
Mutagénesis dirigida.....	47
Purificación del clon mutante .....	49
Transfección .....	50
Manejo de las muestras de transfección .....	52
Secuencias consenso de poblaciones de VSV.....	53
<b>ENSAYOS DE COMPETENCIA</b> .....	<b>54</b>
Creación de un fenotipo marcador .....	54
Ensayos de competencia .....	55
<b>ESTIMACIÓN DE PARÁMETROS GENETICO-POBLACIONALES</b> .....	<b>57</b>
Cálculo de las eficacias relativas .....	57
Obtención de las eficacias relativas expresadas en generaciones .....	57
Tamaño poblacional efectivo y número de generaciones .....	58
Estimación de la tasa de mutación.....	59
<b>RÉPLICAS Y BLOQUES EXPERIMENTALES</b> .....	<b>61</b>

<b>RESULTADOS</b> .....	<b>63</b>
¿Son las mutaciones únicas? .....	65
Mutaciones letales .....	66
Proporción de mutaciones deletéreas, neutrales y beneficiosas .....	67
Eficacias y coeficientes de selección.....	68
Características generales de las distribuciones de efectos .....	70
Distribución de efectos deletéreos .....	71
Distribución de efectos beneficiosos .....	74
Cambios previamente descritos en clones y en secuencias consenso .....	75
Grado de restricción funcional en diferentes genes.....	76
<b>TASA DE MUTACIÓN</b> .....	<b>77</b>
Prueba de fluctuación de Luria-Deibrock .....	77
<b>EPISTASIAS ENTRE PARES DE SUSTITUCIONES NUCLEOTÍDICAS</b> .....	<b>80</b>
Combinaciones de mutaciones aleatorias .....	80
Epistasias entre pares de mutaciones beneficiosas .....	82
Correlación entre el efecto epistático y el coeficiente de selección .....	83
<b>RECUPERACIÓN DE LA EFICACIA BIOLÓGICA</b> .....	<b>84</b>
Reversiones y evolución adaptativa .....	84
Evidencia de evolución compensatoria .....	86
Compensación de mutaciones simples y dobles .....	88
<b>DISCUSIÓN</b> .....	<b>93</b>
Robustez genética de los virus de RNA .....	95
Patrones de robustez genética .....	97
Determinación de los coeficientes de selección .....	98
Distribución de efectos mutacionales .....	99
Efectos deletéreos y mutaciones previamente descritas .....	101
<b>CONSECUENCIAS EVOLUTIVAS DE LAS MUTACIONES DELETÉREAS</b> .....	<b>103</b>
Carga genética de las poblaciones.....	103
Umbral de error .....	104
Tasa de mutación óptima.....	106
Trinquete de Muller.....	109
<b>MUTACIONES BENEFICIOSAS Y EVOLUCIÓN ADAPTATIVA</b> .....	<b>111</b>
Proporción de mutaciones beneficiosas .....	111
Tasas de adaptación e interferencia clonal.....	112
Efectos beneficiosos y mutaciones fijadas .....	112
<b>EFFECTOS EPISTÁTICOS</b> .....	<b>114</b>
Génesis de la epistasia .....	114
Patrones de epistasia .....	115
Epistasia y robustez genética.....	117
<b>IMPLICACIONES EVOLUTIVAS DE LA EPISTASIA</b> .....	<b>120</b>
Efecto sobre el trinquete de Muller .....	120
Evolución compensatoria .....	121
Signo de la epistasia y adaptación.....	123
Epistasia y recombinación .....	125
<b>REFERENCIAS</b> .....	<b>129</b>
<b>APÉNDICES</b> .....	<b>1455</b>



## PRÓLOGO

La biología evolutiva ha sido una disciplina históricamente teórica, debido a que la escala temporal evolutiva en los organismos que acaparaban la atención de la investigación dificultaba la experimentación. La obtención de datos no iba más allá de los estudios comparados y del incompleto registro fósil. Sin embargo hoy en día, el panorama ha cambiado gracias al estudio de sistemas bacterianos y virales, pero también de eucariotas unicelulares e incluso de invertebrados. Además de la principal ventaja de los sistemas microbianos, que reside en su rápida evolución, éstos son fáciles de manejar, permiten realizar gran cantidad de réplicas y pueden ser almacenados durante largos periodos, permitiendo comparar directamente las formas ancestrales con las evolucionadas. Finalmente, el avance de las técnicas moleculares facilita las comprobaciones directas mediante la manipulación genética. Así, muchas de las cuestiones evolutivas esenciales, como por ejemplo el origen de la diversidad y la complejidad, las bases genéticas de la adaptación, la naturaleza gradual o puntual de la evolución, la selección de grupo o el comportamiento social, están siendo actualmente abordadas mediante el diseño de experimentos controlados. Por otro lado, en la era de la genómica y la secuenciación, se produce una gran cantidad de información puramente descriptiva, que da pie al desarrollo de nuevas hipótesis y teorías explicativas.

Los virus de RNA son los sistemas biológicos con las mayores tasas de mutación conocidas. A esto se une su elevada velocidad de replicación, sus altos tamaños poblacionales y su escasa complejidad genómica, que los convierten en unos de los sistemas biológicos que más rápidamente evolucionan, por lo que resultan especialmente atractivos para la evolución experimental. El inicio de los estudios evolutivos con virus de RNA, que data de los años 80, se basaba en constatar su extraordinaria variabilidad genética y capacidad de adaptación. Más adelante, se ha demostrado que los virus de RNA son excelentes modelos para comprobar predicciones de la Genética de Poblaciones, como por ejemplo la acumulación de mutaciones deletéreas en ausencia de selección y recombinación, el principio de exclusión competitiva, o la selección dependiente de frecuencias génicas entre otros. Para ello se han usado modelos como los bacteriófagos  $\phi 6$  y  $Q\beta$ , el virus de la fiebre aftosa, el

de la estomatitis vesicular, el de la inmunodeficiencia humana de tipo 1, el de la rabia, el de la gripe y el del sarampión, pero también los de los mosaicos del tabaco y del pepino. Por lo tanto, el interés de estudiar los virus de RNA es también práctico, ya que constituyen un serio problema sanitario y económico mundial. Su éxito se debe en gran parte a su capacidad para evadir las vacunas, los tratamientos antivirales y al sistema inmunológico, precisamente debido a sus altas tasas de mutación y replicación. Un atractivo adicional de los virus de RNA es que, a diferencia de los microorganismos celulares y parte de los virus de DNA, su genoma es lo bastante pequeño (no supera habitualmente las 12 kilobases) como para ser analizado en su totalidad. Sin embargo, hasta hace una década, los virus de RNA presentaban el inconveniente de no poder ser manipulados por ingeniería genética, debido a la inestabilidad bioquímica del RNA. Este problema ha sido superado gracias al desarrollo de las técnicas de genética reversa, que han permitido que su genoma completo sea retrotranscrito e insertado en un plásmido, en lo que se denomina un clon infeccioso de cDNA.

Esta tesis se enmarca dentro del campo de la evolución experimental, que pretende contrastar las predicciones formuladas por la Teoría de la Evolución, así como aportar observaciones que permitan elaborar nuevas hipótesis. Utilizaremos como modelo experimental un virus de RNA, basando nuestros experimentos en la manipulación genética del virus y el estudio de su eficacia biológica en el laboratorio. Antes de entrar en materia, me gustaría dar las gracias a Santiago Elena por haberme ofrecido la posibilidad de realizar esta tesis, pero también por haberme sabido dirigir contra viento y marea, por haberme transmitido parte de sus conocimientos y por implicarse no sólo en el proyecto que ahora termina, sino también en mi futuro profesional. Desde luego, esta tesis no podría haberse llevado a cabo sin la confianza que mi co-director Andrés Moya ha depositado en esta línea de investigación. Gracias por lo tanto también a él. Por supuesto, mis agradecimientos no terminan aquí, ya que son muchas las personas que uno encuentra dispuestas a echar una mano a lo largo de más de cuatro años de trabajo. Me gustaría hacer mención a Cecilio López Galíndez, Encarna Martínez Salas e Isabel Suárez Novella por su ayuda con los protocolos, a Manuel Serra y a Edu por habernos ofrecido su laboratorio cuando nosotros perdimos el nuestro, a los técnicos como Ana y Raquel, cuya labor es muchas veces subestimada, a mis compañeros Paco y

Yolima por soportarme, a José por su buen humor y a Manoli por su sensibilidad, así como a Vicky por ser tan buena discípula. Recordaré las charlas con Borys el lingüista y las conversaciones metafísicas con Miguel, que espero no abandonemos. Tampoco puedo olvidarme de Toni por acogerme en Madrid mientras aprendía algunas técnicas y por su amistad. Desde luego, quiero reconocer lo importantes que han sido para mí los profesores que han sabido enseñarme a aprender, la mayoría de los cuales no veo desde el liceo. Desde antes, ya estaban mi padre y mi madre preocupándose por exigirme y darme lo mejor. Hoy sé que ellos son lo más valioso que tengo, junto con Regina, por compartir conmigo su tiempo y el mejor de los sentimientos.



## **INTRODUCCIÓN**



## MUTACIÓN Y SELECCIÓN

### Efecto de las mutaciones

La mutación es un arma de doble filo. Por un lado, es la fuente última de variación genética y la materia prima indispensable para la adaptación: un organismo que no mutase sería incapaz de adaptarse a los cambios ambientales y por lo tanto estaría condenado a la extinción. Por otro lado, es bien sabido que las mutaciones suelen tener un efecto adverso sobre los organismos. Uno de los principios básicos de la Teoría de la Evolución es que los seres vivos no son capaces de prever los efectos de la mutación (Lenski y Mitller, 1993). Por lo tanto, como la mutación no puede ser dirigida, es imposible beneficiarse de ella sin padecer sus efectos adversos. Además, las mutaciones deletéreas son mucho más abundantes que las beneficiosas (Bell, 1997), de manera que los seres vivos tendrán que enfrentarse primero a las consecuencias negativas de las primeras, apareciendo los beneficios de la segundas tan solo a medio o largo plazo.

La eficacia biológica ( $W$ ) es el número de descendientes con que un individuo contribuye a la siguiente generación. La eficacia se mide sobre los caracteres expresados o fenotipos, que tienen un componente ambiental y un componente genético. Las variaciones en este segundo componente, el genotipo, producidas en último término por la mutación, son las relevantes para la evolución biológica. El cambio de eficacia que experimenta un fenotipo a consecuencia de la selección natural se mide como

$$s = (w - w_0) / w_0 = W - 1, \quad (1)$$

donde  $w_0$  es la eficacia del fenotipo de referencia o silvestre y  $W$  la eficacia del individuo relativa al silvestre. Algunos autores<sup>2</sup> (Johnson, 1999) definen el coeficiente de selección como

$$s = \log W. \quad (2)$$

Esta definición también es conocida como coeficiente de selección Malthusiano.

La robustez genética se define como la capacidad de un fenotipo para mantener su eficacia biológica a pesar de la mutación. Por lo tanto, el coeficiente de selección asociado a una mutación es una medida de la robustez genética. Lógicamente, ha de tenerse conocimiento del tipo de mutación implicada, ya que los efectos de una sustitución no serán los mismos que los de una inserción, una deleción o un reordenamiento génico. Aunque en ciertos



organismos, como es el caso de *Drosophila melanogaster* (Eggleston *et al.*, 1988) o el de los humanos (Sebat *et al.*, 2004), pueden ser abundantes otros tipos de mutaciones, es comúnmente aceptado que son las sustituciones nucleotídicas las más habituales. Esto es especialmente cierto en el caso de los virus de RNA, donde constituyen más del 95% de las mutaciones identificadas (Novella, 2004a).

La estimación del coeficiente de selección asociado a mutaciones aisladas puede proporcionar una visión incompleta de la realidad, puesto que cuando éstas se presentan conjuntamente, surge la posibilidad de interacciones entre ellas, a las que se denomina epistasis. En el caso más sencillo, el de dos *loci* A y B con dos alelos, se define la epistasis ( $\varepsilon$ ) como

$$\varepsilon = W_{AB} - W_A W_B; \quad (3)$$

donde  $W_A$  y  $W_B$  son las eficacias relativas de los mutantes simples en A y B respectivamente y  $W_{AB}$  es la eficacia relativa del doble mutante. Un valor nulo de  $\varepsilon$  corresponde a un comportamiento multiplicativo de las eficacias biológicas. Cuando la eficacia observada del doble mutante ( $W_{AB}$ ) es mayor que la esperada según el modelo multiplicativo ( $W_A W_B$ ), el signo de la epistasis es positivo, mientras que si la eficacia observada es menor que la esperada según el modelo multiplicativo, el signo de la epistasis es negativo (Figura 1). La epistasis puede escribirse de manera equivalente expresando las eficacias en la

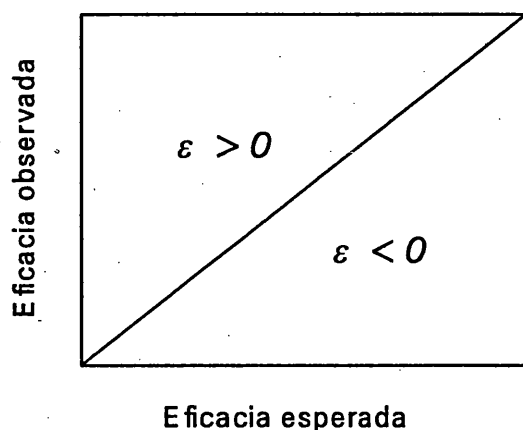


Figura 1. Definición del signo de la epistasis.

ecuación 3 en función del coeficiente de selección definidos según la ecuación 1, quedando



$$\varepsilon = s_{AB} - s_A - s_B - s_A s_B \quad (4)$$

donde  $s_A$  y  $s_B$  son los coeficientes de selección asociados a las mutaciones en  $A$  y  $B$  respectivamente y  $s_{AB}$  es el coeficiente de selección asociado al doble mutante. Existe cierta confusión en los términos sinergismo y antagonismo. Según Wolf *et al.* (2000), cuando se habla de mutaciones beneficiosas, la epistasia es sinérgica si tiene signo positivo y antagonista tiene signo negativo. Por el contrario, cuando hablamos de efectos deletéreos, la epistasia es antagonista si tiene signo positivo y sinérgica si tiene signo negativo. Sin embargo, según Barton y Charlesworth (1998), la epistasia sinérgica es la de signo negativo y epistasia antagonista a de signo positivo, se trate de mutaciones beneficiosas o deletéreas (revisado en Barton y Charlesworth, 1998). Para evitar confusiones, en esta tesis nos limitaremos a hablar de epistasia positiva y negativa.

#### Equilibrio mutación-selección

Para entender la importancia de la mutación en la evolución, es necesario comprender los efectos que ésta tiene sobre las poblaciones. Dado que los cambios con efecto deletéreo son mucho más abundantes que los de efecto beneficioso, la consecuencia primera de la mutación será reducir la eficacia de la población. Los individuos afectados tenderán a ser eliminados de ésta por selección purificadora, pero continuamente estarán apareciendo nuevos mutantes. Cuando el flujo en uno y otro sentido se equilibran, se dice que la población está en el equilibrio mutación-selección (Haldane, 1937). En una población infinita, si se ignoran las interacciones génicas, las mutaciones que reviertan al genotipo silvestre y los efectos pleiotrópicos, la eficacia media de una población haploide de tamaño infinito en el equilibrio viene dada por

$$\bar{W} = e^{-U}; \quad (5)$$

donde  $U$  es la tasa genómica de mutación deletérea (Haldane, 1937; Kimura y Maruyama, 1966). La disminución de eficacia en el equilibrio se denomina carga mutacional o lastre genético, y se define como

$$L = 1 - \bar{W}. \quad (6)$$

Curiosamente, como puede verse en la ecuación 5, bajo las asunciones mencionadas,  $L$  no depende de los efectos mutacionales. Sin embargo, la velocidad con la que se alcanza el equilibrio es inversamente proporcional al

coeficiente de selección Malthusiano promedio y aumenta al aumentar la varianza de  $s$ , siendo pues más lento el ritmo de cambio de eficacia cuanto mayor sea la asimetría de la distribución de efectos mutacionales (Johnson, 1999). En la mayoría de organismos, sin embargo, la forma que adopta esta distribución es aún una incógnita (Keightley y Lynch, 2003).

### **Mutación recurrente**

Una de las asunciones clásicas de la Genética de Poblaciones es que la mutación es un evento raro y que, por lo tanto el destino evolutivo de un genotipo mutante no se verá afectado por mutaciones adicionales. Esta aproximación es válida para tasas de mutación bajas, pero deja de serlo conforme éstas crecen y como consecuencia, la eficacia de un individuo no puede ser medida sin tener en cuenta la de su descendencia. Esta generalización es la que diferencia la Genética de Poblaciones clásica de la teoría de las cuasiespecies (Eigen, 1986; Schuster 1986; Eigen *et al.*, 1988). El léxico de esta teoría resulta algo peculiar para biólogos evolutivos puesto que proviene de la Química Física. Se habla de la concentración de una secuencia en lugar de la frecuencia de un genotipo. La secuencia maestra o silvestre, de máxima eficacia, replica con error, de manera que genera una nube de mutantes a su alrededor, que a su vez, replican también con error. El balance entre la tasa de replicación y la tasa de degradación química determina la eficacia de cada secuencia. Las concentraciones convergen hacia un estado estacionario o equilibrio mutación-selección, denominado cuasiespecie. El término *especie* no ha de ser entendido en su sentido biológico sino químico, mientras que el prefijo *cuasi* hace referencia a la heterogeneidad del conjunto de secuencias.

Más allá de los paralelismos semánticos, las predicciones de la teoría de las cuasiespecies son cualitativamente distintas de las de la Genética de Poblaciones clásica. En primer lugar, la eficacia media de la población no depende sólo de la tasa de mutación, sino también del valor medio y de la distribución de los coeficientes de selección de los mutantes (Schuster y Swetina, 1988; Wilke *et al.*, 2001; Krakauer *et al.*, 2002), de modo que la unidad de selección no es el individuo sino la cuasiespecie en su conjunto, lo que, en términos de la Genética de Poblaciones se denomina selección de grupo. En segundo lugar, existe un umbral en la fidelidad de la replicación,

denominado umbral de error, a partir del cual la distribución de mutantes se convierte en uniforme, diciéndose que la cuasiespecie se ha deslocalizado o entrado en catástrofe de error. El umbral de error depende de la fidelidad de la replicación a nivel genómico y de la superioridad de la secuencia maestra, es decir de la distribución de efectos, de la siguiente manera:

$$l_{\max} = \frac{\log \sigma_0}{\mu}; \quad (7)$$

donde  $l_{\max}$  es a longitud máxima del genoma,  $\mu$  la tasa de mutación por nucleótido y  $\sigma_0$  la superioridad de la secuencia maestra sobre su cuasiespecie, que se puede calcular como

$$\sigma_0 \approx 1 + \left\langle \frac{W_0 - W_i}{W_i} \right\rangle_{\log}; \quad (8)$$

donde  $\langle \rangle_{\log}$  denota la media geométrica (Eigen *et al.*, 1988). El umbral de error, expresado en términos de tasa de mutación genómica máxima ( $U_{\max}$ ), en lugar de tamaño genómico máximo, es

$$U_{\max} = \log \sigma_0. \quad (9)$$

Los efectos de las cuasiespecies comienzan a ser notorios para valores de  $NU > 10$  (van Nimwegen *et al.*, 1999), siendo  $N$  el tamaño poblacional. Es por ello que su utilidad práctica se manifestará en organismos con altos tamaños poblacionales y altas tasas de mutación. En gran medida, los virus de RNA cumplen estos requisitos. Su tasa de mutación está comprendida entre  $10^{-4}$  y  $10^{-6}$  sustituciones por sitio nucleotídico y evento de replicación. A nivel genómico, esto significa entre 0.01 y 1 sustituciones por genoma y ronda de replicación, lo que los destaca por encima del resto de sistemas biológicos (Tabla 1). La causa bioquímica de esta baja fidelidad es la ausencia tanto de mecanismos de prueba de lectura en la RNA polimerasa, es decir de actividad 3' exonucleasa (Steinhauer *et al.*, 1992), como de mecanismo de reparación del material genético post-replicación. Como consecuencia, la naturaleza de las poblaciones de virus de RNA es extremadamente variable (Domingo *et al.*, 1985; Steinhauer *et al.*, 1989). Además, los tamaños poblacionales, bien en cultivo o en hospedadores naturales, alcanzan con frecuencia los cientos o miles de millones (Holland *et al.*, 1982; Neumann *et al.*, 1998). Debido a estas peculiaridades, conceptos de la teoría de las cuasiespecies han encontrado gran aplicación entre los virus de RNA. Algunos experimentos han sugerido que las

Grupo	Especie	$U_r$
Ribovirus	VSV	1.11
	Rinovirus humano tipo 16	0.67
	Polio tipo 1	0.18
Retrovirus	Sarcoma de Rous	0.43
	HIV-1	0.22
	Leucemia bovina	0.03
Virus de DNA	Bacteriófago $\lambda$	0.0038
	Bacteriófago T2	0.0040
	Bacteriófago M13	0.0046
PROCARIOTAS	<i>Escherichia coli</i>	0.0025
Eucariotas	<i>Saccharomyces cerevisiae</i>	0.0027
	<i>Neurospora crassa</i> *	0.0030
	<i>Caenorhabditis elegans</i> *	0.004
	<i>D. melanogaster</i> *	0.005
	<i>Mus musculus</i> *	0.014

**Tabla 1.** Tasa de mutación genómica por evento de replicación,  $U_r$ , para diferentes sistemas biológicos. \*En regiones codificantes y reguladoras. Modificada de Drake *et al.* (1998) y Drake (1999).

tasas de mutación a las que los virus de RNA replican está cerca del umbral de error (Holland *et al.*, 1990), la existencia del cual ha sido demostrada experimentalmente (Crotty *et al.*, 2001). Sin embargo, otras importantes predicciones de la teoría, como es la selección de grupo, han sido verificadas en simulaciones con organismos digitales (Wilke *et al.*, 2001), pero no en virus de RNA.

#### Epistasia, mutación y selección

Otra de las asunciones realizadas a menudo por la Genética de Poblaciones es la ausencia de epistasias. Las consecuencias de la presencia de interacciones génicas dependen de la naturaleza asexual o sexual de la población. En una población asexual, la epistasia no modifica la eficacia media de la población en el equilibrio. Si la epistasia es negativa, los genotipos con más un cambio supondrán un mayor lastre genético que en el caso de efectos multiplicativos, por lo que la selección será más eficiente eliminándolos, así su frecuencia será menor y finalmente, el lastre genético se verá inalterado.

Análogamente, en el caso de epistasia positiva, los genotipos con más de un cambio supondrán un menor lastre genético de lo esperado, pero serán eliminados con mayor dificultad, dejando el lastre también inalterado en el equilibrio (Crow, 1970; Chao, 1988; Charlesworth, 1990; Shnol y Kondrashov, 1993).

El cambio en las frecuencias genotípicas respecto a lo esperado puede cuantificarse mediante el desequilibrio de ligamiento. En un esquema de dos *loci*, con alelos *A/a* y *B/b*, éste se escribe como

$$D = p_{AB} + p_{ab} - p_{Ab} - p_{aB}; \quad (10)$$

donde  $p_{AB}$  es la frecuencia del genotipo silvestre,  $p_{Ab}$  y  $p_{aB}$  la de los mutantes simples en cada uno de los *loci* y  $p_{ab}$  la del doble mutante.  $D < 0$  indica que los alelos mutantes tienden a estar separados y  $D > 0$  que, por el contrario, aparecen agrupados más de lo esperado por azar. El efecto de la recombinación es disminuir el valor absoluto de este desequilibrio. En el equilibrio mutación-selección, el valor del desequilibrio de ligamiento es aproximadamente proporcional al de la epistasia (Barton, 1995).

En una población sexual, la epistasia modifica el lastre genético (Crow, 1970; Chao, 1988; Shnol y Kondrashov, 1993; Charlesworth, 1990; Barton, 1995; Lenormand y Otto, 2000). Para entender por qué, situémonos de nuevo en el ejemplo de dos *loci*, con un alelo mutante deletéreo ( $s < 0$ ) en cada uno. Si la epistasia es negativa, tendremos  $D < 0$  y el efecto inmediato del sexo y la recombinación será crear genotipos dobles mutantes, lo cual incrementará el lastre genético si  $s_a + s_b > s_{ab}$  es decir, si la epistasia es lo bastante intensa como para cumplir  $\varepsilon < -s_a s_b$  según la ecuación 4. En el caso de epistasia positiva, tendremos  $D > 0$  y la recombinación aumentará el lastre genético si  $s_a + s_b < s_{ab}$ , condición implícita en  $\varepsilon > 0$  por lo que ineludiblemente, el lastre genético se verá aumentado. Así pues, en el equilibrio mutación-selección, la recombinación sólo incrementará la eficacia media de la población si la epistasia es negativa y débil, tal que  $0 > \varepsilon > -s_a s_b$ . El efecto reductor del lastre genético es proporcional al cuadrado de la tasa de mutación deletérea  $U$  y por lo tanto, puede ser importante si  $U \gg 1$  (Kondrashov, 1988; Charlesworth 1990). Cuando se tienen en cuenta  $m$  *loci*, la epistasia de orden  $m$  puede tener un efecto reductor de la carga mutacional proporcional a  $U^m$  (Hansen y Wagner, 2001).

La aparición y mantenimiento de la reproducción sexual plantea un problema evolutivo fundamental. En términos de tasa de crecimiento, la eficacia biológica de una población se ve reducida a la mitad por el mero hecho de ser sexual, ya que se requieren dos individuos para la reproducción (Maynard-Smith, 1971). Esta y otras desventajas del sexo (Crow, 1994) plantean la necesidad de buscar beneficios que, a corto plazo, expliquen su predominancia. Como hemos visto, un escenario de epistasia negativa débil y altas tasas de mutación permite que la recombinación reduzca el lastre genético de las poblaciones, siempre que esta se halle en el equilibrio mutación-selección. La relación entre epistasia y carga genética como posible explicación para la evolución del sexo fue primeramente propuesta por Crow (1970) y formalizada por Kondrashov (1988), en lo que se denomina la hipótesis determinista de la mutación.

La definición tradicional de sexo es la alternancia de singamia (fusión de dos núcleos haploides en un cigoto diploide) y meiosis en la que, tras la recombinación de los cromosomas por entrecruzamiento homólogo y la segregación genética, éstos se separan de nuevo en dos núcleos haploides. Esta definición no se aplica a los virus. Sin embargo, la función clave que la recombinación tiene en la reproducción sexual ha llevado a definir el sexo, en sentido amplio, como el intercambio recíproco de material genético, condición que podrá ser cumplida por los virus siempre que haya coinfección de una célula por más de una partícula (Chao, 1988). En virus de DNA existen enzimas encargadas de catalizar el corte y la ligación del DNA durante la recombinación, mientras que esos enzimas están ausentes en virus de RNA. En este último caso, el intercambio de material genético puede hacerse de dos maneras: por salto de la polimerasa de una a otra hebra de RNA o por segmentación genómica. En el caso del salto de molde, la recombinación puede suponer un coste en eficacia, ya que es posible que se ralentice la replicación o que se generen cadenas de RNA incompletas, denominadas partículas defectivas interferentes, que son capaces de completar su ciclo de infección en presencia de genomas completos por complementación genética (Moreno *et al.*, 1997). En el caso de la segmentación, garantizar de manera activa que cada partícula contenga la colección completa de segmentos supondrá presumiblemente un coste. De no ser así, se producirán partículas incapaces de infectar por sí solas, lo que en sí se traduce también en una pérdida de eficacia media. En definitiva,

al igual que en los organismos superiores, la evolución del sexo en virus de RNA requiere de una ventaja que compense sus costes.

## EVOLUCIÓN ADAPTATIVA

### El proceso de adaptación

El primer modelo formal de adaptación fue el modelo geométrico de Fisher (1930), que se basa en que los cambios de pequeño efecto tienen una probabilidad mucho mayor de resultar beneficiosos que los cambios mayores. Fisher concluyó que, como consecuencia de esto, la evolución debería ocurrir de manera gradual, mediante la fijación de mutaciones de pequeño efecto que acercarán a las poblaciones al óptimo adaptativo. Sin embargo, toda mutación beneficiosa tiene una frecuencia inicial pequeña, con lo que está sujeta a perderse por azar. La probabilidad  $p$  de superar esta fase inicial y llegar a la fijación depende de su coeficiente de selección, y cumple la igualdad

$$1 - p = e^{-p(1+s)} \quad (11)$$

(Peck, 1994; Wilke, 2003) que para valores de  $s$  pequeños es aproximadamente

$$p = 2s \quad (12)$$

(Haldane, 1927). En el contexto de las cuasispecies,  $p$  es algo mayor, porque no se calcula la probabilidad de que todos los individuos acaben teniendo dicha mutación, sino de que el mutante beneficioso acabe siendo el progenitor de toda la población (Wilke, 2003). En cualquier caso, como la probabilidad de fijación de una mutación depende de su coeficiente de selección, se favorece la fijación de los cambios de grandes efectos beneficiosos. Añadiendo esta objeción al razonamiento de Fisher, Kimura (1985) concluyó que la probabilidad de fijación máxima debería corresponder a las mutaciones de efecto intermedio.

Inicialmente, en una población poco adaptada se podrán fijar algunos cambios de efecto intermedio o relativamente grande, de acuerdo con el razonamiento de Kimura. Conforme estos cambios se fijan, el nivel de adaptación crecerá, y la proporción de mutaciones ventajosas será menor. Además, al acercarse la población al óptimo, sólo podrán fijarse cambios de pequeño efecto. En suma, el ritmo de adaptación se hará cada vez más lento. Finalmente, la población residirá cerca de un óptimo o pico adaptativo,

tendiendo de nuevo al equilibrio selección-mutación. No obstante, ese máximo es un límite asintótico, ya que siempre puede encontrarse un cambio que produzca un aumento de eficacia infinitesimalmente pequeño. De esa manera, si tenemos en cuenta el recorrido de la población a lo largo del proceso de adaptación, los cambios de efecto pequeño serán los más frecuentes, y las mutaciones fijadas seguirán una distribución aproximadamente exponencial (Orr, 1998).

El estudio de dinámicas de evolución experimental en *E. coli* (Lenski *et al.*, 1991) y VSV (Elena *et al.*, 1998) ha demostrado que, en efecto, la tasa de adaptación es mayor cuando la población está escasamente adecuada a su ambiente, y decrece conforme ésta se acerca al óptimo, aunque sin llegar a anularse (Lenski *et al.*, 1991). En concordancia con la teoría, se ha observado que el aumento de eficacia debe ocurrir mediante un número elevado de cambios de poco efecto (Burch y Chao, 1999). Sin embargo, cambios extremos de eficacia debidos a la fijación de mutaciones de gran efecto han sido también observados (Elena y Lenski, 1996; Bull *et al.*, 2000; Novella, 2004a).

#### Tasa de adaptación en asexuales

En un momento dado del proceso adaptativo, la tasa a la que crece la eficacia biológica depende lógicamente de la probabilidad de fijación de las mutaciones beneficiosas, pero también de que éstas ocurran en genotipos libres de mutaciones deletéreas. La frecuencia de esos genotipos es

$$q_0 = e^{U/s_H}, \quad (13)$$

donde  $s_H$  es la media armónica ( $s_H < 0$ ) de los efectos mutacionales deletéreas (Haigh, 1978; Orr, 2000). Así pues, la probabilidad de que una mutación beneficiosa de efecto  $s$  llegue a fijarse es aproximadamente, de acuerdo con la ecuación 12,  $p = 2q_0s$ . Estos cambios aparecerán en la población a una tasa  $NUp_b$ , donde  $p_b$  es la proporción de mutaciones que son beneficiosas. Por lo tanto, la tasa de sustitución adaptativa esperada es

$$E[k] = 2NUs_b p_b e^{U/s_H}, \quad (14)$$

siendo  $s_b$  el efecto promedio de las mutaciones beneficiosas (Orr, 2000). Por último, la tasa de adaptación, medida como aumento del logaritmo de la eficacia, puede ser obtenida a partir de la tasa de sustitución esperada mediante la expresión



$$d \log W / dt = E[k] \log(1 + s_b) \quad (15)$$

(Johnson y Barton, 2002). Estas ecuaciones muestran que la tasa de sustitución será mayor cuanto mayor sea el tamaño poblacional, pero no necesariamente cuanto mayor sea la tasa de mutación. En realidad, el valor de  $U$  para el que se maximiza la tasa de sustitución es

$$U = s_H \quad (16)$$

(Orr, 2000). Además, la tasa de sustitución aumentará para valores de  $s_b$  y  $p_b$  elevados, situación que se dará cuando las poblaciones estén escasamente adaptadas, permitiendo así entender la deceleración que la tasa de adaptación sufre al acercarse la población al óptimo adaptativo.

Las ecuaciones (13) y (14) predicen que el logaritmo de la eficacia aumentará linealmente con el tamaño poblacional. Sin embargo, esto no es del todo cierto, ya que conforme dicho tamaño crece, aumenta la probabilidad de que dos o más mutaciones beneficiosas estén presentes en la población de manera simultánea en distintos individuos. Este fenómeno de competencia entre mutaciones beneficiosas en poblaciones asexuales se denomina interferencia clonal (Gerrish y Lenski, 1998). En esta situación, las mutaciones menos beneficiosas se perderán y por lo tanto, la tasa de adaptación se verá frenada respecto a la predicha por la ecuación 15, mientras que el efecto de las mutaciones beneficiosas fijadas será mayor que  $s_b$ . Es posible hallar expresiones para la tasa de sustitución y adaptación en presencia de interferencia clonal (Gerrish y Lenski, 1998; Orr, 2000; Wilke, 2004) y se ha comprobado que, en efecto, su efecto es observable experimentalmente, tanto en bacterias (de Visser *et al.*, 1999) como en virus de RNA (Miralles *et al.*, 1999, 2000). Los efectos de la interferencia clonal serán notables cuando

$$N \log N > U p_b q_0 / 2 \quad (17)$$

(Wilke, 2004), y, por lo tanto las expresiones que ignoran sus efectos serán buenas aproximaciones si se cumple esta condición.

#### Tasa de adaptación y recombinación

El destino de los cambios adaptativos dependerá de si estamos frente a una población sexual o asexual. En una población asexual, si varias mutaciones beneficiosas aparecen simultáneamente en diferentes individuos, sólo una de ellas, previsiblemente la de mayor eficacia, se fijará. Habrá que esperar a que el



resto de cambios ocurran de nuevo sobre el genotipo que ha sufrido la mutación de máxima eficacia para que éstos puedan alcanzar la fijación. Como hemos visto, la interferencia clonal puede reducir visiblemente la tasa de adaptación de las poblaciones asexuales. Aunque las poblaciones sexuales pueden padecer un problema análogo, llamado "problema de tráfico" (Kirby y Stephan, 1996), su influencia será mucho menor, ya que éstas tendrán la ventaja de poder combinar en un mismo individuo mutaciones beneficiosas que, de otra manera, permanecerían confinadas al genotipo en que aparecieron. A esta explicación para la evolución del sexo se la conoce como el argumento de Fisher-Muller, (revisado en Barton y Charlesworth, 1998). Lógicamente, esta ventaja será mayor cuanto mayor sea la probabilidad de encontrar varios mutantes beneficiosos simultáneamente en la población y será más dramática si se tienen en cuenta simultáneamente las mutaciones deletéreas, al depender la fijación de un cambio ventajoso en una población asexual de que éste ocurra en un genotipo libre de mutaciones deletéreas (Peck, 1994; Lenski, 2001; Rice y Chippindale, 2001).

Sin embargo, la ventaja evolutiva de la recombinación es fuertemente dependiente de la epistasia. En presencia de selección direccional, se genera un desequilibrio de ligamiento de mismo signo que el de la epistasia (Lenormand y Otto, 2000). Por lo tanto, si ésta es predominante negativa, los alelos ventajosos tenderán a estar separados, y la recombinación será favorable a largo plazo, puesto que tenderá a unirlos. Sin embargo, como vimos previamente, habrá un coste a corto plazo en forma de un aumento del lastre genético, a no ser que se cumpla la condición  $\epsilon > -s_a s_b$ . El espacio de parámetros para los que los beneficios se sobreponen a los costes crece cuanto más baja sea la intensidad de la epistasia negativa, más altos los sean los coeficientes de selección y menor sea la tasa de recombinación (Barton, 1995; Lenormand y Otto, 2000). En el caso de epistasia positiva, se generará un desequilibrio de ligamiento positivo y el sexo no será ventajoso a largo plazo debido a que separará asociaciones favorables. Además, como estamos hablando de mutaciones beneficiosas,  $s_a$  y  $s_b$  tienen el mismo signo y por lo tanto, necesariamente  $s_a + s_b > s_{ab}$ , por lo que existirá un aumento del lastre genético a corto plazo causado por la recombinación. En ese escenario pues, la recombinación no podrá evolucionar.

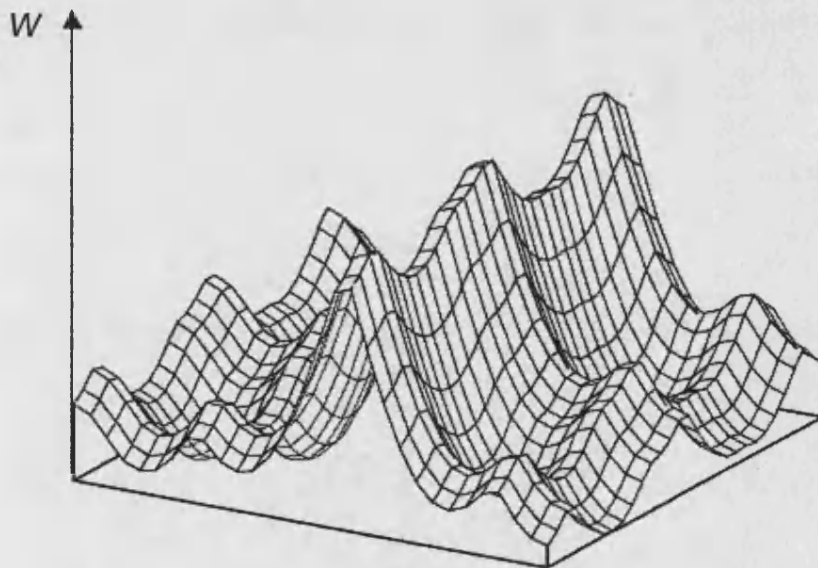
Sin embargo, en situaciones reales, el signo de la epistasia no tiene porqué coincidir con el del desequilibrio de ligamiento. De manera general, con la condición  $s_a s_b > 0$ , la recombinación será siempre favorable a corto plazo cuando los signos de  $D$  y  $\varepsilon + s_a s_b$  sean opuestos, ya que disminuirá el lastre genético. Si estamos en el caso  $D > 0$ , la desventaja a largo plazo habrá de ser compensada por una epistasia negativa fuerte en relación a los coeficientes de selección (Barton, 1995). Si por el contrario, nos encontramos con la situación  $D < 0$  y  $\varepsilon > 0$ , la recombinación será siempre favorable. En síntesis, la ventaja evolutiva del sexo necesita de la condición

$$D(\lambda_{s,\rho} - \varepsilon) > 0 \quad (18)$$

en la que  $\lambda_{s,\rho} = -2s_a s_b (1/\rho - 1)$  y  $\rho$  hace referencia a la tasa de recombinación (Barton, 1995).

### Paisajes adaptativos

Con frecuencia, se visualiza el proceso de adaptación con ayuda de la imagen de los paisajes adaptativos, originalmente propuesta por Wright (1931,1932). La epistasia juega un papel importante en la génesis de estos



**Figura 2.** Paisaje adaptativo rugoso. El eje vertical representa la eficacia biológica ( $W$ ), mientras que los ejes horizontales representan los valores de los caracteres. De manera general, la eficacia depende de más de dos caracteres, por lo que los dos ejes horizontales no son más que una representación bidimensional de un espacio multidimensional.

paisajes, ya que en ausencia total de interacción genética, sólo podrá haber un único óptimo o pico adaptativo, mientras que en presencia de ésta, podrán aparecer varios picos (Wolf *et al.*, 2000). Imaginemos el caso simple de dos sitios nucleotídicos con sus cuatro alelos A, T, G, y C. Si no hay epistasia, el alelo óptimo en el primer sitio (*e.g.* A) será independiente del otro (*e.g.* G) de manera que habrá un único genotipo óptimo (en este caso, AG). Si hay epistasia entre los sitios, el efecto de cada alelo no puede conocerse sin saber qué alelo está presente en el otro *locus*. Así, por ejemplo, podría ocurrir que el nucleótido A en la posición 1 sólo fuera ventajoso en presencia una T en el sitio 2. Igualmente, la G en posición 2 podría requerir por ejemplo, de una C en el sitio 1. Así habría dos combinaciones favorables, AT y CG, no necesariamente igual de eficaces. Este ejemplo ilustra como la epistasia produce paisajes adaptativo rugosos (Figura 2). En general, la población no puede pasar de un pico a otro por selección de mutaciones sucesivas, puesto que necesariamente, se ha de pasar por una región de baja eficacia o valle adaptativo. La selección sólo permitiría dicho cambio mediante una macromutación que evitase el paso por el valle. Incluso en el ejemplo de dos *loci*, se necesitarían dos sustituciones simultáneas, lo cual es poco probable. Las poblaciones permanecerán pues atrapadas en picos adaptativos locales durante largos periodos, pese a que existan otras regiones de mayor eficacia. Las epistasias explican pues la importancia de las contingencias históricas. Este tipo de restricción adaptativa ha sido estudiada en poblaciones de VSV. Se ha observado que algunas mutaciones beneficiosas aparecen de manera recurrente en poblaciones independientes (Cuevas *et al.*, 2002; Novella 2004a), sugiriendo que las soluciones adaptativas son limitadas. Además, algunos de estos cambios aparecen asociados con otros, lo que es una señal de la existencia de interacciones epistáticas.

## DERIVA GENÉTICA

### Tamaño poblacional efectivo

La deriva genética, definida como el muestreo aleatorio de los genotipos entre generación y generación, es el modo en que la evolución progresa en ausencia de selección natural (Kimura, 1985). En principio, la cantidad de

deriva genética depende únicamente del tamaño poblacional. Sin embargo, en la realidad, existen múltiples factores implicados, como la historia demográfica, el modo de reproducción, el desequilibrio de ligamiento o la propia selección natural. El parámetro que resume todos estos factores y permite cuantificar la cantidad de deriva que sufre una población es el tamaño poblacional efectivo,  $N_e$ , que se define como el tamaño poblacional bajo el que, en ausencia de cualquier otro factor, existiría la misma cantidad de deriva que la que hay en la población real.

En una situación de neutralidad completa, la probabilidad de fijación de cada mutación en una población de tamaño  $N_e$  será simplemente  $1/N_e$ , y el aporte de mutantes por generación valdrá  $UN_e$ , llegándose a la conclusión de que la tasa de evolución de las mutaciones neutrales por deriva es

$$E[k] = N_e U / N_e = U \quad (19)$$

(Kimura, 1968). Pero como la acción de la deriva genética no se restringe únicamente a las mutaciones neutrales, sino que se extiende también a aquellas situaciones en las que las fluctuaciones estocásticas de las frecuencias genotípicas ocurran a una tasa no despreciable frente a la tasa de cambio adaptativo, ésta tendrá mayor influencia sobre la evolución cuando menores sean los tamaños poblacionales efectivos y los coeficientes de selección (Ohta, 1992) como puede apreciarse al comparar las ecuaciones 14 y 19.

### **Trinquete de Muller**

En una población infinita, el efecto negativo de la mutación queda limitado por el equilibrio mutación-selección. Estadísticamente, se espera que una población finita alcance ese mismo equilibrio (Haigh, 1978). No obstante, el número de individuos libres de mutación será menor cuanto menor sea el tamaño poblacional y entonces, es posible que por deriva, esa clase de individuos se pierda, lo que conlleva una disminución de la eficacia media, que será irreversible en poblaciones asexuales, si se ignoran las mutaciones reversas y compensatorias. Este proceso se conoce como trinquete de Muller (1964). Cuando la eficacia biológica ha sido disminuida como consecuencia del trinquete de Muller, es esperable que la fecundidad se vea a su vez afectada, reduciéndose pues el tamaño poblacional. Se puede establecer así un bucle entre eficacia y tamaño poblacional que podría acelerar la pérdida de eficacia, llegándose finalmente a la extinción (Lynch *et al.*, 1993).

El número de mutaciones deletéreas ( $s < 0$ ) fijadas por el trinquete viene determinada por la cantidad

$$Q_0 = Ne^{U/s} \quad (20)$$

Haigh, 1978; Stephan *et al.*, 1993). Cuanto menor es  $Q_0$ , menor es la capacidad de la selección para eliminar las mutaciones deletéreas y mayor es la deriva, por lo que más mutaciones aleatorias se fijan. En la ecuación 20 queda patente que la pérdida de eficacia de una población como consecuencia del trinquete será mayor cuanto menor sea el tamaño poblacional y mayor la tasa de mutación. No obstante, la relación entre el coeficiente de selección y la pérdida de eficacia no es tan sencilla. A menor valor de  $s$ , mayor número de mutaciones podrán acumularse, pero menor será la caída de eficacia por cada una. Por ello, la pérdida de eficacia debida al trinquete es máxima para un valor intermedio del efecto de las mutaciones (Felsenstein, 1974; Gabriel *et al.*, 1993; Lynch *et al.*, 1993; Butcher 1995). El coeficiente de selección que permite una máxima pérdida de eficacia crece al aumentar el tamaño poblacional y disminuir la tasa de mutación (Butcher, 1995).

El razonamiento de Muller ignora las mutaciones beneficiosas. A primera vista, éstas deberían anular el avance del trinquete, más aún si se tiene en cuenta que la magnitud y abundancia de las mutaciones beneficiosas debería ser mayor al reducirse la eficacia biológica de la población. A pesar de esto, debe existir un tamaño poblacional crítico a partir del cual los efectos de la fijación de mutaciones deletéreas no puedan ser contrapesados por las mutaciones beneficiosas, reversiones o cambios compensatorios, llevando a la población a la extinción. Este tamaño depende del efecto medio de las mutaciones, de la distribución de efectos mutacionales y de la proporción de cambios beneficiosos (Whitlock, 2000; Whitlock *et al.*, 2003).

Existen innumerables evidencias experimentales que demuestran que las poblaciones que sufren repetidos cuellos de botella repetidos pierden eficacia biológica debido al trinquete de Muller. Este fenómeno ha sido observado en organismos tan diversos como *D. melanogaster* (Mukai *et al.*, 1972), *C. elegans* (Keightley y Caballero, 1997), *Aspergillus nidulans* (Bruggeman *et al.*, 2003), *E. coli* (Elena y Lenski, 1997) y se ha demostrado que, aunque con menor fuerza, el efecto del trinquete también es patente en poblaciones sexuales (Bruggeman

*et al.*, 2003). Por supuesto, al tratarse de un fenómeno estocástico, existe gran variabilidad entre réplicas, e incluso pueden observarse casos aislados de ganancia de eficacia (Shaw *et al.*, 2000). Además, estos experimentos de acumulación de mutaciones han resultado útiles para obtener medidas de los valores promedio de  $s$  y de  $U$  (Mukai *et al.*, 1972; Ohnishi, 1977). Sin embargo, los resultados son altamente dependientes de la distribución subyacente de los efectos mutacionales, la cual precisamente, se desconoce (Keightley y Lynch, 2003). Las condiciones bajo las que actúa el trinquete se cumplen especialmente en los virus de RNA, por sus altas tasas de mutación y la frecuencia con la que sufren cuellos de botella durante su transmisión de un hospedador a otro (Elena *et al.*, 2001). En el laboratorio, la acción del trinquete ha sido puesta de manifiesto en varios virus de RNA, como el bacteriófago  $\phi 6$  (Chao, 1990), los virus de la estomatitis vesicular (VSV) (Duarte *et al.*, 1992; Clarke *et al.*, 1993), de la fiebre aftosa (FMDV) (Escarmís *et al.*, 1996), de la inmunodeficiencia humana tipo 1 (HIV-1) (Yuste *et al.*, 1999), del mosaico del pepino (CMV) (Li y Roossinck, 2004).

### **Epistasia y trinquete de Muller**

De igual modo que lo hacen las mutaciones simples, la epistasia influye sobre el trinquete de Muller. Si la eficacia de un genotipo con varias mutaciones es mayor de lo esperado bajo el modelo multiplicativo ( $\varepsilon < 0$ ), la selección en contra de las mutaciones deletéreas será más eficiente y el avance del trinquete tenderá a ralentizarse, aunque no el descenso de la eficacia poblacional, ya que por cada mutación fijada será mayor el efecto (Charlesworth *et al.*, 1998). No obstante, si la epistasia negativa se hace más fuerte, puede llegar a detener el trinquete, deteniendo por consiguiente la caída de eficacia (Kondrashov, 1994). Para entender este punto, situémonos por un momento en el caso extremo de epistasia negativa: la selección truncada. Imaginemos que los individuos que componen una determinada población que pierde eficacia debido al trinquete de Muller son capaces de acumular hasta  $n$  mutaciones, siendo inviables los individuos con  $n+1$  cambios. El trinquete quedaría necesariamente detenido, a no ser que por azar, no naciese ningún individuo libre de esa mutación adicional. Este ejemplo ayuda a entender porqué la epistasia negativa puede detener el trinquete. Sin embargo, puede objetarse a este razonamiento que los efectos mutacionales no son fijos, sino que siguen una distribución

probabilística de manera que, a pesar de la epistasia, siempre habrá mutaciones con un efecto lo bastante pequeño como para permitir el avance del trinquete. El ritmo de avance de éste se frena, al haber menor disponibilidad de mutaciones poco deletéreas, pero no se detiene. Sólo podría hacerlo si, eventualmente, las mutaciones con efecto lo bastante débil como para asegurar su avance aparecieran con una frecuencia menor o igual que las mutaciones beneficiosas (Butcher, 1995).

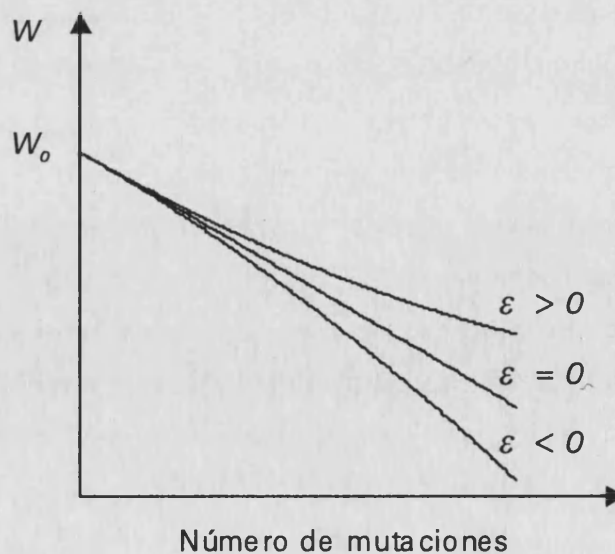
Históricamente, los experimentos de acumulación de mutaciones han sido orientados a la detección de la epistasias entre mutaciones deletéreas (Mukai *et al.*, 1972). Bajo el modelo multiplicativo, el logaritmo de la eficacia debe caer linealmente en función del número de mutaciones acumuladas  $n$ , tal que

$$W_n = W_0(1+s)^n; \quad (21)$$

es decir,

$$\log W_n = \log W_0 + n \log(1+s). \quad (22)$$

En principio, habrá epistasia si la pendiente de la caída de eficacia difiere de  $\log(1+s)$  o si la caída no es lineal. En la práctica, los valores de los coeficientes de selección son desconocidos, de manera que la única evidencia de epistasia proviene de una caída no lineal del logaritmo de la eficacia con el número de mutaciones. Si la curva de caída es cóncava, la epistasia es negativa, mientras que si es convexa, la epistasia es positiva (Figura 3).



**Figura 3.** Patrones hipotéticos de pérdida de eficacia como consecuencia de la acumulación de mutaciones deletéreas.



El problema de los experimentos de acumulación de mutaciones es, por un lado, que la varianza en los efectos de las mutaciones afecta a la potencia estadística del método y por otro, que una tendencia lineal puede ser atribuida a la ausencia de epistasia, pero también a la presencia de gran cantidad de interacciones de distinto signo que se compensan (Elena y Lenski, 1997).

### Deriva, epistasia y recombinación

El daño que el trinquete de Muller ocasiona a una población puede medirse como

$$L_r = 1 - W_0 / W_r \quad (23)$$

donde  $L_r$  se define como la carga del trinquete,  $W_r$  es la eficacia media de la población que sufre el trinquete y  $W_0$  la eficacia del genotipo silvestre (Kondrashov, 1994). El sexo podría conferir una ventaja evolutiva al reducir el valor de  $L_r$ . Siempre que esa ventaja compense los costes de la recombinación, la segmentación, o la producción de machos. Una vez más, el balance entre costes y beneficios queda influido por la epistasia. Contrariamente a la hipótesis determinista de la mutación, la hipótesis que explica la evolución del sexo como un mecanismo para evitar los efectos adversos de la mutación en presencia de deriva genética es compatible con una predominancia de la epistasia de signo positivo, que acentúe la carga del trinquete y por lo tanto, la ventaja del sexo. Sin embargo, en esta ocasión, el sexo puede conferir una ventaja selectiva también en ausencia de epistasia, ya que permite aumentar la frecuencia en la población del genotipo libre de mutación, limitando pues la efectividad del trinquete (Chao, 1988). Además, también se acepta la posibilidad de epistasia negativa, siempre que sea lo bastante débil para no reducir significativamente el avance del trinquete.

Las frecuencias alélicas podrán alejarse más del equilibrio selección-mutación en poblaciones reducidas, al entrar en acción la deriva genética. En principio, el muestreo aleatorio que supone la deriva no debería alterar las frecuencias esperadas, pero esto deja de ser así al tenerse en cuenta la acción añadida de la selección. Cuando la deriva genere por azar un desequilibrio de ligamiento positivo, la selección podrá eliminar de manera eficiente los genotipos extremos, disipando así el desequilibrio. Por el contrario, si por azar los alelos deletéreos se encuentren preferentemente en fondos genéticos beneficiosos ( $D < 0$ ), la selección será menos eficiente a la hora de reducir la

frecuencia de estos genotipos. Así pues, la deriva en presencia de selección tiende a generar un desequilibrio de ligamiento negativo, lo que se conoce como efecto Hill-Robertson (Hill y Robertson, 1966; revisado en Otto y Lenormand, 2002). Al generar un desequilibrio de ligamiento negativo, este efecto tiende a favorecer la evolución de la recombinación. Pero para que esto ocurra, se requerirá una epistasia de signo positivo, ausencia de epistasia o, en todo caso negativa siempre que sea débil, llegándose pues a las mismas conclusiones que en el párrafo anterior.

### **Deriva y adaptación**

La manera en que las poblaciones recuperan eficacia tras la acción del trinquete de Muller depende varios factores. De manera general, como han puesto de manifiesto experimentos con virus de RNA, a mayor pérdida de eficacia, más rápida es la recuperación (Elena *et al.*, 1998). Pueden distinguirse dos mecanismos fundamentales para esta recuperación. Primero, es posible que las mutaciones deletéreas acumuladas simplemente reviertan y segundo, pueden fijarse cambios en otros lugares del genoma que compensen el efecto negativo producido por las mutaciones acumuladas. Este segundo caso, denominado evolución compensatoria, implica la existencia de epistasia positiva (Moore *et al.*, 2000). La epistasia en forma de mutaciones compensatorias, al ser una posibilidad adicional a la reversión, aumenta la capacidad de diversificación de las poblaciones tras la acción del trinquete de Muller (Escarmís *et al.*, 1999, 2002).

La diferencia entre la reversión y la compensación está pues ligada a la abundancia de las interacciones génicas, lo cual puede también entenderse mediante la imagen de los paisajes adaptativos. Como hemos visto, en ausencia de selección, las poblaciones tienden a acumular mutaciones de manera aleatoria y por lo tanto, a divergir. De ese modo, las poblaciones pueden atravesar valles adaptativos. Cuando los tamaños poblacionales vuelven a ser elevados, se retorna a un régimen de selección, que conduce a una recuperación de la eficacia es decir, al ascenso de un pico adaptativo. Pero en esta ocasión, no habrá posibilidad de cruzar de nuevo el valle, de manera que las poblaciones podrán quedar atrapadas en picos adaptativos distintos del que originalmente provenían. Diferentes poblaciones, aisladas las unas de las otras, podrán ascender a picos diferentes, incrementándose así la variabilidad

genética. Finalmente, si éstas entran en contacto, aquélla que haya ascendido al pico más elevado desplazará a las demás. La conclusión que puede deducirse de este proceso es que, la acción combinada de la deriva genética y la selección natural permite una mayor adaptación que la acción única de la selección. La validez de esta teoría, denominada del equilibrio cambiante (Wright, 1982), depende críticamente de la frecuencia e intensidad de la epistasia, ya que de ella depende, a su vez, la topografía del paisaje adaptativo.

## **EL VIRUS DE LA ESTOMATITIS VESICULAR**

### **Características generales**

En esta tesis, utilizaremos VSV como modelo experimental, el cual ha demostrado ser útil para contrastar gran cantidad de hipótesis (revisado en Novella, 2003). Además de las ya mencionadas ventajas que los virus de RNA presentan para la evolución experimental, VSV es especialmente adecuado para los estudios en el laboratorio, por su facilidad para infectar cultivos celulares, la rapidez y la eficiencia con que completa el ciclo de infección y su nivel moderado de riesgo biológico. También cabe mencionar que su estructura genética es muy sencilla, y que no presenta recombinación. Estas dos características facilitan mucho la interpretación de los resultados experimentales, aunque dificultan su extrapolación a otros virus u organismos más complejos. Finalmente, las técnicas de genética reversa están disponibles, aunque su aplicación no es tan sencilla como en otros virus de RNA.

De acuerdo con la Comisión Internacional de Taxonomía Viral (Francki *et al.*, 1991), los virus de RNA se agrupan en cuatro grandes categorías: virus de doble cadena, de cadena sencilla positiva, de cadena sencilla negativa [ss(-) RNA] y retrovirus. VSV pertenece al grupo ss(-) RNA. Esto significa que el virión contiene una sola cadena de RNA, la cual no puede servir como molde para la traducción de la información genética ya que previamente, necesita ser transcrita a un RNA de cadena positiva. Una característica esencial de los virus ss(-) RNA es la ausencia casi total de recombinación homóloga (Chare *et al.*, 2003).

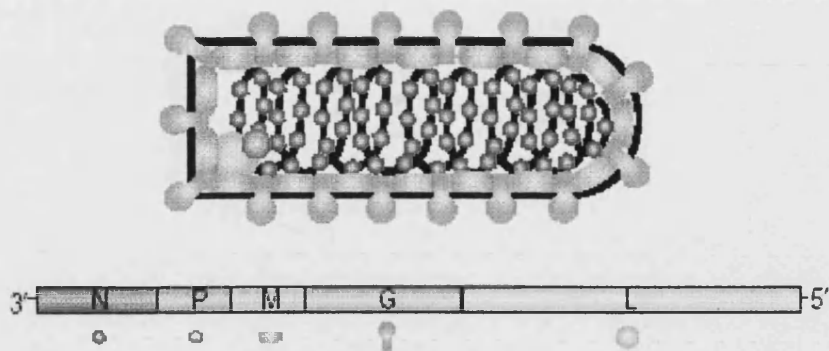
VSV pertenece al orden Mononegavirales, en el cual se encuentra la familia *Rhabdoviridae*, que se subdivide en cuatro géneros de virus animales y

dos de virus vegetales. De ellos, mencionemos el género *Lyssavirus*, cuyo principal exponente es el virus de la rabia, y el género *Vesiculovirus*. En este último, existen varios serotipos o especies, los más importantes de los cuales son VSV-New Jersey y VSV-Indiana. Nuestro modelo experimental pertenece al serotipo Indiana, del que existen varias cepas o aislados de focos de infección, como son las cepas Orsay y San Juan.

La estomatitis vesicular es una enfermedad con graves consecuencias económicas, puesto que afecta a la ganadería porcina y bovina, así como a los caballos, produciendo ulceraciones o aftas en la lengua y tejidos orales. Además, se han encontrado anticuerpos anti-VSV en gran variedad de mamíferos (bóvidos, cánidos, camélidos, cérvidos, quirópteros y primates, incluidos los humanos). La enfermedad se manifiesta mediante brotes periódicos en el sureste de EEUU, sur de Méjico, América Central y la parte norte de América del Sur (Rodríguez *et al.*, 2000). Los *Vesiculovirus* infectan también a insectos, que actúan como vectores de transmisión de la enfermedad (Letchworth *et al.*, 1999).

### **Estructura del genoma y del virión**

El genoma, con una longitud de 11161 nucleótidos, contiene únicamente cinco genes, denominados *N*, *P*, *M*, *G* y *L*, ordenados en dirección 3'-5' (Figura 4). El 95.2% de la secuencia es codificante, lo que indica que se trata de un genoma altamente compacto. Además, las cuatro regiones no codificantes, que separan cada uno de los genes, así como las regiones líder y trailer de los extremos 3' y 5', son funcionalmente esenciales, ya que están implicadas en la regulación de la expresión génica y del balance entre transcripción y replicación. De hecho, mutaciones en estas zonas pueden tener efectos fuertemente negativos en la eficacia biológica del virus (Bracho, 2003). El resto de la variación no codificante, es decir los cambios sinónimos, también está restringida, debido a la importancia de las estructuras secundarias del RNA (Novella, 2004b). Los genes más variables en poblaciones de laboratorio son *G* y *M*, mientras que *N* y *L* son los más conservados (Cuevas *et al.*, 2002). Éste último, que codifica para parte de la RNA polimerasa, ocupa aproximadamente el 60% del genoma.



**Figura 4.** (adaptada de Lichty *et al.*, 2004). Estructura del virión de VSV y de su genoma. En la parte de bajo, se muestra la cadena de polaridad positiva o antígenómica, con los cinco genes y sus correspondientes proteínas. En la parte de arriba, se muestra cómo se estructuran esas proteínas junto con el RNA genómico y la membrana.

El virión de VSV, como puede observarse en la Figura 4, tiene forma de bala, con aproximadamente 180 nm de longitud y 75 nm de anchura, y está compuesto de dos partes principales: núcleo y la envoltura. Comparte con la mayoría de los miembros de la familia *Rhabdoviridae* las características estructurales y funcionales, conservándose el número y orden de los genes, salvo en las especies que infectan plantas, en las que existe un gen adicional. Parte de estas características también son compartidas por la familia *Paramixoviridae*.

El RNA genómico está fuertemente unido a la nucleoproteína *N*, de 422 residuos aminoacídicos, la cual tiene por función principal empaquetar, estabilizar y proteger el RNA. A diferencia de los virus de cadena positiva, el RNA de cadena negativa no es infeccioso por sí mismo, ya que no sirve de molde para la traducción. Por este motivo, se requiere la presencia en el núcleo viral de la proteína *L* (2109 residuos), que desempeña el papel esencial de replicasa y transcriptasa y la fosfoproteína *P*, de tamaño menor (265 residuos), pero indispensable para que la primera sea funcional. Aún así, el núcleo, formado por el RNA y las proteínas *N*, *L* y *P*, ha de poder penetrar en la célula hospedadora para ser infeccioso. Para que la entrada en la célula sea posible, es necesaria la envoltura, que está compuesta de una membrana bifosfolipídica capturada de la célula, en la que se inserta la glicoproteína *G* (511 residuos). Por su situación externa, esta última es el principal determinante antigénico, lo que explica su alta variabilidad. Además, es una proteína de interés práctico, puesto que es allí donde

se encuentran marcadores fenotípicos de resistencia a anticuerpos (Vandepol *et al.*, 1986). En el interior de la membrana se encuentra la proteína de matriz *M* (692 residuos), que parece actuar de puente entre el núcleo y la membrana, permitiendo mantener la estructura completa del virión.

### Ciclo de infección

Los acontecimientos que permiten completar un ciclo de infección a nivel celular son la adsorción, endocitosis, fusión de la envuelta y liberación del núcleo, transcripción, traducción, procesado post-traduccional, replicación, ensamblaje y liberación de las partículas infecciosas (Figura 5). La adsorción o contacto con el receptor celular es llevada a cabo por la glicoproteína *G*, aunque la diana de reconocimiento celular es aún desconocida (Coil y Miller, 2004). Seguidamente, la glicoproteína *G* cataliza la fusión de las membranas celular y viral, permitiendo así la penetración del virus en vesículas endocíticas, donde la

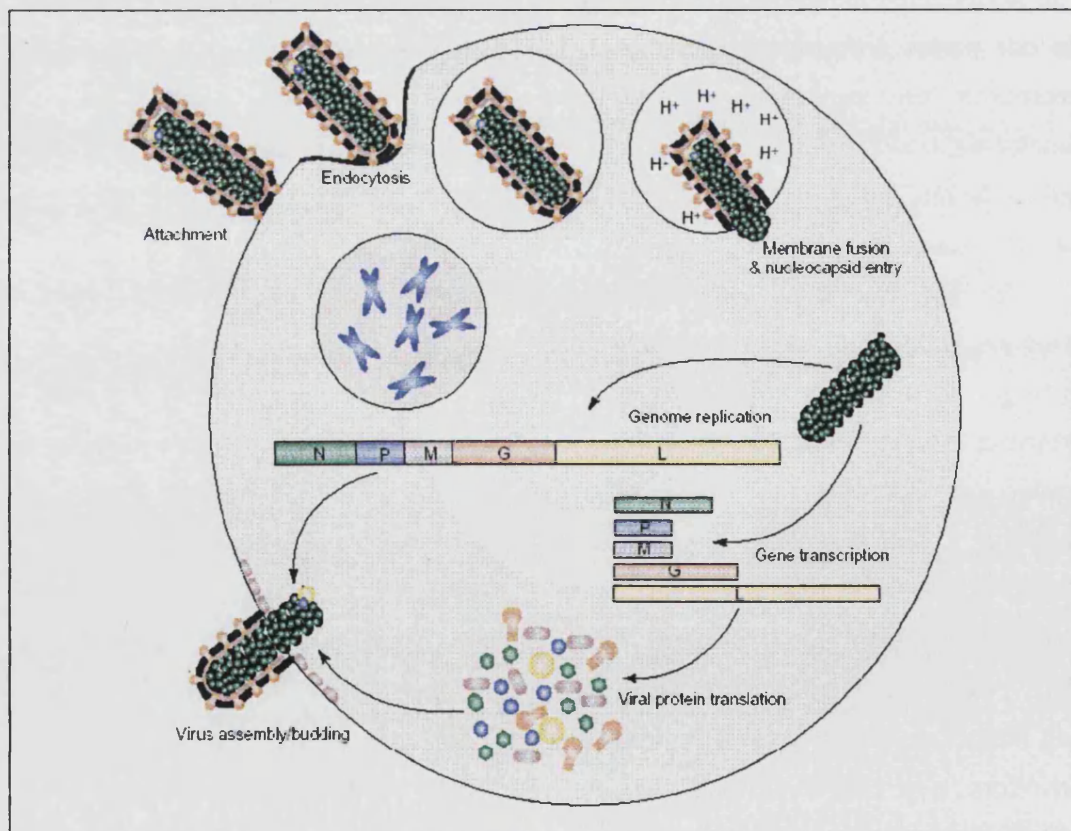


Figura 5. Ciclo infeccioso de VSV (adaptado de Lichty *et al.*, 2004).

envuelta es hidrolizada y el núcleo liberado al citosol. Las proteínas *L* y *P* se unen al extremo 3' del ss(-) RNA donde inician la transcripción secuencial de los cinco RNA mensajeros (mRNA). Las regiones intergénicas, ricas en uracilo, juegan aquí un importante papel, ya que al parecer sirven de molde para la adición de la cola poli-A al final de cada mRNA. Como la transcriptasa tiene una cierta probabilidad de separarse del molde tras la adenilación de cada mRNA y como únicamente puede iniciarse la transcripción desde el extremo 5' del ss(-) RNA, la expresión de los genes sigue un gradiente decreciente en la dirección 5' → 3'. La traducción acontece acoplada a la transcripción de cada mRNA, y las proteínas resultantes son procesadas en el retículo endoplásmico y el aparato de Golgi. La proteína *N*, de carga positiva, se une al extremo 5' del ss(-) RNA e impide que la transcripción se interrumpa, dando lugar a cadenas completas de polaridad positiva. Este proceso, cuya influencia será creciente conforme se acumulen los productos de la traducción, marca el inicio de la replicación, que termina con la producción de cadenas de polaridad negativa. A este ss(-) RNA, revestido de la proteína *N*, se unen las proteínas *L* y *P*, formando núcleos virales en el citosol. Por otro lado, la proteína *G* ha sido glicosilada en el aparato de Golgi y migra en vesículas exocíticas hacia la membrana celular. Finalmente, la proteína de matriz *M*, que tiene afinidad tanto por el núcleo viral como por la proteína *G*, permite el ensamblaje de los viriones, que son liberados al medio extracelular por gemación. La muerte de la célula acontece por lisis, aunque la proteína *M* parece estar implicada en la inducción de la apoptosis (Kopecky *et al.*, 2001).





## **OBJETIVOS**



1. Conocer el valor promedio del coeficiente de selección de mutaciones puntuales y su distribución estadística. Para ello crear una colección de clones de VSV caracterizados por incorporar una única sustitución nucleotídica respecto a un mismo genotipo de referencia y determinar su eficacia biológica relativa.
2. Estudiar la frecuencia, intensidad y signo de las epistasis. Para ello, crear una colección de clones de VSV caracterizados por incorporar diferentes combinaciones de dos de las sustituciones nucleotídicas previamente analizadas y determinar su eficacia biológica.
3. Llevar a cabo la evolución experimental de un subgrupo de mutantes deletéreos con el fin de determinar la importancia de la eficacia inicial, el número de mutaciones, la existencia y tipo de epistasis entre éstas y el tamaño poblacional sobre la tasa de evolución, así como determinar si ésta se produce preferentemente por reversión genética o por compensación.
4. Utilizando los datos obtenidos a partir de la colección de clones mutantes, realizar inferencias acerca de los parámetros genético-poblacionales que determinan la evolución de VSV, así como sobre la modulación de éstos por la epistasia.



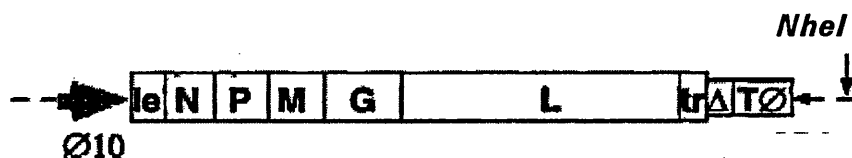
## **MATERIAL Y MÉTODOS**



## EL CLON INFECCIOSO DE VSV

### Estructura

El clon infeccioso de VSV que vamos a utilizar, fabricado y cedido por Whelan *et al.* (1995), contiene el genoma completo de VSV en forma de cDNA, insertado en el vector pGEM-3<sup>®</sup> (Promega<sup>®</sup>). Dicho vector, de 2.87 kb, está dotado de lugares de inicio de transcripción y replicación de los fagos SP6 y T7, así como de un sitio de clonación múltiple que interrumpe el gen de la  $\beta$ -lactamasa. El genoma de VSV, insertado en el sitio de clonación múltiple, está orientado de manera tal que la RNA polimerasa T7 transcriba la cadena de polaridad positiva (antigenómica). En el extremo 3' de dicha cadena se halla la secuencia del ribozima del virus  $\Delta$  de la hepatitis (HDV), que corta el transcrito en el punto exacto donde termina el RNA viral (Figura 6).



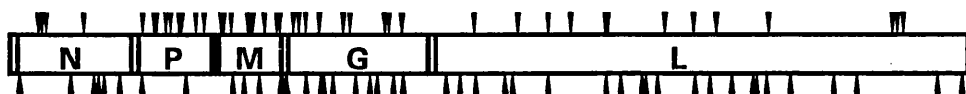
**Figura 6.** Estructura del clon infeccioso pVSV(+). Se muestra el plásmido en su forma lineal cortado mediante el enzima de restricción *NheI*. Ø10: secuencia promotora de T7; *le*: secuencia líder del genoma de VSV; *N*, *P*, *M*, *G* y *L*: distintos genes de VSV; *tr*: secuencia remolque del genoma de VSV;  $\Delta$ : ribozima de HDV. TØ: secuencia terminadora de T7 (adaptado de Whelan *et al.*, 1995).

El genoma de este clon infeccioso pertenece al serotipo Indiana, pero presenta varias peculiaridades. En primer lugar, el gen *G* proviene de la cepa Orsay, mientras que el resto de la secuencia es de la cepa San Juan. Además, los 516 nucleótidos del extremo 3' del genoma, que corresponden a la región terminal del gen *L* y al remolque, provienen de una partícula defectiva interferente (Pattnaik *et al.*, 1992). En esa región, existe una inserción de una adenina en la posición 11140, de manera que la longitud total del genoma es

de 11162 nucleótidos. A este tipo de genomas, contruidos a partir segmentos de diferente origen, se les denomina quimeras genéticas.

### Selección de las mutaciones a introducir

Hemos elegido 91 dianas en las que introducir sustituciones nucleotídicas. En un principio, se pretendía que todas las dianas, así como el cambio a introducir, fueran elegidos al azar. Sin embargo, la falta de datos previos directos lo suficientemente claros acerca del efecto de ese tipo de mutaciones sobre la eficacia biológica de VSV nos llevó a tomar la decisión de dividir la colección en dos bloques; uno de 48 mutantes en el que las mutaciones serán efectivamente aleatorias y otro de 43 mutantes en el que se reproducirán cambios aminoacídicos previamente descritos. De este modo, nos aseguramos de que una fracción importante de los mutantes será viable. Además, ambos grupos podrán ser comparados a la hora de interpretar los resultados. En la Figura 7 se muestra la localización de las mutaciones elegidas.



**Figura 7.** Localización de las 91 sustituciones nucleotídicas sobre el genoma de VSV. En la parte de arriba se muestran las mutaciones previamente descritas y en la parte de bajo se representan los cambios elegidos al azar.

El Apéndice 1 recoge la información básica sobre los cambios previamente descritos, mientras que el Apéndice 2 proporciona información más detallada acerca de esos cambios. Básicamente, se pueden distinguir dos tipos de mutaciones. En primer lugar, están aquéllas que han sido detectadas en clones moleculares y que corresponden pues a genomas individuales de los cuales además, se desconocía la frecuencia en la población de origen. En algunos casos, no existe garantía de que esos cambios no sean fruto de errores cometidos por la polimerasa empleada durante la amplificación por PCR (Schubert *et al.*, 1984; Bracho *et al.*, 1998). En segundo lugar, están aquéllas que han sido detectadas en secuencias consenso de poblaciones es decir, por amplificación directa del cDNA. En este caso, las sustituciones estarán fijadas



en la población, o al menos se hallarán en alta frecuencia. Además, la probabilidad de que se trate de errores en la amplificación por PCR es mínima. Los cambios pueden clasificarse también según su origen natural o de laboratorio, aunque cada mutación presenta una historia particular. La eficacia biológica esperada asociada a cada una de ellas dependerá en gran medida de ésta ya que por ejemplo, a altas multiplicidades de infección (*moi*) se espera que puedan fijarse mutaciones deletéreas, mientras que en pases masivos a baja *moi* se espera que se fijen cambios beneficiosos.

## TÉCNICAS BÁSICAS

### Cultivo celular

Se emplearán fibroblastos de riñón de embrión de hámster chino (BHK21), obtenidos de la Colección Americana de Cultivos Tipo (ATCC), para todos los procedimientos salvo en el caso particular de la obtención del stock de virus Vaccinia. BHK21 es una línea celular con una vida aproximada de 60 divisiones, que crece en monocapa y sufre una división cada ~18 horas en las condiciones de cultivo mencionadas a continuación.

Las células se cultivarán en medio DMEM (la modificación de Dulbecco del medio mínimo de Eagle; Invitrogen®), completado con una batería de aminoácidos no esenciales (Sigma®) y esterilizado por filtración a través de membranas con un tamaño de poro de 0.22 µm. El medio se completará también con suero fetal bovino (FCS; Invitrogen®) a una concentración final de 10% v:v, así como con los antibióticos gentamicina, penicilina y estreptomicina y el antifúngico *n*-butil-*p*-hidroxibenzoato. Las concentraciones de todos estos aditamentos son las habitualmente empleados en nuestro laboratorio (Elena, 1996).

Se utilizarán placas de plástico pretratadas para cultivos celulares en monocapa (Becton-Dickinson®) de 100 mm de diámetro (p100) para el mantenimiento de los cultivos y de 60 mm (p60), 12 pocillos (w12), 24 pocillos (w24) y 96 pocillos (w96) para otros usos.

Los incubadores (Forma®) se mantendrán a 37°C en una atmósfera con el 5% de CO<sub>2</sub> y un 95% de humedad relativa (hr).



Durante los subcultivos de las líneas celulares, con el fin de resuspender las células en el medio de cultivo fresco, las monocapas son lavadas con una solución salina en tampón fosfato (PBS) y despegadas de la placa aplicando tripsina (Invitrogen®) diluida en EDTA. Detalles particulares de este protocolo están descritos en Elena (1996). La densidad del cultivo celular puede ser estimada resuspendiendo una fracción de éste en un volumen conocido de PBS y contando las células en una cámara cuentaglóbulos (Neubauer®). Utilizando esta técnica, podemos estimar que una monocapa BHK21 al 100% de confluencia alcanza una densidad de  $2 \times 10^5$  células/cm<sup>2</sup>. La cantidad de células total presentes en cada tipo de placa de cultivo empleada se muestra en la Tabla 2.

Placa	Area (cm <sup>2</sup> )	Número de células BHK21
w96	0.3	$6 \times 10^4$
w24	2	$4 \times 10^5$
w12	4	$8 \times 10^5$
p60	20	$4 \times 10^6$
p100	56	$1 \times 10^7$

**Tabla 2.** Número estimado de células BHK21 al 100% de confluencia para cada tipo de placa utilizado. En el caso de las placas multipocillo, se muestran los valores por pocillo.

Para su almacenaje, las células pueden ser congeladas, resuspendiendo el contenido de una p100 en 1 mL de FCS con *di*-metil-sulfóxido (DMSO; Sigma®) al 10% v:v y pasándolo a un tubo criogénico (Nunc®). La congelación puede realizarse lentamente dentro de un arcón de  $-80^{\circ}\text{C}$  (Forma®) trasladándose los tubos a un tanque de nitrógeno líquido (Praxair®) tras 24h.

La descongelación puede hacerse por inmersión y agitación en un baño de  $37^{\circ}\text{C}$  (Selecta®) y eliminación del DMSO por centrifugación (Eppendorf®) durante 5 min a 5000 rpm, tras lo que las células se resuspenden en medio de cultivo completo.

### **Infección con VSV en líquido**

Este tipo de infección es el empleado habitualmente en los experimentos de evolución y los ensayos de competencia (Holland *et al.*, 1991). La técnica es muy sencilla, puesto que sólo se requiere decantar el medio de cultivo y añadir un inóculo de pequeño volumen (50 a 200  $\mu$ L dependiendo de la placa empleada) a la monocapa con la cantidad deseada de virus, a la que llamaremos  $N_0$ . Seguidamente, se incuba en las condiciones habituales (37°C, 5% CO<sub>2</sub>, 95% hr) durante 45 min y se añade medio de cultivo completo, reduciendo la cantidad de FCS al 2% v:v.

Las infecciones en líquido se usarán tanto para los ensayos de competencia (en pocillos w24) como para el experimento de recuperación de la eficacia biológica (en pocillos w12).

### **Titulación de VSV**

Cuando el medio es solidificado con agar con el fin de evitar la difusión de los virus, la infección progresa en forma de aros de lisis, cada uno de ellos originado a partir de una única partícula infecciosa (PFU). El título de una suspensión de virus se define como la cantidad de partículas infecciosas por unidad de volumen.

La infección en sólido, o plaqueo, es igual que la infección en líquido, salvo que el medio que se aplica tras la incubación del inóculo lleva agar (Sigma®) al 0.4-0.5% y DEAE-dextrano (Sigma®) al 0.01%. Para que los aros de lisis, también llamados calvas, sean visibles y contables, se utilizan placas p60, las cuales se inoculan con 100  $\mu$ L de la suspensión viral, de manera que la aparición de 100 calvas en la placa indicaría un título de 10<sup>3</sup> PFU/mL. Dado que habitualmente los títulos alcanzados por VSV son mayores, se ha de realizar la dilución que corresponda para observar un número contable de calvas.

Para poder ver los aros de lisis con facilidad, se incuba la infección durante 18-24 h, tras lo cual se fija la monocapa con una solución de formaldehído (Scharlau®) al 3% en agua, que además permite despegar el medio sólido de la placa. Seguidamente, se tiñe la monocapa con una solución de cristal violeta al 5% en el fijador y se enjuaga. La monocapa queda de color azul violáceo intenso mientras que los aros de lisis no se tiñen. El recuento de éstos se hace mediante un contador de calvas (Stuart Scientific®).

### **Producción de anticuerpo monoclonal I<sub>1</sub> (AcM I<sub>1</sub>)**

La línea de hibridoma de ratón productora del AcM I<sub>1</sub> (Lefrancois y Lyles, 1982a, 1982b, 1983; Vandepol *et al.*, 1986), cedida por el profesor J. J. Holland (University of California, San Diego), crece en suspensión a una velocidad parecida a BHK21 y puede ser pasada indefinidamente. Su medio de cultivo se obtiene añadiendo nutrientes (Elena, 1996) al DMEM completo.

Para estimular la producción de anticuerpo, ha de añadirse a los cultivos ya saturados OptiMab® (Invitrogen®) e incubarse durante 10 días, tras los que las células se separan del medio por centrifugación. El sobrenadante, que contiene el AcM I<sub>1</sub>, se almacena a -20°C.

Para evaluar la actividad neutralizante del AcM I<sub>1</sub> producido, se realizan plaques de VSV en los que el medio de infección sólido contiene proporciones crecientes de AcM I<sub>1</sub>. En nuestras condiciones de crecimiento una proporción de 20% v:v de AcM I<sub>1</sub> es habitualmente suficiente para impedir por completo la formación de aros de lisis 24 horas post-inoculación (hpi).

### **Stock e infecciones con el virus Vaccinia vTF-7**

El virus Vaccinia es un potente vector de expresión (Mackett *et al.*, 1982). El recombinante vTF-7, que expresa el gen de la RNA polimerasa del bacteriófago T7 (Pattnaik *et al.*, 1992; Whelan *et al.*, 1995), cedido por la profesora G. W. Wertz (University of Alabama, Birmingham), será necesario en grandes cantidades para la posterior obtención de partículas de VSV a partir del clon infeccioso.

El stock de virus Vaccinia puede obtenerse infectando varias placas de células HeLa (ATCC), las cuales se manipulan de manera idéntica a las BHK21 aunque su tasa de división es algo menor. Tras 48 h, las placas se someten a tres ciclos de congelación a -80°C y descongelación con el objetivo de lisar las células. Seguidamente el medio se centrifuga (5 min a 5000 rpm) y el sobrenadante de todos los tubos es transferido a una botella para homogeneizarlo, tras lo cual se alicuota en tubos de 1.5 mL. Por último, con el fin de desagregar las partículas de vTF-7, los tubos se introducen en un baño de ultrasonidos a 4°C durante 5 min e inmediatamente son almacenados a -80°C.

La titulación de vTF-7 ha de realizarse en células HeLa mediante el mismo procedimiento descrito anteriormente para VSV, excepto que el tiempo para

absorción del inóculo es de 1 h, la concentración de agar de 0.2% y el tiempo de incubación necesario para la formación de aros de lisis visibles es de 72 h. El título del stock que hemos obtenido es  $(7.8 \pm 0.8) \times 10^8$  PFU/mL.

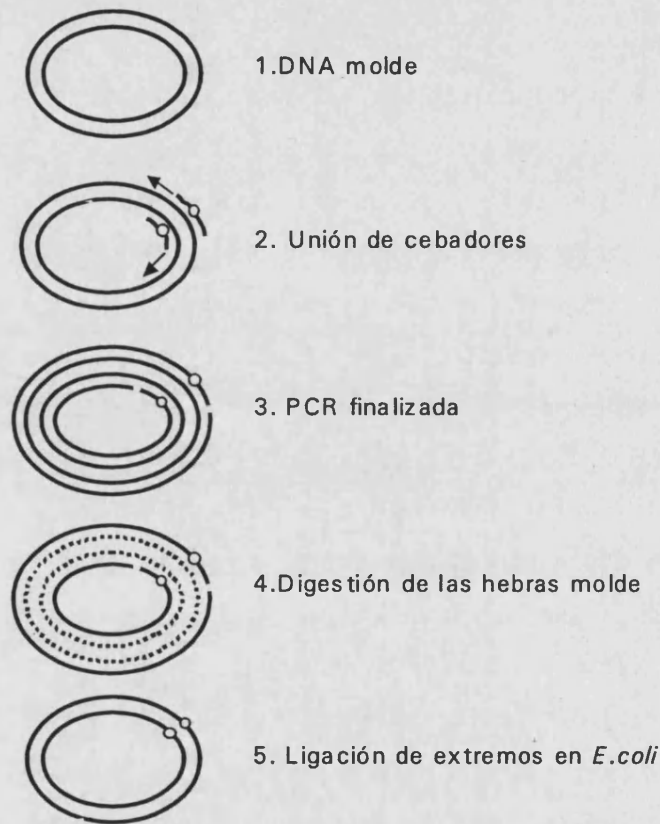
Las infecciones en líquido serán utilizadas para los experimentos de transfección.

## ORTENCIÓN DE LOS VIRUS MUTANTES

### Mutagénesis dirigida

El principio de esta técnica consiste en realizar una reacción en cadena de la polimerasa (PCR) en la que los cebadores incorporan la mutación deseada. No obstante, esta PCR tiene algunas particularidades. En primer lugar, usaremos la polimerasa de alta fidelidad *Pfu* (Promega®), cuya tasa de error es aproximadamente 10 veces menor que la de la polimerasa *Taq* habitualmente empleada en reacciones de PCR (Bracho *et al.*, 1998). Así, minimizamos el riesgo de aparición de mutaciones adicionales no deseadas. En segundo lugar, los cebadores (Invitrogen® o Roche®) son algo más largos de lo habitual (25-40 bases) y deben tener un contenido en GC superior al 40%, una temperatura de unión superior a 75°C y estar los más centrados posible en el sitio diana de mutagénesis. De este modo, podrán unirse específicamente a la región molde del cDNA a pesar de tener un desapareamiento. En tercer lugar, la reacción es lenta, ya que la velocidad de polimerización de la *Pfu* es de 2 kb/min, y el clon infeccioso tiene una longitud total de aproximadamente 14kb. La reacción es interrumpida tras 18 ciclos, ya que la cantidad de DNA obtenido en ese momento es suficiente para la clonación y sobre todo, se reduce el riesgo de introducir cambios no deseados.

Tras la reacción de PCR, no todas las moléculas de DNA llevarán incorporada la mutación, ya que las hebras molde permanecen en solución. Para deshacerse de ellas, se utiliza la endonucleasa de restricción *DpnI* (Promega®), cuya actividad es dependiente de la metilación de DNA, de manera que sólo corta si el DNA se encuentra en su forma metilada. Las cadenas molde de la reacción provienen directamente de una cepa de *E. coli* (XL-Gold®, ver debajo) que metila su DNA, mientras que las hebras mutantes son fruto de la PCR y por lo tanto, carecen de metilación. Por ello, *DpnI* sólo cortará las hebras



**Figura 8.** Estrategia de introducción de mutaciones puntuales en el clon infeccioso de VSV mediante mutagénesis dirigida.

molde. Además, la diana de corte de *DpnI* es de sólo cuatro bases, lo que garantiza que el enzima encuentre varios puntos de corte por puro azar a lo largo de la secuencia.

Aproximadamente 3  $\mu\text{L}$  del DNA resultante son empelados para transformar 25  $\mu\text{L}$  de una suspensión de la cepa ultracompetente XL-Gold® de *E. coli* (Stratagene®) por el método de choque térmico estándar (Ausubel *et al.*,

1999). Notemos que el DNA resultante de la PCR está en forma lineal lo que, junto al elevado tamaño del plásmido, justifica el uso de células ultracompetentes, preparadas para capturar DNA de tamaño superior a las 10 kb. Las células XL-Gold® poseen actividad ligasa que permite circularizar el DNA. Por otro lado, carecen de enzimas de restricción dependientes de metilación, de manera que el DNA no es digerido, aunque poseen los enzimas encargados de metilar el DNA.

La Figura 8 resume la estrategia empleada en este protocolo de mutagénesis.

### **Purificación del clon mutante**

Como consecuencia del gran tamaño del plásmido y de su forma lineal, la eficiencia de la transformación es baja. De hecho, no suelen aparecer más de 200 colonias transformadas por reacción, por lo que cada una de esas colonias habrá sido transformada, con casi total seguridad, con una única molécula de plásmido y constituye como consecuencia, un clon. El siguiente paso es seleccionar una de esas colonias e inocularla en medio de cultivo Luria-Bertani (Ausubel *et al.*, 1999) con el fin de amplificar el clon. Las condiciones más adecuadas para el crecimiento son 37°C, 250 rpm de agitación y con un volumen de aire igual o mayor que el de medio de cultivo. Tras 15h, se elabora un glicerinado y se procede a la extracción del DNA plasmídico.

La extracción, a pequeña escala y de alta pureza, se hace utilizando columnas de intercambio aniónico (Marligen®). El volumen de cultivo necesario (10 mL) excede con creces lo habitual ya que, aunque pGEM-3® es un plásmido de alto número de copias, el elevado tamaño del inserto hace que su replicación sea poco eficiente. El protocolo de alta pureza es algo más lento que el estándar (Ausubel *et al.*, 1999), ya que se añade a éste último un paso de precipitación del eluido con iso-propanol y cloruro sódico, pero es conveniente para la posterior transfección. Además, también se requiere que la solución final de DNA sea estéril, por lo que a partir de la precipitación con iso-propanol, todos los pasos se realizarán en campana de flujo laminar.

El DNA plasmídico resultante, además de estéril, ha de estar libre de proteínas en la medida de lo posible, de modo que el cociente de absorbancia  $A^{260}/A^{280}$  sea cercano a 2 y su concentración ha de ser al menos 100 ng/ $\mu$ L. Para medir la concentración, empleamos tanto el método de medición de la

absorbancia  $A^{260}$  como la visualización de la banda bajo luz UV en presencia de bromuro de etilo tras una electroforesis en gel de agarosa al 0.8%. Este último procedimiento nos permite también asegurar que el DNA está en forma predominantemente circular y que no hay DNA contaminante. Cuando la concentración es mayor que 100 ng/ $\mu$ L, diluimos el DNA en agua ultrapura milli-Q (Millipore®) hasta alcanzar dicha concentración, de modo que las diferentes transfecciones se hagan en condiciones lo más homogéneas posible.

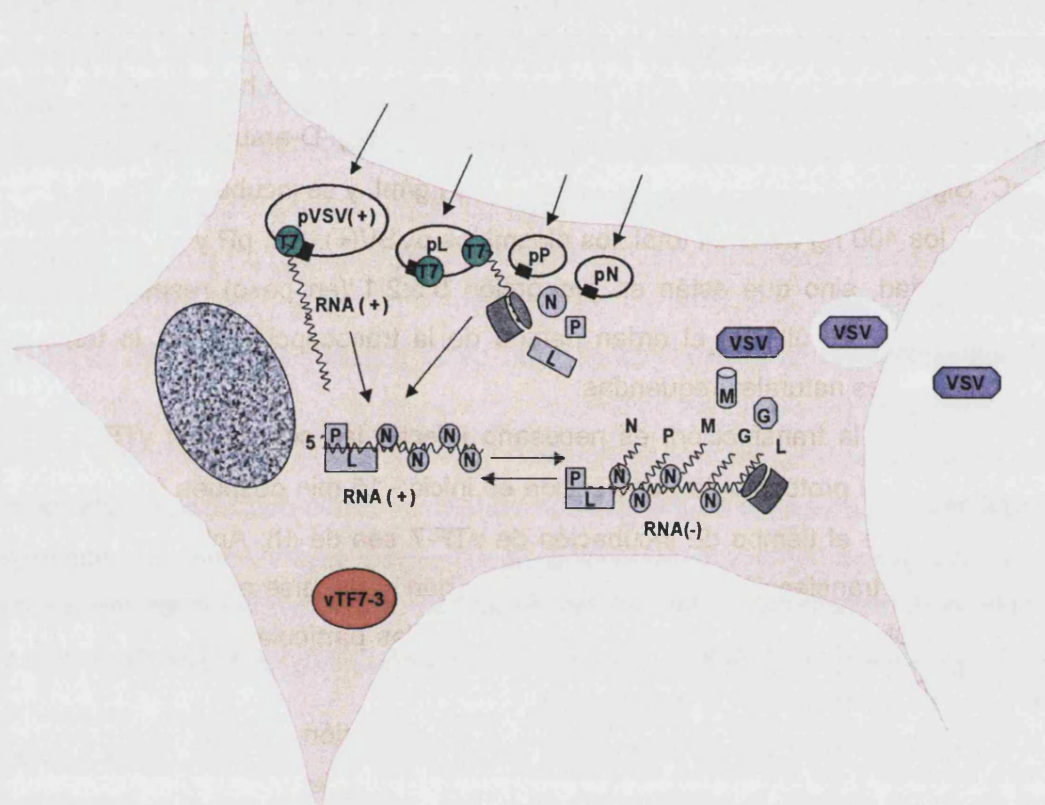
El último paso antes de la transfección es asegurarse de que la mutación deseada ha sido en efecto introducida. Para ello, se secuencia la región de interés, usando aproximadamente 150 ng de plásmido para la reacción de secuenciación, que se llevará a cabo automáticamente por el método de los terminadores marcados con fluoróforos (Ausubel *et al.*, 1999) en un equipo ABI 310 (Applied Biosystems®). Idealmente, habría que proceder a la secuenciación de todo el inserto, e incluso parte del propio vector pGEM-3®, para poder asegurar que no han sido introducidas mutaciones indeseadas, pero esto no es fácilmente realizable en la práctica ya que implica secuenciar una cantidad excesiva de DNA.

### Transfección

El objetivo de la transfección es obtener viriones de VSV a partir del clon de cDNA. Para ello se requiere la introducción en células BHK21 de pVSV(+), pero también de tres plásmidos adicionales pN, pP y pL que llevan insertados respectivamente los genes *N*, *P* y *L*. Estos cuatro plásmidos serán transcritos por la RNA polimerasa del bacteriófago T7 expresada por vTF-7, con el que las células BHK21 han sido infectadas previamente a una *moi* lo bastante elevada (10-30) como para asegurar la presencia de varias partículas de vTF-7 por célula. Por un lado, la transcripción del clon infeccioso da lugar a una cadena anigenómica completa de polaridad positiva. Por otro lado, los genes *N*, *P* y *L* son traducidos por la maquinaria celular, dando lugar a las correspondientes proteínas, que se unen a las cadenas antígenómicas, formando así núcleos virales con capacidad para completar la replicación, y que darán lugar a cadenas de ss(-) RNA genómico completas. A partir de este momento, el ciclo de infección de VSV sigue su curso natural. La Figura 9 sintetiza el proceso de obtención de las partículas infecciosas a partir de los cDNA.



Teóricamente, lo más efectivo sería realizar la transfección con un clon infeccioso de polaridad negativa, ahorrándose el paso de replicación de la cadena positiva. Sin embargo, esto no es posible, ya que las secuencias intergénicas de la cadena genómica mimetizan la señal natural de terminación de la transcripción de la RNA polimerasa T7, impidiendo que ésta complete la transcripción (Whelan *et al.*, 1995).



**Figura 9.** Obtención de partículas infecciosas de VSV a partir del clon infeccioso pVSV(+). El dibujo representa a una célula BHK21 en la que se transfectan los cuatro plásmidos pVSV(+), pN, pP y pL, la transcripción de los cuales es controlada por el promotor T7, cuya RNA polimerasa es expresada por el virus vTF-7.

La co-transfección de los cuatro plásmidos puede llevarse a cabo utilizando el reactivo LipofectAMINE Reagent® (Invitrogen®), compuesto de una mezcla, en forma de liposomas, de un lípido catiónico con un lípido neutro, reactivo que es aconsejable usar en conjunción con el Plus Reagent® (Invitrogen®). Aunque LipofectAMINE Reagent® realiza la transfección, Plus Reagent® permite, de acuerdo con nuestros resultados, que ésta sea exitosa en un mayor rango de

condiciones, compensando así el error experimental. Brevemente, se diluyen 400 ng de DNA total en un volumen total de 21  $\mu$ L de medio DMEM sin complemento alguno (los antibióticos están ausentes en todo el proceso, ya que podrían interferir con los reactivos de transfección). Se añaden 4  $\mu$ L de Plus Reagent®, se mezcla y se deja incubar a temperatura ambiente durante 15 min, tras lo cual se añaden 25  $\mu$ L de LipofectAMINE Reagent® y se mezcla, dejándose incubar durante otros 15 min. Esa mezcla se diluye en 200  $\mu$ L de DMEM, y se añade la dilución a un pocillo w24 con células BHK21 jóvenes (de menos de 25 pases) y 90-95% confluentes, pasadas el día anterior en ausencia de antibióticos. Tras 3-5 h, se suple el medio con FCS a una concentración final del 10% v:v y 1- $\beta$ -D-arabinofuranosilcitosina (araC; Sigma®) a una concentración final de 50  $\mu$ g/mL y se incuba durante 4 días.

En los 400 ng de DNA total, los plásmidos pVSV(+), pN, pP y pL no están en igual cantidad, sino que están en proporción 5:5:2:1 (en peso) respectivamente, reflejando los tres últimos el orden natural de la transcripción y por lo tanto las concentraciones naturales requeridas.

Antes de la transfección, es necesario infectar las células con vTF-7 a *moi* especificada. El protocolo de transfección se inicia ~15 min después de este paso, de manera que el tiempo de incubación de vTF-7 sea de 1h. Antes de añadir los complejos de transfección a las células, éstas han de lavarse al menos tres veces con medio DMEM sin suplementos para eliminar las partículas vTF-7 del medio y que éstas no interfieran en la transfección.

El reactivo araC es un bloqueador de la replicación del DNA, cuya función aquí es la de impedir la proliferación de vTF-7, permitiendo así que las células hospedadoras permanezcan vivas el tiempo suficiente como para ser infectadas por las partículas nacientes de VSV. En algunos de nuestros ensayos preliminares, hemos observado que el título de VSV en una infección de rutina se ve reducido en ~10 veces por añadir araC al medio, y ~1000 veces por preinfectar las células con el virus vTF-7. Sin embargo, cuando la infección con VSV se lleva a cabo en presencia de araC y vTF-7, el título sólo se ve reducido ~100 veces.

### **Manejo de las muestras de transfección**

Los resultados de las transfecciones se evaluarán plaqueando una dilución 1/100 del sobrenadante. Tanto por la experiencia de otros investigadores (Dr. E.

Martínez Salas, comunicación personal) como por nuestra propia experiencia, sabemos que en las condiciones estándar de plaqueo, vTF7 es incapaz de formar aros de lisis visibles tras 24 hpi, por lo que éstos son un signo inequívoco de la obtención de partículas infecciosas de VSV. En cualquier caso, todos los controles negativos han mostrado ausencia completa de calvas en la dilución 1/100. En las transfecciones positivas, por el contrario, las placas aparecen habitualmente arrasadas, siendo necesario realizar una dilución 1/10000 del sobrenadante para poder titular el virus recuperado. Los títulos de VSV obtenidos suelen estar en el rango  $10^5$ - $10^7$  PFU/mL, es decir que existe una apreciable variación entre muestras.

El vTF-7 se elimina del medio mediante filtración a través de filtros con un tamaño de poro de 0.22  $\mu$ m, tras haber llevado a cabo una centrifugación suave (5 min a 5000 rpm) que permita separar el sobrenadante de los restos celulares, evitando así la obturación del filtro. Este procedimiento afecta al título de VSV, ya que en ocasiones, éste se agrega tanto a restos celulares de gran tamaño como al propio vTF-7. Para minimizar este problema, los sobrenadantes pueden someterse a una agitación con vórtex durante 30 s antes de la centrifugación. Aún así, se observa una reducción en el título de VSV de ~10 veces tras la eliminación de vTF-7.

Para pasos posteriores, es importante no sólo que los títulos de los diferentes mutantes de VSV sean homogéneos, sino que el número de generaciones transcurridas sea aproximadamente el mismo en todos los casos. Para ello, se infecta un pocillo w96 con 25  $\mu$ L del filtrado sin diluir, y se deja incubar durante 48 h. En aquellas muestras en las que el título inicial sea elevado, esta infección apenas supondrá un ciclo adicional de infección, mientras que aquéllas con un título inicial escaso podrán igualarse a las primeras. El volumen final (100  $\mu$ L) se diluye en 1 mL total de DMEM, se divide en alícuotas y se almacena a  $-80^{\circ}\text{C}$ . Tras este procedimiento, el título de todas las muestras queda comprendido entre  $5 \times 10^6$  y  $10^7$  PFU/mL.

### **Secuencias consenso de poblaciones de VSV**

Para comprobar si las mutaciones introducidas persisten en la población de VSV, el RNA viral puede extraerse mediante el uso de un sistema comercial (Roche®). Seguidamente, se obtiene el cDNA antígenómico utilizando la transcriptasa reversa MoMuLV (Promega®), un cebador específico del extremo

5´ de la cadena genómica y una batería de hexámeros aleatorios (Promega®), además del resto de componentes habituales en una reacción de retrotranscripción (Ausubel *et al.*, 1999). El siguiente paso es amplificar por PCR la región del genoma en la que se halla la mutación de interés utilizando la *Taq* polimerasa (Amersham®) y precipitar en etanol absoluto con cloruro sódico el amplificado con el fin de eliminar los restos de cebadores y monómeros. Finalmente, se usan unos 4 ng del DNA precipitado para llevar a cabo la secuenciación automática por el método de los terminadores marcados con fluoróforos (Ausubel *et al.*, 1999; Applied Biosystems®).

## ENSAYOS DE COMPETENCIA

### Creación de un fenotipo marcador

En los ensayos de competencia, dos poblaciones de distinto genotipo crecen conjuntamente, siendo pues necesario disponer de un marcador fenotípico que permita diferenciarlos. En nuestro caso, vamos a usar como marcador la resistencia al AcM I<sub>1</sub>, el cual ha sido utilizado durante largo tiempo en los ensayos de eficacia para VSV (Holland *et al.*, 1991). Se han descrito varios genotipos de VSV resistentes al AcM I<sub>1</sub> (Vandepol *et al.*, 1986), a los que se denomina MARM (de sus siglas inglesas *monoclonal antibody resistant mutant*), de los cuales resultan interesantes para los ensayos de competencia los que muestren una eficacia similar al genotipo silvestre (marcadores neutrales). Uno de los que cumple esta condición es MARM C (Holland *et al.*, 1991; Quer *et al.*, 2001), cuyo fenotipo resistente es adquirido por la sustitución A → C en la posición 3853 de la cadena positiva, lo que corresponde a la sustitución D259A en el gen de la glicoproteína G (Vandepol *et al.*, 1986).

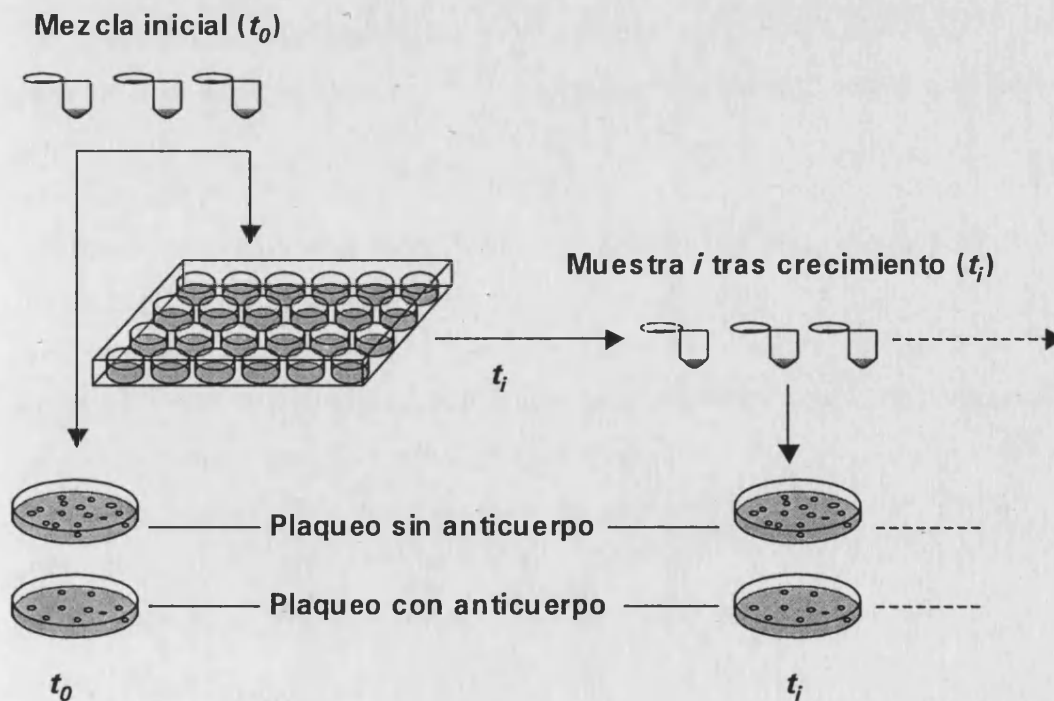
Como ya hemos expuesto, el genoma del clon infeccioso que vamos a utilizar difiere sustancialmente de cualquiera de los genotipos de referencia empleados en experimentos con VSV. Por ello, con el fin de obtener un marcador aproximadamente neutral, es necesario introducir la mutación de resistencia en el fondo genético de nuestro clon infeccioso. Así pues, el primer paso a dar es la obtención del mutante A → C en la posición 3853 mediante mutagénesis dirigida sobre el cDNA y recuperación por transfección en BHK21.

A este virus mutante se le denominará MARM RSV para diferenciarlo del virus obtenido a partir del clon infeccioso intacto, que será nuestro silvestre (*wt*). El cDNA correspondiente a MARM RSV servirá como molde para crear todos los demás mutantes.

Para averiguar la eficacia de MARM RSV relativa al silvestre hemos llevado a cabo 60 ensayos de competencia (ver debajo), obteniéndose un valor de  $W_{RSV|wt} = 0.997 \pm 0.009$ , el cual demuestra que el marcador fenotípico es efectivamente neutral ( $t_{59} = 0.405$ ,  $p = 0.687$ ).

### Ensayos de competencia

El principio del ensayo de competencia consiste en mezclar dos poblaciones en una proporción conocida, dejarlas crecer y medir el cambio en sus frecuencias a lo largo del tiempo, con el fin de obtener una medida de su eficacia biológica relativa. El esquema experimental de las competencias se muestra en la Figura 10. Se mezcla el virus de genotipo mutante, que posee el marcador de resistencia al AcM  $I_1$ , con el virus de genotipo silvestre en proporción 1:1 aproximadamente. Para conocer dicha proporción con exactitud,



**Figura 10.** Esquema experimental de las competencias. Las infecciones en líquido pueden hacerse en placas de 24 pocillos, permitiendo el análisis de varios mutantes simultáneamente. Algunos de los pocillos se reservan para hacer la competencia de referencia entre el MARM RSV y el silvestre. Los plaqueos se realizan en placas de p60.

se realiza una titulación de esa mezcla en presencia y ausencia del AcM I<sub>1</sub>. Por otro lado, se realiza una dilución de la mezcla para infectar un pocillo w24 tal que la cantidad de virus iniciales  $N_0 \approx 5 \times 10^3$  PFU. A las 4, 6, 8 y 10 hpi, se toma una muestra del sobrenadante y se titula de nuevo. Habitualmente, el título alcanzado por el virus silvestre tras esos tiempos ronda las  $10^4$ ,  $10^5$ ,  $10^6$  y  $10^7$  PFU/mL respectivamente.

La elección del tiempo de recogida de las muestras requiere una justificación, ya que lo habitual no es esperar sólo hasta un máximo de 10 hpi, sino hasta que el efecto citopático sea completo, lo cual ocurre generalmente tras ~2 hpi. Además, las muestras recogidas tras ese tiempo suelen ser utilizadas para iniciar una segunda infección de manera que finalmente, se acaba por tener estimaciones de las proporciones en tres tiempos (Holland *et al.*, 1991; Elena *et al.*, 1998; Miralles *et al.*, 1999). Esta manera de estimar la eficacia es válida cuando las poblaciones han sido sometidas a largos períodos de evolución experimental (típicamente 100 generaciones o más), lo que permite asumir que la eficacia va a permanecer constante durante todo el ensayo de competencia. Sin embargo, ésta no es nuestra situación por varios motivos. Primero, el virus con el que trabajamos no ha sufrido dinámica de evolución experimental alguna y segundo, se trata de una quimera genética. Por estas razones, es esperable que su eficacia biológica sea baja y por lo tanto que su tasa de evolución sea rápida (Elena *et al.*, 1998). Además, es nuestro objetivo estudiar mutaciones puntuales aisladas, el efecto de las cuales quedaría fácilmente enmascarado por cualquier otro cambio adicional que ocurriese durante las competencias. Todo ello justifica que nuestros ensayos se realicen en el intervalo de tiempo más breve posible. Mencionemos además que los títulos alcanzados típicamente por el silvestre al completarse el efecto citopático son  $\sim 10^9$  PFU/mL, por lo que el tamaño poblacional tras 10 hpi no excede el 1% del máximo alcanzable. Ello garantiza que el ensayo de competencia sea realizado durante la fase de crecimiento exponencial, lo cual simplifica en gran medida la interpretación de las eficacias biológicas puesto que asegura la ausencia de complementación genética.

## ESTIMACIÓN DE PARÁMETROS GENÉTICO-POBLACIONALES

### Cálculo de las eficacias relativas

En un régimen de crecimiento exponencial, el tamaño poblacional en función del tiempo es

$$N_t = N_0 e^{rt}, \quad (24)$$

donde  $N_0$  es el tamaño poblacional inicial y  $r$  es la tasa de crecimiento o parámetro Malthusiano. En esas condiciones, la eficacia biológica de un genotipo  $i$  vale

$$\omega_i = e^{rt} \quad (25)$$

y la eficacia relativa al silvestre será por lo tanto

$$W_{i|wt} = \omega_i / \omega_{wt} = e^{r_i - r_{wt} t}. \quad (26)$$

Así pues, basta con obtener una estimación de la tasa de crecimiento de cada genotipo para conocer su eficacia biológica relativa. A su vez, las tasas de crecimiento son fáciles de obtener ajustando la forma lineal de la ecuación 24

$$\log N_t = \log N_0 + rt \quad (27)$$

a los datos mediante una regresión lineal en la que sólo estimaremos  $r$ , ya que el parámetro  $N_0$  es obtenido directamente de los datos. En el caso de que sólo se utilicen dos tiempos ( $t_0$  y  $t_1$ ) para estimar  $r$ , sencillamente

$$r = \frac{\log(N_1 / N_0)}{t_1 - t_0}. \quad (28)$$

Aunque el genotipo MARM RSV sea neutral, cada mutante se diferencia del silvestre no sólo por la mutación o mutaciones que queramos analizar, sino también por la presencia del cambio que confiere el fenotipo MARM. Por ello, la medida más correcta de eficacia relativa será la relativa al MARM RSV, es decir

$$W_i = W_{i|wt} / W_{RSV|wt} \quad (29)$$

donde  $W_{RSV|wt}$  es la eficacia del mutante MARM RSV respecto al silvestre.

### Obtención de las eficacias relativas expresadas en generaciones

La eficacia biológica relativa depende de la unidad de tiempo considerada, ya que  $r$  tiene unidades de 1/tiempo, unidad que en nuestras competencias será la hora. Para obtener resultados comparables con otros estudios, es importante



transformar los datos obtenidos en tiempo absoluto en valores donde el tiempo esté expresado en generaciones. En un virus, puede decirse que una generación transcurre cuando se completa un ciclo de infección, ya que éste es el mínimo tiempo necesario para que puedan expresarse todos los caracteres. Definamos  $K$  como el número de virus producidos por célula. Expresando el tiempo en generaciones ( $g$ ), podemos describir la ecuación de crecimiento exponencial como

$$N_g = N_0 K^g \quad (30)$$

por lo que

$$g = \frac{\log(N_g / N_0)}{\log K} \quad (31)$$

y sustituyendo con la ecuación 24, el tiempo medido en generaciones es

$$g = \frac{rt}{\log K} \quad (32)$$

donde  $r$  y  $t$  están expresados en tiempo absoluto. Así pues, basta con dividir  $r$  por el factor  $r/\log K$  para obtener la tasa de crecimiento en generaciones, quedando pues la eficacia en generaciones como  $W_{i(g)} = K_i$  y por lo tanto, la eficacia relativa será  $W_{i(g)} = K_i/K_C$  donde  $K_C$  donde el subíndice  $C$  hace referencia al genotipo MARM RSV. Sin embargo, esta expresión es equívoca, puesto que ignora que el tiempo de generación es a su vez un componente de eficacia (Pianka, 1970; Bulmer, 1994; Cuevas *et al.*, 2005). Para solventar este problema, definamos como tiempo de generación de referencia el de MARM RSV, con lo que finalmente tenemos

$$W_{i(g)} = W_{i(t)}^{\log K_C / r_C} \quad (33)$$

$$W_{i(g)} = W_{i(t)}^{\frac{\log K_C}{r_C}} \quad (34)$$

donde  $r_C$  es la tasa de crecimiento de MARM RSV en tiempo absoluto. La estimación de  $r_C$  puede realizarse utilizando las ecuaciones (27) o (28), mientras que la estimación de  $K_C$  se realizará mediante el método de las curvas clonales de Cuevas *et al.* (2005).

### Tamaño poblacional efectivo y número de generaciones

En una dinámica poblacional fluctuante, como es el caso de las infecciones seriadas, el tamaño poblacional efectivo  $N_e$  puede estimarse como



$$N_e = N_0 g \quad (35)$$

(Lenski *et al.*, 1991). Si en una dinámica de evolución experimental permitimos que se complete el efecto citopático y por tanto ya no sean posibles más ciclos de infección, podemos estimar el número de virus producidos por célula infectada como

$$K = \log\left(\frac{N_f - N_0}{C_0} + 1\right), \quad (36)$$

donde  $N_f$  es la cantidad de virus obtenidos al final de la infección y  $C_0$  la cantidad inicial de células disponibles para la infección (Miralles *et al.*, 2000). Si queremos averiguar qué tamaño de inóculo  $N_0$  necesitamos para que la población evolucione bajo un determinado tamaño poblacional efectivo  $N_e$ , estimaremos  $N_0$  a partir de la ecuación

$$N_e = N_0 \frac{\log(N_f / N_0)}{\log\left(\frac{N_f - N_0}{C_0} + 1\right)} \quad (37)$$

(Miralles *et al.*, 2000) por métodos numéricos.

### Estimación de la tasa de mutación

La prueba de fluctuación de Luria-Delbrück fue propuesta como un método simple para estimar la tasa de mutación bacteriana (Luria y Delbrück, 1943). Para llevarla a cabo, es necesario disponer de un marcador fenotípico, como puede ser la resistencia a un antibiótico o en nuestro caso, a un anticuerpo monoclonal; es decir, a un agente selectivo que define el *locus* sobre el que se estima la tasa de mutación. La prueba permite saber la tasa con la que un fenotipo silvestre muta al fenotipo resistente. La idea consiste en dejar crecer un gran número de poblaciones independientes en ausencia del agente selectivo y averiguar qué cantidad de genotipos mutantes han aparecido en cada población, realizando un recuento en presencia de dicho agente.

En nuestro caso, se infecta una placa w96 con  $\sim 500$  PFU de VSV silvestre por pocillo y se deja incubar hasta que se alcance un tamaño poblacional de  $\sim 3 \times 10^4$  PFU por pocillo. Se usan 84 pocillos para la búsqueda de mutantes de resistencia a AcM I<sub>1</sub>, para lo cual se plaquea el volumen total de cada pocillo (100  $\mu$ L) en una p60 añadiendo 25% v:v de AcM I<sub>1</sub> al medio y se tife tras 28 h de incubación. Los restantes 12 pocillos se usan para plaquear

una dilución 1:300 del sobrenadante en ausencia de AcM I<sub>1</sub> con el fin de obtener una estimación precisa del título.

El método estadístico de la clase nula (Luria y Delbrück, 1943) es probablemente el más robusto a la hora de estimar la tasa de mutación. Únicamente asume que el número de veces que ocurre un evento de mutación tras el tiempo  $t$  de crecimiento sigue una distribución Poisson de parámetro  $2\mu_r(N_t - N_0)$ , siendo  $\mu_r$  la tasa de mutación en el *locus* estudiado por evento de replicación y  $2(N_t - N_0)$  el número eventos de replicación transcurridos. El factor 2 está aquí incluido específicamente para VSV, ya que para completar la replicación, la cadena genómica ha de copiarse a una cadena de polaridad positiva intermediaria. De acuerdo con la distribución Poisson, la probabilidad  $P_0$  de que no aparezca mutante alguno en el cultivo es

$$P_0 = e^{-2(N_t - N_0)\mu_r} \quad (38)$$

por lo que

$$\mu_r = -\frac{1}{2} P_0 \log(N_t - N_0) \quad (39)$$

Si estamos interesados en obtener la tasa de mutación por generación, es decir por ciclo de infección, hemos de asumir que en el interior de cada célula, cada genoma tan sólo pasa una vez por el estado intermediario de polaridad positiva, algo de lo cual no hay evidencia. Entonces, la tasa de mutación individual por generación  $\mu$  es sencillamente

$$\mu = 2\mu_r \quad (40)$$

Existe un método alternativo para estimar la tasa de mutación, pensado específicamente para virus (Drake *et al.*, 1998; Drake y Holland, 1999). Se basa en que, en el interior de la célula, la cadena genómica del virus es copiada múltiples veces a una cadena de polaridad opuesta, y ésta a su vez es copiada múltiples veces a su complementaria genómica. Si esto sólo ocurre una vez en cada célula, se dice que el modo de replicación es lineal y entonces, en cada ciclo o generación se cumplirá que  $\mu = m$ , siendo  $m$  la frecuencia observada del fenotipo mutante. Similarmente, para cualquier número de ciclos,  $c$ ,

$$\mu = m/c; \quad (41)$$

mientras que si estamos interesados en la tasa de mutación por evento de replicación, ésta puede obtenerse a partir de la ecuación 40. Una consecuencia del modo de replicación lineal es que el número observado de mutantes es igual al número de mutaciones ocurridas, por lo que el método puede ser refinado

ajustando a la distribución observada del número de mutantes una distribución Poisson (Chao *et al.*, 2002). Es importante mencionar que, en cualquier caso, el método asume que el fenotipo marcador es neutral respecto al silvestre.

En el otro extremo, asumiendo que el modo de replicación es estrictamente binario, puede estimarse la tasa de mutación como

$$\mu \equiv \mu_r = m / \log(N_t / N_0) \quad (42)$$

(Drake *et al.*, 1998). Igual que antes, si se quiere aprovechar más la información proporcionada por la prueba de fluctuación, puede utilizarse toda la distribución del número de mutantes observados. En esta ocasión, el ajuste no ha de realizarse a una distribución de Poisson sino a la distribución denominada de Luria-Delbrück (Sniegowski *et al.*, 1997). Este método asume también la neutralidad del marcador fenotípico.

En primer lugar, la tasa de mutación por *locus* ha de ser dividida por seis puesto que además del cambio A → G en la posición 3853 de la cadena antigenómica, se han descrito otras cinco sustituciones nucleotídicas que confieren el fenotipo de resistencia a AcM I<sub>1</sub> (Holland *et al.*, 1991). Además, ha de ser multiplicada por tres para convertirla en una tasa de mutación por sitio nucleotídico, ya que sólo uno de los tres cambios nucleotídicos posibles confiere el fenotipo resistente en cada una de las seis posiciones. Finalmente, para obtener una tasa de mutación genómica, habrá que multiplicar este valor por la longitud del genoma. Así pues, la tasa de mutación genómica por generación es

$$U = 3/6 \times 11162 \times \mu. \quad (43)$$

Notemos que el valor así obtenido es una subestima de la tasa de mutación real, puesto que no puede descartarse que existan otras sustituciones capaces de conferir el fenotipo resistente además de las seis descritas.

## RÉPLICAS Y BLOQUES EXPERIMENTALES

Los ensayos de competencia de los mutantes simples serán realizados en 5 bloques independientes, incluyéndose una réplica de competencia de cada mutante y 3 réplicas para el genotipo MARM RSV por cada bloque. Se tomarán muestras a las 4, 6, 8 y 10 hpi para estimar las tasas de crecimiento según la ecuación 27 y la eficacia biológica relativa al silvestre. Los 15 valores (5

bloques  $\times$  3 réplicas) obtenidos en las competencias entre el genotipo MARM RSV y el silvestre serán promediados para obtener una mejor estima del valor de  $W_{RSV|wt}$  con el que calcular los 5 valores de eficacia relativa a MARM RSV para cada mutante usando la ecuación 29.

Para los 10 mutantes simples de mayor eficacia, se harán 5 bloques adicionales de competencias de manera idéntica a como se han realizado los anteriores, obteniéndose así un total de 10 valores de eficacia para cada uno de ellos.

Los ensayos de competencia de los mutantes dobles serán realizados también en 5 bloques independientes, incluyéndose una réplica de competencia de cada mutante y 6 réplicas para el genotipo MARM RSV por cada bloque. Se tomarán muestras a las 10 hpi para estimar las tasas de crecimiento según la ecuación 28. Los 30 valores (5 bloques  $\times$  6 réplicas) obtenidos en las competencias entre el genotipo MARM RSV y el silvestre serán promediados para obtener una mejor estima del valor de  $W_{RSV|wt}$  con el que calcular los 5 valores de eficacia relativa a MARM RSV.

Los experimentos de recuperación de la eficacia biológica serán llevados a cabo en 3 bloques paralelos para cada uno de los 2 valores de  $N_e$  explorados, incluyéndose en cada bloque la totalidad de los genotipos evolucionados. Los ensayos de competencia serán realizados en 3 bloques independientes, incluyendo éstos una competencia para cada una de las 6 réplicas de evolución (3 bloques  $\times$  2  $N_e$ ) de cada genotipo. Se tomarán muestras a las 10 hpi para estimar la tasa de crecimiento según la ecuación 28. Los 3 valores de eficacia obtenidos para cada población evolucionada serán promediados y divididos por la eficacia biológica promedio del mutante ancestral correspondiente relativa al silvestre, con el fin de obtener 3 estimaciones de la tasa de evolución de cada genotipo para cada uno de los 2 tamaños poblacionales.

Todos los análisis estadísticos se realizarán con la versión 10.0 del programa SPSS®, a excepción de la obtención de las probabilidades esperadas según diferentes distribuciones para el coeficiente de selección de los mutantes simples, que puede ser obtenida introduciendo la definición matemática de cada distribución en EXCEL (Microsoft Office 2000®), así como de la generación de pseudo-muestras aleatorias por *bootstrap*, que se llevó a cabo mediante el uso del paquete estadístico R de libre distribución.

## **RESULTADOS**



## EFFECTO DE SUSTITUCIONES NUCLEOTÍDICAS SOBRE LA EFICACIA

### ¿Son las mutaciones únicas?

Hemos obtenido una colección de 91 cDNAs en los que hemos introducido una única sustitución nucleotídica respecto al clon MARM RSV, a los que denominamos mutantes simples. La pregunta que puede surgir es qué grado de certeza tenemos de que realmente, esa sea la única mutación incorporada. Una vez creado el clon mutante de cDNA, se obtienen las partículas infecciosas por transfección y seguidamente, se amplifican hasta alcanzar un título aproximado de  $10^7$  PFU/mL. Asumiendo que durante la reacción de mutagénesis no se ha incorporado ningún cambio indeseado, el clon inicial generará gran cantidad de variantes desde los momentos tempranos de la infección debido a la elevada tasa de mutación de VSV, con lo que en el momento de las competencias no estaremos estimando la eficacia biológica de un individuo, sino de una población heterogénea. Sin embargo, esto no debe preocuparnos, puesto que se trata de la manera en que intrínsecamente los virus de RNA replican.

Nuestra preocupación se centra en la posibilidad de que, durante el tiempo transcurrido entre el ensamblado de las primeras partículas tras la transfección y los ensayos de competencia, se haya fijado en la población alguna mutación que altere el valor de la eficacia media de ésta. El número de generaciones relativas al MARM RSV transcurridas en este intervalo puede ser estimado mediante la ecuación 31. En la fase de transfección, conocemos  $N_r$ , pero no tenemos información precisa acerca de  $N_0$ , ya que ignoramos qué proporción de las moléculas de cDNA han dado lugar a partículas infecciosas. Usando el valor mínimo  $N_0 = 1$  PFU, podemos afirmar que en ningún caso habrán transcurrido más de 4.2 generaciones durante esta fase. El número de generaciones durante la fase de amplificación depende negativamente del título final alcanzado en la transfección, de manera que más generaciones hayan transcurrido durante la primera fase, menos podrán transcurrir durante la segunda. Podemos estimar finalmente que durante el tiempo total previo a los ensayos de competencia, han transcurrido en promedio  $\sim 4$  generaciones, período en el cual no parece probable que hayan podido fijarse un número significativo de sustituciones.



En cualquier caso, para cuantificar la distorsión que este problema podría causar sobre los datos, hemos elegido al azar cuatro de los genotipos (1323, 3950, 5597, 7461) más el MARM RSV y para cada uno de ellos, hemos realizado en paralelo el procedimiento de transfección y amplificación cuatro veces, llevándose a cabo cinco ensayos de competencia para cada amplificado. Por un lado, como era de esperar, el valor obtenido de eficacia depende fuertemente del genotipo (ANOVA encajado:  $F_{4,15}=470.614$ ,  $p<0.001$ ). Por otro lado, si se hubieran fijado mutaciones adicionales antes de los ensayos de competencia, esperaríamos también observar diferencias entre amplificados. Sin embargo, de acuerdo con los datos, tales diferencias no son estadísticamente significativas (ANOVA encajado:  $F_{15,80}=0.975$ ,  $p=0.489$ ).

Por último, hemos obtenido la secuencia consenso del genoma completo para uno de los amplificados de cada genotipo. En tres de los cinco casos, no hay cambio inesperado alguno, mientras que en dos casos se ha detectado un cambio sinónimo inesperado el cual, a la vista de los resultados de las competencias, no produce una distorsión apreciable de la eficacia.

### **Mutaciones letales**

Las transfecciones fallidas y las mutaciones letales producen un mismo resultado aparente, que es la ausencia de partículas infecciosas. Para cada uno de los genotipos, en caso de obtener un resultado de la transfección negativo, ésta se ha repetido hasta lograrse la consecución de partículas infecciosas. Sin embargo, tras diez intentos, en 24 de los 91 casos, los resultados han seguido siendo negativos. A lo largo de esos intentos, el cDNA correspondiente al genotipo silvestre ha sido transfectado 39 veces, obteniéndose partículas infecciosas en 22 casos, por lo que la tasa de transfección fallida es igual a 0.436. Así pues, la probabilidad de no obtener partículas tras diez intentos es  $0.436^{10}=2.5\times 10^{-4}$ , y en una muestra de 91 genotipos,  $2.5\times 10^{-4}\times 91=0.02$ . Por lo tanto, los 24 casos en los que no se han podido obtener partículas infecciosas no son explicables por errores de transfección, llegándose a la conclusión de que, o bien los cDNA correspondientes no están en condiciones adecuadas, o bien las mutaciones que hemos introducido son letales. Esta última posibilidad adquiere más fuerza cuando se tiene en cuenta que en esos 24 genotipos, 19 de las sustituciones introducidas son no sinónimas, tres generan codones de parada y una produce la interrupción del codón de inicio del



gen *G*, mientras que por el contrario, sólo hay un caso de mutación sinónima, situada 53 nucleótidos antes del final del gen *M*. Además, el 40% de las mutaciones aleatorias (19/48) son posibles letales, mientras que esta fracción desciende al 12% entre las mutaciones previamente observadas, pudiéndose afirmar que los posibles letales no están repartidos igualmente entre ambos grupos (prueba exacta de Fisher,  $p < 0.010$ ).

### **Proporción de mutaciones deletéreas, neutrales y beneficiosas**

Para cada uno de los 67 genotipos viables, se han obtenido cinco estimaciones independientes de la eficacia biológica definida en la ecuación 29. Con el fin de expresar estas eficacias en generaciones, utilizamos la ecuación 34, donde  $r_c = 0.642$  ha sido estimado mediante la ecuación 27 y  $\log K_c = 3.828$  ha sido obtenido por el método de Cuevas *et al.* (2005). Los valores medios de estas eficacias se muestran en el Apéndice 3. Para averiguar si las mutaciones introducidas son deletéreas, neutrales o beneficiosas, basta con evaluar la hipótesis nula  $H_0: W_i = 1$  para cada genotipo *i* mediante una prueba *t* de una muestra. De esta manera, se concluye que 31 de las mutaciones introducidas son aparentemente neutrales, mientras que en 32 casos las mutaciones son deletéreas y en 4 casos las mutaciones son beneficiosas.

Estos resultados pueden haberse visto afectados por dos tipos de error estadístico. Primero, es posible que la hipótesis nula de neutralidad haya sido rechazada siendo cierta en realidad (error de tipo I) y segundo, puede haber sido aceptada siendo falsa (error de tipo II). Si todas las mutaciones fueran neutrales, se esperaría uno o dos ( $67 \times 0.025$ ) falsos efectos deletéreos y la misma cantidad de falsos beneficiosos como consecuencia del error de tipo I. Claramente, este no sería un problema a la hora de estimar la proporción de mutaciones deletéreas. En cuanto a las mutaciones beneficiosas, podríamos aplicar una corrección de pruebas múltiples, pero eso haría aumentar ostensiblemente el error de tipo II. En lugar de eso, hemos preferido realizar cinco réplicas adicionales para los 10 genotipos de máxima eficacia, en los que los cuatro posibles mutantes ventajosos están incluidos. Después de estas repeticiones, los cuatro permanecen significativamente beneficiosos y otros cuatro más, antes clasificados como neutrales, pasan a serlo también, habiendo finalmente un total de 8 mutaciones beneficiosas.

La Tabla 3 resume las proporciones estimadas de mutaciones letales, deletéreas, neutrales y beneficiosas para los cambios introducidos al azar y los previamente descritos. Notemos que las estimaciones de las proporciones de mutaciones deletéreas y beneficiosas pueden ser subestimadas de las proporciones reales dado que algunos de los cambios clasificados como neutrales podrían en realidad tener un efecto sobre la eficacia demasiado pequeño para ser detectado en nuestros experimentos (error de tipo II).

	Aleatorias	Descritas	Total
<b>Letales</b>	39.6% (19)	11.6% (5)	26.4% (24)
<b>Deletéreas</b>	29.2% (14)	41.9% (18)	35.2% (32)
<b>Neutrales</b>	27.1% (13)	32.6% (14)	29.7% (27)
<b>Beneficiosas</b>	4.2% (2)	14.0% (6)	8.8% (8)
<b>Total</b>	100% (48)	100% (43)	100% (91)

**Tabla 3.** Proporción y número (entre paréntesis) observados de cambios letales, deletéreos, neutrales y beneficiosos para las mutaciones aleatorias y previamente descritas

En análisis posteriores, debemos distinguir entre mutaciones aleatorias y previamente descritas, ya que no solamente el procedimiento del muestreo difiere entre ambos grupos, sino también su significado biológico. En el primer caso, los valores obtenidos deben ser únicamente el reflejo de los efectos mutacionales, mientras que en el segundo, estos son además atribuibles a la acción de la deriva genética o la selección natural.

#### **Eficacias y coeficientes de selección**

La eficacia media de los genotipos que incorporan mutaciones aleatorias es  $0.445 \pm 0.064$ , mientras que la de los genotipos con mutaciones previamente descritas es  $0.759 \pm 0.069$ , siendo la diferencia entre los dos grupos estadísticamente significativa (prueba de Mann-Whitney de una cola:  $Z = 2.973$ ,  $p = 0.003$ ). En Genética de Poblaciones, suele hacerse necesaria la distinción entre el coeficiente de selección de las mutaciones deletéreas y de las mutaciones beneficiosas. Esta división, que se hace sencillamente en función del signo de dicho coeficiente sin tener en cuenta si los valores son o no

significativamente distintos de cero es la que aplicaremos en lo sucesivo. Excluyendo los letales, la eficacia biológica media de los genotipos que incorporan mutaciones de efecto deletéreo es  $0.663 \pm 0.065$ , mientras que para los genotipos con mutaciones beneficiosas esta vale  $1.224 \pm 0.069$ . La Tabla 4 muestra la eficacia biológica y el coeficiente de selección medios para estos dos grupos, distinguiendo entre mutaciones aleatorias y previamente descritas. Los resultados indican que la intensidad de la selección (valor absoluto del coeficiente de selección) en contra de las mutaciones aleatorias deletéreas es notoriamente mayor que la intensidad de la selección a favor de las mutaciones beneficiosas (prueba de Mann-Whitney:  $Z=3.211$ ,  $p=0.001$ ). Sin embargo, no puede decirse lo mismo de las mutaciones previamente observadas (prueba de Mann-Whitney:  $Z=1.127$ ,  $p=0.260$ ).

Como era de esperar, la eficacia biológica relativa de los genotipos que han incorporado cambios sinónimos es mayor ( $W=0.759 \pm 0.131$ ) que la de aquellos que han incorporado mutaciones aleatorias no sinónimas ( $W=0.374 \pm 0.072$ ; prueba de Mann-Whitney de una cola:  $Z=2.418$ ,  $p=0.009$ ). Resulta interesante que el efecto promedio de los cambios sinónimos es deletéreo, aunque la evidencia estadística a favor de esta afirmación es cuestionable (prueba  $t$  de una muestra y una cola:  $t_9=1.844$ ,  $p=0.049$ ).

	Aleatorias		Descritas	
	Deletéreas	Beneficiosas	Deletéreas	Beneficiosas
<b>W</b>	$0.623 \pm 0.060$	$1.091 \pm 0.034$	$0.693 \pm 0.042$	$1.317 \pm 0.108$
<b>W - 1</b>	$-0.377 \pm 0.060$	$0.091 \pm 0.034$	$-0.307 \pm 0.042$	$0.317 \pm 0.108$
<b>logW</b>	$-0.656 \pm 0.159$	$0.085 \pm 0.030$	$-0.422 \pm 0.066$	$0.250 \pm 0.074$

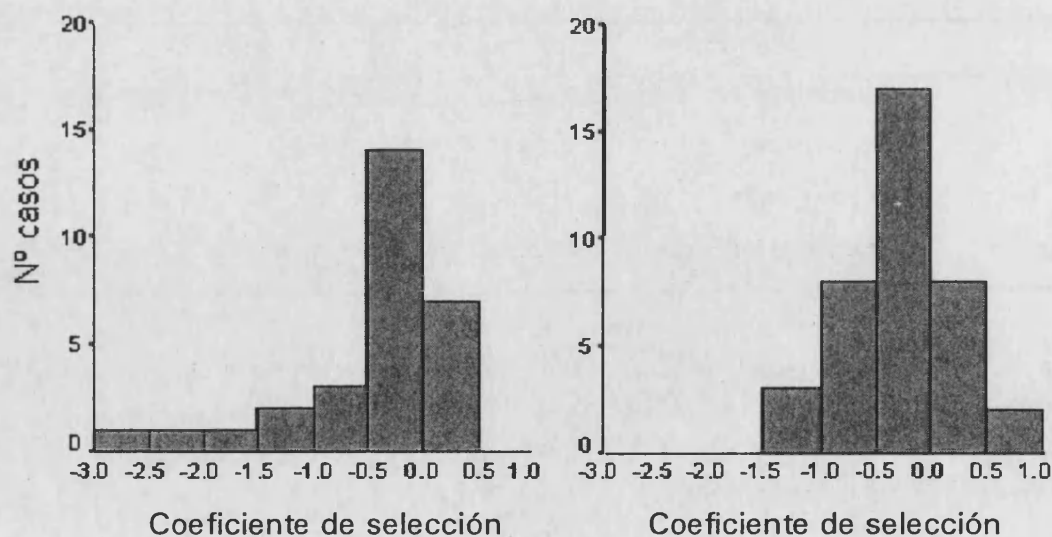
**Tabla 4.** Eficacia biológica, coeficiente de selección y coeficiente de selección Malthusiano para las mutaciones deletéreas (excluyendo las letales) y beneficiosas, en función del criterio de elección de éstas. Se muestran los valores medios y su error asociado.

También de acuerdo con lo esperado, el efecto de las mutaciones deletéreas aleatorias es mayor que el de las previamente observadas. Para hacer esta comparación de manera adecuada, hemos de excluir los cambios sinónimos

del primer grupo (recordemos que en el segundo grupo sólo se han introducido cambios no sinónimos), tras lo cual la eficacia biológica media se convierte en  $0.549 \pm 0.069$ , siendo ésta significativamente menor en las mutaciones aleatorias que para las previamente descritas (prueba de Mann-Whitney de una cola:  $Z=1.705$ ,  $p=0.044$ ). Del mismo modo, tras excluir los cambios sinónimos, la eficacia biológica media asociada a las mutaciones aleatorias beneficiosas es  $1.068 \pm 0.028$ , valor significativamente menor que en el caso de las mutaciones previamente descritas (prueba de Mann-Whitney de una cola:  $Z=1.980$ ,  $p=0.024$ ).

### Características generales de las distribuciones de efectos

Los valores mínimo (0.059) y máximo (0.981) de eficacia en el caso de las mutaciones deletéreas cubren prácticamente todo el intervalo de definición  $[0,1[$  de ésta. Lógicamente, en el caso de las mutaciones beneficiosas, las eficacias no pueden cubrir su intervalo de definición, ya que éste es  $]1, +\infty[$ . Esta asimetría en la definición de la eficacia dificulta la comparación directa de los coeficientes de selección deletéreos y beneficiosos definidos según la ecuación 1. Es por ello, que a la hora de realizar esta comparación, así como el ajuste de los valores a distribuciones probabilísticas, vamos a utilizar los coeficientes de selección Malthusianos, definidos según la ecuación 2.



**Figura 11.** Distribución observada del coeficiente de selección Malthusiano para mutaciones aleatorias (izquierda) y previamente descritas (derecha).

La distribución observada del coeficiente de selección para mutaciones aleatorias y previamente descritas se muestra en la Figura 11. Para las mutaciones aleatorias, ésta difiere de una distribución Normal por ser claramente asimétrica ( $g_1 = -1.909$ ,  $t_{28} = -4.399$ ,  $p < 0.001$ ) y por tener una curtosis positiva ( $g_2 = 3.484$ ,  $t_{28} = 4.123$ ,  $p < 0.001$ ), lo que significa que las observaciones se concentran en valores extremos. Por lo tanto, es conveniente explorar modelos probabilísticos diferentes de la distribución Normal, para lo cual vamos a distinguir entre mutaciones deletéreas y beneficiosas.

Estas conclusiones no son igualmente válidas en el caso de las mutaciones previamente descritas, ya que ni la asimetría ( $g_1 = -0.163$ ,  $t_{37} = 0.426$ ,  $p = 0.336$ ) ni la curtosis son significativas ( $g_2 = -0.297$ ,  $t_{37} = 0.396$ ,  $p = 0.347$ ). Sin embargo, vamos a explorar los mismos modelos probabilísticos que en el caso de las mutaciones aleatorias, con el fin de mantener unas observaciones de referencia.

#### Distribución de efectos deletéreos

En este apartado, vamos a ajustar diferentes modelos probabilísticos a la distribución acumulada observada de los valores absolutos de los coeficientes de selección Malthusianos. Para ello ha empleado el método de ajuste de datos a modelos no-lineales de Levenberg-Marquard implementado en SPSS®. La distribución más sencilla que permite describir la distribución de efectos mutacionales es la Exponencial, cuya función de densidad de probabilidad (*fdp*) es

$$Ex(s | \lambda) = \lambda e^{-\lambda s}, \quad (44)$$

donde  $\lambda$  es el único parámetro de la distribución. En base al criterio de los mínimos cuadrados, la función de probabilidad acumulada de esta distribución se ajusta a las observaciones de manera relativamente satisfactoria, tanto para las 22 mutaciones aleatorias como para las 28 previamente descritas, explicando más del 95% de la varianza observada (Tabla 5).

El modelo exponencial puede ser generalizado por una distribución Gamma, cuya *fdp* es

$$Ga(s | \alpha, \beta) = C s^{\alpha-1} e^{-\beta s}, \quad (45)$$



donde  $C$  es una constante normalizadora, mientras que  $\alpha$  y  $\beta$  son los dos parámetros de la distribución. Notemos que si  $\alpha=0$ , estamos en el caso de la distribución Exponencial. De acuerdo con el criterio de información de Akaike, tanto en el caso de las mutaciones aleatorias como en el de las previamente descritas, la ventaja en términos de aumento de precisión que supone la introducción de un parámetro adicional no compensa la desventaja en términos de una mayor complejidad del modelo (Tabla 5).

La distribución Weibull, cuya *fdp* viene dada por la expresión

$$We(s | \alpha, \beta) = \alpha \beta^{-\alpha} s^{\alpha-1} e^{-(s/\beta)^\alpha} \quad (46)$$

representa otra generalización de la distribución Exponencial (la Weibull es igual a una Exponencial para  $\alpha=1$ ). En esa ocasión, la mejora en el ajuste que proporciona la adición de un parámetro compensa el aumento de complejidad, en lo que a las mutaciones aleatorias se refiere, mientras que el modelo Exponencial sigue siendo el más adecuado en el caso de las mutaciones previamente descritas de acuerdo con el criterio de Akaike (Tabla 5).

La distribución de Pareto, también llamada ley de potencia, y cuya *fdp* es

$$Pa(s | \alpha, \beta) = \alpha \beta^\alpha s^{-(1+\alpha)}, \quad (47)$$

es un modelo relacionado las distribuciones que acabamos de introducir. Su

	Aleatorias		Descritas	
	$r^2$	<i>AIC</i>	$r^2$	<i>AIC</i>
Exponencial	0.960	9.785	0.966	2.652
Gamma	0.961	9.962	0.974	3.094
Weibull	0.963	9.536	0.972	3.192
Pareto	0.973	8.784	0.882	6.011
log-Normal	0.974	8.087	0.951	3.722
Semi-Normal	0.924	27.177	0.938	3.571

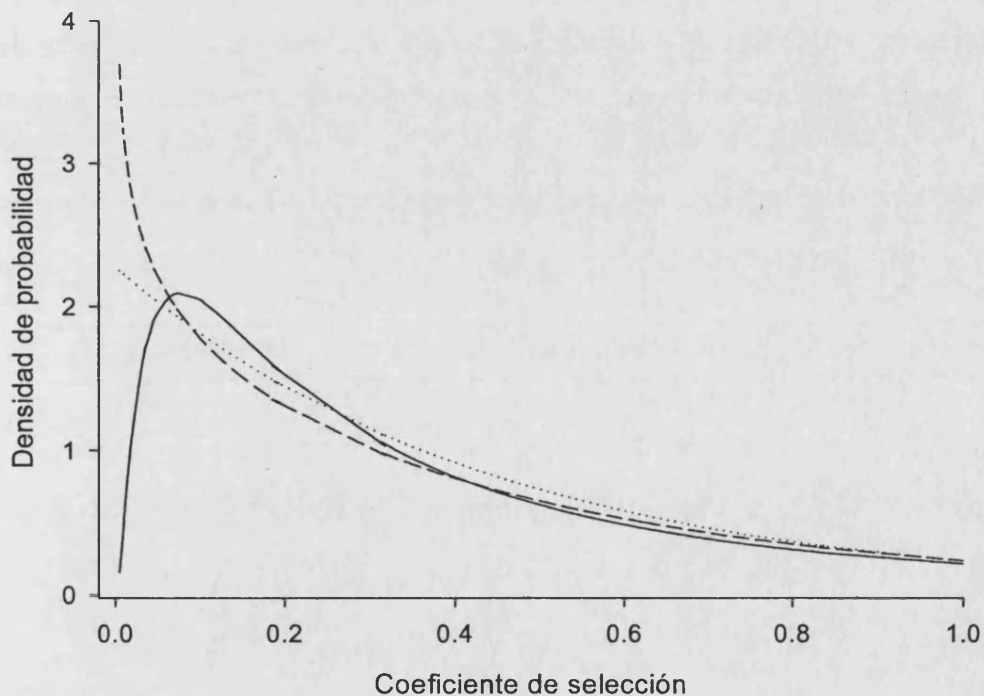
**Tabla 5.** Ajuste de diferentes modelos probabilísticos a la distribución acumulada observada del valor absoluto de los coeficientes de selección para mutaciones deletéreas. Se muestra el cuadrado del coeficiente de correlación ( $r^2$ ), que indica la proporción de la varianza explicada por el modelo, así como el estadístico de Akaike (*AIC*), que permite ordenar los modelos de mejor (menor *AIC*) a peor (mayor *AIC*).

ajuste es mejor que el de la Gamma y la Weibull para las mutaciones aleatorias y también es preferible a la exponencial de acuerdo con el criterio de Akaike (Tabla 5), mientras que el modelo Exponencial sigue siendo el más adecuado para las mutaciones previamente observadas.

Finalmente, el mejor ajuste para las mutaciones aleatorias se logra con una distribución log-Normal, la cual se define como aquella en la que el logaritmo de la variable se distribuye de manera Normal, y por lo tanto su *fdp* es

$$\text{Log}N(s|\alpha, \beta) = \frac{1}{\alpha\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(s/\beta)}{\alpha}\right)^2\right]. \quad (48)$$

Los valores estimados para los parámetros de esta distribución son  $\alpha = 1.203 \pm 0.068$  y  $\beta = 0.327 \pm 0.013$ , con lo que el valor medio o esperanza de los efectos deletéreos (en valor absoluto) según este modelo es



**Figura 12.** Funciones de probabilidad Exponencial (punteado) y Gamma (línea discontinua) que describen el efecto de las mutaciones deletéreas previamente descritas sobre la eficacia, así como función de probabilidad log-Normal (línea continua) que describe el efecto de las mutaciones aleatorias.

$E[s] = e^{\log \beta + \alpha/2} = 0.674$  y su varianza es  $Var[s] = e^{2 \log \beta + 2\alpha} - e^{2 \log \beta + \alpha} = 1.477$ . En cuanto a las mutaciones previamente descritas, de acuerdo con la distribución Exponencial,  $E[s] = 1/\lambda = 0.440$  y  $Var[s] = 1/\lambda^2 = 0.193$ . Las distribuciones de efectos inferidas para los dos conjuntos de datos se muestran en la Figura 12.

En el caso de que los efectos observados fueran una mera consecuencia del error experimental aditivo, éstos se ajustarían a una Normal de media cero, y por lo tanto los efectos deletéreos observados seguirían la semi-Normal asociada. Sin embargo, este modelo está entre los que más pobremente se ajusta a los datos (Tabla 5), quedando pues descartada esta posibilidad.

#### Distribución de efectos beneficiosos

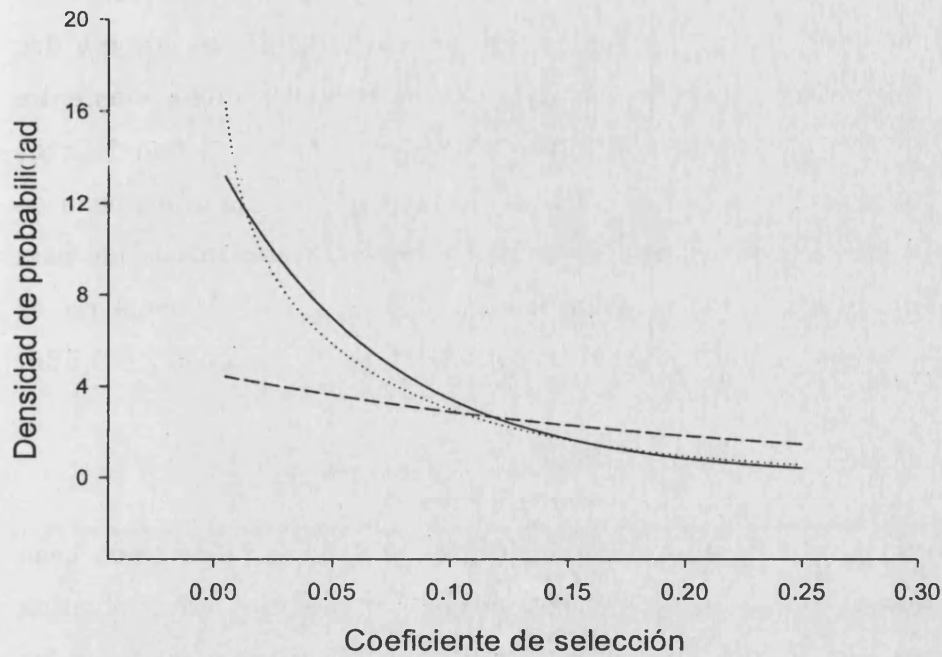
La contrastación de los diferentes modelos probabilísticos en el caso de las mutaciones beneficiosas es menos resolutive, debido al escaso tamaño de las muestras. Los ajustes correspondientes, mostrados en la Tabla 9, indican que para las 7 mutaciones aleatorias, el mejor modelo es una distribución Exponencial de parámetro  $\lambda = 14.130$ , con lo que la esperanza y la varianza de los efectos beneficiosos aleatorios son  $E[s] = 0.070$  y  $Var[s] = 0.005$ , aunque el ajuste logrado con una Gamma es mejor y el criterio de Akaike indica que el coste de incluir el parámetro adicional está cerca de ser compensado. Para las 10 mutaciones previamente descritas, los datos no permiten afirmar que

	Aleatorias		Descritas	
	$r^2$	AIC	$r^2$	AIC
Exponencial	0.932	-23.585	0.941	1.339
Gamma	0.950	-23.066	0.941	2.514
Weibull	0.948	-23.025	0.942	2.714
Pareto	0.738	-18.536	0.920	1.789
log-Normal	0.933	-22.735	0.917	3.635
Semi-Normal	0.887	-21.647	0.942	3.480

**Tabla 6.** Ajuste de diferentes modelos probabilísticos a la distribución acumulada del valor absoluto de los coeficientes de selección observados en mutaciones beneficiosas.



ninguno de los modelos de dos parámetros sean mejores que la Exponencial. El valor del parámetro estimado para este último es  $\lambda=4.535$ , por lo que se estima  $E[s]=0.220$  y  $Var[s]=0.049$ . Las distribuciones de efectos inferidas



**Figura 13.** Funciones de probabilidad Exponencial (continua) y Gamma (punteada) que describen el coeficiente de selección Malthusiano de las mutaciones beneficiosas previamente descritas sobre la eficacia, así como función de probabilidad Exponencial (línea discontinua) que describe el efecto de las mutaciones aleatorias.

para los dos conjuntos de datos se muestran en la Figura 13.

Por último, los efectos observados no pueden ser una mera consecuencia del error experimental aditivo, ya que la distribución semi-Normal no está entre las que mejor se ajusta a los datos.

**Cambios previamente descritos en clones y en secuencias consenso**

Las mutaciones previamente descritas pueden separarse en dos grupos principales. En primer lugar, están aquéllas que han sido detectadas en clones moleculares, que corresponden pues a genomas individuales de los cuales se desconoce la frecuencia en la población de origen y además, en algunos casos, no existe garantía de que esos cambios no sean fruto de errores cometidos

durante la amplificación por PCR. Por estos motivos no es de extrañar que los efectos de esos cambios sobre la eficacia bien puedan ser deletéreos. En segundo lugar, están aquéllas que han sido detectadas en secuencias consenso de poblaciones, de las cuales se espera que, al menos el efecto sobre la eficacia sea menos deletéreo que en el caso de las primeras. Sin embargo, de las cinco mutaciones letales observadas, tres pertenecen al primer grupo y dos al segundo y en cuanto a las no letales, tampoco parece haber diferencias entre ambos grupos en la frecuencia de cambios deletéreos, neutrales y beneficiosos (prueba de una cola,  $\chi^2 = 1.519$ ,  $p = 0.234$ ). Por último, aunque la mediana de los efectos de las mutaciones aisladas en clones es  $-0.301$  mientras que para aquéllas halladas en secuencias consenso es  $-0.107$ , esta diferencia no es estadísticamente significativa (prueba de Mann-Whitney de una cola:  $Z = 0.565$ ,  $p = 0.294$ ).

#### **Grado de restricción funcional en diferentes genes**

En principio, cabe esperar que los efectos mutacionales deletéreos sean distintos en los diferentes genes de VSV debido a diferencias en el nivel de restricción funcional. Utilizando sólo las observaciones correspondientes a los cambios aleatorios, es posible detectar heterogeneidad del efecto mutacional entre genes (prueba de Kruskal-Wallis: ( $\chi^2 = 7.946$ ,  $p = 0.047$ ), siendo el gen *N* el que más restricción funcional muestra, con un efecto mediano de  $-2.170$ , seguido por el *G* ( $-0.459$ ), el *L* ( $-0.196$ ) y el *P* ( $-0.136$ ), mientras que en el *M* hay tan sólo una observación. Con el fin de incluir a este último y en general, debido al escaso tamaño muestral en cada gen, puede ser conveniente analizar conjuntamente las mutaciones aleatorias y las previamente descritas. Aun así, sigue existiendo heterogeneidad de los efectos mutacionales entre genes ( $\chi^2 = 12.600$ ,  $p = 0.013$ ), siendo el gen *N* de nuevo el que más restricción funcional muestra, con un efecto mediano de  $-0.622$ , seguido por el *G* ( $-0.462$ ) y el *L* ( $-0.406$ ), mientras que el *M* ( $-0.096$ ) el y el *P* ( $-0.074$ ) muestran la menor restricción funcional.

A partir de las clases establecidas en la Tabla 3, es posible reelaborar una clasificación de las mutaciones según si son por un lado, deletéreas o letales y por otro lado, neutrales o beneficiosas. El estudio de la proporción de cada uno de estas dos categorías en los diferentes genes nos da una idea de la restricción funcional a la que están sometidos. Como puede apreciarse en la

Tabla 7, en los genes *L*, *G* y *N* las mutaciones tienden a ser perjudiciales con más frecuencia que en los genes *P* y *M*, siendo las diferencias entre genes significativas ( $\chi^2 = 16.186$ ,  $p = 0.006$ ).

	Letal + deletérea	Neutral + beneficiosa
<b>N</b>	2 + 4	2 + 3
<b>P</b>	1 + 1	8 + 1
<b>M</b>	3 + 2	5 + 2
<b>G</b>	6 + 10	4 + 1
<b>L</b>	12 + 14	7 + 1

**Tabla 7.** Número de cambios letales o deletéreos por un lado y neutrales o beneficiosos por otro lado, en los cinco genes de VSV. En la primera columna, se indica el número de mutaciones letales seguido del número de deletéreas, mientras que en la segunda columna se hace esto mismo con las neutrales y beneficiosas, en este orden.

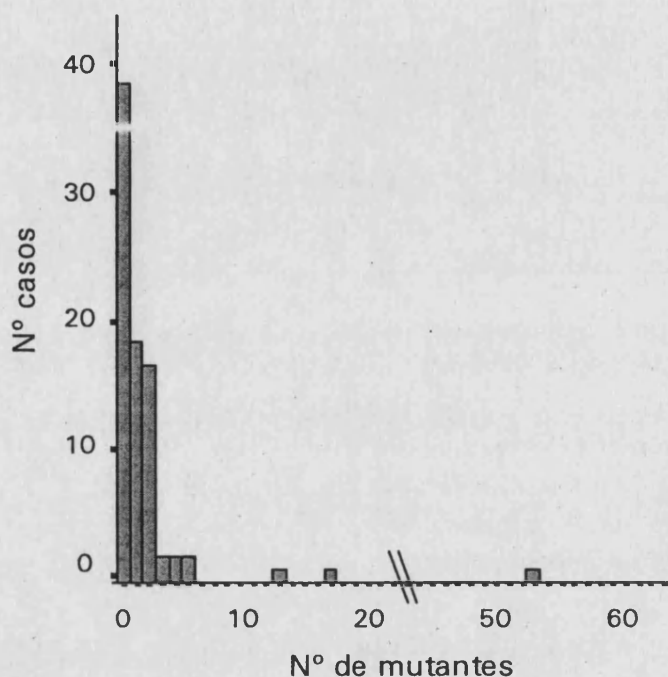
## TASA DE MUTACIÓN

### Prueba de fluctuación de Luria-Delbrück

Para estimar la tasa de mutación del fenotipo silvestre, hemos llevado a cabo 6 pruebas de fluctuación de Luria-Delbrück independientes usando el anticuerpo AcM I<sub>1</sub> como agente selectivo. Para cada uno de estos ensayos, se infectan 84 pocillos y se averigua el número de mutantes MARM que aparecen por pocillo, obteniéndose así 6 distribuciones observadas del número de mutantes, una de las cuales se muestra en la Figura 14.

Si se pretende utilizar la frecuencia total de mutantes aparecidos en cada prueba de fluctuación o incluso la propia distribución observada para obtener una estimación de la tasa de mutación en este carácter, es necesario tener conocimiento del modo de replicación de VSV. Si el modo de replicación fuera binario, podríamos utilizar la ecuación 42 para estimar la tasa de mutación, o incluso ajustar la distribución observada del número de mutantes a la

distribución de Luria-Delbrück. Sin embargo, es bien sabido que VSV no replica de manera binaria, sino que cada cadena molde es copiada múltiples veces.



**Figura 14.** Distribución representativa del número de mutantes MARM aparecidos en una de las pruebas de fluctuación de Luria-Delbrück.

En el extremo opuesto, si en cada ciclo de infección la cadena que entra en la célula hospedadora fuera el único molde utilizado en la replicación, de manera que todas las partículas liberadas en un ciclo son derivadas directamente de ésta (replicación lineal), la tasa de mutación podría ser obtenida mediante la ecuación 41 o también ajustando la distribución observada del número de mutantes a una Poisson. Este modelo es en principio más realista que el de la replicación binaria. En cualquier caso, podemos evaluar su validez comprobando si la distribución Poisson se ajusta realmente a la distribución observada. Tal y como se puede observar en la Tabla 8, el modelo

Poisson es incompatible con los datos en cuatro de los seis casos, por lo que difícilmente puede asumirse un modo de replicación lineal.

Nº de mutantes	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5	Prueba 6
0	31	40	50	32	39	39
1	22	22	19	23	27	19
2	19	12	9	17	9	17
3	5	6	3	5	2	2
4	3	0	2	3	1	2
5	2	1	0	2	0	2
6	1	0	0	1	1	0
>6	1	3	1	1	5	3
	(14)	(17, 27,100)	(17)	(17)	(8, 16,31,36 , 104)	(12, 16, 53)
<i>p</i> -valor	0.173	<0.001	0.020	0.104	<0.001	<0.001

**Tabla 8.** Distribución observada (número de casos) de la cantidad de mutantes en cada una de las 6 pruebas de Luria-Delbrück. Cuanto ésta es mayor que 6, los valores entre paréntesis indican la cantidad de mutantes obtenidos. Se muestra también la probabilidad de que la distribución sea Poisson, de acuerdo con una prueba de Kolmogorov-Smirnov.

El método de la clase nula desaprovecha gran parte de información proporcionada por la prueba de fluctuación, ya que sólo utiliza la frecuencia con la que no aparece mutante alguno. Sin embargo, tiene la ventaja fundamental de no hacer asunciones acerca del modo de replicación, por lo que a la vista de los resultados obtenidos, es el más adecuado para estimar la tasa de mutación. Utilizando la ecuación 39, obtenemos  $\mu = (2.473 \pm 0.595) \times 10^{-5}$  mutaciones por *locus* MARM y generación, con lo cual la tasa por sitio nucleotídico es  $\mu = (1.200 \pm 0.298) \times 10^{-5}$  y la tasa de mutación genómica  $U = 0.138 \pm 0.033$ .

## EPISTASIAS ENTRE PARES DE SUSTITUCIONES NUCLEOTÍDICAS

### Combinaciones de mutaciones aleatorias

Combinando al azar las 29 mutaciones aleatorias viables, hemos construido 47 cDNAs que incorporan dos de esas mutaciones, con el fin de poder comparar la eficacia biológica observada en cada caso con la esperada en ausencia de epistasia (modelo multiplicativo).

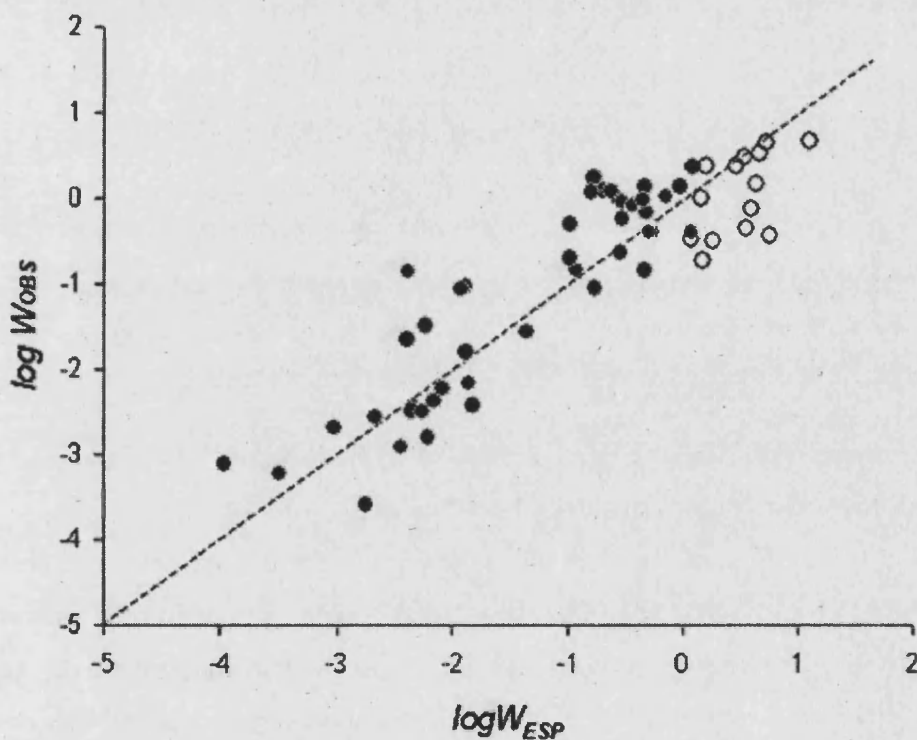
Hemos podido recuperar partículas infecciosas a partir de 44 de los 47 cDNAs tras cinco intentos de transfección. A lo largo de esos intentos, el cDNA correspondiente al genotipo silvestre ha sido transfectado 28 veces, obteniéndose partículas infecciosas en 23 casos, por lo que la tasa de transfección fallida es igual a 0.179. Así pues, la probabilidad de no obtener partículas tras diez intentos es  $0.179^5 = 1.8 \times 10^{-4}$  por lo que en una muestra de 47 genotipos, será  $1.8 \times 10^{-4} \times 47 = 0.009$ . Por consiguiente y de igual manera que en el caso de los mutantes simples, los tres mutantes dobles para los que no hemos podido recuperar partículas infecciosas son candidatos a ser genotipos letales. En esta ocasión, sin embargo, la presencia de posibles letales es más llamativa, ya que las dos mutaciones son viables por separado. A esta clase de genotipos, que son un tipo particular de letales condicionales, se les denomina letales sintéticos, y representan un caso extremo de epistasia negativa.

En la Figura 15 se muestra la eficacia biológica observada para cada mutante doble frente al valor esperado en ausencia de epistasia, mientras que en el Apéndice 4 se proporciona información más detallada acerca de cada genotipo, así como la estimación del efecto epistático en cada caso, obtenido a partir de la ecuación 3. Tras aplicar pruebas *t* para muestras independientes, encontramos evidencia de epistasia positiva en 21 casos, mientras que ésta es significativamente negativa en cuatro casos (excluyendo los tres posibles letales sintéticos). Aplicando la corrección de Bonferroni para pruebas múltiples de la misma hipótesis, tan sólo permanecen significativos 7 casos, siendo en todos ellos la epistasia de signo positivo.

El efecto medio epistático por generación es  $\bar{\epsilon} = 0.147 \pm 0.036$ , valor que es significativamente positivo ( $t_{43} = 4.109$ ,  $p < 0.001$ ). Podría no obstante argumentarse que, dado que las eficacias biológicas están definidas en el

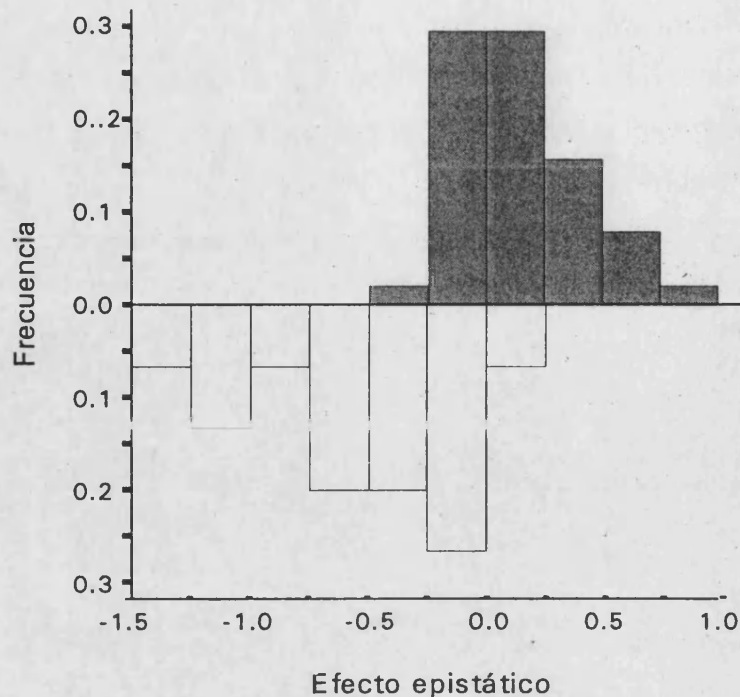
intervalo  $[0, +\infty[$ , fluctuaciones puramente aleatorias en la intensidad de la epistasia pueden traducirse en una distribución asimétrica de los valores de  $\varepsilon$ , como de hecho ocurre ( $g_1 = 2.695$ ;  $t_{43} = 2.694$ ,  $p = 0.005$ ) tal y como puede apreciarse en la Figura 16, dando la falsa impresión de una epistasia predominantemente positiva. Este problema puede solucionarse sencillamente si recalculamos la epistasia como

$$\varepsilon'_{ij} = \log W_{ij} - \log(W_i W_j) \quad (49)$$



**Figura 15.** Relación entre la eficacia observada de los mutantes dobles y su eficacia esperada en ausencia de epistasia, en escala logarítmica. En negro se representan las eficacias de los genotipos que combinan dos mutaciones deletéreas, mientras que en blanco se representan las eficacias de aquellos que combinan dos mutaciones beneficiosas.

tras lo cual obtenemos  $\bar{\varepsilon}' = 0.227 \pm 0.077$ , valor que es también claramente positivo ( $t_{43} = 2.891$ ,  $p = 0.006$ ), siendo ahora su distribución efectivamente simétrica ( $g_1 = 0.908$ ;  $t_{43} = 0.381$ ,  $p = 0.353$ ).



**Figura 16.** Histograma de frecuencias para el efecto epistático entre pares de mutaciones aleatorias (gris) y entre pares de mutaciones beneficiosas (blanco).

#### **Epistasis entre pares de mutaciones beneficiosas**

Hemos realizado todas las combinaciones entre las seis mutaciones simples de efecto más beneficioso, todas ellas previamente descritas, obteniendo así 15 mutantes dobles que, según el modelo multiplicativo, deben ser beneficiosos. En esta ocasión, hemos podido recuperar partículas infecciosas a partir de los 15 cDNAs, por lo que no hay caso alguno de letal sintético.

En la Figura 15 se muestra la eficacia biológica observada para cada mutante doble en función del valor esperado en ausencia de epistasia, mientras que en el Apéndice 5 se proporciona información más detallada acerca de los datos para cada mutante. En 14 de los 15 casos, el signo del efecto epistático es negativo, habiendo evidencia estadística de epistasia negativa en 9 casos (8 de ellos permanecen significativos tras aplicar la corrección de Bonferroni). Llama la atención además, que en esos casos, el mutante doble es menos eficaz que cualquiera de los dos mutantes simples correspondientes, caso particular de epistasia al que se denomina epistasia descompensatoria.



El efecto medio epistático por generación para los pares de mutaciones beneficiosas es  $\bar{\varepsilon} = -0.537 \pm 0.120$ , valor que es significativamente negativo ( $t_{14} = 4.487$ ,  $p = 0.001$ ). En esta ocasión, tal y como puede observarse en la Figura 16, no hay evidencia de que la distribución de valores de  $\varepsilon$  sea asimétrica ( $g_1 = -0.425$ ;  $t_{14} = 0.733$ ,  $p = 0.238$ ) y en cualquier caso, el efecto medio epistático calculado tal y como se define en la ecuación 49 sigue siendo claramente negativo ( $t_{14} = 4.468$ ,  $p = 0.001$ ). Sin embargo, en esta ocasión podría argumentarse que los 15 genotipos carecen de la necesaria independencia estadística, ya que se trata de todas las combinaciones posibles de 6 mutaciones simples. El máximo número de valores independientes que pueden extraerse de las  $n(n-1)/2$  combinaciones posibles de dos elementos es  $n-1$ . Así pues, el problema puede ser resuelto remuestreando 5 de los 15 valores un gran número de veces, mediante la técnica del *bootstrap*. Obtenemos 1000 pseudo-muestras de tamaño 5 y sus correspondientes 1000 pseudo-valores del efecto medio epistático  $\bar{\varepsilon}$ . En los 1000 remuestreos, el efecto medio epistático sigue siendo negativo, por lo que claramente, podemos reafirmarnos en la conclusión de que la epistasia entre mutaciones de efecto beneficioso es predominantemente de signo negativo.

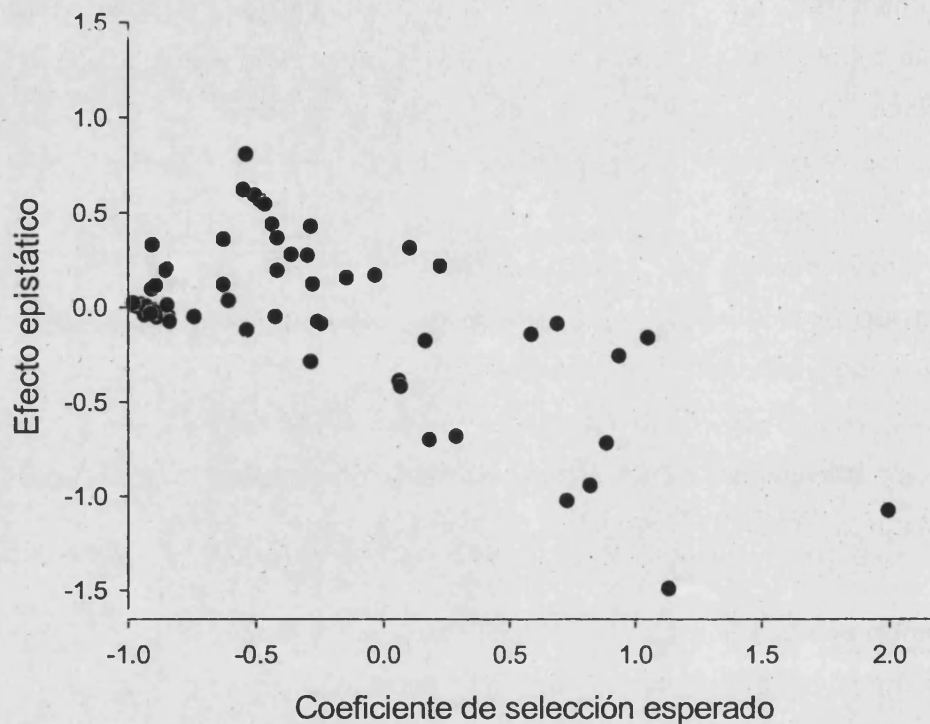
### Correlación entre el efecto epistático y el coeficiente de selección

Como hemos visto, el signo de la epistasia entre mutaciones deletéreas es en promedio positivo, mientras que entre mutaciones beneficiosas es negativo. Este resultado puede ser el reflejo de dos posibles situaciones. En primer lugar, es posible que el cambio de signo de la epistasia se produzca mediante una transición súbita, caracterizándose el genotipo silvestre por estar en el punto de inflexión entre la epistasia negativa y positiva. En segundo lugar, puede que exista un compromiso general entre el efecto de las mutaciones y el valor de la epistasia entre éstas, lo que se traduciría en una correlación negativa entre la eficacia esperada según el modelo multiplicativo y el valor del efecto epistático.

De acuerdo con la definición convencional del coeficiente de selección (ecuación 1) y según el modelo multiplicativo, el valor esperado del coeficiente de selección de un genotipo con las mutaciones  $i$  y  $j$  es

$$E[s_{ij}] = W_{ij} - 1 = s_i + s_j + s_i s_j = s_{ij} - \varepsilon_{ij}. \quad (50)$$

Como puede apreciarse en la Figura 17, hay una clara correlación negativa entre  $E[s]$  y el valor del efecto epistático  $\varepsilon$  (Spearman:  $r=-0.411$ ,  $p=0.001$ ), lo cual nos indica que el cambio de signo que experimenta la epistasia al pasar de mutaciones deletéreas a mutaciones beneficiosas no es más que una consecuencia del compromiso general que existe entre el efecto de las mutaciones y el valor de la epistasia entre éstas.



**Figura 17.** Correlación negativa entre el coeficiente de selección esperado en genotipos dobles mutantes (ecuación 50) y el efecto epistático.

## RECUPERACIÓN DE LA EFICACIA BIOLÓGICA

### Reversiones y evolución adaptativa

Hemos escogido seis mutantes simples y seis mutantes dobles, todos ellos con una eficacia biológica menor que la del silvestre, para llevar a cabo una dinámica de evolución experimental durante aproximadamente 20 generaciones. Cada uno de estos 12 mutantes, cuyas características se detallan en el Apéndice 6, ha sido evolucionado por triplicado a dos tamaños

poblacionales efectivos:  $N_e = 2 \times 10^2$  PFU y  $N_e = 2 \times 10^4$  PFU. Para conseguir hacer variar el tamaño efectivo, han de inocularse las infecciones con un número diferente de partículas iniciales ( $N_0 = 50$  y  $N_0 = 7 \times 10^3$  PFU, respectivamente, de acuerdo con la ecuación 37), siendo el número de generaciones transcurridas por pase  $g = 4.4$  y  $g = 3.1$ , respectivamente (ecuación 35). Es por ello que en el primer caso se han llevado a cabo cuatro infecciones seriadas para completar las 20 generaciones, mientras que en el segundo se han llevado a cabo 6 infecciones seriadas. Tras ese período de evolución experimental, hemos estimado la eficacia biológica de cada población  $i$ -ésima y hemos calculado su ganancia de eficacia como

$$qW_i = W_{i(20)} / W_{i(0)}, \quad (51)$$

donde  $W_{i(0)}$  y  $W_{i(20)}$  son las eficacias biológicas del genotipo ancestral y de la población evolucionada relativa al silvestre expresadas en generaciones, respectivamente.

Hemos comprobado en cada caso si tras la evolución, la mutación ancestral se ha mantenido o ha revertido al estado silvestre. De las 18 líneas derivadas de los seis mutantes simples a un tamaño poblacional  $N_e = 2 \times 10^2$  PFU, tan sólo hemos observado un caso de reversión incompleta, es decir un polimorfismo entre la mutación y el estado silvestre, mientras que para  $N_e = 2 \times 10^4$  PFU, hemos observado cuatro casos en los que la mutación ancestral ha cambiado. Análogamente, de las 18 líneas derivadas de los 6 mutantes dobles a un tamaño poblacional  $N_e = 2 \times 10^2$  PFU, en tan sólo un solo caso una de las dos mutaciones ha sufrido una reversión polimórfica, mientras para  $N_e = 2 \times 10^4$  PFU, se han producido seis reversiones (tres polimórficas y tres fijadas). Estos resultados nos permiten concluir que a lo largo del tiempo de evolución estudiado, las reversiones son poco frecuentes bajo un tamaño poblacional efectivo de  $2 \times 10^2$  PFU (5% de los casos) y que bajo un tamaño  $2 \times 10^4$ , aunque la reversión sigue sin ser mayoritaria (28% de los casos), su frecuencia se ve incrementada ( $\chi^2 = 6.400$ ,  $p = 0.011$ ).

El aumento medio de eficacia biológica relativa de los genotipos revertientes respecto a sus ancestros es  $q_w = 6.887 \pm 0.378$  para  $N_e = 2 \times 10^2$  PFU y  $q_w = 9.073 \pm 0.235$  para  $N_e = 2 \times 10^4$  PFU, incremento que ha de ser consecuencia tanto de las propias reversiones como de la aparición de otro tipo de mutaciones. De hecho, para  $N_e = 2 \times 10^4$  PFU, los genotipos que no han

sufrido reversiones experimentan también un aumento de la eficacia biológica ( $q_w = 7.163 \pm 0.137$ ) aunque como es de esperar, éste es menos intenso que en el primer caso (prueba de Mann-Whitney de una cola:  $U = 5.700$ ,  $p < 0.001$ ). En el caso  $N_e = 2 \times 10^2$  PFU, los genotipos que no han sufrido reversión experimentan también un aumento de la eficacia ( $q_w = 6.576 \pm 0.096$ ), aunque no parece haber diferencias con el grupo de los genotipos revertientes (prueba de Mann-Whitney de una cola:  $U = 0.993$ ,  $p = 0.161$ ). Estos resultados nos permiten concluir que, a lo largo del tiempo de evolución estudiado, las reversiones se ven acompañadas de mutaciones beneficiosas adicionales, pero no nos permiten afirmar nada respecto a la posible existencia de mutaciones compensatorias.

### **Evidencia de evolución compensatoria**

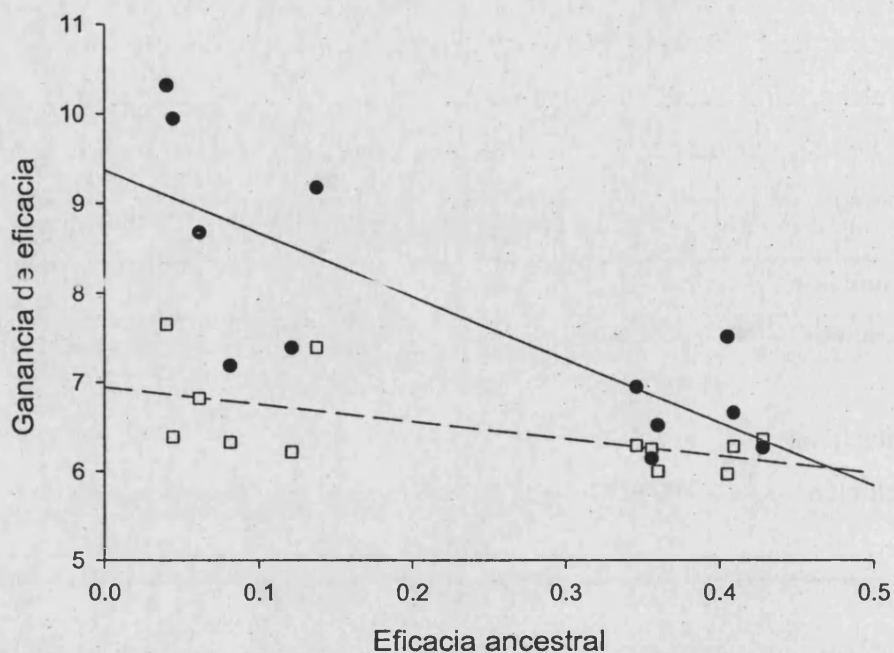
Las mutaciones compensatorias son aquellas que sin ser reversiones, permiten paliar específicamente el efecto de las mutaciones deletéreas total o parcialmente. La ganancia en eficacia biológica que han sufrido las poblaciones a lo largo del experimento de evolución puede haber sido determinada por varios factores. En primer lugar, podemos distinguir entre los genotipos que partían con una mutación de los que partían con dos. En segundo lugar, cada uno de los mutantes simples y dobles puede ver su evolución condicionada en función de la mutación concreta introducida. En tercer lugar, cada genotipo es evolucionado bajo dos tamaños poblacionales distintos, esperándose pues ganancias de eficacia distintas. Por último, para cada genotipo y tamaño poblacional, han sido llevadas a cabo tres líneas de evolución independientes, el resultado de las cuales puede variar por efectos aleatorios. Podemos determinar la influencia de cada uno de estos factores sobre la ganancia eficacia biológica observada mediante un análisis de la varianza (ANOVA) encajado, que se muestra en la Tabla 9. Hemos excluido de este y los subsiguientes análisis todos los casos de reversión, puesto que estamos ahora interesados en la posible existencia de mutaciones compensatorias. Los resultados del ANOVA indican primeramente que un mismo genotipo evolucionado bajo un mismo tamaño poblacional experimenta ganancias de eficacia significativamente distintas en las diferentes líneas, resultado esperable teniendo en cuenta que la aparición de cambios beneficiosos es un fenómeno estocástico y raro. En segundo lugar, a mayor  $N_e$ , mayor es la probabilidad de aparición de

mutaciones beneficiosas, por lo que no es de extrañar que hayamos detectado en el ANOVA un efecto significativo de este factor. Más interesante resulta observar que la ganancia de eficacia está fuertemente determinada por el genotipo de partida. Las diferencias son notorias no sólo entre mutantes simples y dobles, sino también entre los distintos genotipos de cada grupo. Puesto que hemos excluido del análisis todos los casos de reversión, estos resultados ponen

Fuente de variación	SS	g.l.	F	p
Nº de mutaciones	19.144	1	15.231	< 0.001
Genotipo	99.048	10	7.880	< 0.001
Tamaño poblacional	44.398	11	3.211	0.004
Línea de evolución	46.508	37	4.987	< 0.001
Error	30.244	120		

**Tabla 9.** ANOVA anidado de la ganancia de eficacia. Cada factor está anidado en el inmediatamente superior. Todos los factores son fijos a excepción de la línea de evolución. Se muestra la suma de cuadrados residual, los grados de libertad asociados, el valor del estadístico  $F$  y la significación estadística.

de manifiesto que las mutaciones de efecto beneficioso que han aparecido durante la evolución interaccionan con las mutaciones deletéreas ancestrales. Pueden ser múltiples las propiedades de los genotipos ancestrales que están influyendo sobre su capacidad de adaptación, siendo una de ellas su eficacia biológica ancestral. Para explorar esta posibilidad, sustituimos el factor "genotipo" en el ANOVA de la Tabla 9 por su eficacia ancestral, que entra ahora como covariable en el análisis. Su efecto es altamente significativo ( $F_{1,167} = 47.652$ ,  $p < 0.001$ ), lo que demuestra que la eficacia ancestral está determinando en gran medida el potencial adaptativo de los genotipos. De hecho, a menor eficacia inicial, mayor es la ganancia de eficacia (correlación parcial controlando para  $N_e$ :  $r = -0.442$ ,  $p < 0.001$ ), tal y como se muestra en la Figura 18. De la presencia de mutaciones beneficiosas, la ausencia de reversiones, la presencia de interacciones epistáticas y la correlación negativa entre la eficacia ancestral y la ganancia de eficacia se deduce la existencia de mutaciones compensatorias.

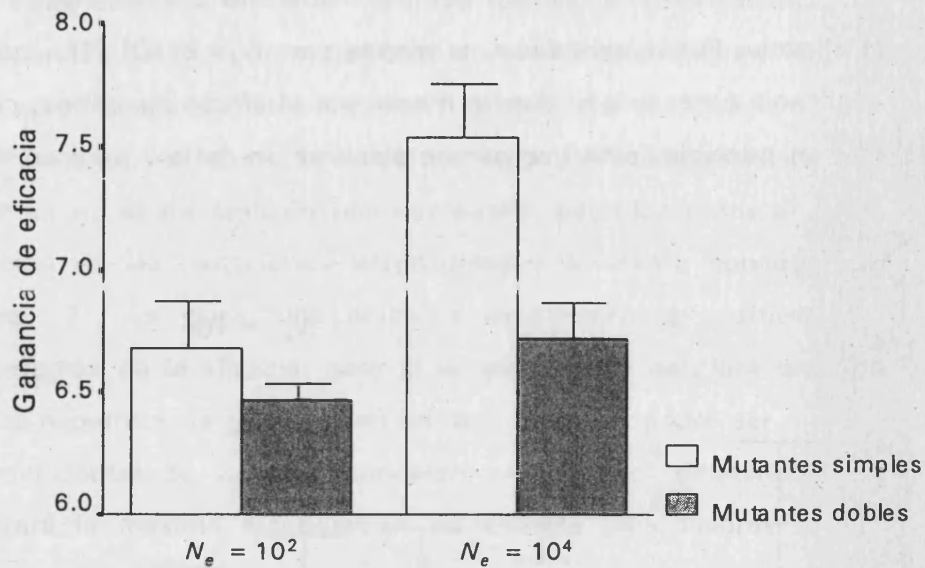


**Figura 18.** Correlación negativa entre la eficacia ancestral y la ganancia de eficacia, para la cual se muestra la media de las tres líneas de evolución. Los círculos negros y la línea de regresión continua corresponden al tamaño poblacional  $N_e = 2 \times 10^4$  PFU, mientras que los cuadrados blancos y la línea de regresión discontinua corresponden a  $N_e = 2 \times 10^2$  PFU.

### Compensación de mutaciones simples y dobles

De acuerdo con los resultados mostrados en la Figura 18 y dado que hemos hecho un esfuerzo para que los grupos de mutantes simples y dobles tengan aproximadamente la misma eficacia inicial, sería de esperar que en promedio los mutantes dobles experimentasen la misma recuperación de eficacia que los simples. Sin embargo, como demuestra el ANOVA de la Tabla 9, esto no es así y en realidad los mutantes dobles recuperan eficacia más lentamente que los simples, tal y como puede observarse en la Figura 19. Estos resultados son esperables si durante el experimento de evolución no ha transcurrido el tiempo suficiente para que puedan compensarse las dos mutaciones ancestrales.

Puesto que gran parte de los genotipos dobles mutantes tan sólo habrá podido compensar una de sus mutaciones durante el tiempo de evolución

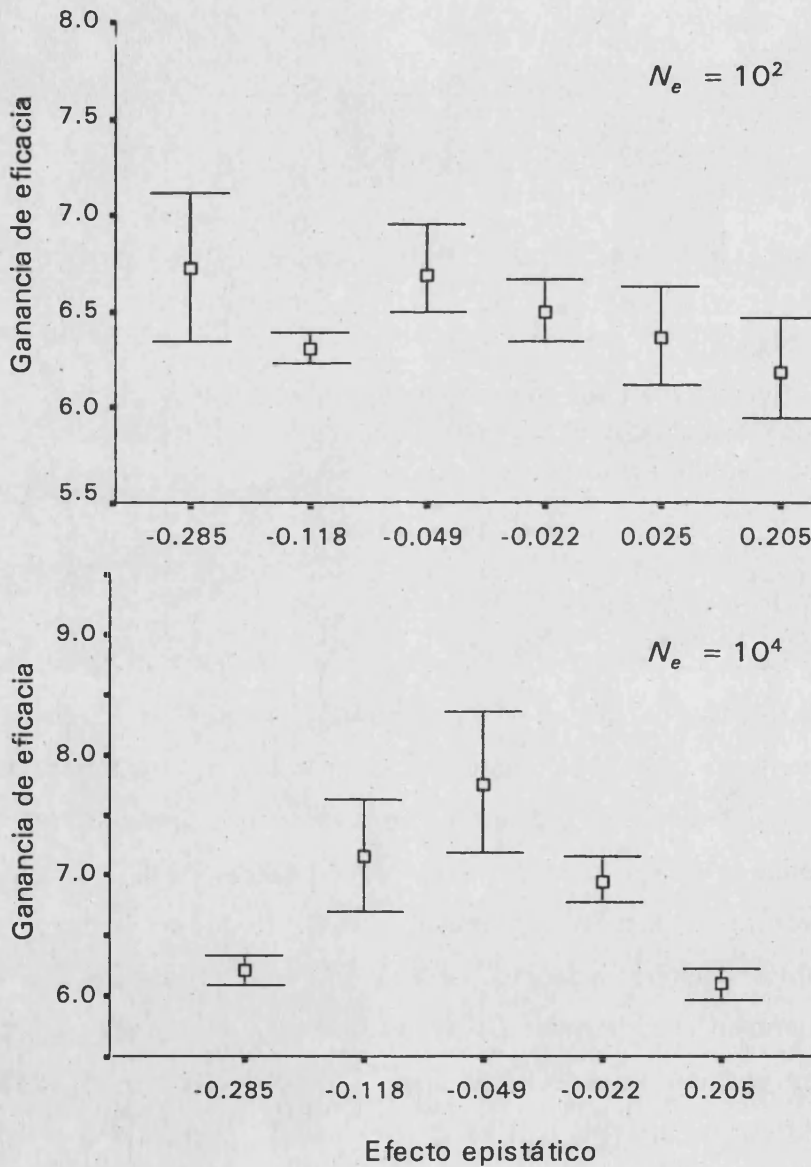


**Figura 19.** Recuperación de eficacia en mutantes simples y dobles bajo los dos tamaños poblacionales explorados. Se muestra el valor medio y su error asociado.

experimental, esperamos que los efectos epistáticos entre esas mutaciones ancestrales tengan una influencia sobre la recuperación de eficacia observada. Si la epistasia entre ellas es negativa, la compensación de una de las dos supondrá un aumento de la eficacia mayor que en ausencia interacción alguna, ya que no sólo se reducirá el efecto deletéreo de la mutación, sino que también el término epistático. A su vez, en ausencia de interacción esperamos una recuperación mayor que en el caso de epistasia positiva, ya que en éste último caso la compensación tenderá a eliminar el efecto deletéreo de la mutación, pero también el efecto epistático beneficioso. Para contrastar esta hipótesis, podemos reemplazar el factor "genotipo" en el ANOVA mostrado en la Tabla 9 no sólo por su eficacia ancestral, como hemos hecho previamente, sino también por el valor correspondiente del efecto epistático ( $\epsilon$ ). Dicho análisis, realizado lógicamente sólo con los mutantes dobles, demuestra que además de la eficacia ancestral ( $F_{1,79} = 4.136$ ,  $p = 0.001$ ), la epistasia también tiene una influencia sobre la ganancia de eficacia observada tras la evolución ( $F_{1,79} = 2.188$ ,  $p = 0.016$ ). De hecho, cuanto más negativa es la epistasia inicial,

mayor es la ganancia de eficacia (correlación parcial controlando para el tamaño poblacional y la eficacia inicial,  $r = -0.311$ ,  $p = 0.004$ ).

La Figura 20 muestra la relación que hay entre el efecto epistático y la ganancia de eficacia. Puede apreciarse, al menos para  $N_e = 2 \times 10^4$  PFU, que la ganancia de eficacia no es mayor cuanto menor sea el efecto epistático, como cabría esperar en principio, sino que parece alcanzar un óptimo para un valor



**Figura 20.** Influencia sobre la recuperación de la eficacia de la epistasia entre las mutaciones introducidas. Se muestran los valores medios y su error asociado. En el caso  $N_e = 10^4$  hay una categoría ausente ya que en las tres líneas de evolución al menos una de las dos mutaciones ha revertido.



intermedio. Este patrón no es más que la consecuencia de nuestros resultados precedentes. En primer lugar, existe una correlación negativa entre el coeficiente de selección de la mutación introducida y la ganancia de eficacia, reflejo de la existencia de mutaciones compensatorias (Figura 18). En segundo lugar como acabamos de mencionar, la ganancia de eficacia se correlaciona negativamente con el efecto epistático entre las mutaciones introducidas. En tercer lugar, existe también una correlación negativa entre el coeficiente de selección de las mutaciones introducidas y el efecto epistático entre ellas (Figura 17). Así pues, una epistasia excesivamente positiva dificultará la recuperación de la eficacia, pero si la epistasia se desplaza en exceso hacia valores negativos, la ganancia en eficacia tampoco podrá ser elevada, ya que los coeficientes de selección tenderán a ser menos deletéreos. Por ello, se alcanzará la máxima recuperación de eficacia para valores intermedios de epistasia.



## **DISCUSIÓN**



## MAGNITUD DE LAS MUTACIONES DELETÉREAS

### Robustez genética de los virus de RNA

Los virus de RNA se distinguen de otros sistemas biológicos por tener una tolerancia muy limitada a la mutación (Domingo *et al.*, 1996). Esta característica fue puesta de manifiesto por primera en el bacteriófago  $\phi 6$  (Chao, 1990) que, tras infecciones a un tamaño poblacional efectivo muy reducido (pases placa-placa), perdía un 22% de su eficacia como consecuencia de la acción del trinquete de Muller. Pero este no es un caso único: FMDV vio reducida su eficacia biológica en más de un 35% tras ser sometido a 30 cuellos de botella consecutivos mediante pases placa-placa (Escarmis *et al.*, 1996). Curiosamente, esas poblaciones debilitadas no fijaron en promedio más de 2 ó 3 mutaciones durante ese período, lo que sugiere que cada mutación debería tener un efecto deletéreo considerable. Especialmente llamativo es el caso de HIV-1, que tras sólo 15 pases placa-placa perdió un 82% de su eficacia (Yuste *et al.*, 1999), aunque en este caso el número de mutaciones fijadas por genoma fue mayor (Yuste *et al.*, 2000).

La respuesta de VSV a las dinámicas poblacionales con cuellos de botella es la habitual entre los virus de RNA. Una dinámica de 20 pases placa-placa mermó su eficacia biológica en aproximadamente un 38 % (Duarte *et al.*, 1992; Clarke *et al.*, 1993; Elena y Moya, 1999). Nuestras estimaciones van más allá y nos permiten proponer que en promedio, una sola mutación reduce en un 55% la eficacia de VSV, es decir que su coeficiente de selección vale  $s=0.555$ . Una parte sustancial de esa reducción es debida a las mutaciones letales, que constituyen hasta un 40% del total. Para entender la diferencia entre nuestros resultados y los obtenidos previamente en VSV y otros virus de RNA, hay que tener en cuenta que en los experimentos de acumulación de mutaciones, hay dos posibles fuentes de error que conducen a subestimar el efecto de las mutaciones. En primer lugar, se tiene la seguridad de estar ignorando los genotipos letales, puesto que estos no formarán aros de lisis visibles. En segundo lugar, por mucho que se reduzca el tamaño poblacional efectivo, no se tiene garantía de que haya ausencia total de selección durante el crecimiento del aro de lisis (Kibota y Lynch, 1996). De hecho, para un mismo tamaño poblacional, cuanto mayores sean los coeficientes de selección, mayor



será la acción de la selección en relación a la deriva y por lo tanto, mayor será la subestima del efecto real de las mutaciones.

En cualquier caso, no hay duda de que la sensibilidad de los virus de RNA a la mutación es mayor que la de cualquier organismo. Usando un método de mutagénesis por transposición, Elena *et al.* (1998b) crearon una colección de 226 mutantes de *E. coli*, cada uno de ellos portador de una inserción aleatoria del transposón mini-Tn10. Los experimentos de competencia indicaron que, aunque al menos el 80% de las mutaciones tienen un efecto deletéreo, su coeficiente de selección promedio es aproximadamente  $s=0.031$ . Mediante un experimento de acumulación de mutaciones en este mismo organismo, Kibota y Lynch (1996) habían obtenido previamente un valor máximo de  $s=0.012$ , algo menor que el obtenido por Elena *et al.*, lo cual no es de extrañar teniendo en cuenta que Elena *et al.* trabajaron sólo con inserciones. Experimentos de inactivación sistemática de genes en *S. cerevisiae* mediante la inserción aleatoria del elemento transponible Ty1 han concluido que casi el 40% de las mutaciones tienen un efecto no detectable en la tasa de crecimiento en diferentes ambientes (Smith *et al.*, 1997), mientras que la reducción de la eficacia sólo era de superior al 25% en el 17% de los casos, algo impensable para un virus de RNA. En el hongo *A. nidulans* en fase asexual, se ha estimado mediante experimentos de acumulación de mutaciones que el coeficiente de selección medio en contra de las mutaciones deletéreas es como máximo  $s=0.0069$  (Bruggeman *et al.* 2003). Keightley y Caballero (1997),

Grupo	Especie	s
Ribovirus	VSV	0.555
	Polio tipo 1	0.230
	$\phi 6$	0.054
Procariotas	<i>E. coli</i>	0.0120
		0.0310
Eucariotas	<i>A. nidulans</i>	0.0069
	<i>D.melanogaster</i>	0.007

Tabla 10. Coeficiente de selección estimado para diferentes sistemas biológicos. Hay que recordar que éstos han sido obtenidos por métodos muy dispares.

llevaron a cabo un experimento de acumulación de mutaciones en líneas paralelas durante 60 generaciones en *C. elegans*. A pesar de que la varianza genética entre líneas aumentó, como es de esperar en presencia de deriva genética, la eficacia promedio no descendió significativamente. Este resultado confirma que o bien los efectos mutacionales o bien la tasa de mutación genómica en este organismo son bastante menores que en virus de RNA. El coeficiente de selección promedio en *D. melanogaster* está entorno a 0.003-0.030 (Mukai, 1972; Ohnishi, 1974; Keightley, 1994) y la fracción de letales es aproximadamente el 1% (Mukai, 1964). En general, en eucariotas multicelulares, muchas mutaciones que inactivan genes muestran un débil o nulo efecto fenotípico (Wang *et al.*, 1996). La Tabla 10 resume los valores del coeficiente de selección estimados en los diferentes estudios mencionados.

#### **Patrones de robustez genética**

La redundancia genética es quizás la manera más directa de alcanzar la robustez genética. A su vez, la forma más trivial de redundancia es la existencia de varias copias de la información genética, sea mediante di- o poliploidía o mediante duplicaciones de genes. Sin embargo, parece que no existe una correlación positiva entre las semejanza filogenética o funcional de los genes duplicados y la robustez genética, medida como el efecto que las mutaciones en esos genes tienen sobre las funciones que desempeñan. Esto es una indicación de que la redundancia genética se alcanza preferentemente mediante el establecimiento de una red de interacciones entre genes y no la mera duplicación de éstos (Wagner, 2000).

La redundancia y la robustez genéticas crecen con la complejidad de los organismos, por lo que es típica de los eucariotas. Por el contrario, la anti-redundancia debería ser más típica de las bacterias y los virus (Krakauer y Plotkin, 2000). Algunos ejemplos de las características propias de esta anti-redundancia son la haploidía, las pautas de lectura solapadas, los elementos de regulación de la expresión compartidos, (Krakauer, 2000), el sesgo en el uso de codones (Berg y Silva, 1997), los cuellos de botella en la transmisión (Bergstrom y Pritchard, 1998), las altas tasas de replicación (Lenski *et al.*, 1999) o la ausencia de mecanismos de reparación del material genético (Steinhauer *et al.*, 1992). Todas estas características han sido puestas de

manifiesto en los virus de RNA (Holland *et al.*, 1982, Steinhauer *et al.*, 1992; Krakauer, 2000; Krakauer y Plotkin, 2002; Jenkins y Holmes, 2003).

La relación entre complejidad y robustez viene apoyada en primer lugar por el análisis del patrón de sustituciones sinónimas y no sinónimas en eucariotas superiores, que permite demostrar la existencia de una correlación negativa entre el tiempo de generación y la intensidad de la selección. Aunque este resultado puede atribuirse a que los organismos con mayor tiempo de generación tienden a tener tamaños poblacionales más pequeños, aumentando pues la acción de la deriva genética en relación a la selección (Keightley y Eyre-Walker, 2000), también puede interpretarse como el reflejo de una correlación positiva entre la robustez genética y la complejidad de los organismos. En segundo lugar, experimentos con organismos digitales han concluido que los genomas más reducidos y sencillos muestran una mayor sensibilidad a la mutación que los genomas grandes y complejos (Lenski *et al.*, 1999; Wilke y Adami, 2001).

#### **Determinación de los coeficientes de selección**

Como hemos visto, el coeficiente de selección es la medida más habitual de robustez genética. Puesto que el principal interés de la determinación de la robustez es poder realizar comparaciones entre diferentes especies, se requiere que los coeficientes de selección estén no sólo correctamente estimados, sino expresados de la manera más general posible. Sin embargo por desgracia, este no suele ser el caso. En primer lugar, en la mayoría de ocasiones, éstos han sido derivados de experimentos de acumulación de mutaciones en los que no se conoce el número ni la naturaleza de las mutaciones fijadas y además, se tiende a subestimar el valor real de los coeficientes de selección, tanto porque se ignoran las mutaciones letales como porque no se puede garantizar la ausencia total de selección. Estos inconvenientes pueden subsanarse mediante el uso de la mutagénesis dirigida. En segundo lugar, es habitual encontrar diferentes medidas de eficacia biológica, fundamentalmente porque éstas son estimadas en unidades de tiempo absoluto. Dada la enorme variabilidad en el tiempo de generación entre diferentes taxones, las unidades empleadas son dispares, dificultando pues las comparaciones. Por ello, sería apropiado expresar las eficacias biológicas de las diferentes especies por generación y no por unidad de tiempo absoluto. En tercer lugar, existen al menos dos maneras de calcular



el coeficiente de selección a partir de la eficacia biológica. La más habitual es  $s = W - 1$  (ecuación 1), pero también existe la versión  $s = \log W$  (ecuación 2). Ambas variantes convergen para efectos mutacionales pequeños, pero divergen fuertemente cuando éstos crecen. La ventaja de la primera reside justamente en su uso habitual no ya en experimentos de evolución experimental, sino en los propios modelos de Genética de Poblaciones. Sin embargo, presenta el serio inconveniente de ser asimétrico respecto al genotipo de referencia silvestre en su intervalo de definición  $[-1, +\infty[$ , por lo que la comparación de los efectos beneficiosos y deletéreos se ve distorsionada.

### Distribución de efectos mutacionales

Si se pretende tener un conocimiento real de los efectos mutacionales, hay que ir más allá de la mera estimación del valor medio del coeficiente de selección y caracterizar su distribución. Aunque este es un objetivo que requiere un gran esfuerzo experimental, existen precedentes en sistemas biológicos como *C. elegans* (Keightley y Caballero, 1997; Keightley y Bataillon, 2000), *D. melanogaster* (Seager y Ayala, 1982; Keightley, 1994, 1996; García-dorado, 1997) *E. coli* (Elena *et al.*, 1998b) y los propios virus de RNA (Elena y Moya, 1999; Lázaro *et al.* 2003). La conclusión general que puede extraerse de estos estudios es que la mayoría de mutaciones tienen un efecto pequeño sobre la eficacia biológica, aunque la distribución presenta una larga cola correspondiente a valores poco frecuentes pero extremadamente deletéreos. Debido a esta última característica, la distribución Exponencial, que es el modelo razonable más sencillo para describir la distribución de efectos (Mukai, 1972), no suele resultar totalmente satisfactorio. Habitualmente, puede ser mejorada mediante generalizaciones como son la Gamma (Keightley, 1994) tal y como propuso Kimura (1979), la Weibull (Lázaro *et al.*, 2003) o incluso mediante modelos compuestos por combinaciones lineales de más de una distribución (Elena *et al.*, 1998b; Elena y Moya, 1999). Una vez más, hemos de decir que las medidas del efecto mutacional empleadas en los diferentes estudios son heterogéneas. Nosotros nos hemos decantado por la definición Malthusiana del coeficiente de selección ya que, además de la ventaja arriba expuesta, permite un uso más correcto de la mayoría de los modelos probabilísticos, ya que éstos suelen estar definidos en el intervalo  $]-\infty, +\infty[$ .

De los modelos probabilísticos que hemos explorado, la distribución log-Normal es la que mejor describe el efecto deletéreo de las mutaciones aleatorias. Muchos procesos biológicos, como son los períodos de latencia en las enfermedades infecciosas, la sensibilidad de los microorganismos a los tratamientos antibióticos, los tiempos de supervivencia en medicina, la presencia de contaminantes en el aire o la abundancia de especies en ecología han sido descritos mediante modelos log-Normales (revisado en Limbert y Stahel, 2001). En general, esta distribución surge cuando la variable de estudio está determinada por efectos múltiples que actúan de manera multiplicativa. Recientemente, Lázaro *et al.* (2003) han mostrado que el patrón de fluctuaciones del título en poblaciones de FMDV bajo el equilibrio mutación-selección sigue una distribución log-Normal. Ese resultado puede ser la consecuencia de los múltiples factores celulares que determinan, de manera multiplicativa, la cantidad de partículas a que el virus puede dar lugar. Sin embargo, en ese sistema experimental dichos factores celulares no podían ser discernidos de los propios efectos mutacionales. Por el contrario, nuestros resultados permiten atribuir el patrón log-Normal a las mutaciones introducidas en el genoma viral. Los virus de RNA se caracterizan por tener genomas extremadamente compactos, de manera que una región genómica dada puede tener múltiples funciones, no sólo como mero soporte de la información genética, sino también como elemento regulador (revisado en Novella, 2003).

A la hora caracterizar el efecto de las mutaciones deletéreas previamente descritas, la distribución Exponencial es la más adecuada de acuerdo con nuestros resultados, aunque seguida de cerca por la Gamma. Esta familia de distribuciones es la que más habitualmente ha sido propuesta para describir los efectos mutacionales en los diferentes sistemas biológicos estudiados, como hemos mencionado. Además, la distribución Gamma ha sido utilizada como modelo de la variación en las tasas de sustitución por codón a lo largo de los genomas en virus de RNA (Yang *et al.*, 2000; Fares y Holmes, 2002). Esta forma de distribución, en la que las mutaciones de pequeño efecto son más frecuentes que en el caso de la log-Normal, puede ser una consecuencia de la acción de la selección negativa, simplemente porque la eficiencia de ésta para eliminar las mutaciones deletéreas de la población será mayor cuanto mayor sea el efecto de dichas mutaciones.

## Efectos deletéreos y mutaciones previamente descritas

Aunque como era de esperar, tanto el efecto mutacional medio como la proporción de letales difiere entre el grupo de mutaciones aleatorias y previamente descritas, llama la atención que el efecto promedio de estas últimas sea deletéreo, incluso con algunos casos de efecto letal. En realidad, esto no es de extrañar para aquellas mutaciones descritas en clones moleculares aislados, ya que cualquier población se encuentra en un equilibrio dinámico entre la entrada de nuevos variantes deletéreos por mutación y su salida debido a la acción de la selección purificadora. Además, algunas de esas mutaciones podrían haber permanecido ocultas a la selección por medio de la complementación genética, siempre que la multiplicidad de infección fuera lo bastante elevada para permitir este fenómeno (Schubert *et al.*, 1984; Wilke y Novella, 2003). Sin embargo, gran parte de los cambios que hemos introducido no fueron descritos en clones sino en secuencias consenso de poblaciones.

Novella *et al.* (1999) secuenciaron la mitad 3' del genoma consenso de poblaciones de VSV evolucionadas tanto en células de mamífero como de insecto, encontrando un total de 13 sustituciones nucleotídicas, dos de las cuales aparecieron en más de una ocasión de manera independiente en líneas evolutivas sometidas a ambientes diferentes. Es interesante reseñar que ambos cambios han resultado ser beneficiosos al ser recreados en nuestro sistema experimental (P120A y L123W, ambos en el gen *M*) y son por lo tanto buenos candidatos a mutaciones de efecto ventajoso generalista. En todas las líneas del mencionado experimento hemos encontrado al menos un cambio de efecto beneficioso pero por otro lado, todas ellas contienen al menos una mutación que muestra un efecto deletéreo en nuestro sistema experimental. Estas últimas podrían haber sido arrastradas hasta la fijación al encontrarse en el mismo genotipo donde se produjo la mutación beneficiosa, en un fenómeno descrito en poblaciones asexuales como autostop genético (Maynard-Smith y Haigh, 1974; Sniegowski *et al.*, 1997).

Cuevas *et al.* (2002) encontraron 25 mutaciones en secuencias consenso de 21 líneas independientes de evolución de VSV, muchas de las cuales ocurrieron recurrentemente en diferentes líneas. Todas las poblaciones fueron adaptadas al mismo tipo celular en el que hemos trabajado y mostraron ganancia neta en eficacia biológica durante el período de evolución. En al menos cuatro de esas poblaciones, todas las sustituciones halladas han

mostrado un efecto deletéreo al ser introducidas en nuestro sistema experimental e incluso una de ellas es, al parecer, letal. Esta aparente discrepancia de resultados puede ser debida a la existencia de fuertes epistasis e interacciones genotipo-ambiente (Elena y de Visser, 2003). Sin embargo, no se puede descartar que sea fruto de errores experimentales.

El hallazgo de las mencionadas discrepancias nos lleva a considerar una situación hasta ahora pasada por alto. Cabría la posibilidad de que durante la mutagénesis dirigida, la DNA polimerasa haya cometido errores indeseados. La probabilidad de que esto ocurra puede ser estimada modificando la ecuación 42 como

$$m = \mu \log 2^c, \quad (52)$$

donde  $m$  es la frecuencia de mutaciones,  $\mu$  la tasa de error por sitio nucleotídico y evento de replicación y  $c$  el número de ciclos de amplificación. La tasa de error de la DNA polimerasa *Pfu* está alrededor de  $10^{-6}$  cambios por ciclo de PCR y nucleótido (Bracho *et al.*, 1998; André *et al.*, 1997; Cline *et al.*, 1996; Lundberg *et al.*, 1991), mientras que el valor de  $c$  en nuestros experimentos de mutagénesis es de 18 ciclos, por lo que tenemos  $m = 10^{-6} \times \log 2^{18}$  es decir, una frecuencia en el plásmido amplificado de  $1.25 \times 10^{-5}$  cambios por nucleótido y por tanto, de 0.14 cambios por genoma. Esto significa que el 87% de los mutantes serán efectivamente simples, mientras que el 13% restante podrían contener al menos un cambio indeseado. Evidentemente, sería apropiado descartar estos últimos, pero ello implicaría secuenciar la totalidad de los 91 cDNAs generados, lo cual hubiera desbordado nuestros objetivos experimentales. Nuestros resultados (Tabla 3) sin tener en cuenta este inconveniente rendían una proporción de letales entre las mutaciones previamente descritas del 11.6%, valor muy cercano al porcentaje de genotipos que podrían incorporar un cambio indeseado. Así pues, sería prudente reemplazar dicha estimación por el intervalo 0-12% para los cambios previamente descritos y 27-40% para las mutaciones aleatorias.

## CONSECUENCIAS EVOLUTIVAS DE LAS MUTACIONES DELETÉREAS

### Carga genética de las poblaciones

Uno de los resultados clásicos de la Genética de Poblaciones es que, en el equilibrio mutación-selección la carga genética es independiente del coeficiente de selección en contra de las mutaciones deletéreas y depende únicamente de la tasa de mutación genómica, tal y como ilustra la ecuación 5 (Haldane, 1937; Kimura y Maruyama, 1966). Utilizando nuestra estimación de tasa de mutación genómica por generación  $U=0.138$ , obtenemos un valor esperado para la eficacia media de la población de  $E(W)=0.871$ , es decir que la mutación representa un lastre genético de aproximadamente 13% de la eficacia. Si en vez de nuestra estimación utilizamos la propuesta por Drake y Holland (1999) (del orden de magnitud de la obtenida por Elena y Moya (1999) a partir de datos de experimentos de acumulación de mutaciones),  $U=1.11$ , entonces obtenemos  $E(W)=0.330$ , es decir que la mutación supone un lastre genético de aproximadamente 67% de la eficacia. Una clara discrepancia existe entre estas dos estimaciones.

Cuando dejan de ignorarse las mutaciones recurrentes, el valor del coeficiente de selección influye también en la eficacia poblacional en el equilibrio (Schuster y Swetina, 1988). Utilizando las expresiones de Krakauer y Plotkin (2002) y nuestras estimaciones de la media del coeficiente de selección ( $s_H=-0.304$ ) y de la tasa de mutación, se obtiene que en la población en equilibrio hay promedio 0.454 mutaciones por individuo y que la eficacia media de ésta vale  $E(W)=0.258$ . Si en vez de nuestra estimación de la tasa de mutación utilizamos la proporcionada por Drake y Holland (1999), se obtiene que el número promedio de mutaciones por individuo es 3.650 y que la eficacia media de la población se reduce hasta  $E(W)=0.080$ . Estos resultados son relativamente robustos a la presencia de epistasia (Krakauer y Plotkin, 2002).

Estos cálculos indican que ignorar las mutaciones recurrentes conduce a grandes sobreestimaciones de un parámetro tan esencial en Genética de Poblaciones como la eficacia media poblacional. Si quisiéramos calcular mediante las expresiones de Krakauer y Plotkin (2002) la eficacia media de una población con un valor relativamente pequeño del coeficiente de selección ( $s=0.012$ , Tabla 10) y la tasa de mutación genómica que hemos estimado en

VSV, obtendríamos  $E(W)=0.010$ , mientras que la estimación clásica obtenida mediante la ecuación 5 sería, como hemos visto,  $E(W)=0.871$ . Si por otro lado, quisiéramos obtener la eficacia media esa misma población utilizando un valor relativamente bajo de la tasa de mutación ( $U=0.0025$ , Tabla 1) y el coeficiente de selección que hemos estimado para VSV, obtendríamos  $E(W)=0.303$  mientras que la estimación clásica sería  $E(W)=0.997$ . En otras palabras, la aproximación clásica de la Genética de Poblaciones, que ignora las mutaciones recurrentes, deja de ser válida, no sólo cuando las tasas de mutación son elevadas, sino también cuando los coeficientes de selección son altos.

Parece poco verosímil que la tasa de mutación en VSV sea tan elevada como proponen Drake y Holland (1999), ya que ello conlleva una reducción de la eficacia en términos de lastre genético de más del triple respecto al valor obtenido a partir de nuestra estimación de la tasa de mutación. Hay que mencionar que en sus cálculos, Drake y Holland asumieron un modo de replicación lineal el cual, como hemos demostrado mediante el ajuste de los datos de las pruebas de Luria-Delbrück a la distribución Poisson, no está justificado. Precisamente, esta asunción produce una sobreestimación de las tasas de mutación. Por otro lado, recordemos que aunque la estimación de la tasa de mutación por evento de replicación utilizando el método de la clase nula es insesgado, a la hora de expresarla en generaciones, hemos tenido que asumir que no hay más que dos eventos de replicación por individuo y generación (replicación lineal), lo que subestima nuestro valor de la tasa de mutación. Finalmente, existe una fuente de incertidumbre adicional tanto en nuestra estimación como en la de Drake y Holland, ya que no puede descartarse que además de las seis sustituciones descritas (Holland *et al.*, 1991), otras mutaciones confieran el fenotipo de resistencia al anticuerpo monoclonal empleado.

### Umbral de error

El umbral de error es un concepto introducido por la teoría de las cuasiespecies (Eigen *et al.*, 1988). Impone un límite superior al número de mutaciones que una población puede incorporar por genoma sin que se produzca la corrupción del mensaje codificado (Schuster, 1986; Nilsson y Snoad, 2000), de manera que predice que la tasa de error máxima soportable

es inversamente proporcional a la longitud del genoma. Utilizando la ecuación 8, podemos obtener una estimación de la tasa de mutación a la cual VSV entraría en catástrofe de error. El valor de la superioridad de la secuencia maestra sobre su nube de mutantes ( $\sigma_0$ ) para los mutantes simples aleatorios, de acuerdo con la ecuación 9 sería  $\sigma_0 = 1.533$ , con lo cual la tasa máxima de mutación por genoma y generación sería  $U_{max} = 0.427$ . Sin embargo, puede objetarse que no estamos teniendo en cuenta la frecuencia en la población de los diferentes mutantes, la cual dependerá de su eficacia. Dicha frecuencia puede estimarse en el equilibrio mutación-selección, donde el número de mutaciones por individuo se distribuye de acuerdo con una Poisson de parámetro  $U/s_H$ , (Kimura y Maruyama, 1966) con lo que la frecuencia en la población de mutantes simples vendrá dada por la expresión

$$p_1 = Ue^{U/s_H} / s_H. \quad (53)$$

Mutaciones por genoma ( $n$ )	$W_n$	$p_n$	$U_{max}$
1	0.696	0.366	0.584
2	0.484	0.166	0.636
3	0.337	0.050	0.676
4	0.235	0.011	0.690
5	0.163	0.002	0.693
6	0.114	$3 \times 10^{-4}$	0.693
7	0.079	$4 \times 10^{-5}$	0.694
8	0.056	$5 \times 10^{-6}$	0.694

**Tabla 11.** Eficacia biológica multiplicativa ( $W_n$ ) y frecuencia de los individuos en la población ( $p_n$ ) función del número de mutaciones por genoma. Se muestra finalmente la estimación de la tasa de mutación máxima ( $U_{max}$ ) cuanto se tiene en cuenta los genotipos portadores de una hasta  $n$  mutaciones.

Una vez hecha esta corrección, el valor estimado de la superioridad de la secuencia maestra es  $\sigma_0 = 1.794$  y el límite de error queda en  $U_{max} = 0.584$ . No obstante, aun puede objetarse que no se han tenido en cuenta los mutantes múltiples para realizar los cálculos. Si asumimos ausencia de epistasia, la

eficacia promedio de los genotipos portadores de  $n$  mutaciones puede deducirse de la eficacia promedio de los mutantes simples mediante el modelo multiplicativo (ecuación 21) y su frecuencia en la población se obtiene de la distribución Poisson como:

$$p_n = e^{-U/s_H} (U/s_H)^n / n!. \quad (54)$$

Puede apreciarse en la Tabla 11 que no es necesario ir más allá de  $n=4$ , ya que las estimaciones del umbral de error convergen rápidamente, pudiéndose decir que la tasa de mutación máxima por genoma y generación es  $U_{max} \approx 0.7$ .

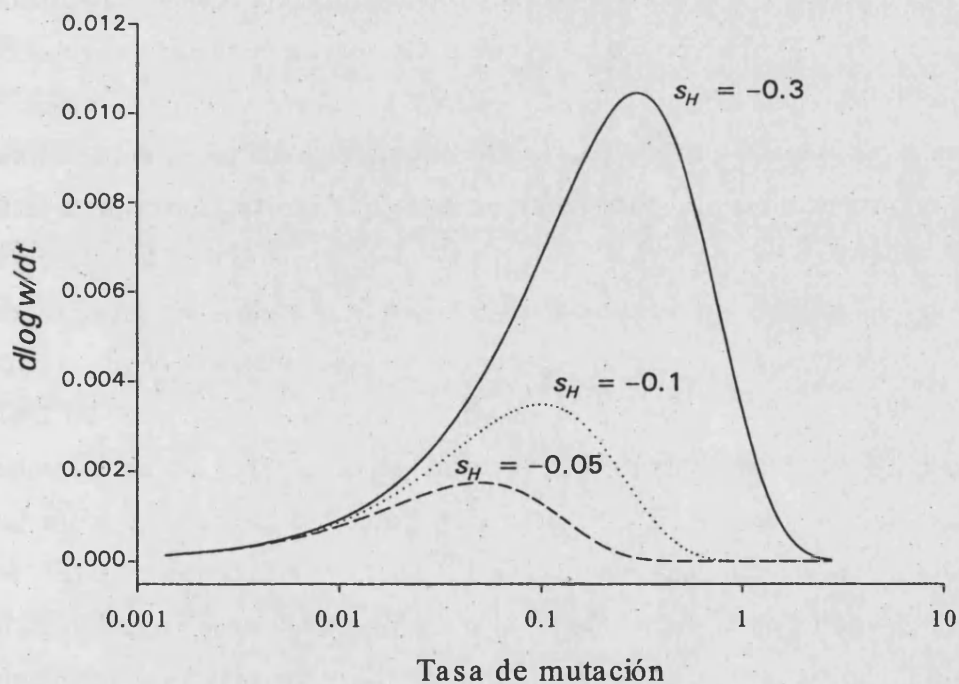
Notemos que el valor de Drake y Holland (1999),  $U=1.11$ , está por encima de esta estimación del umbral de error, mientras que la inferida a partir de nuestros experimentos es tan sólo cinco veces menor que el umbral de error estimado. Aplicando mutágenos químicos tanto a VSV como a FMDV, Holland *et al.* (1990) observaron que un aumento de dos veces en la frecuencia de mutantes conllevaba una reducción del porcentaje de virus viables de más del 90%, apuntando a la posibilidad de que los virus de RNA repliquen cerca del umbral de error. En bacterias generalmente puede aumentarse la frecuencia de mutación varios órdenes (genotipos mutadores) de magnitud sin producir necesariamente un daño excesivo (Goodman y Fygenon, 1998). No obstante, también existen evidencias de cepas mutadoras con graves defectos funcionales (Funchain *et al.*, 2000). Independientemente de que los genomas repliquen o no cerca del umbral de error, mirando la ecuación 8 puede entenderse que el valor de éste último depende del valor promedio de los coeficientes de selección. Como consecuencia, puede suponerse que la escasa robustez genética de los virus de RNA les permitiría tolerar una tasa de mutación superior a la de sistemas biológicos más complejos y robustos.

#### Tasa de mutación óptima

De manera habitual, se argumenta que las altas tasa de mutación permiten a los virus de RNA una mayor capacidad adaptativa (Holland *et al.*, 1982; Novella *et al.* 1995; Domingo y Holland, 1997; Domingo, 2000). Sin embargo, esta hipótesis carece de evidencia experimental a su favor y por otro lado, su fundamento evolutivo es dudoso. Es cierto que a mayor tasa de mutación y a mayor tamaño poblacional efectivo, mayor número de mutaciones adaptativas se generan. Sin embargo, al mismo tiempo, las poblaciones que aumenten su



tasa de mutación sufrirán un descenso en eficacia debido al aumento de la carga genética. Teniendo en cuenta estos dos factores contrapuestos, se llega a una tasa de óptima de mutación que permite maximizar la tasa de adaptación (Figura 21), cuyo valor en poblaciones asexuales es igual a la media armónica del los coeficientes de selección en contra de las mutaciones deletéreas (Orr, 2000). A partir de nuestros resultados con las mutaciones simples aleatorias, podemos estimar  $s_H = -0.304$ , por lo que  $U_{OPT} = 0.304$ . Si tal y como hemos sugerido previamente, corregimos a la baja la proporción de cambios letales y consideramos que éstos representan el 25% del total, obtendríamos  $U_{OPT} = 0.267$ . Estos valores son bastante cercanos a nuestra medida directa de la tasa de mutación,  $U = 0.138$ .



**Figura 21.** Tasa de adaptación predicha por las ecuaciones de Orr (2000) en función de la tasa de mutación para tres valores del coeficiente de selección en contra de las mutaciones deletéreas.  $N = 10^3$ ,  $s_b = 0.07$ ,  $p_b = 0.01$ .

A la vista de los resultados, parece lógico sugerir que la selección natural ha optimizado la tasa de mutación de VSV. Habitualmente, se argumenta que

las altas tasas de mutación típicas de los virus de RNA son consecuencia de la carencia de mecanismos de prueba de lectura. En efecto, está generalmente aceptado que, a diferencia de la mayoría de las DNA polimerasas, las RNA polimerasas virales no poseen actividad 3'-exonucleasa (Steinhauer *et al.*, 1992) ni la posibilidad de reparar mutaciones mediante edición post-replicativa, lo que conlleva que los virus de RNA tengan unas tasas de mutación al menos dos órdenes de magnitud mayores que los de la mayoría de sistemas biológicos basados en el DNA (Drake *et al.*, 1998). Sin embargo, estas características no deben entenderse como la consecuencia de restricciones bioquímicas, ya que existen múltiples evidencias experimentales recientes de que en determinadas condiciones, la selección natural puede favorecer RNA polimerasas con una mayor fidelidad replicativa. Pfeiffer y Kirkegaard (2003) han demostrado que el poliovirus de tipo 1 es capaz de generar resistencia a la ribavirina multiplicando por tres la fidelidad de su replicación, mediante una única sustitución aminoacídica en el gen de su RNA polimerasa. La tasa de mutación del virus de la fiebre amarilla parece ser de  $2 \times 10^{-7}$  sustituciones por sitio nucleotídico y evento de replicación (Pugachev *et al.*, 2004), aunque esos datos deberían ser tomados con mucha cautela, dado que el método que se usó para obtener esta estima es altamente indirecto, basado en pocos datos y sensible a la selección. En HIV-1, se han descrito numerosas mutaciones puntuales capaces de modificar la fidelidad de la transcriptasa reversa (Bakhanashvili *et al.*, 1996; Wisniewski *et al.*, 1999; Lewis *et al.*, 1999; Boyer y Hughes, 2000). En DNA polimerasas, la fidelidad de la replicación está gobernada por un compromiso entre velocidad de replicación y fidelidad, en forma de un balance entre las actividades 5'-polimerasa y 3'-exonucleasa (Goodman y Fyngenson, 1998). En virus de RNA aún no se ha explorado la relación entre fidelidad y velocidad de la replicación. Puesto que no existe actividad 3'-exonucleasa, es razonable suponer que la fidelidad pueda ser dependiente de la constante de polimerización. Esta relación es la predicha por la hipótesis de la prueba de lectura cinética (Hopfield, 1974), la cual ha demostrado ser válida en procesos como la traducción (Ruusala *et al.*, 1982), la transducción de señales celulares (McKeithan, 1995) o el mecanismo de acción de los enzimas implicados en el plegado del DNA (Yan *et al.*, 1999). Así pues, existe suficiente evidencia experimental para postular que la escasa fidelidad de las RNA polimerasas virales no debe entenderse como algo inevitable, sino como una carácter al

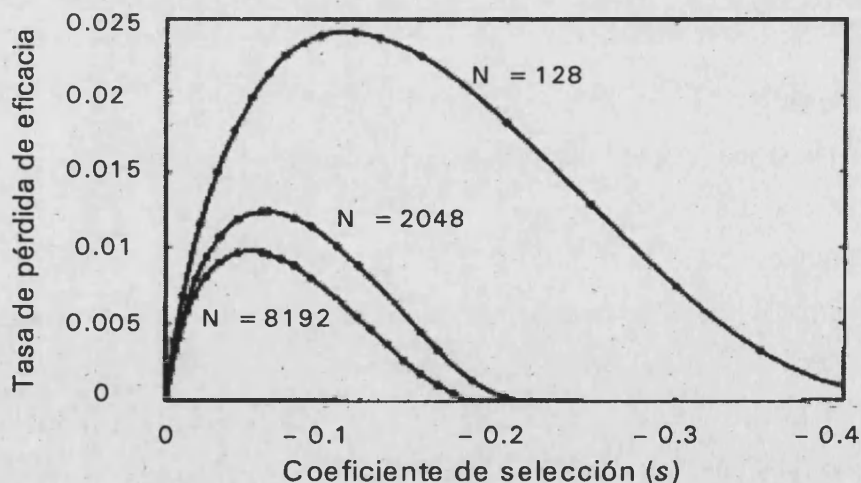
menos parcialmente optimizable, siendo la anti-robustez genética de los virus de RNA la causa de que sus tasas de mutación óptimas sean mayores que las de otros sistemas biológicos (Figura 21).

La predicción de Orr (2000) acerca de la tasa óptima de mutación es válida para diferentes distribuciones de efectos mutacionales y robusta a la presencia de interferencia clonal. Sin embargo, se asienta sobre el supuesto de que el efecto de las mutaciones beneficiosas es siempre menor que el de las mutaciones deletéreas, de manera que un cambio beneficioso sólo podrá proliferar en la población si ha aparecido en un genotipo libre de mutaciones deletéreas. Nuestros datos apoyan la validez de esta suposición, ya que el coeficiente de selección medio de las mutaciones beneficiosas es  $0.092 \pm 0.034$ , mientras que el de las deletéreas es  $-0.666 \pm 0.059$ , incluidas las letales. En el caso de que los cambios ventajosos tuvieran un efecto más fuerte que los deletéreos, se produciría el fenómeno del autostop genético (Maynard-Smith y Haigh, 1974). Este efecto conllevaría que la tasa de mutación óptima fuera mayor que la predicha por Orr, pudiendo genotipos mutadores alcanzar altas frecuencias en la población (Sniegowski *et al.*, 1997; Taddei *et al.*, 1997). Este aumento de la tasa óptima de mutación también podría producirse si, como hemos sugerido previamente, existiera un coste de la fidelidad replicativa (Johnson, 1999).

### **Trinquete de Muller**

En presencia de cuellos de botella poblacionales, la selección queda minimizada frente a la acción de la deriva genética; y las mutaciones deletéreas se fijan en la población, con la consecuente pérdida de eficacia biológica. Como ya hemos comentado, los virus de RNA son especialmente sensibles al trinquete de Muller. El ritmo al que las mutaciones deletéreas se fijan en la población crece conforme la tasa de mutación aumenta, lo que explica en parte esta propiedad. Pero además, el tamaño poblacional crítico a partir del cual la frecuencia de mutaciones beneficiosas es incapaz de frenar el avance del trinquete depende del efecto de las mutaciones deletéreas y de la distribución de estos efectos (Whitlock, 2000). La relación entre la robustez genética y la pérdida de eficacia debida a la acción del trinquete no es sencilla. Por un lado, cuanto menores sean los coeficientes de selección, mayor número de mutaciones podrán acumularse pero por otro lado, también será menor la caída

de eficacia por cada mutación fijada. Por ello, la pérdida de eficacia es máxima para un valor intermedio del efecto de las mutaciones (Felsenstein, 1974; Gabriel *et al.*, 1993; Lynch *et al.*, 1993; Butcher, 1995). Como puede apreciarse en la Figura 22, existe un primer tramo en el que la robustez individual, medida como coeficiente de selección, determina la robustez poblacional al trinquete. Sin embargo, a partir de un valor crítico, la tasa de pérdida de eficacia decrece al aumentar el coeficiente de selección en contra de las mutaciones deletéreas, por lo que se establece un conflicto entre la robustez individual y poblacional. Conforme el tamaño poblacional se reduce, el valor del coeficiente de selección que maximiza la pérdida de eficacia aumenta. Por ello, cuando los tamaños poblacionales sean muy reducidos, las poblaciones formadas por individuos muy poco robustos, como es el caso de los virus de RNA, sufrirán con más intensidad el efecto del trinquete que otros sistemas individualmente más robustos. Por el contrario, si los tamaños poblacionales son algo mayores, serán las poblaciones con mayor robustez individual las que perderán más eficacia, aunque el ritmo de pérdida podría ser lo suficientemente leve como para ser compensado por las mutaciones beneficiosas. Precisamente



**Figura 22.** Tasa de pérdida de eficacia como consecuencia del trinquete de Muller en función del coeficiente medio de selección de las mutaciones deletéreas, para tres tamaños poblacionales (modificada de Butcher, 1995).

es en condiciones de tamaños poblacionales reducidos cuando mayor es la acción del trinquete, lo que explica que los virus de RNA muestren tasas de pérdida de eficacia tan altas.

## MUTACIONES BENEFICIOSAS Y EVOLUCIÓN ADAPTATIVA

### Proporción de mutaciones beneficiosas

En ausencia de interferencia clonal, la tasa de fijación de mutaciones beneficiosas en una población asexual depende linealmente de la proporción de cambios beneficiosos. En nuestros experimentos, parece que uno de cada 15 cambios aleatorios tiene un coeficiente de selección positivo, proporción inusualmente elevada si la comparamos con los valores hallados en otros trabajos y comúnmente aceptados (Miralles *et al.*, 1999; Imhof y Schlötterer, 2001; Keightley y Lynch, 2003). Desde luego, la primera causa a la que se puede atribuir este resultado es la imprecisión experimental. Obtener un valor medio del coeficiente de selección positivo tras cinco réplicas no garantiza en absoluto que el valor real sea efectivamente positivo. Una vez aplicada la correspondiente prueba estadística, tan sólo uno de los 48 mutantes muestra una eficacia significativamente mayor que el silvestre. Este valor podría ser una subestima de la proporción real, ya que la potencia estadística con cinco réplicas no puede ser muy elevada mientras que la magnitud del efecto puede ser ilimitadamente pequeña. Por otro lado, tras aplicar 48 pruebas estadísticas con una especificidad del 95%, es de esperar obtener más de un falso efecto positivo. Así pues, tenemos un alto grado de incertidumbre sobre la proporción real de efectos beneficiosos.

Tras crear 226 mutantes de *E. coli* por inserción de elementos transponibles y utilizando ensayos de competencia parecidos a los nuestros, Elena *et al.* (1998b) no encontraron ningún caso de mutación beneficiosa. Esta diferencia puede deberse al tipo de mutación introducida, pero también al grado de adaptación del genotipo mutado. La proporción de cambios beneficiosos debe depender lógicamente de lo cerca que la población de halle de su óptimo adaptativo. Nuestro clon infeccioso es una quimera genética de las cepas San Juan y Orsay en la que además, el extremo 5' del genoma proviene de una partícula defectiva interferente. Es por ello esperable que existan gran cantidad

de vías para optimizar este genotipo, pudiendo esto explicar, en parte, la existencia de algún cambio beneficioso entre los 48 cambios aleatorios que hemos introducido.

### **Tasas de adaptación e interferencia clonal**

En el experimento de recuperación de la eficacia que hemos realizado, ésta se ha multiplicado aproximadamente por seis a un tamaño poblacional efectivo de  $2 \times 10^2$ , lo que equivale a una tasa de aumento de 0.10 del logaritmo de la eficacia por generación, valor bastante mayor que el predicho por las ecuaciones (14) y (15), que sería aproximadamente de 0.0017. Esta discrepancia puede ser debida a que el tamaño poblacional efectivo es adecuado para medir la deriva genética y la probabilidad de que una mutación beneficiosa aparezca en un fondo libre de mutaciones deletéreas, pero no la probabilidad de aparición de cambios beneficiosos. Para estimar esta última, podría ser más apropiado utilizar el tamaño poblacional real, cuya media geométrica está entorno a  $N = 10^5$ . Con este valor y una proporción de mutaciones beneficiosas de  $10^{-3}$ , los valores observado y esperado serían más o menos coincidentes. Sin embargo, cuando trabajamos con un tamaño poblacional efectivo de  $2 \times 10^4$ , la tasa observada de aumento para el logaritmo de la eficacia es apenas mayor que en el primer caso (0.11), mientras que su valor predicho por las ecuaciones es aproximadamente 0.17. Si corrigiésemos el valor esperado usando el tamaño poblacional real, que en este caso vale  $N \approx 10^6$  y una proporción de cambios beneficiosos de  $10^{-3}$ , la tasa esperada sería de 0.33.

A pesar de la inexactitud de estas estimaciones, parece claro que la tasa de adaptación observada no aumenta proporcionalmente al tamaño poblacional, lo cual es un indicio de la presencia de interferencia clonal. Aplicando la ecuación 17 (Wilke, 2004) a nuestros datos experimentales, la interferencia clonal debería ser notable a partir de  $N \approx 10^2$  si asumimos una proporción de cambios beneficiosos de  $1/50$  y  $N \approx 10^3$  si asumimos una proporción de  $10^{-3}$ , lo que indica que en efecto, esperamos que la interferencia clonal pueda haber entrado en juego.

### **Efectos beneficiosos y mutaciones fijadas**

La distribución empírica de efectos beneficiosos es una incógnita en casi la totalidad de sistemas biológicos, debido a lo escaso de estos cambios. De

manera teórica y basándose en que los efectos beneficiosos son los casos extremos de una distribución cualquiera de efectos mutacionales, Gillespie (1983) y Orr (2003) predijeron que la distribución de efectos beneficiosos debería ser exponencial. Hasta ahora, sólo Bull *et al.* (2000) han estudiado directamente el efecto de las mutaciones beneficiosas. Para ello, aislaron clones de  $\phi$ X174 (bacteriófago de DNA) resistentes a altas temperaturas y encontraron tres mutaciones de efecto extremo, con un coeficiente de selección ( $W-1$ ) entre  $s=4$  y  $s=13$ . Estos resultados demuestran la existencia de mutaciones de gran efecto en este fago, aunque no permiten hacer inferencias sobre la frecuencia de éstos ni su distribución. Nuestros datos apuntan a que una función de probabilidad Exponencial o Gamma podrían ser modelos adecuados en VSV, aunque el resultado no es concluyente debido al escaso tamaño de la muestra. El coeficiente de selección medio de las mutaciones con efecto beneficioso está entorno a  $s=0.1$ .

La distribución de efectos fijados en la población puede deducirse de la distribución de efectos mutacionales. Si la primera es una Exponencial y la probabilidad de que una mutación beneficiosa sobreviva a la deriva es proporcional a su coeficiente de selección, entonces los efectos de las mutaciones que no se pierden por deriva siguen una Gamma con un valor medio mayor que la Exponencial (Orr, 2002). A su vez, la interferencia clonal hará que sólo la mejor de estas mutaciones se fije, de manera que la distribución de los efectos fijados tendrá forma de campana y un valor esperado aún mayor (Rozen *et al.*, 2002). Esta predicción ha sido corroborada experimentalmente propagando 30 líneas de *E. coli* durante 400 generaciones, durante las cuales pudieron identificarse un total de más de 30 eventos de fijación. El efecto promedio de estos cambios estaba entorno a  $s=0.1$  y la distribución de efectos era concordante con la teoría (Rozen *et al.*, 2002). Aunque en un momento dado del proceso de adaptación, la distribución de efectos fijados sea acampanada, ésta será aproximadamente exponencial a lo largo de un proceso completo de adaptación, ya que conforme la población se acerca al óptimo adaptativo, los cambios beneficiosos irán teniendo un efecto cada vez menor (Orr, 1998). A favor de esta predicción están los resultados obtenidos por Imhof y Schlötterer (2001), quienes propagaron 10 líneas paralelas durante aproximadamente 1000 generaciones e identificaron un total de 66 eventos adaptativos. En nuestros experimentos, la distribución de efectos beneficiosos

previamente descritos puede interpretarse como un muestreo de los efectos fijados a lo largo de diferentes procesos de evolución (aunque 3/11 mutaciones fueron aisladas de clones). El coeficiente de selección medio de éstas está entorno a  $s=0.3$ , valor mayor que el de las mutaciones aleatorias tal y como predice la teoría y además, su distribución es compatible con un modelo exponencial.

## EFFECTOS EPISTÁTICOS

### Génesis de la epistasia

Existen diversos mecanismos subyacentes a la epistasia medida sobre la eficacia biológica como la desviación del modelo de efectos independientes. Primeramente, el efecto de una mutación sobre la función enzimática o estructural de una proteína no tiene porqué ser proporcional a su efecto medido sobre el carácter estudiado, como se ha comprobado estudiando los genes *lacZ* y *lacY* de *E. coli* (Dykhuizen, 1987), abriendo así la posibilidad a los efectos no multiplicativos de las mutaciones. Además, la interacción bioquímica que existe entre distintos enzimas por pertenecer a una misma ruta o a rutas relacionadas se traduce en una interacción entre los efectos de las mutaciones en los genes correspondientes (Keightley, 1989). De hecho, a nivel bioquímico, la eficacia biológica viene determinada por el flujo de producto que se establece en una determinada ruta. Cuando se mutan los enzimas implicados, un efecto multiplicativo sobre la energía de unión entre enzima y sustrato puede traducirse en una modificación no proporcional de la actividad enzimática, que a su vez tampoco se traducirá proporcionalmente sobre el flujo de producto (Szathmáry, 1993). Cuando se trate de mutaciones en un mismo enzima, se espera que mayoritariamente, su efecto nocivo se transmita de manera sinérgica sobre el flujo, produciendo una epistasia negativa. Sin embargo, este patrón no aparece en nuestros datos cuando restringimos el análisis del efecto epistático a los casos en que hemos introducido dos mutaciones en uno de los genes *P* o *L*. Cuando se trata de mutaciones sobre enzimas distintos, el tipo de epistasia esperado depende del tipo de selección que se ejerce sobre la ruta bioquímica, como veremos después.

La existencia de epistasia sobre la eficacia no implica necesariamente la presencia de un nexo bioquímico entre los *loci* implicados. Aunque los efectos



de los cambios genéticos sean independientes, la epistasia puede surgir como consecuencia de la selección. Por ejemplo, en el caso de que los alelos de varios genes contribuyan aditivamente al valor de un carácter sometido a selección estabilizadora, el paso a un alelo nulo en un gen podrá verse compensado por la activación de otro de los genes implicados [revisado por Fenster *et al.* (1997) y Whitlock *et al.* (2003)]. Además, efectos multiplicativos sobre la eficacia en ausencia de competencia pueden acabar siendo epistáticos en presencia de competencia. Notablemente, la epistasia negativa puede aparecer cuando la lucha por los recursos limitados ocurre en pequeños grupos (Peck y Waxman, 2000). Finalmente, de manera general, dado que puede suponerse la existencia de un valor máximo y mínimo de eficacia biológica a los que denominamos pico adaptativo y genotipo letal respectivamente, es de esperar que las mutaciones beneficiosas tiendan a mostrar epistasia negativa, mientras que las mutaciones deletéreas tiendan a mostrar epistasia positiva. La epistasia negativa entre mutaciones beneficiosas ha sido hallada tanto en el bacteriofago  $\phi$ X174 (Bull *et al.*, 2000) como en nuestro sistema experimental. Además, hemos encontrado que, en efecto, la epistasia entre mutaciones deletéreas es predominantemente positiva.

### Patrones de epistasia

Hasta hace poco, no podía llegarse a conclusiones generales acerca del signo de la epistasia, aunque parecía que las interacciones negativas leves predominaban en eucariotas mientras que los procariotas y los virus no mostraban tendencia alguna. En todo caso, experimentos con organismos digitales sugerían que la epistasia debería ser más abundante en los organismos complejos que en los simples (Lenski *et al.*, 1999). Las ciencias biomédicas se han valido generalmente de técnicas como la mutagénesis dirigida para el diseño de vacunas atenuadas, en las que se introducen varios cambios, generalmente deletéreos, en la cepa salvaje. Burch *et al.* (2003) aprovecharon estos datos para llevar a cabo un análisis de la epistasia en virus de RNA. Se compilaron datos de 14 especies y muy diversas medidas de eficacia, como la tasa de crecimiento, el título máximo en cultivo o *in vivo*, o la actividad enzimática máxima. Sin embargo, no se encontró tendencia alguna hacia un tipo particular de epistasia y aunque pares concretos de mutaciones mostraron epistasia, estos casos no fueron abundantes. Por el contrario, nuestros datos

demuestran que al menos en VSV, existe una predominancia de la epistasia positiva entre mutaciones deletéreas. Este resultado está en concordancia con la evidencia más reciente hallada en otros virus de RNA. Experimentos de acumulación de mutaciones en el bacteriofago  $\phi 6$  han demostrado que la magnitud del coeficiente de selección decrece conforme el genotipo sobre el cual se acumulan las mutaciones decrece en eficacia (Burch y Chao, 2004). Por otro lado, la pérdida de infectividad del poliovirus 1 en presencia de ribavirina se asocia con la acumulación de mutaciones debido al efecto mutagénico de ésta (Crotty *et al.*, 2001). Un reanálisis de estos datos (Burch y Chao, 2004) permite concluir que la caída del logaritmo de la eficacia no es proporcional al número de mutaciones incorporadas, sino que existe un efecto epistático positivo, aunque no estadísticamente significativo. Bonhoeffer *et al.* (2004) clonaron los genes de la proteasa y la retrotranscriptasa de 9466 aislados de pacientes infectados con HIV-1, secuenciaron cada fragmento para conocer el número de mutaciones que cada genotipo tenía respecto al silvestre y midieron la eficacia de los virus recombinantes resultantes. La conclusión clara es que la epistasia entre mutaciones deletéreas en HIV-1 es predominantemente positiva. No obstante, a este estudio cabe objetarle que utilizaron como medida de eficacia la cantidad de partículas producidas por célula, lo cual no es en realidad más que un componente de eficacia ligado a la selección de tipo *K* (Cuevas *et al.*, 2005).

Los experimentos con otros sistemas biológicos no son aún tan concluyentes. Creando líneas de *Aspergillus niger* con diferentes combinaciones de mutaciones deletéreas conocidas, de Visser *et al.* (1997) determinaron que, aunque la epistasia es frecuente, no existe tendencia significativa hacia valores positivos o negativos. Elena y Lenski (1997) generaron 225 genotipos de *E. coli* portadores de una, dos o tres mutaciones deletéreas y midieron su eficacia. La relación entre el número de mutaciones y el logaritmo de la eficacia era lineal, lo que indica que en promedio, no hay desviación del modelo multiplicativo. Sin embargo, al medir la eficacia en mutantes simples y dobles tal y como hemos realizado nosotros, observaron que la epistasia es frecuente aunque la desviación hacia valores positivos era parecida en magnitud y frecuencia a la desviación hacia valores negativos. Utilizando la asimetría de la distribución de los logaritmos de las eficacias en cruces del alga unicelular *Chlamydomonas moewussii*, de Visser *et al.* (1997) encontraron una predominancia de epistasia

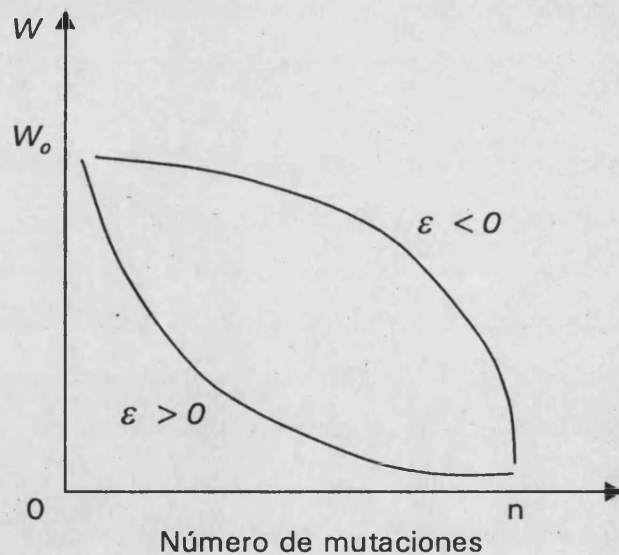
negativa cuando el carácter de eficacia medido era la capacidad de carga ( $K$ ), mientras que si el carácter estudiado era la tasa de crecimiento ( $r$ ), parecía no haber desviación del modelo multiplicativo. Sin embargo, mientras que en el caso de la capacidad de carga la medida correcta del logaritmo de la eficacia es, en efecto,  $\log K$ , en el caso de la tasa de crecimiento la medida correcta es  $r$  y no  $\log r$  como de Visser *et al.* emplearon. Si ese error no se hubiera cometido, es posible que la asimetría hubiera sido significativa y se hubiera detectado epistasia negativa. Withlock y Bourget (2000) abordaron el tema usando un conjunto de marcadores deletéreos en *D. melanogaster*. La productividad, que es la combinación de la fecundidad de las hembras y la supervivencia de la prole, mostró un patrón de epistasia fuertemente negativa, mientras que el componente de eficacia correspondiente al éxito reproductivo de los machos no mostró tendencia alguna. En suma, puede afirmarse que con los datos disponibles, los organismos de DNA se diferencian de los virus de RNA por no mostrar epistasia positiva, sino más bien una aparente ausencia de epistasia o una tendencia hacia la epistasia negativa.

### **Epistasia y robustez genética**

Puesto que la robustez genética se define como la capacidad de mantener la eficacia biológica inalterada frente a la mutación, es intuitivo decir que la epistasia positiva es un signo de robustez. Sin embargo, vamos a argumentar que esta visión no se corresponde con la realidad. Pensemos primeramente en la forma en que la eficacia cae en función del número de mutaciones deletéreas acumuladas. Si la curva es cóncava, la epistasia es negativa mientras que si es convexa la epistasia es positiva. En el primer caso, el silvestre, que es el genotipo de referencia para medir la robustez, tiene mayor capacidad para tolerar los cambios genéticos que aquellos genotipos que ya han sido dañados por la mutación. Por lo tanto, como esgrimen Burch y Chao (2004), puede decirse que existe un tamponamiento del efecto de las mutaciones es decir, una robustez genética. Si por el contrario la curva es convexa y la epistasia es positiva, el genotipo sobre el que estamos midiendo la robustez no está tamponado y no es robusto.

A este razonamiento inicial hay que añadir que la eficacia biológica de los mutantes deletéreos está comprendida entre los valores máximo (eficacia igual al silvestre,  $W=1$ ) y mínimo (genotipo letal,  $W=0$ ). Es ineludible que todo

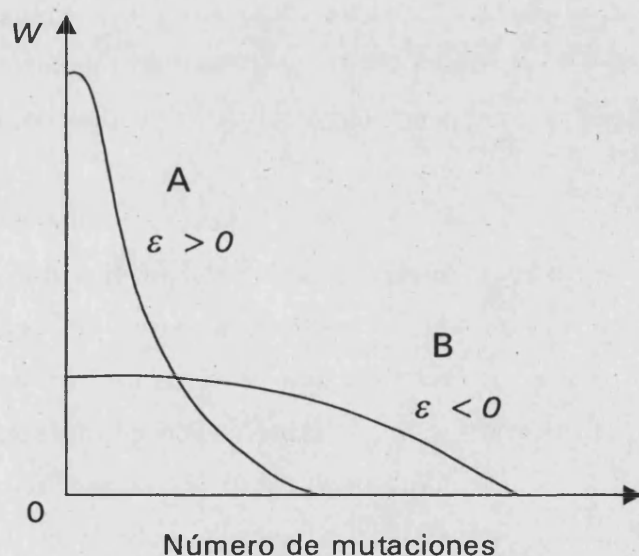
genotipo silvestre pase de uno a otro extremo si se introduce un número creciente ilimitado de mutaciones. Para una eficacia de partida fija y un número de mutaciones fijo tras el que un organismo alcanza la letalidad, la epistasia entre esas mutaciones determinará el grado de robustez genética, tal y como puede observarse en la Figura 23, siendo la epistasia negativa un carácter de robustez y la epistasia positiva de anti-robustez. Podría sin embargo argumentarse que el número de mutaciones tolerables puede ser a su vez función de la epistasia. Esto es lo que pasaría si un genoma aumentase su grado de redundancia. Como ejemplo sencillo podemos plantear el caso de la duplicación de un gen esencial, aunque el razonamiento debería ser válido para cualquier tipo de redundancia. Independientemente del efecto que la duplicación tenga sobre la eficacia biológica, queda claro que el genoma ganará robustez frente a las mutaciones que ocurran en el gen duplicado. Situándonos en el caso ilustrativo de mutaciones que inactiven el gen, el genotipo ancestral sólo tolerará un cambio, mientras que el que ha sufrido la duplicación tolerará al menos dos. La selección será en este caso de tipo truncado, que no es más que un caso extremo de epistasia negativa. Este ejemplo ilustra como la



**Figura 23.** Curvas de pérdida de eficacia alternativas ante la acumulación de mutaciones deletéreas fijando la eficacia inicial (cero mutaciones) y la final ( $n$  mutaciones).

redundancia y la robustez generan epistasia negativa y no positiva. Expresado en otros términos, si un organismo no es redundante, cualquier estrés mutacional tendrá serios efectos adversos sobre las funciones clave del metabolismo. Pero una vez dañadas esas funciones, el estrés adicional tendrá efectos poco importantes (Elena y de Visser, 2003). Por el contrario, si las funciones esenciales para el crecimiento son redundantes de manera que pueden ser realizadas por varios caminos, entonces el organismo será más robusto a la mutación, pero los efectos añadidos tenderán a ser cada vez mayores (Lyman *et al.*, 1996).

La epistasia está fuertemente relacionada con la geometría del espacio fenotípico (Wagner *et al.*, 1998; Wilke y Adami, 2001; You y Yin, 2002). *A priori*, se podría esperar que la epistasia fuera independiente de los efectos de las mutaciones simples. Sin embargo, si fijamos la condición de que la suma de todos los efectos mutaciones se mantenga constante (se denomina a esta constante la constante de neutralidad del espacio genotípico), se deduce por razones puramente geométricas que



**Figura 24.** Curvas de pérdida de eficacia hipotéticas ante la acumulación de mutaciones deletéreas de un genotipo anti-robusto A y un genotipo robusto B respetando la constante de neutralidad del espacio genotípico.

existe necesariamente un compromiso entre la epistasia y la robustez genética. Si por ejemplo, en una región del espacio los efectos mutacionales son fuertes, entonces conforme los cambios se acumulan y nos alejamos de dicha región, los efectos disminuirán y por lo tanto la epistasia será positiva (Wilke y Adami, 2001). La correlación negativa entre la robustez y la epistasia se ha observado en el plegamiento de secuencias simuladas de RNA (Wilke y Adami, 2001), en organismos digitales (Wilke y Adami, 2001) y mediante mutagénesis sobre el fago T7 *in silico* (You y Yin, 2002). Nuestros datos confirman los resultados de todas estas simulaciones.

El modelo de Peck y Waxman (2000) y los resultados obtenidos por de Visser *et al.* (1996) en *Chlamydomonas* sugieren que la epistasia negativa ocurre bajo regímenes de selección de tipo *K*. De hecho, la teoría del flujo metabólico predice que, si la selección actúa favoreciendo el máximo flujo, las mutaciones deletéreas sobre enzimas distintos mostrarán epistasia positiva mientras que por el contrario, si la selección actúa optimizando el flujo o la homeostasis, entonces las mutaciones interaccionarán mediante epistasia negativa. La primera situación debería darse en organismos simples de rápida tasa de reproducción, pequeños genomas y escasos mecanismos de robustez genética, mientras que la segunda debería ser propia en organismos con rutas metabólicas complejas con mayor tamaño y robustez genómicos, en los que la competencia favorezca en máximo aprovechamiento de los recursos. Esta hipótesis explicaría porqué los sistemas menos robustos, notablemente los virus de RNA, muestran una predominancia de epistasia positiva mientras que los sistemas más complejos y robustos muestran ausencia de epistasia o una epistasia negativa. Además, es esperable que los sistemas robustos paguen su mayor capacidad de reproducción ante la acumulación de mutaciones con una menor tasa de reproducción en ausencia de éstas, restricción que permite mantener la constante de neutralidad del espacio genotípico. La Figura 24 ilustra la relación esperable entre tasa de reproducción, robustez y epistasia.

## **IMPLICACIONES EVOLUTIVAS DE LA EPISTASIA**

### **Efecto sobre el trinquete de Muller**

En principio, podría pensarse que la epistasia positiva, al ralentizar el ritmo de pérdida de eficacia consecuencia de la acumulación de mutaciones, conferiría

a las poblaciones una mayor robustez ante el trinquete de Muller. Pero una vez más, esto no es así. Como hemos visto, debe existir, para un tamaño poblacional dado, un valor del coeficiente de selección más allá del cual el avance del trinquete se reduce, pudiéndose incluso detener (imaginemos el caso extremo de la selección truncada) o verse contrarrestado por las mutaciones beneficiosas. En este sentido, la epistasia negativa contribuye a frenar el trinquete de Muller (Charlesworth *et al.*, 1993; Kondrashov, 1994; Butcher, 1995), mientras que su avance es más rápido con epistasia positiva (Butcher, 1995; Hansen y Wagner, 2001). A su vez, el tamaño poblacional crítico que permite la supervivencia (Whitlock, 2000) se verá previsiblemente aumentado como consecuencia de las epistasias positivas, haciendo pues que la población sea más susceptible a la extinción.

El efecto de la epistasia sobre el trinquete de Muller puede contribuir a explicar, junto con otros factores como la tasa de mutación, porqué los virus de RNA sufren unas pérdidas tan acusadas de eficacia bajo un régimen poblacional de cuellos de botella repetidos, observándose incluso la extinción de la población en algunos casos (Yuste *et al.*, 1999), mientras que organismos más complejos como *E. coli* o *C. elegans* sufren pérdidas de eficacia menores (Kibota y Lynch, 1996; Keightley y Caballero, 1997). Sin embargo, carecemos aún de evidencia experimental que permita comprobar esta predicción.

### **Evolución compensatoria**

Tanto las reversiones como las mutaciones compensatorias pueden contribuir a la recuperación de eficacia en virus de RNA (Novella, 2004a). Cuando poblaciones de HIV-1 resistentes al inhibidor de la proteasa evolucionan en el laboratorio, se observa una ganancia de eficacia parcialmente atribuible a la reducción del coste de la resistencia mediante la fijación de pequeños cambios compensatorios y no por reversión de la mutación de resistencia. De igual modo, poblaciones del bacteriofago  $\phi 6$  altamente adaptadas que han sufrido la fijación de un único cambio deletéreo recuperan eficacia mediante la fijación de pequeños cambios, siendo la mayoría de estos pequeños cambios de tipo compensatorio (Burch y Chao, 1999). A su vez, el análisis de los cambios genéticos asociados a la recuperación de la eficacia en poblaciones experimentales de FMDV reveló que existen múltiples soluciones moleculares compensatorias, aunque la reversión también tiene lugar (Escarmís *et al.*,

1999). En el caso de VSV, hemos comprobado que la recuperación de eficacia se produce también de manera preferente por compensación, ya que sólo hemos identificado 12 reversiones entre las 72 líneas de evolución.

En primer lugar, hemos observado que la probabilidad de reversión aumenta al incrementarse el tamaño poblacional. La reversión constituye el modo más eficaz de recuperar eficacia, pero es menos probable en términos mutacionales que la compensación. Por ello es esperable que la probabilidad de que la evolución se produzca por reversión se correlacione positivamente con el tamaño poblacional y el tiempo de evolución (Levin *et al.*, 2000), lo cual explicaría porqué Novella (2004a) encontró en VSV una proporción de reversiones muy superior a la que nosotros hemos observado. El fenómeno de supercompensación, por el cual la eficacia del genotipo compensado es mayor que las del ancestral (Whitlock *et al.*, 2003) sería la excepción a esta regla. Aunque hemos observado que la eficacia de los genotipos de VSV tras la evolución es sistemáticamente mayor que la del silvestre, esto es atribuible a la fijación de mutaciones de efecto beneficioso inespecífico. En segundo lugar, nuestro sistema experimental permite contrastar la hipótesis de Novella (2004a) según la cual aquellos genotipos con mayor número de mutaciones deletéreas tenderían a recuperar eficacia por reversión, mientras que aquellos con un menor número de mutaciones tenderían a fijar cambios compensatorios. Nuestros datos no apoyan tal hipótesis, puesto que en ambos tamaños poblacionales examinados, la frecuencia de reversiones entre los mutantes simples no es menor que entre los mutantes dobles.

La importancia relativa de la compensación y la reversión en la recuperación de la eficacia biológica ha sido explorada más allá de los virus de RNA. En el bacteriófago T7, la recuperación de la eficacia tras la delección del gen de la ligasa se produce en su totalidad por cambios compensatorios (Rokyta *et al.* 2002), resultado más que previsible ya que la probabilidad de reversión ante una mutación tan grande es realmente nula. Más interesante es el caso del experimento de Crill *et al.* (2000) con el bacteriófago  $\phi$ X174, donde se observó que la recuperación de la eficacia en un genotipo debilitado por la fijación de tres sustituciones puntuales se producía preferentemente por reversión. En *E. coli*, se ha demostrado que el coste de la resistencia a estreptomicina es compensada por mutaciones secundarias y que además, éstas son epistáticas con las propias mutaciones primarias de resistencia, de manera que la reversión



de éstas últimas pasa a ser deletérea en el nuevo fondo genético (Schrag *et al.*, 1997). Estos resultados fueron generalizados por Moore *et al.* (2000) quienes, tras introducir una serie de inserciones, siguieron el patrón de recuperación de la eficacia, observando que la compensación genética en *E. coli* es altamente frecuente. De manera semejante, cepas de *Salmonella typhimurium* resistentes a antibióticos muestran inicialmente un fenotipo no virulento en ratón en ausencia del antibiótico, pero evolucionan de nuevo hacia la virulencia mediante la fijación de mutaciones compensatorias y no por reversión de la resistencia (Björkman *et al.*, 1998). En *C. elegans*, también se ha observado una rápida recuperación tras la acumulación de mutaciones deletéreas (Estes y Lynch, 2003), aunque en este caso no pudo determinarse la fracción de reversiones y de cambios compensatorios.

Con este conjunto de resultados en la mano, parece claro que salvo excepciones, la recuperación de la eficacia biológica tras la acumulación de mutaciones se produce fundamentalmente por la fijación de cambios compensatorios y que, por lo tanto, existe un alto grado de irreversibilidad en la evolución. Esto pone de manifiesto que las epistasis no sólo son abundantes, sino que tienen una fuerte influencia en la evolución. Aunque el hecho de que la adaptación en estos casos proceda por pequeños pasos compensatorios proporciona evidencia a favor del modelo geométrico de Fisher (1930) y el modelo exponencial de Orr (1998), la existencia misma de los cambios compensatorios da una gran validez al modelo de paisajes adaptativos rugosos de Wright (1982). Así pues, parece que la conjunción de deriva genética, selección y epistasia permite explicar en parte la diversificación de las poblaciones a al vez que el mantenimiento de las contingencias históricas.

### **Signo de la epistasia y adaptación**

Nuestros resultados apuntan claramente hacia una epistasia negativa entre las mutaciones beneficiosas, en concordancia con las observaciones de Bull *et al.* (2002) en  $\phi$ X174. Este fenómeno puede ser atribuido a la existencia de un techo adaptativo a pocos pasos mutacionales del silvestre, de manera que los pares de mutaciones beneficiosas resulten parcialmente redundantes. La presencia de epistasia negativa contrasta con el signo de la interacción entre las mutaciones deletéreas, tanto en VSV como en el resto de virus de RNA. Como hemos visto, este cambio de signo no debe entenderse como una

transición súbita que ocurre en el genotipo silvestre, sino como la consecuencia de la correlación negativa general entre el efecto epistático y el coeficiente de selección esperado en ausencia de interacción, atribuible a su vez a la conservación de la constante de neutralidad del espacio genotípico.

Es importante resaltar que a diferencia de los resultados obtenidos por Bull *et al.* (2002), en VSV parece predominar un caso particular de epistasia negativa entre las mutaciones beneficiosas, que se caracteriza por que la eficacia del doble mutante es menor que la de cualquiera de los mutantes simples a partir de los cuales está construido y a la que se denomina epistasia descompensatoria (Wolf *et al.*, 2000). Puesto que el genotipo portador de dos mutaciones beneficiosas no obtiene beneficio alguno, el techo adaptativo queda reducido. Además, los linajes portadores de las mutaciones simples tendrán que competir por la fijación, acentuando la interferencia clonal. Por otro lado, este tipo de epistasia fue propuesto como mecanismo de especiación alopátrica (Dobzhansky, 1936; Muller, 1939) y en todo caso, apoya la idea de un paisaje adaptativo rugoso para VSV.

Si bien la epistasia descompensatoria reduce necesariamente la capacidad adaptativa, no resulta tan intuitivo el efecto que pueda tener sobre la tasa de adaptación la epistasia convencional es decir, aquélla en que los rangos de las eficacias se corresponden con el número de mutaciones presentes. En nuestros experimentos, hemos explorado directamente el efecto de la epistasia sobre la tasa de adaptación de VSV midiendo la magnitud de la recuperación de la eficacia en genotipos portadores de mutaciones dobles que muestran epistasia de diferente signo e intensidad. Este análisis tiene un precedente en el trabajo de Moore *et al.* (2000) quienes concluyeron que en *E. coli*, la compensación es más eficiente cuando las mutaciones ancestrales introducidas interactúan mediante epistasia negativa, seguida por aquéllas que no muestran epistasia y finalmente por las que muestran epistasia positiva. Este resultado es el esperado si tenemos en cuenta que en cuanto a la recuperación de la eficacia concierne, existe un máximo adaptativo bien definido por el silvestre. En el caso de la epistasia negativa, si se compensa o revierte uno de los cambios deletéreos, la ganancia de eficacia resultante es no sólo la disipación del efecto deletéreo individual de dicho cambio, sino también del término epistático, siendo contraria la situación en el supuesto de epistasia positiva. Este fenómeno puede ser visualizando leyendo de derecha a izquierda la Figura 23.

Como hemos visto, puede predecirse en poblaciones asexuales la existencia de una tasa de mutación óptima a la cual se maximiza la tasa de adaptación. Dicha tasa de mutación podrá ser mayor cuanto mayor sea el coeficiente de selección en contra de las mutaciones deletéreas (Orr, 2000). Este razonamiento se basa en que las únicas mutaciones beneficiosas con opciones de fijarse en la población son aquellas que ocurran en genotipo libres de mutaciones deletéreas. La frecuencia de éstos últimos se ve reducida como consecuencia de la epistasia positiva (Chao, 1988) de manera que la tasa de mutación óptima también se verá reducida. Esta es una de las posibilidades para explicar por qué la tasa de mutación que hemos medido en nuestro genotipo silvestre es algo menor que la esperada por el modelo de Orr (2000), el cual ignora la influencia de las epistasias.

De manera paralela al razonamiento sobre el modelo de Orr, la lectura de la Figura 23 permite también entender que (imponiendo la condición de que exista un techo adaptativo), la epistasia negativa entre mutaciones beneficiosas acelera la respuesta adaptativa. Este tipo de epistasia es precisamente el que hemos observado en VSV que, como el conjunto de virus de RNA, muestra una rápida tasa de adaptación (Holland *et al.*, 1982; Novella *et al.*, 1995). Notemos que el efecto de la epistasia sobre la tasa de adaptación es coincidente con el del coeficiente de selección multiplicativo, ya que cuanto mayor es éste, menor es la tasa de adaptación (Figura 18). Si además se conserva la constante de neutralidad del espacio genotípico, existirá un compromiso entre dicho coeficiente de selección y la epistasia, de manera que lo esperado ya no es que la tasa de adaptación decrezca monótonamente con el valor de epistasia, sino que se maximice en un punto, tal y como se aprecia en la Figura 20 y en discordancia con los resultados de Moore *et al.* (2000) en los que posiblemente, el rango de valores explorado o el peso relativo de la epistasia y los coeficientes de selección impidió llegar a tal conclusión.

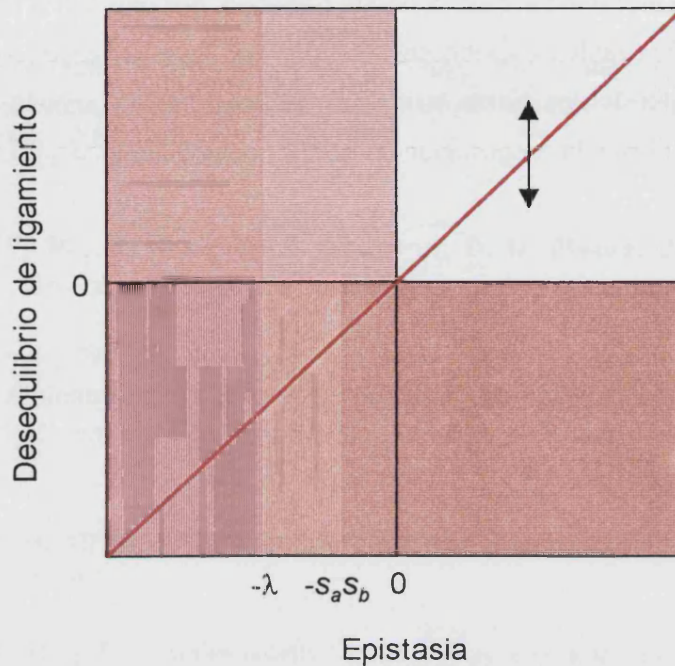
### **Epistasia y recombinación**

Gran parte de las investigaciones realizadas acerca de la epistasia han sido motivadas por su posible relación con la evolución y el mantenimiento de y la reproducción sexual y la recombinación. De entre las muchas hipótesis planteadas para explicar la evolución del sexo, la hipótesis determinista de la mutación (Kondrashov, 1988) ha sido quizás la que más atención ha acaparado

dentro de la Genética de Evolutiva, debido a su generalidad y su facilidad para ser contrastada experimentalmente. Como hemos visto, la idea se basa en que en el equilibrio mutación-selección, la reproducción sexual permite reducir la carga genética de las poblaciones siempre que la epistasia entre mutaciones deletéreas sea en promedio débil y negativa y que la tasa de mutación sea mayor de un cambio por genoma y generación. Aunque se han descrito ciertos casos de epistasia negativa en organismos sexuales (de Visser *et al.*, 1997; Withlock y Bourget, 2000) y nuestros resultados confirman la predominancia de la epistasia positiva entre mutaciones deletéreas en un genoma asexual como es el de VSV, los recientes hallazgos de epistasia positiva en virus de RNA con reproducción sexual contradicen esta hipótesis (Burch y Chao, 2004; Bonhoeffer *et al.*, 2004). A esto se añade que la condición de epistasia débil y negativa no es suficiente, sino que se requiere que la varianza de la epistasia sea pequeña (Otto y Feldman, 1997), lo cual tampoco parece de acuerdo con la evidencia empírica (Elena y Lenski, 1997; de Visser *et al.*, 1997; Withlock y Bourget, 2000). Por último, utilizando el hecho de que bajo neutralidad estricta, la tasa de evolución es igual a la tasa de mutación (Kimura, 1983) y asumiendo que los cambios silenciosos en zonas codificantes son estrictamente neutrales, Keightley y Eyre-Walker (2000) estimaron que la tasa de mutación deletérea en algunos sexuales obligados es hasta 20 veces menor de lo requerido por la hipótesis determinista de la mutación.

La relación entre epistasia y recombinación puede ser generalizada mediante la ecuación 18 (Barton, 1995). En el caso particular de que la población se encuentre en el equilibrio mutación-selección, el valor de la epistasia y el desequilibrio de ligamiento son proporcionales (Barton, 1995), por lo que la epistasia ha de ser débil y negativa, tal y como establece la hipótesis determinista de la mutación. Sin embargo, fuera del equilibrio la epistasia podrá estar desacoplada del desequilibrio de ligamiento, por ejemplo en ambientes temporal o espacialmente heterogéneos. En un modelo de selección fluctuante, las condiciones para la recombinación serán favorables cuando la epistasia cambie de signo con rapidez, cada 2 - 5 generaciones (Barton, 1995). Esta idea podría ser relevante para virus de RNA ya que se han descrito patrones de evolución fluctuantes en presencia de partículas defectivas interferentes (de Polo *et al.*, 1987). Otra manera de desacoplar la epistasia del desequilibrio de ligamiento es mediante ambientes espacialmente heterogéneos y más

concretamente, por migración. La selección sobre diferentes *loci* podría estar positivamente correlacionada entre ambientes, de manera que si al cambiar de ambiente, la intensidad de la selección se hace, por ejemplo, más fuerte en un



**Figura 25.** Importancia de la epistasia y el desequilibrio de ligamiento en la evolución de la recombinación en un modelo de dos *loci*. La línea roja indica el equilibrio selección-mutación. La flecha simboliza los procesos que pueden romper ese equilibrio. La trama roja horizontal indica una ventaja de la recombinación a corto plazo, mientras que la trama roja vertical indica una ventaja de la recombinación a largo plazo. La zona sombreada corresponde a las condiciones favorables para la evolución de la recombinación.

*locus*, también tenderá a ser así para los demás *loci*. Entonces se generaría un desequilibrio de ligamiento positivo, por lo que se requerirá una fuerte epistasia negativa entre las mutaciones para favorecer la recombinación (Lenormand y Otto, 2000). La situación habitualmente más favorable para la recombinación se daría en presencia de una correlación negativa de la selección a través del espacio (Lenormand y Otto, 2000). El efecto de la migración puede entenderse como un

desplazamiento de la línea mutación-selección en la Figura 25. Si la selección está positivamente correlacionada entre ambientes, la línea mutación-selección se desplaza hacia arriba. Si, por el contrario, dicha correlación es negativa, la migración desplaza la línea hacia abajo. Finalmente, las poblaciones pueden alejarse del equilibrio al entrar en acción la deriva genética, por ejemplo a través de los cuellos de botella frecuentes experimentados por los virus en su transmisión. Entonces, el desequilibrio de ligamiento tenderá a ser negativo de acuerdo con el efecto Hill-Robertson (1966) y la recombinación podrá evolucionar si la epistasia no es negativa y fuerte.

## REFERENCIAS

- Andre, P., A. Kim, K. Khrapko y W. G. Thilly. 1997. Fidelity and mutational spectrum of *Pfu* DNA polymerase on a human mitochondrial DNA sequence. *Genome Res.* 7:843-852.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith y K. Struhl. 1999. Short protocols in molecular biology. Wiley, New York.
- Bakhanashvili, M., O. Avidan y A. Hizi. 1996. Mutational studies of human immunodeficiency virus type 1 reverse transcriptase: the involvement of residues 183 and 184 in the fidelity of DNA synthesis. *FEBS Lett.* 391:257-262.
- Barton, N. H. 1995. A general model for the evolution of recombination. *Genet. Res.* 65:123-145.
- Barton, N. H. y B. Charlesworth. 1998. Why sex and recombination? *Science* 281:1986-1990.
- Bell, G. 1997. Selection, the Mechanism of Evolution. Chapman & Hall, New York.
- Berg, O. G. y P. J. Silva. 1997. Codon bias in *Escherichia coli*: the influence of codon context on mutation and selection. *Nucleic Acids Res.* 25:1397-1404.
- Bergstrom, C. T. y J. Pritchard. 1998. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics* 149:2135-2146.
- Bilsel, P. A. y S. T. Nichol. 1990. Polymerase errors accumulating during natural evolution of the glycoprotein gene of vesicular stomatitis virus Indiana serotype isolates. *J. Virol.* 64:4873-4883.
- Björkman, J., D. Hughes y D. I. Andersson. 1998. Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* 95:3949-3953.
- Bonhoeffer, S., C. Chappey, N. T. Parkin, J. M. Whitcomb y C. J. Petropoulos. 2004. Evidence for positive epistasis in HIV-1. *Science* 306:1547-1550.
- Boyer, P. L. y S. H. Hughes. 2000. Effects of amino acid substitutions at position 115 on the fidelity of human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.* 74:6494-6500.

**Bracho, M. A., A. Moya y E. Barrio.** 1998. Contribution of *Taq* polymerase-induced errors to the estimation of RNA virus diversity. *J. Gen. Virol.* **79**:2921-2928.

**Bracho, M. A.** 2003. Caracterizació molecular i anàlisi de la dinàmica poblacional de recuperació d'eficàcia del virus de l'estomatitis vesicular. Tesis doctoral, Universidad de Valencia.

**Bruggeman, J., A. J. Debets, P. J. Wijngaarden, J. A. deVisser y R. F. Hoekstra.** 2003. Sex slows down the accumulation of deleterious mutations in the homothallic fungus *Aspergillus nidulans*. *Genetics* **164**:479-485.

**Bull, J. J., M. R. Badgett y H. A. Wichman.** 2000. Big-benefit mutations in a bacteriophage inhibited with heat. *Mol. Biol. Evol.* **17**:942-950.

**Bulmer, M.** 1994. *Theoretical Evolutionary Ecology*. Sinauer, Sunderland.

**Burch, C. L. y L. Chao.** 1999. Evolution by small steps and rugged landscapes in the RNA virus  $\phi 6$ . *Genetics* **151**:921-927.

**Burch, C. L., P. E. Turner y K. A. Hanley.** 2003. Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design. *J. Evol. Biol.* **16**:1223-1235.

**Burch, C. L. y L. Chao.** 2004. Epistasis and its relationship to canalization in the RNA virus  $\phi 6$ . *Genetics* **167**:559-567.

**Butcher, D.** 1995. Muller's ratchet, epistasis and mutation effects. *Genetics* **141**:431-437.

**Chao, L.** 1988. Evolution of sex in RNA viruses. *J. Theor. Biol.* **133**:99-112.

**Chao, L.** 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* **348**:454-455.

**Chao, L., C. U. Rang y L. E. Wong.** 2002. Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage  $\phi 6$ . *J. Virol.* **76**:3276-3281.

**Chare, E. R., E. A. Gould y E. C. Holmes.** 2003. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J. Gen. Virol.* **84**:2691-2703.

**Charlesworth, B.** 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**:199-221.

**Charlesworth, B.** 1993. Directional selection and the evolution of sex and recombination. *Genet. Res.* **61**:205-224.



**Charlesworth, B.** 1998. The effect of synergistic epistasis on the inbreeding load. *Genet. Res.* **71**:85-89.

**Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena , E. Domingo y J. J. Holland.** 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J. Virol.* **67**:222-228.

**Cline, J., J. C. Braman y H. H. Hogrefe.** 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* **24**:3546-3551.

**Coil, D. A. y A. D. Miller .** 2004. Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. *J. Virol.* **78**:10920-10926.

**Crill, W. D., H. A. Wichman y J. J. Bull.** 2000. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**:27-37.

**Crotty, S., C. E. Cameron y R. Andino.** 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl. Acad. Sci. U. S. A* **98**:6895-6900.

**Crow, J. F.** 1970. Genetic loads and the cost of natural selection, p. 128-177. *Mathematical topics in Population Genetics.* Springer-Verlag, New York.

**Crow, J. F.** 1994. Advantages of sexual reproduction. *Dev. Genet.* **15**:205-213.

**Cuevas, J. M., S. F. Elena y A. Moya.** 2002. Molecular basis of adaptive convergence in experimental populations of RNA viruses. *Genetics* **162**:533-542.

**Cuevas, J. M., A. Moya y R. Sanjuan.** 2005. Following the very initial growth of biological RNA virus clones. *J. Gen. Virol*, en prensa.

**de Visser, J. A. G. M., R. F. Hoekstra y H. van den Ende.** 1996. The effect of sex and deleterious mutations on fitness in *Chlamydomonas*. *Proc. R. Soc. Lond. B* **263**:193-200.

**de Visser, J. A. G. M., R. F. Hoekstra y H. van den Ende.** 1997. Test of interaction between genetic markers that affect fitness in *Aspergillus nidulans*. *Evolution* **51**:1499-1505.

**de Visser, J. A. G. M., C. Zeyl, P. J. Gerrish, J. L. Blanchard y R. E. Lenski.** 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* **283**:404-406.

**DePolo, N. J., C. Giachetti y J. J. Holland.** 1987. Continuing coevolution of virus and defective interfering particles and of viral genome sequences during

undiluted passages: virus mutants exhibiting nearly complete resistance to formerly dominant defective interfering particles. *J. Virol.* **61**:454-464.

**Dobzhansky, T.** 1936. Studies on hybrid sterility II. Localization of sterility factors in *Drosophila pseudoobscura*. *Genetics* **21**:113-135.

**Domingo, E., E. Martínez-Salas, F. Sobrino, J. C. de la Torre, A. Portela, J. Ortín, C. López-Galíndez, P. Pérez-Brena, N. Villanueva y R. Nájera.** 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance--a review. *Gene* **40**:1-8.

**Domingo, E., C. Escarmís, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. S. Novella y J. J. Holland.** 1996. Basic concepts in RNA virus evolution. *FASEB J.* **10**:859-864.

**Domingo, E. y J. J. Holland.** 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**:151-178.

**Domingo, E.** 2000. Viruses at the edge of adaptation. *Virology* **270**:251-253.

**Drake, J. W., B. Charlesworth, D. Charlesworth y J. F. Crow.** 1998. Rates of spontaneous mutation. *Genetics* **148**:1667-1686.

**Drake, J. W. y J. J. Holland.** 1999. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**:13910-13913.

**Duarte, E., D. Clarke, A. Moya, E. Domingo y J. Holland.** 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc. Natl. Acad. Sci. U. S. A* **89**:6015-6019.

**Dykhuisen, D. E., A. M. Dean y D. L. Hartl.** 1987. Metabolic flux and fitness. *Genetics* **115**:25-31.

**Eggleston, W. B., D. M. Johnson-Schlitz y W. R. Engels.** 1988. P-M hybrid dysgenesis does not mobilize other transposable element families in *D. melanogaster*. *Nature* **331**:368-370.

**Eigen, M.** 1986. The physics of molecular evolution. *Chem. Scrip.* **26B**:13-26.

**Eigen, M., J. McCaskill y P. Schuster.** 1988. Molecular quasi-species. *J. Phys. Chem.* **92**:6881-6891.

**Elena, S. F., V. S. Cooper y R. E. Lenski.** 1996. Punctuated evolution caused by selection of rare beneficial mutations. *Science* **272**:1802-1804.

**Elena, S. F.** 1996. Evolución de la eficacia biológica en poblaciones experimentales del virus de la estomatitis vesicular. Tesis doctoral, Universidad de Valencia.

Elena, S. F. y R. E. Lenski. 1997. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* 390:395-398.

Elena, S. F., L. Ekunwe, N. Hajela, S. A. Oden y R. E. Lenski. 1998. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. *Genetica* 102/103:349-358.

Elena, S. F., M. Dávila, I. S. Novella, J. J. Holland, E. Domingo y A. Moya. 1998. Evolutionary dynamics of fitness recovery from the debilitating effects of Muller's ratchet. *Evolution* 52:309-314.

Elena, S. F. y A. Moya. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. *J. Evol. Biol.* 12:1078-1088.

Elena, S. F., R. Sanjuán, A. V. Bordería y P. E. Turner. 2001. Transmission bottlenecks and the evolution of fitness in rapidly evolving RNA viruses. *Infect. Genet. Evol.* 1:41-48.

Elena, S. F. y J. A. de Visser. 2003. Environmental stress and the effects of mutation. *J. Biol.* 2:12.

Escarmís, C., M. Dávila, N. Charpentier, A. Bracho, A. Moya y E. Domingo. 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. *J. Mol. Biol.* 264:255-267.

Escarmís, C., M. Dávila y E. Domingo. 1999. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* 285:495-505.

Escarmís, C., Gómez-Mariano, G., Dávila, M., Lázaro, E. y Domingo, E. 2002. Resistance to extinction of low fitness virus subjected to plaque-to-plaque transfers: diversification by mutation clustering. *J. Mol. Biol.* 315, 647-661.

Estes, S. y M. Lynch. 2003. Rapid fitness recovery in mutationally degraded lines of *Caenorhabditis elegans*. *Evolution Int. J. Org. Evolution* 57:1022-1030.

Fares, M. A. y E. C. Holmes. 2002. A revised evolutionary history of hepatitis B virus (HBV). *J. Mol. Evol.* 54:807-814.

Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.

Fenster, C. B., L. F. Galloway y L. Chao. 1997. Epistasis and its consequences for the evolution of natural populations. *TREE* 12:282-286.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.

**Francki, R. I. B., C. M. Fauquet, D. D. Knudson y F. Brown.** 1991. Fifth report of the International Committee on Taxonomy of Viruses. *Arch. Virol.* **2**:1-450.

**Funchain, P., A. Yeung, J. L. Stewart, R. Lin, M. M. Slupska y J. H. Miller.** 2000. The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. *Genetics* **154**:959-970.

**Gabriel, W., M. Lynch y R. Bürger.** 1993. Muller's ratchet and mutational meltdown. *Evolution* **47**:1744-1757.

**Gallione, C. J., J. R. Greene, L. E. Iverson y J. K. Rose.** 1981. Nucleotide sequences of the mRNA's encoding the vesicular stomatitis virus N and NS proteins. *J. Virol.* **39**:529-535.

**García Dorado, A.** 1997. The rate and distribution of viability mutation in *Drosophila*: minimum distance estimation. *Evolution* **51**:1130-1139.

**Gerrish, P. J. y R. E. Lenski.** 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* **102-103**:127-144.

**Gillespie, J. H.** 1983. A simple stochastic gene substitution model. *Theor. Popul. Biol.* **23**:202-215.

**Goodman, M. F. y K. D. Fygenon.** 1998. DNA polymerase fidelity: from genetics toward a biochemical understanding. *Genetics* **148**:1475-1482.

**Haihg, J.** 1978. The accumulation of deleterious genes in a population- Muller's ratchet. *Theor. Pop. Biol.* **14**:251-267.

**Haldane, J. B. S.** 1927. A mathematical theory of natural and artificial selection. *Proc. Camb. Phil. Soc.* **23**:838-844.

**Haldane, J. B. S.** 1937. The effect of variation on fitness. *Am. Nat* **37**:337-349.

**Hansen, T. F. y G. P. Wagner.** 2001. Epistasis and the mutation load: a measurement-theoretical approach. *Genetics* **158**:477-485.

**Hill, W. G. y A. Robertson .** 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**:269-294.

**Holland, J. J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol y S. Vandepol.** 1982. Rapid evolution of RNA genomes. *Science* **215**:1577-1585.

**Holland, J. J., E. Domingo , J. C. de la Torre y D. A. Steinhauer.** 1990. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J. Virol.* **64**:3960-3962.

Holland, J. J., J. C. de la Torre, D. K. Clarke y E. Duarte. 1991. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *J. Virol.* **65**:2960-2967.

Hopfield, J. J. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U. S. A* **71**:4135-4139.

Hudson, L. D., C. Condra y R. A. Lazzarini. 1986. Cloning and expression of a viral phosphoprotein: structure suggests vesicular stomatitis virus NS may function by mimicking an RNA template. *J. Gen. Virol.* **67**:1571-1579.

Imhof, M. y C. Schlötterer. 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Proc. Natl. Acad. Sci. USA* **98**:1113-1117.

Jenkins, G. M. y E. C. Holmes. 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. *Virus Res.* **92**:1-7.

Johnson, T. 1999. The approach to mutation-selection balance in an infinite asexual population, and the evolution of mutation rates. *Proc. R. Soc. Lond B Biol. Sci.* **266**:2389-2397.

Johnson, T. y N. H. Barton . 2002. The effect of deleterious alleles on adaptation in asexual populations. *Genetics* **162**:395-411.

Keightley, P. D. 1989. Models of quantitative variation of flux in metabolic pathways. *Genetics* **121**:869-876.

Keightley, P. D. 1994. The distribution of mutation effects on viability in *Drosophila melanogaster*. *Genetics* **138**:1315-1322.

Keightley, P. D. 1996. Nature of deleterious mutation load in *Drosophila*. *Genetics* **144**:1993-1999.

Keightley, P. D. y A. Caballero. 1997. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **94**:3823-3827.

Keightley, P. D. y T. M. Bataillon. 2000. Multigeneration maximum-likelihood analysis applied to mutation-accumulation experiments in *Caenorhabditis elegans*. *Genetics* **154**:1193-1201.

Keightley, P. D. y A. Eyre-Walker. 2000. Deleterious mutations and the evolution of sex. *Science* **290**:331-333.

Keightley, P. D. y M. Lynch. 2003. Toward a realistic model of mutations affecting fitness. *Evolution Int. J. Org. Evolution* **57**:683-685.

**Kibota, T. T. y M. Lynch.** 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* **381**:694-696.

**Kimura, M. y T. Maruyama.** 1966. The mutational load with epistatic gene interactions in fitness. *Genetics* **54**:1337-1351.

**Kimura, M.** 1968. Evolutionary rate at the molecular level. *Nature* **217**:624-626.

**Kimura, M.** 1979. Model of effectively neutral mutations in which selective constraint is incorporated. *Proc. Natl. Acad. Sci. USA* **79**:3440-3444.

**Kimura, M.** 1985. The neutral theory of molecular evolution. *New Scientist*. **11**:41-46.

**Kirby, D. A. y W. Stephan.** 1996. Multi-locus selection and the structure of variation at the white gene of *Drosophila melanogaster*. *Genetics* **144**:635-645.

**Kondrashov, A. S.** 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**:435-440.

**Kondrashov, A. S.** 1994. Muller's ratchet under epistatic selection. *Genetics* **136**:1469-1473.

**Kopecky, S. A., M. C. Willingham y D. S. Lyles.** 2001. Matrix protein and another viral component contribute to induction of apoptosis in cells infected with vesicular stomatitis virus. *J. Virol.* **75**:12169-12181.

**Krakauer, D. C.** 2000. Stability and evolution of overlapping genes. *Evolution Int. J. Org. Evolution* **54**:731-739.

**Krakauer, D. C. y J. B. Plotkin.** 2002. Redundancy, antiredundancy, and the robustness of genomes. *Proc. Natl. Acad. Sci. U. S. A* **99**:1405-1409.

**Lazaro, E., C. Escarmis, J. Perez-Mercader, S. C. Manrubia y E. Domingo.** 2003. Resistance of virus to extinction on bottleneck passages: study of a decaying and fluctuating pattern of fitness loss. *Proc. Natl. Acad. Sci. U. S. A* **100**:10830-10835.

**Lefrancios, L. y D. S. Lyles.** 1982. The interaction of antibody with the major surface glycoprotein of vesicular stomatitis virus. I. Analysis of neutralizing epitopes with monoclonal antibodies. *Virology* **121**:157-167.

**Lefrancois, L. y D. S. Lyles.** 1982. The interaction of antibody with the major surface glycoprotein of vesicular stomatitis virus. II. Monoclonal antibodies of nonneutralizing and cross-reactive epitopes of Indiana and New Jersey serotypes. *Virology* **121**:168-174.

**Lefrancois, L. y D. S. Lyles.** 1983. Antigenic determinants of vesicular stomatitis virus: analysis with antigenic variants. *J. Immunol.* **130**:394-398.

**Lenormand, T. y S. P. Otto .** 2000. The evolution of recombination in a heterogeneous environment. *Genetics* **156**:423-438.

**Lenski, R. E., M. R. Rose, S. C. Simpson y S. C. Tadler.** 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations. *Am. Nat.* **138**:1315-1341.

**Lenski, R. E. y J. E. Mittler.** 1993. The directed mutation controversy and neo-Darwinism. *Science* **259**:188-194.

**Lenski, R. E., C. Ofria, T. C. Collier y C. Adami .** 1999. Genome complexity, robustness and genetic interactions in digital organisms. *Nature* **400**:661-664.

**Lenski, R. E.** 2001. Come fly, and leave the baggage behind. *Science* **294**:533-534.

**Letchworth, G. J., L. L. Rodriguez y J. Del carrera.** 1999. Vesicular stomatitis. *Vet. J.* **157**:239-260.

**Levin, B. R., V. Perrot y N. Walker.** 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* **154**:985-997.

**Lewis, D. A., K. Bebenek, W. A. Beard, S. H. Wilson y T. A. Kunkel.** 1999. Uniquely altered DNA replication fidelity conferred by an amino acid change in the nucleotide binding pocket of human immunodeficiency virus type 1 reverse transcriptase. *J. Biol. Chem.* **274**:32924-32930.

**Li, H. y M. J. Roossinck.** 2004. Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J. Virol.* **78**:10582-10587.

**Lichty, B. D., A. T. Power , D. F. Stojdl y J. C. Bell.** 2004. Vesicular stomatitis virus: re-inventing the bullet. *Trends Mol. Med.* **10**:210-216.

**Limbirt, E., W. A. Stahel y M. Abbt.** 2001. Log-normal distributions across sciences: keys and clues. *BioScience* **51**:341-352.

**Lundberg, K. S., D. D. Shoemaker, M. W. Adams, J. M. Short, J. A. Sorge y E. J. Mathur.** 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**:1-6.

**Luria, S. E. y M. Delbrück.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**:491-511.

**Lyman, R. F., F. Lawrence, S. V. Nuzhdin y T. F. C. Mackay.** 1996. Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**:277-292.

**Lynch, M., R. Bürger, D. Butcher y W. Gabriel.** 1993. The mutational meltdown in asexual populations. *J. Heredity* **84**:339-344.

**Mackett, M., G. L. Smith y B. Moss.** 1982. Vaccinia virus: a selectable eukaryotic cloning and expression vector. *Proc. Natl. Acad. Sci. U. S. A* **79**:7415-7419.

**McKeithan, T. W.** 1995. Kinetic proofreading in T-cell receptor signal transduction. *Proc. Natl. Acad. Sci. U. S. A* **92**:5042-5046.

**Miralles, R., P. J. Gerrish, A. Moya y S. F. Elena.** 1999. Clonal interference and the evolution of RNA virus. *Science* **285**:1745-1747.

**Miralles, R., A. Moya y S. F. Elena.** 2000. Diminishing returns of population size in the rate of RNA virus adaptation. *J. Virol.* **74**:3566-3571.

**Moore, F. B., D. E. Rozen y R. E. Lenski.** 2000. Pervasive compensatory adaptation in *Escherichia coli*. *Proc. R. Soc. Lond B Biol. Sci.* **267**:515-522.

**Moreno, I. M., J. M. Malpica, E. Rodríguez-Cerezo y F. García-Arenal.** 1997. A mutation in tomato aspermy cucumovirus that abolishes cell-to-cell movement is maintained to high levels in the viral RNA population by complementation. *J. Virol.* **71**:9157-9162.

**Morita, K., R. Vanderoef y J. Lenard.** 1987. Phenotypic revertants of temperature-sensitive M protein mutants of vesicular stomatitis virus: sequence analysis and functional characterization. *J. Virol.* **61**:256-263.

**Mukai, T.** 1964. The genetic structure of natural populations of *Drosophila melanogaster*. I. spontaneous mutation rate of polygenes controlling viability. *Genetics* **30**:1-19.

**Mukai, T., S. I. Chigusa, L. E. Mettler y J. F. Crow.** 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**:335-355.

**Muller, H. J.** 1939. Reversibility of evolution considered from the standpoint of genetics. *Biol. Rev. Camb. Philos. Soc.* **14**:261-280.

**Muller, H. J.** 1964. The relation of recombination to mutational advance. *Mut. Res.* **1**:2-9.

**Neumann, A. U., N. P. Lam, H. Dahari, D. R. Gretch, T. E. Wiley, T. J. Layden y A. S. Perelson.** 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon- $\alpha$  therapy. *Science* **282**:103-107.



- Nickels, M. S. y D. M. Hunt.** 1994. Identification of an amino acid change that affects N protein function in vesicular stomatitis virus. *J. Gen. Virol.* **75**:3591-3595.
- Nilsson, M. y N. Snoad.** 2000. Error thresholds for quasispecies on dynamic fitness landscapes. *Phys. Rev. Lett.* **84**:191-194.
- Novella, I. S., C. L. Hershey, C. Escarmis, E. Domingo y J. J. Holland.** 1999. Lack of evolutionary stasis during alternating replication of an arbovirus in insect and mammalian cells. *J. Mol. Biol.* **287**:459-465.
- Novella, I. S., J. Quer, E. Domingo y J. J. Holland.** 1999. Exponential fitness gains of RNA virus populations are limited by bottleneck effects. *J. Virol.* **73**:1668-1671.
- Novella, I. S.** 2003. Contributions of vesicular stomatitis virus to the understanding of RNA virus evolution. *Curr. Opin. Microbiol.* **6**:399-405.
- Novella, I. S. y Ebendick-Corp.** 2004. Molecular basis of fitness loss and fitness recovery in vesicular stomatitis virus. *J. Mol. Biol.* **342**:1423-1430.
- Novella, I. S., S. Zarate, D. Metzgar y Ebendick-Corp.** 2004. Positive selection of synonymous mutations in vesicular stomatitis virus. *J. Mol. Biol.* **342**:1415-1421.
- Ohnishi, O.** 1977. Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effect of polygenic mutants. *Genetics* **87**:529-545.
- Ohta, T.** 1992. The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* **23**:263-286.
- Orr, H. A.** 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* **52**:935-949.
- Orr, H. A.** 2000. The rate of adaptation in asexuals. *Genetics* **155**:961-968.
- Orr, H. A.** 2002. The population genetics of adaptation: the adaptation of DNA sequences. *Evolution Int. J. Org. Evolution* **56**:1317-1330.
- Orr, H. A.** 2003. The distribution of fitness effects among beneficial mutations. *Genetics* **163**:1519-1526.
- Otto, S. P. y M. W. Feldman.** 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. *Theor. Popul. Biol.* **51**:134-147.
- Otto, S. P. y T. Lenormand .** 2002. Resolving the paradox of sex and recombination. *Nat Rev. Genet.* **3**:252-261.

**Pattnaik, A. K., L. A. Ball, A. W. LeGrone y G. W. Wertz.** 1992. Infectious defective interfering particles of VSV from transcripts of a cDNA clone. *Cell* **69**:1011-1020.

**Peck, J. R.** 1994. A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. *Genetics* **137**:597-606.

**Peck, J. R. y D. Waxman.** 2000. Mutation and sex in a competitive world. *Nature* **406**:399-404.

**Pfeiffer, J. K. y K. Kirkegaard.** 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc. Natl. Acad. Sci. U. S. A* **100**:7289-7294.

**Pianka, E. R.** 1970. On *r*- and *K*-selection. *Am. Nat.* **104**:592-597.

**Pugachev, K. V., F. Guirakhoo, S. W. Ocran, F. Mitchell, M. Parsons, C. Penal, S. Girakhoo, S. O. Pougatcheva, J. Arroyo, D. W. Trent y T. P. Monath.** 2004. High fidelity of yellow fever virus RNA polymerase. *J. Virol.* **78**:1032-1038.

**Quer, J., C. L. Hershey, E. Domingo, J. J. Holland y I. S. Novella.** 2001. Contingent neutrality in competing viral populations. *J. Virol.* **75**:7315-7320.

**Rice, W. R. y A. K. Chippindale.** 2001. Sexual recombination and the power of natural selection. *Science* **294**:555-559.

**Rodriguez, L. L., T. A. Bunch, M. Fraire y Z. N. Llewellyn.** 2000. Re-emergence of vesicular stomatitis in the western United States is associated with distinct viral genetic lineages. *Virology* **271**:171-181.

**Rokyta, D., M. R. Badgett, I. J. Molineux y J. J. Bull.** 2002. Experimental genomic evolution: extensive compensation for loss of DNA ligase activity in a virus. *Mol. Biol. Evol.* **19**:230-238.

**Rozen, D. E., J. A. de Visser y P. J. Gerrish.** 2002. Fitness effects of fixed beneficial mutations in microbial populations. *Curr. Biol.* **12**:1040-1045.

**Ruusala, T., M. Ehrenberg y C. G. Kurland.** 1982. Is there proofreading during polypeptide synthesis? *EMBO J.* **1**:741-745.

**Schrag, S. J., V. Perrot y B. R. Levin.** 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proc. R. Soc. Lond. B* **264**:1287-1291.

**Schubert, M., G. G. Harmison y E. Meier.** 1984. Primary structure of the vesicular stomatitis virus polymerase (L) gene: evidence for a high frequency of mutations. *J. Virol.* **51**:505-514.

**Schuster, P.** 1986. The physical basis of molecular evolution. *Chemica. Scripta.* **26B**:27-41.

Schuster, P. y J. Swetina. 1988. Stationary mutant distributions and evolutionary optimization. *Bull. Math. Biol.* **50**:635-660.

Seager, R. D. y F. J. Ayala. 1982. Chromosome interactions in *Drosophila melanogaster*. I. viability studies. *Genetics* **102**:467-483.

Sebat, J., B. Lakshmi, J. Troge, J. Alexander, J. Young, P. Lundin, S. Maner, H. Massa, M. Walker, M. Chi, N. Navin, R. Lucito, J. Healy, J. Hicks, K. Ye, A. Reiner, T. C. Gilliam, B. Trask, N. Patterson, A. Zetterberg y M. Wigler. 2004. Large-scale copy number polymorphism in the human genome. *Science* **305**:525-528.

Shaw, R. G., D. L. Byers y E. Darms. 2000. Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. *Genetics* **155**:369-378.

Shnol, E. E. y A. S. Kondrashov. 1993. The effect of selection on the phenotypic variance. *Genetics* **134**:995-996.

Smith, J. M. 1971. What use is sex? *J. Theor. Biol.* **30**:319-335.

Smith, J. M. y J. Haigh. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**:23-35.

Smith, V., K. N. Chou, D. Lashkari, D. Botstein y P. O. Brown. 1996. Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science* **274**:2069-2074.

Sniegowski, P. D., P. J. Gerrish y R. E. Lenski. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**:703-705.

Steinhauer, D. A., J. C. de la Torre y J. J. Holland. 1989. High nucleotide substitution error frequencies in clonal pools of vesicular stomatitis virus. *J. Virol.* **63**:2063-2071.

Steinhauer, D. A., E. Domingo y J. J. Holland. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* **122**:281-288.

Stephan, W., L. Chao y J. G. Smale. 1993. The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion. *Genet. Res., Camb.* **61**:225-231.

Szathmary, E. 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. *Genetics* **133**:127-132.

Taddei, F., M. Radman, J. Maynard-Smith, B. Toupance, P. H. Gouyon y B. Godelle. 1997. Role of mutator alleles in adaptive evolution. *Nature* **387**:700-702.

van Nimwegen, E., J. P. Crutchfield y M. Huynen. 1999. Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. U. S. A* **96**:9716-9720.

Vandepol, S. B., L. Lefrancois y J. J. Holland. 1986. Sequences of the major antibody binding epitopes of the Indiana serotype of vesicular stomatitis virus. *Virology* **148**:312-325.

Wagner, A. 2000. Robustness against mutations in genetic networks of yeast. *Nature Genet.* **24**:355-361.

Wagner, G. P., M. D. Laubichler y H. Bagheri-Chaichian. 1998. Genetic measurement of theory of epistatic effects. *Genetica* **102-103**:569-580.

Wang, Y., P. N. Schnegelsberg, J. Dausman y R. Jaenisch. 1996. Functional redundancy of the muscle-specific transcription factors Myf5 and myogenin. *Nature* **379**:823-825.

Whelan, S. P., L. A. Ball, J. N. Barr y G. T. Wertz. 1995. Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones. *Proc. Natl. Acad. Sci. U. S. A* **92**:8388-8392.

Whitlock, M. C. y D. Bourguet. 2000. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution Int. J. Org. Evolution* **54** :1654-1660.

Whitlock, M. C. 2000. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution Int. J. Org. Evolution* **54**:1855-1861.

Whitlock, M. C., C. K. Griswold y A. D. Peters. 2003. Compensating for the meltdown: the critical effective population size of a population with deleterious and compensatory mutations. *Ann Zool Fennici* **40**:169-183.

Wilke, C. O., J. L. Wang, C. Ofria, R. E. Lenski y C. Adami. 2001. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412**:331-333.

Wilke, C. O. y C. Adami. 2001. Interaction between directional epistasis and average mutational effects. *Proc. R. Soc. Lond B Biol. Sci.* **268**:1469-1474.

Wilke, C. O. 2003. Probability of fixation of an advantageous mutant in a viral quasispecies. *Genetics* **163**:467-474.

Wilke, C. O. y I. S. Novella. 2003. Phenotypic mixing and hiding may contribute to memory in viral quasispecies. *BMC. Microbiol.* **3**:11.

Wilke, C. O. 2004. The speed of adaptation in large asexual populations. *Genetics* **167**:2045-2053.

**Wisniewski, M., C. Palaniappan, Z. Fu, S. F. Le Grice, P. Fay y R. A. Bambara.** 1999. Mutations in the primer grip region of HIV reverse transcriptase can increase replication fidelity. *J. Biol. Chem.* **274**:28175-28184.

**Wolf, J. B., E. D. Brodie III y M. J. Wade.** 2000. Epistasis and the evolutionary process. Oxford University Press, Oxford.

**Wright, S.** 1931. Evolution in mendelian populations. *Genetics* **16**:97-159.

**Wright, S.** 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Intern. Congr. Genet.* 356-366.

**Wright, S.** 1982. The shifting balance theory and macroevolution. *Annu. Rev. Genet.* **16**:1-19.

**Yan, J., M. O. Magnasco y J. F. Marko.** 1999. A kinetic proofreading mechanism for disentanglement of DNA by topoisomerases. *Nature* **401**:932-935.

**Yang, Z., R. Nielsen, N. Goldman y A. M. Pedersen.** 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**:431-449.

**You, L. y J. Yin.** 2002. Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage t7. *Genetics* **160**:1273-1281.

**Yuste, E., S. Sánchez-Palomino, C. Casado, E. Domingo y C. López-Galíndez.** 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *J. Virol.* **73**:2745-2751.

**Yuste, E., C. López-Galíndez y E. Domingo.** 2000. Unusual distribution of mutations associated with serial bottleneck passages of human immunodeficiency virus type-1. *J. Virol.* **74**:9546-9552.





## **APÉNDICES**





Posición	Gen	Codón	Referencia
406T	N	115	Novella <i>et al.</i> (1999) <sup>a</sup>
447A	N	129	De Polo <i>et al.</i> (1987) <sup>b</sup>
527G	N	155	Nickels <i>et al.</i> (1994) <sup>c</sup>
904C	N	282	Nickels <i>et al.</i> (1994) <sup>c</sup>
1544A	P	50	Gallione <i>et al.</i> (1981) <sup>d</sup>
1622C	P	76	Gallione <i>et al.</i> (1981) <sup>d</sup>
1663A	P	90	Rodríguez <i>et al.</i> (2000) <sup>e</sup>
1724C	P	110	Gallione <i>et al.</i> (1981) <sup>d</sup>
1820C	P	142	Cuevas <i>et al.</i> (2002) <sup>f</sup>
1846T	P	151	Hudson <i>et al.</i> (1986) <sup>g</sup>
1879G	P	162	Rodríguez <i>et al.</i> (2000) <sup>e</sup>
2033A	P	213	Rodríguez <i>et al.</i> (2000) <sup>e</sup>
2104G	P	237	Gallione <i>et al.</i> (1981) <sup>d</sup>
2296C	M	16	Cuevas <i>et al.</i> (2002) <sup>f</sup>
2344A	M	32	Morita <i>et al.</i> (1987) <sup>d</sup>
2409C	M	54	Morita <i>et al.</i> (1987) <sup>g</sup>
2607T	M	120	Novella <i>et al.</i> (1999) <sup>a</sup>
2608A	M	120	Cuevas <i>et al.</i> (2002) <sup>f</sup>
2617G	M	123	Novella <i>et al.</i> (1999) <sup>a</sup>
2743A	M	165	Cuevas <i>et al.</i> (2002) <sup>f</sup>
2892G	M	215	Morita <i>et al.</i> (1987) <sup>h</sup>
2936C	M	230	Novella <i>et al.</i> (1999) <sup>a</sup>
3120T	G	15	Bilsel <i>et al.</i> (1990) <sup>i</sup>
3183C	G	36	Bilsel <i>et al.</i> (1990) <sup>i</sup>
3277C	G	67	Novella <i>et al.</i> (1999) <sup>a</sup>
3441G	G	122	Novella <i>et al.</i> (1999) <sup>a</sup>
3709C	G	211	Cuevas <i>et al.</i> (2002) <sup>f</sup>
3802G	G	242	Cuevas <i>et al.</i> (2002) <sup>f</sup>
4111A	G	345	Cuevas <i>et al.</i> (2002) <sup>f</sup>
4179G	G	368	Cuevas <i>et al.</i> (2002) <sup>f</sup>
4394G	G	440	Cuevas <i>et al.</i> (2002) <sup>f</sup>
5202G	L	157	Cuevas <i>et al.</i> (2002) <sup>f</sup>
5733C	L	334	Cuevas <i>et al.</i> (2002) <sup>f</sup>
6131C	L	467	Schubert <i>et al.</i> (1984) <sup>j</sup>
6376C	L	549	Schubert <i>et al.</i> (1984) <sup>j</sup>
6741C	L	670	Schubert <i>et al.</i> (1984) <sup>j</sup>
7461C	L	910	Schubert <i>et al.</i> (1984) <sup>j</sup>
7808C	L	1026	Schubert <i>et al.</i> (1984) <sup>j</sup>
8775C	L	1348	Schubert <i>et al.</i> (1984) <sup>j</sup>
9278G	L	1516	Cuevas <i>et al.</i> (2002) <sup>f</sup>
10739G	L	2003	Schubert <i>et al.</i> (1984) <sup>k</sup>
10809C	L	2026	Schubert <i>et al.</i> (1984) <sup>k</sup>
10892G	L	2054	Schubert <i>et al.</i> (1984) <sup>k</sup>

**Apéndice 1.** Localización, gen y codón afectado, cambio aminoacídico y referencia bibliográfica de cada uno de los 44 cambios previamente descritos que hemos elegido para ser introducidos en el clon infeccioso. El superíndice tras cada referencia conduce al Apéndice 2.

<b>Código</b>	<b>Aislado</b>	<b>Origen</b>	<b>Información adicional disponible</b>
<i>a</i>	Consenso	Laboratorio	Pasado alternativamente en células de mamífero (BHK21) y de insecto (LL-5). Ceba Mudd-Summers, en un régimen típico de ganancia de eficacia biológica.
<i>b</i>	Consenso	Laboratorio	Aislado a alta m.o.i., resistente a partículas defectivas interferentes. Ceba Glasgow.
<i>c</i>	Consenso	Laboratorio	Presente en mutantes termosensibles ( <i>ts</i> ).
<i>d</i>	Clon	Laboratorio	Ceba San Juan.
<i>e</i>	Consenso	Natural	Focos de Veracruz, Nuevo Méjico, Colorado y Utah (1997-1998) en ganado bovino y equino.
<i>f</i>	Consenso	Laboratorio	Pases masivos en células BHK, en un régimen típico de ganancia de eficacia biológica.
<i>g</i>	Consenso	Laboratorio	Virus silvestre.
<i>h</i>	Clon	Laboratorio	Encontrado <i>in vivo</i> en ratones y en virus mutantes <i>ts</i> .
<i>i</i>	Consenso	Natural	Norte América
<i>j</i>	Clon	Laboratorio	Ceba San Juan. Posibles mutantes deletéreos o errores de la <i>Taq</i> polimerasa en el protocolo de PCR para la secuenciación.
<i>k</i>	Clon	Laboratorio	Ceba San Juan o Mudd-Summers. Posibles mutantes deletéreos o errores de la <i>Taq</i> polimerasa en el protocolo de PCR para la secuenciación.

**Apéndice 2.** Contexto en el que fueron descritas las 43 mutaciones elegidas para ser introducidas en el clon infeccioso. El código se refiere a las referencias citadas en el Apéndice 1

Gen	Sitio mutado	Cambio	Aleatorio	Eficacia ( $w_i$ )	
N	112C	K17Q	sí	Letal	
	406T	G115C	no	$1.165 \pm 0.042$ *	
	447A	S129R	no	$0.595 \pm 0.032$ **	
	527G	K155R	no	$0.967 \pm 0.063$	
	636A	Y192Stop	sí	Letal	
	877C	D272H	sí	$0.537 \pm 0.059$ *	
	904C	F282L	no	$1.172 \pm 0.056$ *	
	906C	sinónimo	sí	$1.268 \pm 0.094$ *	
	1044C	sinónimo	sí	$1.010 \pm 0.037$	
	1073G	A337G	sí	$0.029 \pm 0.030$ **	
	1323A	F421L	sí	$0.040 \pm 0.012$ **	
	P	1544A	R50K	no	$1.151 \pm 0.073$
		1585A	S64T	sí	$0.879 \pm 0.042$ *
1622C		V76A	no	$1.275 \pm 0.102$ *	
1663A		G90R	no	$0.929 \pm 0.083$	
1724C		Q110P	no	$0.939 \pm 0.065$	
1820C		I142T	no	$0.374 \pm 0.036$ **	
1846T		H151Y	no	$0.930 \pm 0.044$	
1879G		S162A	no	$0.958 \pm 0.080$	
2018T		Q208L	sí	$0.867 \pm 0.053$ *	
2033A		S213N	no	$1.027 \pm 0.105$	
M		2104G	N237D	no	Letal
		2296C	K16T	no	$0.908 \pm 0.054$
		2344A	S32N	no	$1.137 \pm 0.047$ *
	2409C	Y54H	no	Letal	
	2607T	P120A	no	$1.354 \pm 0.189$	
	2608A	P120G	no	$0.433 \pm 0.027$ **	
	2613C	V122L	sí	$1.014 \pm 0.067$	
	2617G	L123W	no	$1.893 \pm 0.240$ *	
	2743A	G165D	no	$0.890 \pm 0.020$ *	
	2767C	L173P	sí	Letal	
	2877T	sinónimo	sí	Letal	
	2892G	K215E	no	$0.954 \pm 0.043$	
	2936C	K230N	no	$0.981 \pm 0.074$	
Intergénico	2969T	intergénico	sí	$0.964 \pm 0.071$	
	2969A	intergénico	sí	$0.839 \pm 0.031$ *	
G	3079A	M1K	sí	Letal	
	3120T	N15Y	no	$0.719 \pm 0.079$ *	
	3183C	N36H	no	$0.570 \pm 0.053$ *	
	3192C	Y39H	sí	$0.137 \pm 0.011$ **	
	3277T	A67V	no	$0.668 \pm 0.041$ *	
	3327G	C84G	sí	Letal	
	3441G	N122D	no	$0.344 \pm 0.021$ **	
	3446C	Q124P	sí	$0.876 \pm 0.081$	
	3615A	S180T	sí	$0.633 \pm 0.056$ *	
	3625C	W183S	sí	Letal	
	3688C	F204S	sí	Letal	
	3709C	L211P	no	$0.405 \pm 0.043$ **	



	3802G	H242R	no	1.980 ± 0.246 *
	3950G	sinónimo	sí	0.732 ± 0.046 *
	4111A	R345K	no	Letal
	4163C	M363I	sí	Letal
	4175G	sinónimo	sí	1.090 ± 0.087
	4179G	T368A	no	0.360 ± 0.028 **
	4394G	sinónimo	no	0.886 ± 0.057
	4398C	F441L	sí	0.630 ± 0.061 *
	4459T	W461L	sí	0.121 ± 0.037
L	4824A	M31K	sí	Letal
	4984G	N85K	sí	1.099 ± 0.122
	5124A	I131N	sí	0.427 ± 0.035 **
	5202G	L157W	no	1.021 ± 0.139
	5496A	L255Q	sí	1.030 ± 0.073
	5597C	V289L	sí	1.130 ± 0.046 *
	5733C	F334S	no	0.736 ± 0.027 **
	6010C	sinónimo	sí	0.823 ± 0.053 *
	6131C	D467H	no	Letal
	6376C	L549F	no	0.360 ± 0.048**
	6676A	Y649Stop	sí	Letal
	6741C	R670P	no	0.624 ± 0.034 **
	6907T	K726N	sí	Letal
	7123C	sinónimo	sí	0.829 ± 0.072 *
	7128T	C799F	sí	0.254 ± 0.032 **
	7287T	P852L	sí	0.296 ± 0.017 **
	7461C	F910S	no	0.350 ± 0.049 **
	7739A	L1003I	sí	0.848 ± 0.036 *
	7808C	A1026P	no	Letal
	7918C	sinónimo	sí	0.900 ± 0.090
	8188A	Y1153Stop	sí	Letal
	8374C	K1215N	sí	Letal
	8426C	S1232P	sí	Letal
	8459T	S1243C	sí	0.638 ± 0.047 *
	8775C	G1348A	no	0.697 ± 0.052 *
	8787T	N1352I	sí	0.822 ± 0.124
	9278G	I1516V	no	0.717 ± 0.057 *
	9315C	L1528R	sí	Letal
	9535C	W1602C	sí	Letal
	10595A	P1955T	sí	Letal
	10739G	I2003V	no	0.836 ± 0.062 *
	10809C	I2026T	no	0.409 ± 0.031**
	10892G	K2054E	no	0.443 ± 0.057 **
	10941C	W2070S	sí	Letal

**Apéndice 3.** Posición y cambio nucleotídico, cambio aminoacídico y eficacia biológica para cada uno de los mutantes simples creados. Se muestra el valor medio de eficacia y su error asociado. \*: significativamente diferente de  $w_i=1$  sin corrección por pruebas múltiples. \*\*: significativamente diferente de  $w_i=1$  utilizando la corrección de Bonferroni.

Mutación 1	Mutación 2	Eficacia observada	Efecto epistático
877C	7123C	0.433 ± 0.068	0.039 ± 0.079
1044C	2969T	1.418 ± 1.003	0.315 ± 0.655
1044C	3615A	0.528 ± 0.107	-0.047 ± 0.169
1044C	7128T	0.114 ± 0.012	-0.041 ± 0.032
1323A	906C	0.068 ± 0.009	0.019 ± 0.007*
1323A	7123C	0.040 ± 0.003	0.009 ± 0.003*
1585A	4398C	0.346 ± 0.044	-0.118 ± 0.095
1585A	4984G	0.671 ± 0.067	-0.083 ± 0.151
2018T	906C	0.846 ± 0.081	0.125 ± 0.063*
2018T	5124A	0.356 ± 0.044	0.205 ± 0.029**
2613C	3950G	1.000 ± 0.082	0.440 ± 0.082**
2969A	6010C	1.084 ± 0.094	0.567 ± 0.063**
2969A	7123C	1.089 ± 0.095	0.594 ± 0.091**
2969A	7128T	0.093 ± 0.013	-0.022 ± 0.018
2969A	7739A	1.270 ± 0.383	0.810 ± 0.247*
2969T	3950G	0.780 ± 0.093	0.197 ± 0.095*
2969T	7739A	1.142 ± 0.179	0.429 ± 0.132*
3192C	2969A	0.190 ± 0.077	0.099 ± 0.049*
3192C	3446C	0.107 ± 0.010	-0.017 ± 0.015
3615A	877C	0.209 ± 0.013	-0.048 ± 0.041
3615A	3192C	0.027 ± 0.005	-0.037 ± 0.011*
3615A	7123C	0.496 ± 0.067	0.122 ± 0.052*
4175G	2613C	Letal	
4175G	2969A	0.951 ± 0.110	0.368 ± 0.112*
4459T	5597C	0.087 ± 0.008	-0.009 ± 0.021*
4459T	7128T	Letal	
4459T	7918C	0.077 ± 0.009	0.008 ± 0.016
T4984G	2613C	0.667 ± 0.053	-0.074 ± 0.078
5124A	3950G	0.343 ± 0.027	0.198 ± 0.021**
5496A	8459T	Letal	
6010C	906C	1.138 ± 0.447	0.173 ± 0.247
6010C	1585A	0.428 ± 0.044	-0.285 ± 0.093*
6010C	5597C	1.012 ± 0.220	0.157 ± 0.146
7123C	4459T	0.423 ± 0.035	0.331 ± 0.029**
7123C	4984G	0.918 ± 0.083	0.281 ± 0.164
7128T	3950G	0.082 ± 0.005	-0.013 ± 0.013
7128T	5496A	0.061 ± 0.006	-0.049 ± 0.012*
7128T	8787T	0.081 ± 0.004	-0.022 ± 0.013
7287T	1585A	0.165 ± 0.023	0.013 ± 0.025
7287T	3446C	0.088 ± 0.006	-0.073 ± 0.039*
7287T	3615A	0.054 ± 0.005	-0.032 ± 0.024
7287T	3950G	0.224 ± 0.039	0.118 ± 0.030*
7739A	2018T	1.074 ± 0.090	0.623 ± 0.063**
7739A	3446C	0.977 ± 0.120	0.276 ± 0.111*
7739A	8787T	1.080 ± 0.224	0.546 ± 0.157*
8459T	6010C	0.736 ± 0.195	0.362 ± 0.112*
8787T	1323A	0.044 ± 0.013	0.025 ± 0.007*

**Apéndice 4.** Eficacia de los mutantes dobles aleatorios y valor del efecto epistático con sus errores asociados. \*: significativamente diferente de  $\varepsilon=0$  sin corrección por pruebas múltiples. \*\*: significativamente diferente de  $\varepsilon=0$  utilizando la corrección de Bonferroni.

Mutación 1	Mutación 2	Eficacia observada	Efecto epistático
904C	406T	0.734 ± 0.076	-0.388 ± 0.096**
2344A	406T	1.034 ± 0.067	-0.700 ± 0.130**
2607T	406T	0.723 ± 0.051	-0.683 ± 0.087**
2617G	406T	1.169 ± 0.105	-0.718 ± 0.114**
3802G	406T	0.940 ± 0.050	-0.945 ± 0.182**
2344A	904C	0.603 ± 0.055	-0.417 ± 0.092**
2607T	904C	0.672 ± 0.080	-0.175 ± 0.107
2617G	904C	1.062 ± 0.071	-1.026 ± 0.151**
3802G	904C	1.512 ± 0.176	-0.143 ± 0.219
2607T	2344A	1.194 ± 0.146	0.220 ± 0.158
2617G	2344A	1.636 ± 0.386	-0.257 ± 0.309
3802G	2344A	2.202 ± 0.259	-0.087 ± 0.404
2617G	2607T	0.648 ± 0.057	-1.491 ± 0.244**
3802G	2607T	1.824 ± 0.457	-0.162 ± 0.389
3802G	2617G	2.065 ± 0.330	-1.074 ± 0.452*

**Apéndice 5.** Eficacia de los mutantes dobles que combinan mutaciones de efecto beneficioso y valor del efecto epistático. \*: significativamente diferente de  $\varepsilon=0$  sin corrección por pruebas múltiples. \*\*: significativamente diferente de  $\varepsilon=0$  utilizando la corrección de Bonferroni.

Mutación 1	Mutación 2	Eficacia ancestral	Ganancia de eficacia
1323A	-	0.040 ± 0.012	7.989 ± 0.431 9.910 ± 0.360
3192C	-	0.137 ± 0.011	7.721 ± 0.590 9.155 ± 0.267
4459T	-	0.121 ± 0.365	6.217 ± 0.204 7.735 ± 0.197
3709C	-	0.405 ± 0.043	5.926 ± 0.080 7.528 ± 0.240
6376C	-	0.360 ± 0.048	6.052 ± 0.095 6.379 ± 0.094
10809C	-	0.409 ± 0.031	6.428 ± 0.155 6.652 ± 0.110
7128T	8787T	0.081 ± 0.004	6.494 ± 0.119 7.458 ± 0.273
8787T	1323A	0.044 ± 0.013	6.370 ± 0.138 9.952 ± 0.172
7128T	5496A	0.061 ± 0.006	6.689 ± 0.138 8.450 ± 0.468
1585A	4398C	0.346 ± 0.044	6.309 ± 0.061 7.085 ± 0.248
2018T	5124A	0.356 ± 0.044	6.189 ± 0.209 6.107 ± 0.085
6010C	1585A	0.428 ± 0.044	6.728 ± 0.302 6.219 ± 0.111

**Apéndice 6.** Eficacia ancestral y ganancia de eficacia de los mutantes elegidos para el experimento de evolución. Para cada eficacia ancestral, se indica la ganancia de eficacia media y su error asociado para los tamaños poblacionales  $N_e = 2 \times 10^2$  (arriba) y  $N_e = 2 \times 10^4$  (debajo).





