

VNIVERSITAT
E VALÈNCIA



Facultad de
Ciencias Biológicas

DEPARTAMENTO DE GENÉTICA

INSTITUTO DE BIOLOGÍA INTEGRATIVA DE SISTEMAS (UV-CSIC)

Programa de Doctorado de Biomedicina y Biotecnología

REGULACIÓN DE LA TASA DE MUTACIÓN EN VIRUS DE DNA



TESIS DOCTORAL

JENNIFER RISSO BALLESTER

Director: RAFAEL SANJUÁN VERDEGUER

Septiembre 2019



UNIVERSITAT DE VALÈNCIA

INSTITUTO DE BIOLOGÍA INTEGRATIVA DE SISTEMAS- I2Sysbio

REGULACIÓN DE LA TASA DE MUTACIÓN EN VIRUS DE DNA

JENNIFER RISSO BALLESTER

Septiembre 2019

Programa de Doctorado en Biomedicina y Biotecnología

Director: **Rafael Sanjuán Verdeguer**



El Dr. Rafael Sanjuán Verdeguer, profesor e investigador titular de la Universitat de València, CERTIFICA Que D.^a Jennifer Risso Ballester ha realizado en el Laboratorio de Evolución Experimental en Virus del Instituto de Biología Integrativa de Sistemas de la Universitat de València y el Consejo superior de Investigaciones Científicas (CSIC), bajo su dirección, el trabajo experimental que ha llevado a la redacción de la presente memoria de Tesis Doctoral, titulada "Regulación de la tasa de mutación en virus de DNA". Revisado el trabajo, autorizo su presentación para optar al grado de Doctor.

A handwritten signature in blue ink, consisting of a large, stylized loop that encircles the name.

Firmado: Rafael Sanjuán Verdeguer.

Yo, Jennifer Risso Ballester, declaro que soy autora del presente trabajo de investigación y que lo he realizado en el Laboratorio de Evolución Experimental en Virus del Instituto de Biología Integrativa de Sistemas, bajo la dirección del Dr. Rafael Sanjuán Verdeguer, profesor e investigador titular de la Universitat de València.

Firmado: Jennifer Risso Ballester

A handwritten signature in blue ink, written in a cursive style.

A mis padres

ÍNDICE

PRÓLOGO	7
INTRODUCCIÓN	13
1. LOS VIRUS Y SU EVOLUCIÓN.....	15
1.1 Mecanismos evolutivos.....	15
1.2 Tasas de mutación y fuentes de variación genética en virus .	23
1.3 Importancia de estimar las tasas de mutación en virus.....	33
2. MODELO EXPERIMENTAL: ADENOVIRUS.....	40
2.1 Clasificación	40
2.2 Estructura del virión	41
2.3 Estructura y organización del genoma	43
2.4 Ciclo de infección	46
2.4.1 Unión, entrada y tráfico intracelular	46
2.4.2 Replicación del DNA y regulación de la expresión génica viral .	49
2.4.3 Expresión de genes tardía y ensamblaje de nuevos viriones	50
3. LOS VIRUS Y EL SISTEMA DE REPARACIÓN DE DNA	52
3.1 El sistema de reparación de DNA celular: vías de señalización	52
3.2 Activación e inhibición del sistema de reparación de DNA celular en respuesta a la infección.....	56
OBJETIVOS	63
METODOLOGÍA.....	67
1. Bases de datos de secuencias del género <i>Mastadenovirus</i>	69
2. Alineamiento de secuencias	70
3. Árbol filogenético.....	70
4. Recombinación y selección	71
5. Diversidad genética	72
6. Células	74
7. Virus	75
8. Infección en medio líquido	75
9. Stock viral.....	76
10. Titulación en placa	77

11. Dilución límite.....	78
12. Cálculo del número de partículas virales infecciosas producidas por célula (“ <i>burst size</i> ”).....	79
13. Purificación de DNA viral	80
13.1 Método de purificación utilizando un detergente suave.....	81
13.2 Método de purificación utilizando DNAsa I.....	82
14. Estimación del ratio DNA viral:celular.....	83
15. Secuenciación dúplex (DS).....	85
16. PCR cuantitativa a tiempo real	87
17. Análisis de la relación entre tasa de mutación y diversidad <i>in vivo</i>	89
18. Activación e inhibición de la DDR: compuestos químicos.....	90
19. Ensayos de citotoxicidad	91
20. Inmunofluorescencia	92
21. Evolución experimental.....	94
22. Curvas de crecimiento	96
23. Identificación de deleciones: métodos no convencionales.....	96

RESULTADOS Y DESARROLLO ARGUMENTAL99

1. Estudio de la presión de selección y la diversidad genética en el género <i>Mastadenovirus</i>.....	101
1.1 Banco de datos de secuencias y filogenia.....	103
1.2 Diversidad en el género <i>Mastadenovirus</i>	107
1.3 Huellas de selección en los genes de <i>Mastadenovirus</i>	115
2. Estimación de la tasa de mutación de HAdV5	121
2.1 Identificación de mutaciones en HAdV5.....	123
2.2 Espectro mutacional y tasa de mutación espontánea	132
2.3 Distribución de mutaciones a lo largo del genoma	135
2.4 <i>Hot spots</i> mutacionales.....	139
2.5 Relación entre la tasa de mutación y diversidad de secuencias <i>in vivo</i>	141
3. Estudio del posible papel del sistema de reparación del DNA sobre la tasa de mutación de HAdV5	145
3.1 Inhibición de la DDR y evolución experimental de HAdV5 en células MRC-5.....	147

3.2 Estudio de las poblaciones evolucionadas mediante DS	159
3.3 Identificación de deleciones mediante métodos alternativos	165

4. DISCUSIÓN GENERAL..... 169

4.1 La diversidad a lo largo del genoma de adenovirus no es uniforme	169
4.2 Tasa de mutación de HAdV5.....	176
4.3 Probabilidad de mutación alta en dinucleótidos CpG.....	181
4.4 La DDR y la tasa de mutación de HAdV5.....	183

CONCLUSIONES FINALES 191

BIBLIOGRAFÍA 197

INFORMACIÓN SUPLEMENTARIA..... 243

PRÓLOGO

Agradecimientos

No quisiera empezar, sin antes agradecer a todas las personas que han hecho posible, de una manera u otra, esta tesis doctoral. En primer lugar, quisiera dar las gracias a mi director Rafael Sanjuán por darme esta oportunidad y abrirme las puertas del laboratorio. Gracias por tu paciencia, especialmente en esta última fase de la tesis.

Me acuerdo como si fuera hoy de mis comienzos en el grupo. El primer día, me presentaron a José, cuya mesa estaba desordenada y repleta de artículos y que debía guiarme durante mis primeros días. Para ti solo tengo palabras de agradecimiento, gracias por tus consejos y por calmarme a veces. Gracias por guiarme en el camino, incluso todavía a día de hoy que ya no compartimos el día a día en el lab. Te deseo lo mejor en este nuevo camino como líder de grupo, sé que tendrás mucho éxito. Allí también estaba Silvia, que siempre me acompañó en todo momento y que me dejó muy triste cuando se fue. Gracias a ti también por siempre comprenderme y apoyarme. Quisiera agradecer de una forma especial a Cristina. Hizo que los comienzos fueran fáciles y la echo mucho de menos, gracias por todo amiga. Nunca me dejaste sola. Por aquel entonces también estaba Ron, con el que coincidí poco tiempo. Menos mal que volvimos a encontrarnos en el i2Sysbio. Ron es un gran científico pero también una gran persona, siempre está ahí para lo que necesites. Aprendí mucho gracias a ti. Te deseo lo mejor, *man* (aunque sé que no te hará falta).

PRÓLOGO

Al poco tiempo llego María D, símbolo de esfuerzo. Siempre bajo el microscopio, cerca de la campana de gases y con la sonrisa puesta. Gracias a ti también por tus consejos y tu apoyo durante esta etapa de mi vida. A las nuevas generaciones, Alejandra, Inés, Laila, Salomé y María C os deseo mucha suerte en este nuevo camino que empezáis en el grupo.

Pero esta tesis doctoral no hubiera sido posible sin la ayuda de mis dos amigos JV y IAM a los que quiero agradecer su apoyo incondicional durante gran parte de esta andadura. Ellos conocen cada línea de esta tesis y me han aguantado como nadie, sobre todo en esos momentos en los que tenía ganas de abandonar, que no fueron pocos. Gracias por apoyarme en mis momentos de bajón, gracias por sacarme siempre una sonrisa, gracias por ayudarme en todo y preocuparos por mí. Si tuviera que contar todas las que hemos liado, no acabaría nunca. Os voy a echar mucho de menos, pero sé que nos vamos a reencontrar en el camino. No puedo permitirme perder vuestra amistad. Y qué decir de Ernesto, el cantarín que se define como un poco antisocial. Qué pena que nos hayamos conocido tan tarde. Sé que te tengo para lo que necesite y quiero que sepas que tú a mí también. Iba a desearte suerte, pero esa tesis va viento en popa. Cuídame a esos dos. Gracias a ti también, Vale, mi libro de frases inteligentes y de sabios consejos. He aprendido muchas cosas de ti y sé que me llevo una amiga. Gracias por tu cariño en los momentos más difíciles y por darme un abrazo cuando tanta falta me hacía.

Víctor de VirBio, también de los últimos en llegar, cuánto me has hecho reír. Esa sala de la P2 se nos quedaba cada vez más pequeña pero no importaba porque era un gusto trabajar con vosotros. Pilar, Flo, Ahmed y Mireia gracias también por vuestros consejos. Siempre habéis estado cuando os he necesitado y espero que así siga siendo. Esther y Rebeca, gracias también a vosotras por vuestra compañía y por los buenos momentos que hemos pasado juntas. Esos ratitos entre experimento y experimento lo hacían todo mucho más fácil.

Quisiera también agradecer a mi familia por su apoyo incondicional. Ellos han trabajado muy duro para conseguir que estudiara y estaré eternamente agradecida por ello. Ellos me han enseñado lo que es el esfuerzo y la constancia. Para mí son un ejemplo a seguir en todos los ámbitos en la vida. Por supuesto, gracias a Toni, por nunca soltarme la mano a pesar de las adversidades. Tú eres quien realmente me ha aguantado durante todo este tiempo, especialmente durante el período de escritura. El periodo más duro para mí en estos cuatro años y que parecía no tener fin. Menos mal que te tenía a mi lado en todo momento. No sabes lo importante que era para mí escuchar que era capaz de esto y más. Pero no te conformas solo con estar siempre que te necesito, sino que te esfuerzas en comprender el sacrificio y el esfuerzo vocacional que normalmente conlleva la carrera científica, estando dispuesto a dejarlo todo por mí. Nunca podré agradecértelo lo suficiente. Por ahora solo puedo dedicarte esta tesis. Prometo recompensarte el resto de mi vida, que sé que la pasare a tu lado.

PRÓLOGO

Artículos publicados a partir de los datos y resultados obtenidos en la presente tesis:

- **Risso-Ballester J.**, Cuevas J.M. and Sanjuán R. 2016. Genome-wide estimation of the spontaneous mutation rate of human adenovirus 5 by high-fidelity deep sequencing. ***PLoS pathogens*** 12: e1006013.
- **Risso-Ballester J.** and Sanjuán R. High fidelity deep sequencing reveals no effect of cellular DNA damage response on adenovirus mutation rate. ***Viruses***.

En revisión

Otras publicaciones relacionadas:

- Sanjuán R., Pereira-Gómez M. and **Risso J.** (2016). Genome instability in DNA viruses, in Genome Stability. Elsevier, pp. 37–47.

INTRODUCCIÓN

1. LOS VIRUS Y SU EVOLUCIÓN

1.1 Mecanismos evolutivos

Los virus son un grupo de parásitos genéticos obligados muy heterogéneo que experimenta cambios genéticos a través de varios mecanismos. Entre ellos, se encuentra la mutación que es la última fuente de variabilidad genética y la materia prima de la evolución. Numerosas estimaciones de tasas de mutación han sido realizadas haciendo uso de métodos bioquímicos y genéticos. Las tasas de mutación en virus varían desde 10^{-4} a 10^{-8} sustituciones nucleotídicas por base por infección celular (s/n/c) (Drake y col. 1998; Duffy y col. 2008; Sanjuan y col. 2010). Estas tasas de mutación son muy altas en comparación con las obtenidas para otros microbios como *Escherichia coli* o *Saccharomyces cerevisiae* ($<10^{-9}$) (Lang y Murray 2008; Jee y col. 2016).

Desde el punto de vista evolutivo, existen tres tipos de mutaciones: beneficiosas (mejoran el fenotipo, individuos favorecidos por la selección natural), neutrales (no afectan a la eficacia biológica) y deletéreas (tienden a ser eliminadas de la población). La fijación o la pérdida de una mutación y por consiguiente, la cantidad de diversidad genética en las poblaciones virales está determinada por la acción de la deriva genética, la selección natural, la recombinación y el flujo genético. En consecuencia, la evolución se basa en el origen de variación genética y la posterior fijación de los cambios.

INTRODUCCIÓN: Los virus y su evolución

Las frecuencias alélicas pueden variar de forma aleatoria fruto de la deriva genética. Este proceso estocástico está determinado por el tamaño poblacional efectivo que, a su vez, es dependiente de otros factores como los cuellos de botella durante la transmisión o la estructura espacial. Generalmente, a medida que el tamaño poblacional es más pequeño, la deriva genética aumenta, pues el error de muestreo es mayor. Una de las consecuencias de la deriva genética es la pérdida de diversidad, pues solo un subconjunto de la población contribuye a la siguiente generación. A pesar de sus grandes tamaños poblacionales, los virus experimentan deriva genética. Durante su transmisión, sufren cuellos de botella importantes de manera que, cada vez que se inicia una nueva infección, esta puede ser iniciada por una o unas pocas partículas virales (Butcher 1995; Gutiérrez y col. 2012). Asimismo, al terminar el ciclo, la distribución de la progenie no es uniforme (Ellegren y Galtier 2016; Sanjuán 2018). Muchas partículas virales serán defectuosas y sin éxito reproductivo mientras que unas pocas contribuirán con grandes cantidades de material genético para que tenga lugar la siguiente generación. Cuando la deriva actúa de manera reiterada, se van perdiendo sucesivamente genotipos y puesto que la mayoría de mutaciones son deletéreas, se producirá una pérdida de eficacia en la población, lo que se conoce como Trinquete de Müller. La existencia de este fenómeno en poblaciones virales se demostró por primera vez en 1990 por Lin Chao (Chao 1990). En este estudio, se sometió al bacteriófago $\phi 6$ a sucesivos cuellos de botella observándose una disminución significativa de la eficacia biológica poblacional.

INTRODUCCIÓN: Los virus y su evolución

Más adelante, estos resultados se extendieron a otros bacteriófagos (MS2) (de la Peña y col. 2000) y a otros virus como el virus de la estomatitis vesicular (VSV) (Elena y col. 1996) y el virus de la inmunodeficiencia humana (HIV-1) (Yuste y col. 1999).

Además de la deriva genética, la celeridad y la probabilidad de que la frecuencia de un alelo cambie en la población dependen del tipo y la presión de la selección natural. Cuando el tamaño de la población es grande, a diferencia de la deriva genética, la selección natural modifica las frecuencias alélicas de forma no aleatoria permitiendo así adaptar la población al ambiente. Por ende, la selección natural es el proceso por el que tiene lugar la supervivencia de unos individuos frente a otros en base a su eficacia biológica. El término *fitness* o eficacia biológica define al número de descendientes producidos por un individuo que contribuirán a la siguiente generación. Este parámetro es relevante a la hora de estudiar los mecanismos evolutivos y generalmente se calcula en relación a un genotipo de referencia en un ambiente dado.

La selección positiva describe la selección de alelos beneficiosos, cuya frecuencia se espera que vaya aumentando a lo largo de las sucesivas generaciones hasta que estos lleguen a fijarse en la población (Weigand y Leese 2018). Los alelos beneficiosos y que están bajo selección positiva pueden disminuir su frecuencia debido a la deriva genética, principalmente cuando su frecuencia es baja. En contraposición, la selección purificadora captura el proceso de selección frente a mutaciones perjudiciales o deletéreas.

INTRODUCCIÓN: Los virus y su evolución

Los alelos que supongan un efecto nulo o pequeño sobre la eficacia y que estén ligados genéticamente a una mutación seleccionada positivamente se espera que sufran un cambio simultáneo en su frecuencia, proceso denominado *hitchhiking* (Kaplan y col. 1989). Al mismo tiempo, la frecuencia a la que cambian los alelos neutrales también puede verse afectada por la acción de la selección purificadora frente a las mutaciones deletéreas, proceso que se conoce como *background selection* (Charlesworth y col. 1993; Good y col. 2014; Ewing y Jensen 2016). Este ligamiento puede romperse mediante recombinación. En ausencia de recombinación, si dos mutaciones beneficiosas aparecen en diferentes linajes nunca se combinarán en un único genotipo (Muller 1964; Felsenstein 1974), sino que se desencadenará una competencia entre ambas descendencias. Este proceso recibe el nombre de interferencia clonal y puede llegar a la extinción de uno de los dos linajes (Gerrish y Lenski 1998; Park y Krug 2007).

La recombinación también determina la evolución de los virus haciendo posible la fijación de genotipos y la purga de mutaciones deletéreas. El proceso de recombinación se produce cuando al menos dos genomas virales coinfectan la misma célula huésped y tiene lugar el intercambio de segmentos genéticos. En virus, generalmente el intercambio genético no es recíproco, como en los organismos diploides es decir, el receptor de una determinada porción genómica no actúa como donante de la porción reemplazada en la fuente original.

INTRODUCCIÓN: Los virus y su evolución

Dos tipos de recombinación han sido identificadas en virus (Becher 2012; Scheel y col. 2013): la recombinación homóloga (Lai 1992), en la que se intercambian dos segmentos similares y la recombinación no homóloga o ilegítima (Troels y col. 2013) en la que se intercambian segmentos dispares y que frecuentemente, da lugar a estructuras aberrantes. Un mecanismo particular de recombinación en el que se produce la reordenación de segmentos genómicos completos tiene lugar en virus con genomas segmentados (Pérez-Losada y col. 2014). Aunque la recombinación es un proceso generalizado en los virus y puede tener un gran impacto en su evolución (Martin y col. 2011; Simon-Loriere y Holmes 2011), las frecuencias de recombinación son muy dispares entre virus.

La recombinación no homóloga es poco frecuente en los virus de RNA mientras que en virus de DNA puede ocurrir a frecuencias similares que la recombinación homóloga (Robinson y col. 2011). No obstante, la recombinación homóloga parece ser el mecanismo más frecuente entre los virus de dsDNA. Un ejemplo de ello lo encontramos en los α -herpesvirus, en los que la recombinación está directamente relacionada con los mecanismos de reparación y de replicación y por tanto, es capaz de prevenir la acumulación progresiva de mutaciones dañinas (Thiry y col. 2005, 2006). Aunque se ha postulado que la recombinación facilita la transmisión de material genético entre especies, su relevancia real es bastante incierta en algunos virus de RNA.

INTRODUCCIÓN: Los virus y su evolución

La frecuencia de recombinación es alta en los retrovirus, especialmente en HIV (Jetzt y col. 2000), donde se ha observado que la tasa de recombinación es más elevada que la tasa de mutación. No obstante, las frecuencias de recombinación entre el resto de grupos de virus de RNA son muy variables. Comúnmente, se le encuentra poco sentido evolutivo a la recombinación en los virus de RNA. Se piensa que la razón de su evolución reside principalmente en sus altas tasas de mutación y grandes tamaños poblacionales (Pérez-Losada y col. 2015).

Por último, el flujo genético, también conocido como migración, es la transferencia de alelos de una población a otra. La alta migración favorece la uniformidad genética entre poblaciones y por tanto, disminuye la diversidad genética global (Moya y col. 2004).

En biología evolutiva, se recurre a los paisajes adaptativos, introducidos por primera vez por Sewall Wright en 1932 (Wright 1932), para entender las relaciones entre los genotipos y su éxito reproductivo (Figura 1). Es decir, para comprender el proceso adaptativo de los organismos a su entorno. Según el modelo, en el paisaje adaptativo los organismos se sitúan en el plano X y la altura (eje Y) corresponde al nivel de adaptación alcanzado, es decir a su eficacia biológica. El movimiento de los organismos a través del paisaje, va a depender de la aparición de mutaciones en su genoma.

INTRODUCCIÓN: Los virus y su evolución

La aparición de mutaciones beneficiosas hará que los organismos aumenten su eficacia biológica y en consecuencia, generen una mayor descendencia y por tanto, su frecuencia aumentará en la población. La selección positiva favorecerá a estas variantes y estas irán escalando hasta alcanzar la cima adaptativa. Una vez alcanzada la cima, las mutaciones que aparezcan solo podrán empeorar su eficacia, provocando su descenso hacia los valles. La selección purificadora se encargará de eliminar estas variantes deletéreas.

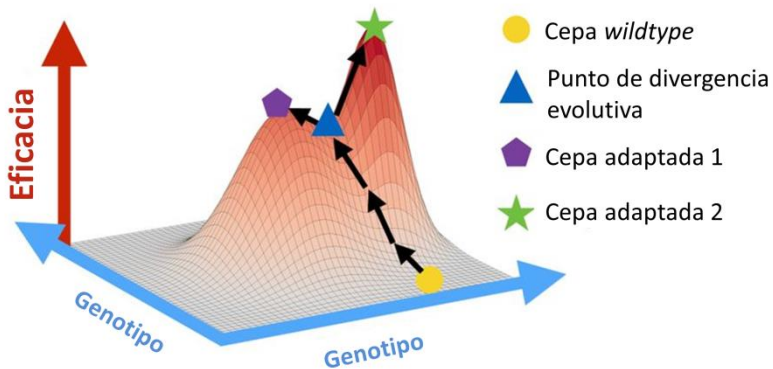


Figura 1. Evolución en un paisaje adaptativo de Wright. Según el modelo de Wright, se representa un paisaje rugoso con diferentes picos de adaptación en el que en el eje X se sitúan los genotipos y en el eje Y la eficacia biológica. Se ejemplifican dos trayectorias evolutivas de un único genotipo (círculo amarillo). Estas trayectorias divergen en un punto (triángulo azul) y terminan en dos puntos de adaptación máxima (cimas) distintos (estrella verde y pentágono morado) (adaptado de Nichol y col. 2018).

INTRODUCCIÓN: Los virus y su evolución

En virus, se ha observado que cuando se inocula un aislado en un nuevo huésped, el virus experimenta un proceso rápido de adaptación a este nuevo ambiente. Sin embargo, la velocidad con la que aumenta la eficiencia disminuye progresivamente. Esto refleja una adaptación discontinua y limitada (Novella y col. 1995; Bull y col. 1997). Además, cuando este experimento se repite varias veces, aunque se observa un comportamiento similar, el nivel máximo de adaptación alcanzado no siempre es el mismo (Domingo-Calap y Sanjuán 2011). Ambas observaciones se ajustan perfectamente al modelo anteriormente descrito. Las dos fases observadas durante la infección corresponden al periodo de escalada y al tiempo que permanece la variante en el pico de máxima adaptación. Por otro lado, debido a que la mutación es un proceso aleatorio, los organismos pueden alcanzar picos adaptativos de diferente altura en el paisaje evolutivo. Esto también explicaría las diferencias observadas entre réplicas en experimentos con virus.

Asimismo, la evolución paralela también es un proceso muy frecuente en las poblaciones virales. La evolución paralela tiene lugar cuando dos o más poblaciones, que evolucionan de manera independiente y en ambientes separados, experimentan cambios similares a lo largo del tiempo. Este fenómeno ha sido observado en virus de RNA, tanto de forma natural como bajo condiciones de laboratorio, y en bacteriófagos (Crill y col. 2000; Gutierrez y col. 2019).

INTRODUCCIÓN: Los virus y su evolución

Por ejemplo, en HIV se observan de manera recurrente y en distintos pacientes las mismas mutaciones relacionadas con la resistencia a la terapia frente al virus (Wensing 2016). El proceso de adaptación a un nuevo ambiente (hospedador nuevo, sistema inmune, antiviral, etc.) genera cambios en el virus que, a su vez, causan la aparición de fuertes presiones selectivas que favorecen la evolución de determinados genotipos (los más aptos). Este fenómeno explica que tras un ensayo de evolución experimental, aparezcan las mismas variantes beneficiosas en los diferentes linajes o réplicas experimentales.

1.2 Tasas de mutación y fuentes de variación genética en virus

El concepto de tasa de mutación hace referencia al número de nuevos cambios genéticos (mutaciones puntuales, inserciones y deleciones) que aparecen aleatoriamente en el genoma por unidad de tiempo (Ge y col. 2007), por generación (Chao y col. 2002) o por ronda de replicación (Drake 1991a; Drake y col. 1998). Los virus son los sistemas biológicos para los que se ha observado la mayor variación en tasas de mutación, mostrando un rango entre aproximadamente 10^{-8} y 10^{-4} s/n/c (Duffy y col. 2008; Sanjuán y col. 2010). Las diferencias más grandes se encontraron entre los virus de RNA (10^{-6} – 10^{-4}) y DNA (10^{-8} – 10^{-6} ; Figura 2).

INTRODUCCIÓN: Los virus y su evolución

Al igual que las tasas de mutación, las tasas de evolución en virus varían de manera considerable a lo largo de las siete clases de Baltimore. Se obtuvieron valores entre 10^{-6} – 10^{-2} sustituciones nucleotídicas por posición por año. Los virus de ssRNA y RT son los virus que evolucionan más rápido, seguidos por los de dsRNA y ssDNA. En último lugar, se encuentran los virus de dsDNA, pues se ha observado que evolucionan más lentamente en promedio. En un estudio llevado a cabo por Sanjuán (Sanjuán 2012a), se observó una correlación positiva entre las tasas de mutación y las tasas de evolución a lo largo de los grupos de Baltimore. Esto confirmó las diferencias observadas entre virus de RNA y DNA (Domingo 2006; Duffy y col. 2008). Además, puso de manifiesto nuevamente que las altas tasas de mutación observadas en los virus de RNA se traducen en una mejor adaptación a nuevos ambientes y por tanto, en una rápida evolución. Este estudio también reveló que los virus que presentan genomas monocatenarios evolucionan más rápido que los que presentan genomas bicatenarios, independientemente de si su material genético es DNA o RNA.

El amplio rango de tasas de mutación observadas en virus se corresponde aproximadamente con la fidelidad de las polimerasas utilizadas durante la replicación. Las RNA polimerasas dependientes de RNA (RdRps) y las retrotranscriptasas (RTs) presentan las tasas de error más elevadas. Después encontraríamos las RNA polimerasas dependientes de DNA (RdDps) (Flint y col. 2004) y, en último lugar, las DNA polimerasas, poseedoras de un alto grado de

INTRODUCCIÓN: Los virus y su evolución

precisión. Las DNA polimerasas y algunas RdDps presentan dominios de corrección de errores o *proofreading* lo que reduce, al menos en un orden de magnitud, la tasa de mutación durante la replicación (García-Díaz y Bebenek 2007). En general, las RdRps y las RTs carecen de actividad 3' exonucleasa (Steinhauer y col. 1992; Černý y col. 2014). Hasta la fecha, únicamente las RdRps de coronavirus han mostrado actividad *proofreading* (Smith y col. 2013; Smith y col. 2014).

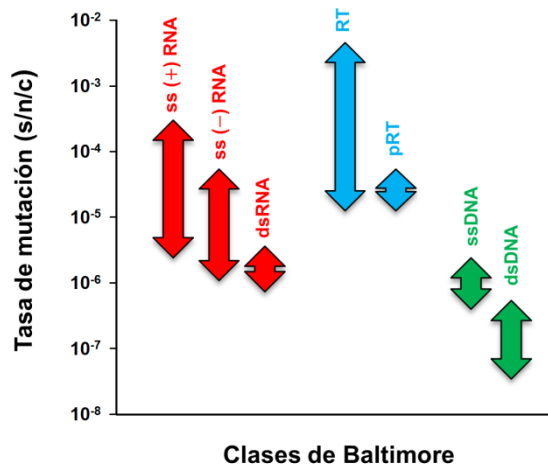


Figura 2. Variabilidad en las tasas de mutación estimadas a lo largo de las diferentes clases de virus de Baltimore. (ss: *single-strand*, ds: *double-strand*; +/- polaridad genómica, RT retrovirus, pRT para-retrovirus (adaptado de Sanjuán y Domingo-Calap 2016).

INTRODUCCIÓN: Los virus y su evolución

Esta función 3' exonucleasa también está ausente en las RTs y en consecuencia, se observan tasas de mutación similares entre los retrovirus y la mayoría de los virus de RNA (Hong y col. 2004; Menéndez-Arias 2009). Además, para el dsDNA existen sistemas de reparación de errores mediante escisión de bases desapareadas (García-Díaz y Bebenek 2007). Las diferencias bioquímicas existentes entre las replicasas virales las convierten en una posible fuente de variación de las tasas de mutación en virus. De hecho, la ausencia de dominios *proofreading* en las RdRps y RTs refleja las altas tasas de mutación observadas en los virus de RNA.

Por otro lado, existen virus de DNA que utilizan las polimerasas de la célula hospedadora durante la replicación como por ejemplo, los poliomavirus (Mesters y col. 2006). Estos virus de dsDNA codifican proteínas capaces de inactivar el dominio 3' exonucleasa de las polimerasas celulares. Consecuentemente, la evasión del sistema de corrección de errores durante la replicación podría constituir un mecanismo a través del cual determinados virus de DNA aumentan la diversidad en la población.

Los virus de ssDNA muestran tasas de mutación más altas que los virus de dsDNA, pero los mecanismos que subyacen esta observación son poco conocidos. Por el contrario, entre los virus de RNA no se encuentran grandes diferencias a lo largo de la clasificación de Baltimore.

INTRODUCCIÓN: Los virus y su evolución

Una de las causas que podría explicar las diferencias observadas entre virus de ssDNA y dsDNA podría ser su mayor susceptibilidad al daño endógeno y espontáneo (Lindahl 1993), pues el ssDNA es más propenso a lesión que el dsDNA (Chan y col. 2012).

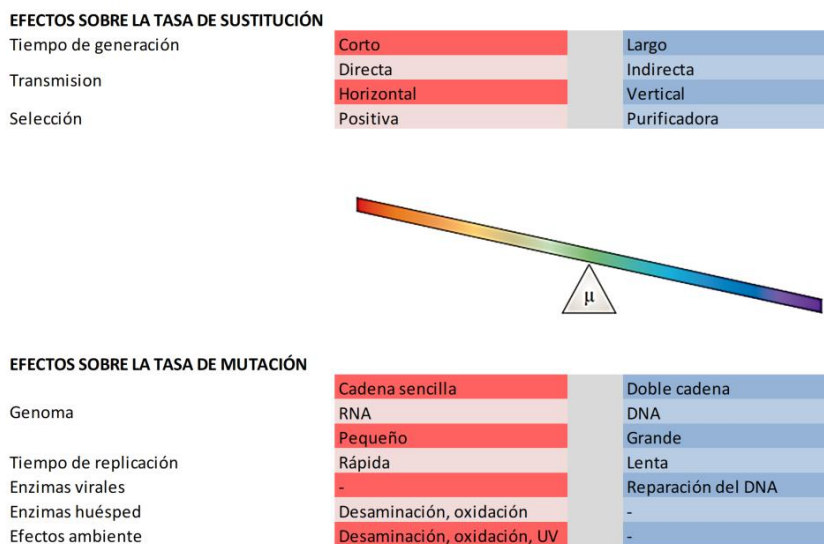


Figura 3. Factores que afectan a las tasas de sustitución y mutación en virus. La parte superior de la figura muestra los factores que aumentan (rojo) y disminuyen (azul) la tasa de sustitución. En la parte inferior se exponen los factores que alteran la tasa de mutación. Cabe destacar que muchos de los factores expuestos no son independientes. Por ejemplo, los virus que presentan un genoma pequeño, tienden a replicar más rápido o los virus que se transmiten de manera vertical en ocasiones se mantienen en un estado latente en el genoma del hospedador y en consecuencia, experimentan menos generaciones que los virus líticos, los cuales se transmiten horizontalmente. UV: radiación ultravioleta (adaptado de Duffy y col. 2008).

INTRODUCCIÓN: Los virus y su evolución

Otra causa podría ser el acceso de los virus de dsDNA al sistema post-replicativo celular. El sistema post-replicativo celular es un sistema altamente eficiente encargado de restaurar tanto los errores que se han originado durante la replicación así como de reparar el DNA dañado. El desarrollo de estrategias por parte de los virus para evadir estos sistemas de reparación del DNA daría como resultado un aumento de la frecuencia de mutación. Este fenómeno ha sido observado en el bacteriófago ϕ X174. En *E.coli*, el *DNA Mismatch Repair* (MMR) es llevado a cabo por el sistema Dam/MutHLS (Jiricny 2013). Las sustituciones nucleotídicas, inserciones y deleciones son reconocidas por MutS que posteriormente, interacciona con MutL, lo que causa la activación de la endonucleasa MutH. Esta endonucleasa reconoce la cadena parental por la presencia de un grupo metilo presente en la adenosina de motivos GATC, previamente añadido por la metilasa Dam. Esto hace que se ponga en marcha la reparación del DNA dañado (Li 2008; Fukui 2010). Sorprendentemente, se ha descubierto que la frecuencia de sitios GATC en el bacteriófago ϕ X174 está muy por debajo de lo esperado por azar. Se piensa que el virus ha evolucionado para no contener motivos GATC, evitar al sistema de reparación de DNA de la célula hospedadora y en consecuencia, aumentar su tasa de mutación (Pereira-Gómez y Sanjuán 2015). Sin embargo, no hay evidencias que nos impidan valorar la posibilidad de que un proceso similar al anterior pudiese ocurrir también en virus de dsDNA.

INTRODUCCIÓN: Los virus y su evolución

En consecuencia, este fenómeno podría no ser una explicación a las diferencias en tasas de mutación observadas entre virus de ssDNA y dsDNA. A diferencia de ϕ X174, la posible relación entre la tasa de mutación y el sistema post-replicativo celular no ha sido esclarecida en los virus de DNA eucariotas, que también son capaces de desregular las vías de reparación del DNA (Luftig 2014a).

Por otro lado, según la conocida regla de Drake, existe una relación inversa entre el tamaño genómico y la tasa de mutación (Figura 4). Esta regla se aplica a microorganismos basados en DNA (virus, bacterias y eucariotas unicelulares) (Drake 1991; Lynch 2010). Aunque el rango de tamaño en los virus de RNA es más pequeño y esto dificulta el análisis, parece existir una relación similar para los virus de RNA. De manera indirecta, la observación de que la mayoría de los virus de DNA, para los que se han obtenido elevadas tasas de mutación, presentan un tamaño genómico reducido apoya a esta hipótesis (Sanjuán 2012). A pesar de ello, las causas subyacentes a esta regla son todavía desconocidas. Nuevamente, la interacción de los virus de DNA y las vías de reparación de DNA celular podría explicar las diferencias observadas entre virus de DNA de tamaños distintos.

Se ha observado que algunos virus de DNA como los poliomavirus, papilomavirus o los parvovirus requieren de la activación del sistema de reparación de DNA (DDR, por sus siglas en inglés, *DNA damage response*) para que la infección tenga lugar de manera eficiente (Dahl y col. 2005; Gillespie y col. 2012; Luo y Qiu 2013).

INTRODUCCIÓN: Los virus y su evolución

Sin embargo, no se han realizado todavía estudios que nos permitan conocer si la activación o represión de la DDR por parte de los virus de DNA de estructura y tamaño distinto se traduce en tasas de mutación diferentes. La DDR comprende DNA polimerasas con elevadas tasas de error para llevar a cabo a síntesis *de novo* de las cadenas escindidas. Por tanto, la participación de estas polimerasas en la replicación del genoma del virus podría aumentar su tasa de mutación y en consecuencia, la diversidad en la población (Ciccia y Elledge 2010).

El concepto de que los virus de RNA evolucionan más rápido que los virus de DNA por presentar altas tasas de mutación está cambiando (Duffy y col. 2008). Aunque numerosos estudios corroboran esta fuerte división entre virus de DNA y RNA, estudios recientes han mostrado que, en términos evolutivos, esta división no es tan evidente como se pensaba. Se ha observado que algunos virus de RNA como por ejemplo, el virus espumoso de los simios (SFV) (Switzer y col. 2005) y el virus linfotrópico humano de células T tipo II (HTLV-II) (Salemi y col. 1999; Lemey y col. 2005) exhiben tasas bajas, 1.7×10^{-8} y 10^{-7} sustituciones/sitio/año, respectivamente.

En el caso de SFV, se observó diversidad genética a lo largo de diferentes aislados por lo que se descartó la posibilidad de que la RT hubiera desarrollado algún mecanismo de corrección de errores. Se pensó que una posible explicación a estas observaciones podrían ser los largos periodos de latencia ya que suponen bajas tasas de

INTRODUCCIÓN: Los virus y su evolución

replicación. En el caso de HTLV-II también se propuso como posible explicación su baja tasa de transmisión inter-hospedador. Sin embargo, poblaciones de HTLV-II, bajo una situación de transmisión epidémica rápida, muestran tasas cercanas a 10^{-4} sustituciones/sitio/año (Lemey y col. 2005). Bajo estas circunstancias, estos virus necesitan replicar de forma activa para iniciar la infección de nuevos hospedadores, lo que causa un aumento de la tasa de mutación considerable.

Otro ejemplo más controvertido es el de virus de RNA que utilizan RdRps para replicar su genoma. Entre ellos se encuentran los hantavirus y los flavovirus de roedores, para los que se ha estimado una tasa de 10^{-7} sustituciones/sitio/año (Suzuki y col. 1999; Plyusnin y Morzunov 2001), y algunos virus de RNA que infectan plantas (García-Arenal y col. 2001). En el caso de ambos virus animales, esta lenta evolución se asoció fundamentalmente a su tipo de infección crónica y a la codivergencia con sus hospedadores (Charrel y col. 1999; Hughes y Friedman 2000). Al mismo tiempo, estudios filogenéticos proponen que determinados virus de DNA evolucionan rápidamente (Umemura y col. 2002; Shackelton y col. 2005; Duffy y Holmes 2008). De hecho, se han estimado tasas de sustitución anormalmente altas para algunos virus de DNA. Entre ellos se encuentran los parvovirus (CPV-2 y B19), el virus de la panleucopenia felina (FPV), los circovirus y anelovirus SEN-V, y el geminivirus del rizado amarillo del tomate (Shackelto y col. 2005; Shackelton y Holmes 2006; Duffy y Holmes 2008).

INTRODUCCIÓN: Los virus y su evolución

Para estos virus de ssDNA, se han obtenido tasas de mutación similares a las estimadas para los virus de RNA. Se pensaba que estos virus evolucionaban lentamente debido a que utilizan las polimerasas de la célula hospedadora (Flint y col. 2004). También se ha demostrado que algunos virus de dsDNA evolucionan rápidamente (Firth y col. 2010).

1.3 Importancia de estimar las tasas de mutación en virus

Cabe destacar que la tasa de mutación no debe confundirse con la frecuencia de mutación, que hace referencia a la frecuencia a la que se encuentran las mutaciones en una población determinada. Este parámetro nos da información sobre la diversidad genética poblacional que, a su vez, es dependiente de otros procesos como la deriva genética aleatoria, la recombinación y la selección natural. Una mayor tasa de mutación implica mayor diversidad genética, sin embargo, no siempre es posible calcular la tasa de mutación a partir de datos de frecuencias mutacionales (Sanjuán y col. 2010).

Del mismo modo, las tasas de mutación tampoco deben confundirse con las tasas de evolución. Aunque la teoría neutral de la evolución propone una relación lineal entre ambas, la última se refiere a la fijación de nuevos alelos en las poblaciones, mientras que la mutación es un proceso meramente bioquímico/genético (Duffy y col. 2008; Sanjuán 2012).

INTRODUCCIÓN: Los virus y su evolución

La estimación de las tasas de mutación virales y el estudio de los procesos que las subyacen es de gran importancia para comprender y al mismo tiempo subsanar determinados problemas de salud globales. La diversidad genética de los virus está determinada en última instancia por la tasa de mutación. Un ejemplo de ello es el virus de la hepatitis B (HBV), en el que una serie de mutaciones puntuales han sido asociadas a la evasión del sistema inmune y a la falta de éxito en los procesos de vacunación (Coppola y col. 2015).

Otro ejemplo muy conocido es el del virus de la influenza, que constantemente sufre cambios antigénicos, por lo que anualmente se necesita actualizar las vacunas (Schotsaert y García-Sastre 2014).

Los virus que mutan rápido también desarrollan resistencia a antivirales o fármacos. La historia de los tratamientos desarrollados contra el HIV ilustra perfectamente la importancia de las tasas de mutación virales. La azidotimidina (AZT) fue el primer medicamento antirretroviral (ARV) aprobado en 1987 por su efecto en la supresión de la replicación viral. Desafortunadamente, la rápida aparición de variantes resistentes hizo que este medicamento no tuviera éxito (Perelson 2002). Más tarde, se desarrolló una nueva terapia más exitosa basada en la combinación de diferentes fármacos con el objetivo de reducir al máximo las posibilidades de que apareciese una mutación que provocará resistencia.

INTRODUCCIÓN: Los virus y su evolución

Por tanto, el estudio de los mecanismos responsables de la generación de diversidad en las poblaciones virales podría mejorar la comprensión y el manejo de la resistencia a los medicamentos, la evasión del sistema inmune, vacunación, patogénesis e incluso la aparición de nuevas enfermedades. En este sentido, existen tres aproximaciones distintas para medir las tasas de mutación en virus: ensayos libres de células (*in vitro*), ensayos en cultivos celulares (*ex vivo*) o tomando muestras directamente de los huéspedes (*in vivo*) (Sanjuán y col. 2012).

Las estrategias *in vitro* se basan fundamentalmente en estudiar la incorporación de bases correctas o incorrectas durante la replicación del genoma viral. Estos ensayos han sido realizados utilizando la RT de HIV y las RdRp de otros virus de RNA y al mismo tiempo, han sido de gran utilidad para obtener más información acerca de la fidelidad de las polimerasas (Steinhauer y col. 1992b; Arnold y col. 2005; Minskaia y col. 2006). Sin embargo, los resultados obtenidos a partir de estas aproximaciones suelen diferir de los obtenidos a partir de ensayos en cultivos celulares. Tal y como se ha discutido en el apartado anterior, la aparición de mutaciones en el genoma viral no solo depende de las polimerasas utilizadas durante la replicación.

Por otra parte, los ensayos *ex vivo* han sido llevados a cabo en numerosos virus y empleando diferentes metodologías (Sanjuán y col. 2010). Por ejemplo, si partimos de un inóculo genéticamente homogéneo y lo propagamos uno o unos pocos pases en cultivo celular, durante este intervalo de tiempo aparecerá variación genética

INTRODUCCIÓN: Los virus y su evolución

en la población. Esta variación es fácilmente detectable mediante PCR, técnicas de clonaje o secuenciación de un número considerable de clones. Finalmente, la tasa de mutación podría obtenerse dividiendo por el número de ciclos de infección transcurridos. Uno de los problemas fundamentales de este tipo de ensayos es la acción de la selección. La selección natural tiende a eliminar las mutaciones deletéreas, lo que introducirá un sesgo en la estimación. Este sesgo fruto de la selección aumentara a lo largo del tiempo, por lo que este problema necesita ser corregido de forma adecuada. La mayoría de las estrategias desarrolladas para abordar este problema consisten en replicar el virus en cultivo y en ausencia de selección. Otra estrategia muy distinta consiste en estudiar las mutaciones letales que aparecen en el último ciclo de infección (Sanjuan 2010). Sin embargo, las mutaciones introducidas por las polimerasas utilizadas durante la amplificación del DNA mediante PCR o secuenciación son otra causa de imprecisión a la hora de estimar la tasa de mutación viral por estos métodos.

Un método alternativo es el llamado test de fluctuación, una prueba clásica desarrollada por Luria y Delbrück, a través de la que demostraron que las mutaciones aparecen al azar y no en respuesta a la presión de selección en bacterias (Luria y Delbrück 1943). Este ensayo se basa en crecer, partiendo de un número inicial de individuos pequeño, varios cultivos en paralelo en un medio permisivo hasta que se alcance un determinado título viral.

INTRODUCCIÓN: Los virus y su evolución

Después de la fase de crecimiento, cada uno de los cultivos se ensaya por separado en presencia de un agente selectivo que permita cuantificar las variantes mutantes. Como agente selectivo suele utilizarse un anticuerpo monoclonal, un antiviral o incluso un tipo celular en el que únicamente determinados mutantes sean capaces de formar calvas. Finalmente, la distribución de mutantes resistentes en cada uno de los cultivos ensayados es utilizada para estimar la tasa de mutación. A través de este ensayo somos capaces de evitar los dos problemas anteriormente mencionados: los errores de PCR y la acción de la selección. Sin embargo, este método fenotípico tiene la desventaja de que la estimación se basa únicamente en una única región genómica, la responsable del fenotipo mutante.

El nacimiento de las técnicas de secuenciación de nueva generación (NGS) ha perfeccionado las estimas de las tasas de mutación, pues ha permitido realizar estimaciones a nivel de todo el genoma. Sin embargo, al igual que los procedimientos anteriores, también tiene sus limitaciones. A diferencia de los virus de RNA, pocas tasas de mutación han sido inferidas mediante secuenciación para virus de DNA (Sanjuán y col. 2016). El problema es básicamente que la frecuencia de mutación es muy baja y en consecuencia, la tasa de error de las plataformas dificulta la detección de variantes poco frecuentes. A los errores de secuenciación también debemos sumar las mutaciones introducidas artificialmente durante la amplificación previa mediante PCR.

INTRODUCCIÓN: Los virus y su evolución

Esto se traduce en aproximadamente, la identificación incorrecta del 1% de variantes, dependiendo de la plataforma utilizada (Kanagawa 2003; Shendure y Ji 2008). Aunque esta tasa de error puede ser tolerada para llevar a cabo otras aproximaciones, supone un problema en el caso de la estimación de tasas de mutación en virus de DNA. De hecho, en algunos casos, esta tasa de error puede llegar a ser órdenes de magnitud superior a la tasa de mutación que se desea estimar (Quiñones-Mateu y col. 2014).

Por tanto, el conocimiento sobre tasas de mutación en virus de DNA es más limitado que para virus de RNA, seguramente debido a estas limitaciones técnicas. Al igual que en los casos anteriores, se están implementando y desarrollando nuevas tecnologías de secuenciación de alta fidelidad para poner solución a estas limitaciones. Entre estas prometedoras alternativas a la clásica NGS, encontramos la secuenciación dúplex (DS) (Schmitt y col. 2012; Kennedy y col. 2014), a la que recurriremos durante esta tesis doctoral para estimar la tasa de mutación del adenovirus humano tipo 5 (HAdV5), y la secuenciación circular (CirSeq) (Acevedo y col. 2014).

La secuenciación dúplex ha mostrado una mejora en la precisión de alrededor de cuatro órdenes de magnitud en comparación con los otros métodos de NGS convencionales. Esta reciente metodología se basa en la ligación de etiquetas dúplex moleculares degeneradas a los fragmentos de dsDNA, de manera que cada fragmento de DNA estará marcado con una secuencia específica. Por consiguiente, mediante el análisis bioinformático correspondiente, seremos capaces

INTRODUCCIÓN: Los virus y su evolución

de localizar para cada molécula de dsDNA original cada una de las lecturas procedentes de cada una de sus cadenas complementarias. Este análisis permitirá el descarte de los artefactos de secuenciación, ya que estos sólo estarán presentes en una de las dos hebras (Schmitt y col. 2012). A diferencia del DS, la secuenciación CirSeq y otras estrategias como la secuenciación de células individuales, a pesar de suponer una mejora en la precisión, se basan en la secuenciación de una única hebra procedente del dsDNA original, lo que puede llevar a identificar erróneamente daños en el DNA ocurridos en una sola hebra como mutaciones.

Finalmente, muy pocas estimaciones de tasas de mutación virales han sido realizadas *in vivo* (Cuevas y col. 2009, 2015; Gago y col. 2009; Ribeiro y col. 2012). Esta aproximación es complicada debido al desconocimiento del número de generaciones o ciclos de infección y a la eliminación de mutaciones deletéreas, ya sea por selección y deriva genética, entre otras. Una aproximación útil reside en el estudio de la aparición de mutaciones letales para inferir la tasa de mutación. Debido a que estos virus no van a ser capaces de generar progenie, este método se centra en el estudio de un único ciclo de infección y en consecuencia, evita la mayoría de los factores no controlados (Sanjuán y col. 2012). Sin embargo, los virus portadores de mutaciones letales podrían completar sucesivos ciclos de replicación gracias a la complementación genética, tal y como se ha observado en virus de RNA (Aaskov 2006).

2. MODELO EXPERIMENTAL: ADENOVIRUS

2.1 Clasificación

El aislamiento del primer adenovirus humano (HAdV) tuvo lugar en el año 1953 por Rowe y col. a partir del cultivo de células adenoides y fue posteriormente caracterizado por dos grupos independientes (Rowe y col. 1953; Hilleman y Werner 1954). Desde entonces, se han identificado más de 100 miembros de la familia *Adenoviridae* capaces de infectar a un amplio rango de vertebrados.

Tabla 1. Clasificación de Adenovirus humanos.

Especies	Serotipos	Receptor celular	Tropismo	Enfermedad
A	12, 18, 31	CAR	Entérico, respiratorio	Gastroenteritis
B1	3, 7, 16, 21, 50	CD46, CD80, CD86	Respiratorio, ocular	Enfermedad Respiratoria
B2	11, 14, 34, 35, 55	CD46, CD80, CD86, DSG2	Renal, respiratorio, ocular	Enfermedad tracto urinario
C	1, 2, 5, 6, 57	CAR, HSPG, MHC1, SR, VCAM-1	Respiratorio, ocular, linfóide, hepático	Enfermedad Respiratoria
D	8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, 42-49, 51, 53, 54, 56	Ácido siálico, CAR, CD46	Entérico, ocular	Conjuntivitis
E	4	CAR	Respiratorio, ocular	Enfermedad respiratoria Conjuntivitis
F	40, 41	CAR	Entérico	Gastroenteritis
G	52	Ácido siálico, CAR	Entérico	Gastroenteritis

La familia de *Adenoviridae* se divide en cinco géneros principales: *Mastadenovirus* (mamíferos), *Aviadenovirus* (aves), *Atadenovirus* (alto contenido A+T), *Siadenovirus* (aves y réptiles) e *Ichtadenovirus* (peces). Esta clasificación se basa principalmente en estudios bioinformáticos de secuencias y en la especie huésped (Davison y col. 2003). Hasta la fecha, alrededor de sesenta tipos de adenovirus humanos han sido identificados y clasificados en siete especies (A-G) en base a estudios serológicos y al análisis de los genomas virales (Benkö y Harrach 2003) (Tabla 1). Los adenovirus humanos causan con mayor frecuencia enfermedades respiratorias pero, en función del serotipo infectante, pueden llegar a causar otras patologías como la conjuntivitis o la gastroenteritis infantil (Chu y Pavan-Langston 1979; Anderson 2010). Aunque la mayoría de las infecciones son leves, en determinados grupos de riesgo, como por ejemplo los pacientes inmunodeprimidos, adenovirus puede llegar a causar neumonía fatal, hepatitis y/o encefalitis (Kojaoghlanian y col. 2003).

2.2 Estructura del virión

Las partículas adenovirales (70-90 nm de diámetro) se caracterizan por carecer de envuelta lipídica y presentar una forma icosaédrica. Los viriones están constituidos por una cápside proteica, formada al menos por 10 proteínas, que protege al core constituido por el DNA y proteínas.

INTRODUCCIÓN: Modelo experimental

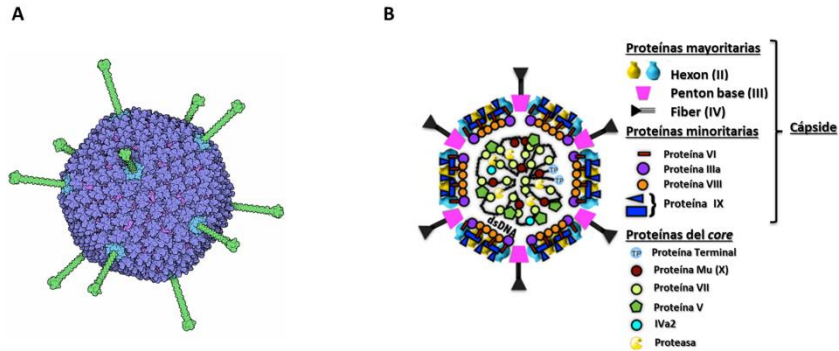


Figura 5. Estructura de adenovirus. **A.** Estructura de la cápside de adenovirus donde se muestran las proteínas mayoritarias *hexon* (morado), *penton* (azul) y *fiber* (verde) y las proteínas minoritarias (rosa) (Fuente: RCSB *Protein Data Bank*). **B.** Representación de las proteínas que constituyen la cápside y el *core* de adenovirus y su organización (Adaptado de Reddy y col. 2010).

La cápside está formada por las tres proteínas mayoritarias conocidas como *hexon*, *penton* y *fiber* (II, III y IV) (Figura 5). *Hexon* es la proteína más abundante y conforma las 20 caras triangulares del icosaedro. En cada uno de los 12 vértices, se encuentra la proteína *penton* a través de la cual se extiende la proteína *fiber*, que se proyecta desde el vértice como si se tratara de una antena. En la parte más distal respecto a la cápside, la proteína *fiber* presenta un dominio globular (*knob*), que es el responsable de la entrada del virus en la célula a través de su unión al receptor celular (Wu y col. 2003).

INTRODUCCIÓN: Modelo experimental

Las otras proteínas minoritarias (VI, IIIa, VIII y IX) que componen la cápside, actúan estabilizando las interacciones entre los *hexones* en la partícula viral (Liu y col. 2010; Reddy y col. 2010) y para alguna de ellas se ha observado su participación en procesos que tiene lugar durante el ciclo de infección (Wiethoff y col. 2005; Smith y col. 2010). El núcleo del virión está constituido por el genoma y las proteínas virales del *core*. Los polipéptidos VII, V y X se encuentran asociados al DNA debido a su alto contenido en residuos básicos y por tanto, hacen posible su condensación (Russell y col. 1968; Hosokawa y Sung 1976; Anderson y col. 1989).

El polipéptido IVa2 tiene un papel fundamental en el empaquetamiento del DNA viral dentro de la cápside durante la generación de nuevos viriones (Russell 2009). La proteína terminal que se encuentra covalentemente unida al extremo 5' del DNA actúa como primer durante la replicación (Rekosh y col. 1977). Por último, la proteasa participa en el procesamiento del virión durante el ensamblaje, desensamblaje y la salida del endosoma (Weber 2007; Russell 2009).

2.3 Estructura y organización del genoma

Los adenovirus presentan un genoma de dsDNA y lineal de 26-45 kb (Davison y col. 2003) (Figura 6). Los genes compartidos por los diferentes adenovirus se encuentran en la zona central del genoma, mientras que los genes más específicos se localizan en los extremos.

INTRODUCCIÓN: Modelo experimental

Los genes homólogos codifican proteínas necesarias para la replicación del DNA y los componentes mayoritarios que constituyen la cápside. En ambos extremos del genoma se localiza una secuencia de 36-200 pb llamada *Inverted Terminal Repeat (ITR)* que funciona como origen de replicación del DNA. En el inicio del genoma, entre el origen de replicación y la primera región codificante (E1A), se encuentra la señal de empaquetamiento (Ψ), imprescindible para la incorporación del DNA en la cápside durante la generación de nuevos viriones (Hearing y col. 1987; Gräble y Hearing 1990; Alba y col. 2011).

La información genética está organizada en unidades de transcripción solapantes localizadas en ambas hebras del DNA. Estas unidades se clasifican en tres grupos *early*, *intermediate* y *late* en función de cuando tiene lugar su expresión durante el ciclo de infección (Davison y col. 2003). La transcripción de la mayoría de las unidades genera múltiples mRNA que experimentan *splicing alternativo* de manera que se generan más de 40 proteínas distintas (Zhao y col. 2014).

Las unidades de transcripción *early* (E1A, E1B, E2, E3 y E4) codifican proteínas que participan en la replicación del DNA y garantizan un entorno celular óptimo para la replicación. La región E3 inhibe la respuesta inmune antiviral garantizando la supervivencia de la célula el tiempo necesario para que tenga lugar la infección.

INTRODUCCIÓN: Modelo experimental

Las unidades de transcripción *intermediate* (IX, IVa2, L4-22K/33K) (Tribouley y col. 1994; Morris y col. 2010), entre otras funciones, activan la expresión de las *late* (L1-L5) bajo el control del promotor llamado *Major Late Promoter* (MLP).

Este último grupo codifica las proteínas estructurales y otras proteínas necesarias para que tenga lugar el ensamblaje y la maduración de los nuevos viriones. En la fase *late*, también se expresa la unidad de transcripción U-exon, cuya función es todavía desconocida y utiliza un promotor propio (Tollefson y col. 2007; Ying y col. 2010).

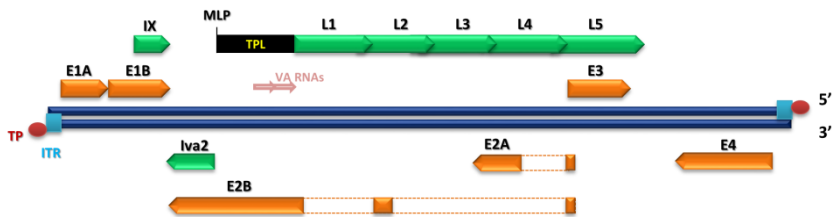


Figura 6. Genoma de adenovirus. Estructura y organización del genoma de adenovirus. La secuencia ITR presente en cada una de las hebras del dsDNA lineal de adenovirus se representa de color azul. La proteína terminal (TP) (rojo) se encuentra unida covalentemente a cada uno de los extremos del genoma. La organización y orientación de cada una de las unidades de transcripción *early* (naranja), *intermediate* y *late* (verde) vienen representadas utilizando flechas.

2.4 Ciclo de infección

2.4.1 Unión, entrada y tráfico intracelular

El primer paso del ciclo de infección consiste en la unión del dominio globular de la proteína *fiber* al receptor presente en la membrana celular. Se ha observado que adenovirus utiliza diferentes receptores celulares (Tabla 1). La proteína *fiber* de las especies de HAdV A, C, D, E y F tiene afinidad por el receptor celular transmembrana *Coxsackie B*, *Adenovirus Receptor* (CAR) mientras que la de los HAdV B y HAdV D37, se une al receptor celular CD46, presente en la membrana plasmática de la mayoría de tipos celulares (Gaggar y col. 2003; Arnberg 2009). Después de la unión del dominio *knob* de la proteína *fiber* al receptor celular, tiene lugar la unión entre los dominios RGB (motivos Arg-Gly-Asp) de la proteína *penton* y las integrinas $\alpha\beta 2$ y $\alpha\beta 5$ situadas en la superficie celular (Wickham y col. 1993; Wickham 1994; Nemerow y col. 2009). Ambas uniones provocan la aparición de fuerzas mecánicas que tensionan la cápside así como cambios conformacionales en ambas proteínas virales. Esto causa la desunión de la proteína *fiber* de la cápside y da comienzo al proceso de desensamblaje del virus (Nakano y col. 2000; Lindert y col. 2009; Burckhardt y col. 2011). La unión de *penton* a las integrinas activa determinadas vías de señalización que dan lugar a la reorganización del citoesqueleto y a la estimulación de la endocitosis del virus en compartimentos endosomales (Meier y Greber 2004; Gastaldelli y col. 2008).

INTRODUCCIÓN: Modelo experimental

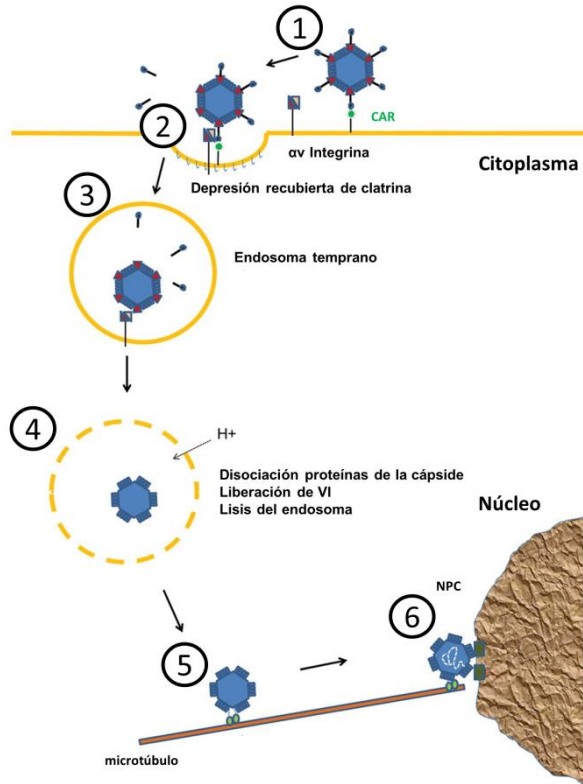


Figura 7. Unión, entrada y transporte hasta el núcleo de adenovirus. El primer paso consiste en la unión de la proteína *fiber* al receptor primario celular (CD46 o CAR) (1). Posteriormente tiene lugar la interacción entre la proteína *penton* y las α v integrinas, dando lugar a la endocitosis mediada por clatrina de adenovirus (2). En el interior del endosoma, tiene lugar la disociación del virión debido a las condiciones de pH (3), al mismo tiempo que se produce la liberación de la proteína viral VI durante su maduración. La liberación de esta proteína causa la lisis del endosoma y la salida de adenovirus al citoplasma (4). El subvirión viaja a través de los microtúbulos (5) hasta el NPC, a través del cual alcanza finalmente el núcleo de la célula (6) (Adaptado de Nemerow y col. 2009).

INTRODUCCIÓN: Modelo experimental

En el caso del HAdV B, se ha observado que esta unión estimula un mecanismo alternativo de endocitosis llamado macropinocitosis (Amstutz y col. 2008; Kalin y col. 2010). Las condiciones de pH en el interior de los endosomas durante su maduración, causan el desprendimiento tanto de las proteínas mayoritarias de la cápside (*penton* y *hexon*) como de las proteínas internas (IIIa, VI y VIII).

La liberación de la proteína VI provoca la lisis de la membrana endosomal y en consecuencia, la salida del virus al citosol (Wiethoff y col. 2005; Maier y Wiethoff 2010; Maier y col. 2010). Una vez fuera del endosoma, los subviriones son transportados hacia el núcleo celular mediante la utilización de los microtúbulos (Mabit y col. 2002; Walkiewicz y col. 2009; Strunze y col. 2011).

Finalmente, el genoma viral asociado a la proteína VII entra al núcleo a través del *Nuclear Pore Complex* (NPC) (Leopold y col. 2000; Mabit y col. 2002). Existen evidencias experimentales de que esta proteína inhibe la detección del genoma viral por parte de la DDR. En ausencia de VII, se ha demostrado que tiene lugar la activación del complejo MRE-11-RAD50-Nbs1 y la posterior degradación del genoma viral, inhibiéndose la replicación (Karen y Hearing 2011; Avgousti y col. 2017).

2.4.2 Replicación del DNA y regulación de la expresión génica viral

Una vez en el interior del núcleo, la proteína VII se disocia del genoma viral al mismo tiempo que se asocia con histonas celulares para que tenga lugar la expresión de los genes *early* (Ross y col. 2011). Esta disociación tiene lugar de forma gradual, pues VII permanecerá más tiempo asociada a aquellas regiones genómicas que se expresan en fases tardías de la infección.

La primera unidad de transcripción en transcribirse es E1A, cuyos productos génicos causan la transactivación de E1B, E2, E3 y E4. Además, E1A promueve la entrada de la célula en fase S mediante la disociación del complejo Rb-E2f (Bandara y La Thangue 1991; Arroyo y Raychaudhuri 1992; Parreño y col. 2001). La liberación de E2f, a su vez, promueve la expresión de E2, que codifica las proteínas necesarias para que tenga lugar la replicación (DNA polimerasa, pTP y DBP) (Kovesdi y col. 1987). En consecuencia, comienza la replicación formándose un complejo constituido por la DNA polimerasa y la PT en el origen de replicación (Hoeben y Uil 2013). A este complejo heterodimérico se unen también dos factores de transcripción celulares (NF1 y OCT), aumentando la eficiencia de inicio de la transcripción (van Breukelen y col. 2003; Mysiak y col. 2004). La unión de la proteína DBP provoca la separación de ambas hebras de DNA y, al mismo tiempo, sirve de protección frente a la acción de las nucleasas (Zijderveld y van der Vliet 1994).

INTRODUCCIÓN: Modelo experimental

El producto de esta primera fase de replicación consiste en un dsDNA, constituido por la cadena naciente y parental, y un ssDNA (cadena parental desplazada). La segunda fase de replicación consiste en la elongación de la cadena parental desplazada durante la fase anterior. Dado que el complejo de iniciación requiere como molde dsDNA, el DNA se circulariza mediante la hibridación de ambos extremos (ITRs) generándose una región de dsDNA de secuencia idéntica a la del final del genoma parental (Girard y col. 1977). A esta región se unirá el complejo de iniciación y tendrá lugar el proceso de elongación.

2.4.3 Expresión de genes tardía y ensamblaje de nuevos viriones

Además de la replicación del genoma, el virus necesita generar las proteínas necesarias para que tenga lugar la encapsidación de los genomas nacientes. La transcripción de estos genes (*late*) está regulada por el promotor MLP. Este promotor se encuentra activo durante la fase *early*, pero solo hace posible la expresión de L1 (Young 2003). La expresión de los genes que codifican los polipéptidos que conforman la cápside y la proteína L4-100k se expresan en el momento que inicia la replicación del DNA.

INTRODUCCIÓN: Modelo experimental

La estimulación de este promotor viene dada por dos proteínas virales codificadas en L4 (IVa2 y L4-22K), cuya transcripción se produce durante la fase intermedia bajo el control de un promotor propio llamado L4P (Farley y col. 2004; Morris y col. 2010b; Ying y col. 2010). La activación del promotor MLP da lugar a la expresión de las unidades de transcripción *late* (L1-L5), generándose una familia de mRNAs que viajan al citoplasma para que la maquinaria de traducción los convierta en proteínas (Cuesta y col. 2001). El ensamblaje de las proteínas *penton* y *hexon* comienza en el citoplasma (Hong y col. 2005). Posteriormente, estas proteínas serán transportadas al núcleo donde tendrá lugar la encapsidación de los genomas nacientes. El mecanismo es poco conocido pero se conocen las proteínas (pIIIa, L4-52K/55K, IVa2 y L4-22K) (Zhang y col. 2001; Ostapchuk y col. 2006; Wohl y Hearing 2008; Ma y Hearing 2011) que dirigen el proceso a través del reconocimiento de la señal de encapsidación y la formación de un complejo (Zhang e Imperiale 2000; Tyler y col. 2007; Wu y col. 2012). El dominio ATPasa de la proteína IVa2 (Ahi y col. 2015) se ha propuesto como motor para la encapsidación del DNA en la procápside, al igual que ocurre en bacteriófagos (Roy y col. 2012; Sun y col. 2012).

La acción de la proteasa sobre determinadas proteínas estructurales culmina el proceso de ensamblaje de nuevos viriones con capacidad infectiva (Mangel y San Martín 2014). Los mecanismos a través de los cuales los nuevos viriones salen de la célula también son poco conocidos.

INTRODUCCIÓN: Los virus y sistema de reparación del DNA

Se ha propuesto que la proteína E3 11.6KD, también conocida como *Adenovirus Death Protein* (ADP), juega un papel fundamental en el proceso de lisis celular (Doronin y col. 2003; Murali y col. 2014). Esta proteína se acumula en el núcleo celular durante las fases tardías de la infección. Se ha propuesto que la función de ADP podría ser la de antagonizar a la proteína viral E1B19K, provocando la muerte de la célula por apoptosis (Subramanian y col. 2006). Asimismo, se ha demostrado que la proteasa viral (L3 23K) escinde la citoqueratina K18 celular (Chen y col. 1993), inhibiéndose su capacidad de polimerizar y causando la desestabilización de la estructura celular haciendo que la célula sea más susceptible a la lisis.

3. LOS VIRUS Y EL SISTEMA DE REPARACIÓN DE DNA

3.1 El sistema de reparación de DNA celular: vías de señalización

El DNA está continuamente en riesgo de ser dañado por numerosos factores, tanto exógenos como endógenos. La estabilidad genómica es fundamental para mantener la homeostasis celular y garantizar la continuidad genética durante la proliferación celular (Ciccia y Elledge 2010).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

En consecuencia, las células han desarrollado un sistema de rutas de señalización complejo capaz de detectar daños en el DNA y tratar de repararlos. La DDR está controlada principalmente por tres serina-treonina quinasas de la familia PIKKs (quinasas similares a la *fosfatidilinositol-3-kinasa*): ATM (mutada en ataxia-telangiectasia), ATR (relacionada con ATM y Rad3) y DNA-PK (subunidad catalítica de la proteína quinasa dependiente de ADN) (Smith y col. 2010; Davis y col. 2014).

Cuando se detecta un daño en el DNA (proteínas sensoras), las PIKKs (proteínas transductoras) fosforilan a sus correspondientes sustratos (proteínas efectoras), lo que favorece el reclutamiento de las proteínas necesarias para garantizar la reparación del daño o, si este es irreparable, promover la entrada de la célula en apoptosis o senescencia (respuesta celular; Figura 8).

Tanto la ruta de ATM como la de DNA-PK responden a una de las lesiones más deletéreas, las roturas de DNA de doble cadena (DSBs, del inglés *double stranded breaks*). Las DSBs son rápidamente detectadas por el complejo MRN (MRE11-RAD50-NBS1) y el heterodímero Ku70-Ku80. Seguidamente, estos reclutan a las proteínas transductoras, ATM y DNA-PK (Lee y Paull 2005; Mahaney y col. 2009), que se activan por autofosforilación.

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

Posteriormente, fosforilan a la histona H2AX en la Ser 139, modificación conocida como γ H2AX (Burma y col. 2001). γ H2AX recluta a MDC1 (mediador del checkpoint de daño en el DNA 1) que amplifica la señal mediante el reclutamiento de más complejos MRN-ATM (Stewart y col. 2003).

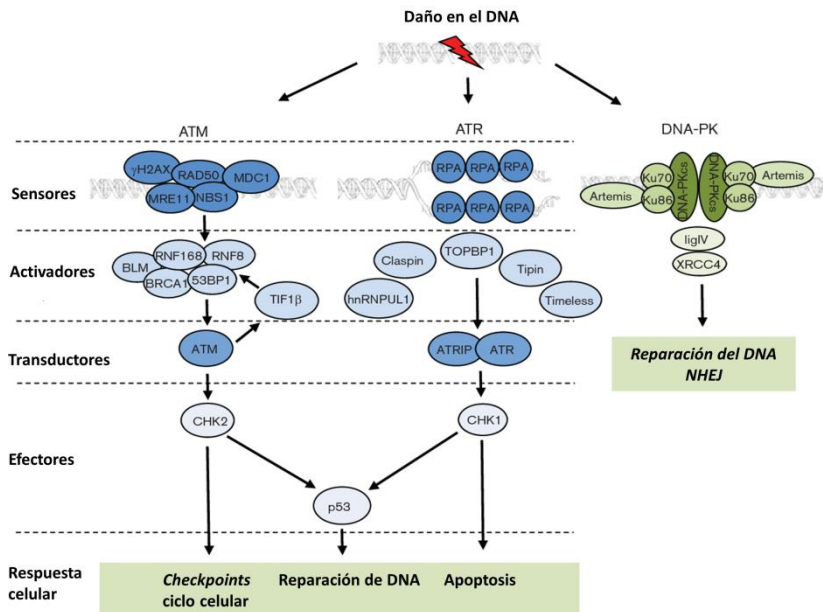


Figura 8. Representación de las rutas de señalización de ATM, ATR y DNA-PK. En la figura se muestra el papel que desempeña cada una de las proteínas que participan en la ruta así como la respuesta celular. Recurrir al texto para más detalles (Adaptado de Turnell y Grand 2012).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

Finalmente, la actividad sostenida de ATM facilita la fosforilación de cientos de las dianas de ATM como Chk2 (quinasa de checkpoint 2), CDC25 y p53, que regulan los puntos de control del ciclo celular y la entrada en apoptosis, respectivamente (Matsuoka y col. 2007). En el lugar de la lesión, el complejo Ku-DNA-PK es el regulador de la reparación de DSBs mediante reparación no homóloga a través del reclutamiento del adaptador XRCC4 y la DNA ligasa IV. Por el contrario, la ruta de ATM promueve la reparación de DSBs mediante recombinación homóloga (Burma y Chen 2004; Meek y col. 2008).

La acumulación de ssDNA por estrés replicativo, también es detectada por la DDR debido a su alto potencial recombinogénico. En este caso, el ssDNA se recubre por la proteína RPA (proteína de replicación A) que, a través de la interacción con su cofactor ATRIP, recluta a ATR (Zou y Elledge 2003). ATR se activa mediante autofosforilación y TOpBP1 (*topoisomerase-binding protein 1*) es reclutado al lugar de la lesión a través del complejo 9-1-1 (Rad9-Rad1-Hus1) (Kumagai y col. 2006; Matsuoka y col. 2007). Este complejo estimula la activación de ATR promoviéndose la fosforilación de diversas proteínas, entre las que se encuentra Claspin. Esta proteína recluta a la proteína efectora Chk1 (quinasa de checkpoint 1) para que ATR promueva su activación. Las dianas de Chk1 son también CDC25 y p53, lo que promueve la activación de los *checkpoints* celulares (Jeong y col. 2003; Kumagai y Dunphy 2003).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

Las proteínas Chk1 y Chk2, además de fosforilar a p53 e interrumpir el ciclo celular, juegan un papel importante en la reparación del DNA dañado ya que, son capaces de fosforilar y activar a proteínas involucradas en las rutas de reparación mediante recombinación homóloga (Stracker y col. 2009). A diferencia de ATM y DNA-PK, que se activan durante todas las fases del ciclo celular, la actividad de ATR está restringida a las fases S y G2 (Zou y Elledge 2003).

3.2 Activación e inhibición del sistema de reparación de DNA celular en respuesta a la infección

Desde el comienzo de la infección, los virus son descubiertos por el sistema inmunitario celular y si estos son de DNA también puede ser detectados por la DDR. Además, el genoma viral, al igual que el celular, está expuesto a fuentes de daño exógeno y endógeno y por consiguiente, puede acumular daños que promuevan la activación de la DDR.

Una vez en el núcleo, la replicación del genoma viral tiene lugar en unos microorgánulos bien definidos conocidos como centros de replicación. Se ha demostrado que durante la replicación de algunos virus como adenovirus (Stracker y col. 2002; Carson y col. 2003; Blackford y col. 2008), parvovirus (Cervelli y col. 2008; Collaco y col. 2009; Schwartz y col. 2009; Adeyemi y col. 2010) o herpesvirus

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

(Taylor y Knipe 2004; Wilkinson y Weller 2004; Lilley y col. 2005; Shirataa y col. 2005; Xiaofei y col. 2014), se produce el reclutamiento de proteínas efectoras de la DDR en esos centros de replicación.

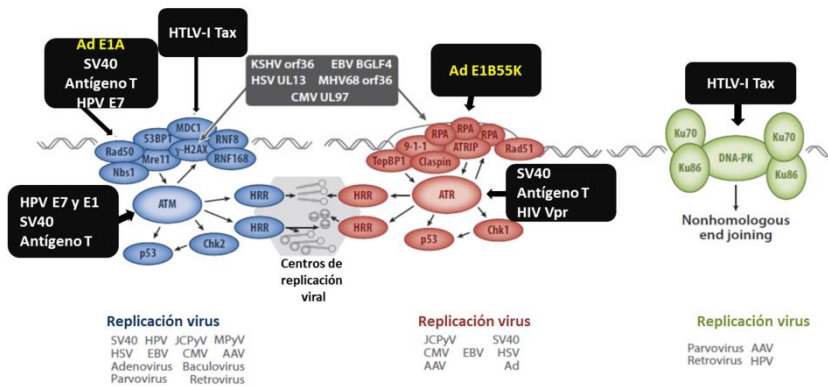


Figura 9. Activación de la DDR como resultado de la replicación viral y proteínas virales involucradas. Debajo de cada una de las vías de señalización se muestran algunos virus para los que se ha demostrado una activación de dicha ruta durante la infección. Los centros de replicación están representados cerca de las rutas de ATM y ATR activas (Adaptado de Luftig 2014a).

Pero no sólo la replicación de los ácidos nucleicos virales *per se* causa la activación de la DDR. Existen evidencias experimentales de que los virus, a lo largo de su evolución, han desarrollado mecanismos para modular la DDR (Lilley y col. 2007; Chaurushiya y Weitzman 2009; Luftig 2014b).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

En el caso de la activación de la DDR, todavía no está bien definido si se produce fruto de la infección *per se* o de manera intencionada por el virus. Si esta activación supone un beneficio real o no para el virus también es un tema controvertido. El papel de la DDR durante la replicación va a depender de muchos factores como el tipo de infección, las estructuras que se forman durante la replicación o incluso el tipo celular (Luftig 2014b). En este sentido, se ha observado que determinadas proteínas virales pueden desencadenar su activación de manera directa e indirecta.

Las oncoproteínas virales que causan la entrada de la célula en fase S para garantizar un entorno de replicación adecuado causan la activación indirecta de la DDR. Ejemplos de oncoproteínas virales son la proteína E7 del papilomavirus humano (HPV), la proteína E1A de adenovirus o el antígeno T del poliomavirus SV40 (Hein y col. 2009; Singhal y col. 2013; Squarzanti y col. 2018).

Otras proteínas virales son capaces de unirse y activar de manera directa los componentes de la DDR. Por ejemplo, la proteína E7 de HPV se une directamente a ATM promoviendo la activación de sus dianas *downstream* (Liang y col. 2006). La proteína Vpr del HIV perturba uno de los puntos de control del ciclo celular causando la activación de la ruta de ATR mediante su interacción con DDB1 (Schröfelbauer y col. 2007).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

Otro mecanismo bien distinto consiste en conseguir la activación de la DDR mediante la acumulación de especies reactivas de oxígeno. Este fenómeno ha sido observado en el virus de Epstein-Barr (EBV) y HTLV-1 mediado por las proteínas virales EBNA1 y Tax, respectivamente (Gruhne y col. 2009; Kinjo y col. 2010).

Se ha postulado que la activación de la DDR podría tener algunos beneficios para el virus. Por ejemplo, el reclutamiento de componentes a los centros de replicación virales podría proporcionarle al virus una herramienta para controlar la calidad de sus genomas antes del ensamblaje. Sin embargo, el papel de los componentes activados durante la infección en la replicación del virus todavía es desconocido.

Por otro lado, aunque parece que la activación de la DDR supone una ventaja para los virus durante la replicación, puede suponer un problema para su supervivencia. En consecuencia, los virus también han desarrollado estrategias para prevenir aquellos efectos de la DDR no deseados. Entre estos mecanismos encontramos la inhibición mediante antagonistas virales, degradación y la relocalización de componentes específicos de la DDR.

La acción de ubiquitin-ligasas virales sobre determinados componentes de la DDR fue por primera vez descrita en un estudio en el que se examinaba la restricción de la replicación viral en HAdV5 (Stracker y col. 2002). En este estudio se demostró que la expresión

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

del complejo E1B55K/E4orf6 promovía la degradación del complejo MRN durante la infección. A partir de ese momento fueron hallados otros procesos similares en otros serotipos de adenovirus. En el serotipo 12 (HAdV12) se observó que la proteína viral E4orf6 degradaba a TopBP1, inhibiéndose la ruta de ATR (Blackford y col. 2010). Otra diana hallada para el complejo E1B55K/E4orf6 fue la DNA ligasa IV, cuyo papel es esencial durante la reparación mediante recombinación no homologa (Baker y col. 2007). Otro ejemplo es el de la degradación de componentes de la DDR como DNA-PK por la ubiquitin-ligasa ICP0 en HSV (Parkinson y col. 1999; Chaurushiya y col. 2012).

Otro mecanismo diferente se basa en cambiar la localización de componentes de la DDR específicos para promover la replicación viral. Durante la infección, en HSV se ha demostrado que RPA, ATRIP y ATR se localizan en los centros replicación, lo que supone la inhibición de esta ruta (Wilkinson y Weller 2006; Mohni y col. 2010). La proteína E4orf3 de adenovirus promueve la agregación de MRN en el citoplasma, inhibiendo las rutas de ATR y ATM (Stracker y col. 2002). La acción del virus sobre proteínas mediadoras de la ruta de ATM es muy común. Por ejemplo, la proteína del core de HCV interacciona con Nbs1 o la proteína Tax de HTLV-1 se une a MDC1, quedando inhibida en ambos casos la ruta de ATM (Belgnaoui y col. 2010; Machida y col. 2010).

INTRODUCCIÓN: Los virus y el sistema de reparación del DNA

Al mismo tiempo, debido a que la activación de las rutas de reparación puede dar lugar a la entrada de la célula en apoptosis, los virus también han desarrollado mecanismos de acción frente a componentes *downstream*. Por ejemplo, algunos virus como adenovirus, EBV o HPV han desarrollado estrategias para prevenir la activación de p53 y así promover la supervivencia celular durante la replicación (Scheffner y col. 1990; Querido y col. 2001; Sato y col. 2009; Pennella y col. 2010; Soria y col. 2010).

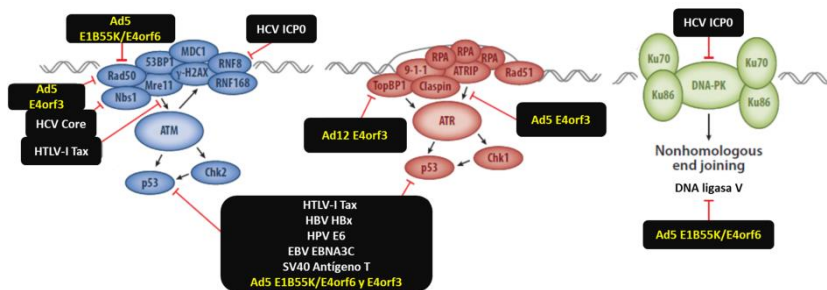


Figura 10. Inhibición de las vías de señalización de la DDR por proteínas virales. Cada una de las proteínas virales antagonistas se muestra en la figura. La mayoría de estas interacciones así como la inhibición de la ruta *downstream* en la DDR, han sido confirmadas bioquímicamente. En color amarillo se muestran las interacciones que tienen lugar durante la infección de adenovirus, nuestro modelo de estudio (Adaptado de Luftig 2014a).

OBJETIVOS

OBJETIVOS

1. Estudio de la presión de selección y la diversidad genética en el género *Mastadenovirus*
2. Estimación de la tasa de mutación de HAdV5
3. Estudio del posible papel del sistema de reparación de DNA sobre la tasa de mutación de HAd5

METODOLOGÍA

1. Bases de datos de secuencias del género *Mastadenovirus*

Se descargaron 618 secuencias de genoma completo de *GenBank* pertenecientes al género *Mastadenovirus*. Además de la secuencia de genoma completo, para cada una de ellas, se descargó un fichero en formato FASTA que contenía las regiones codificantes concatenadas. Utilizando un script de *R* (*Rstudio Version 1.1.419*; (www.r-project.org/), cada uno de las regiones codificantes se extrajo de los correspondientes ficheros y aquellas sin función asignada o desconocida anotadas como hipotéticas fueron eliminadas.

Basándonos en la nomenclatura asignada por el autor y recurriendo a la base de datos *Universal Protein Resource* (UniProt) para la búsqueda de sinónimos, creamos ficheros en formato FASTA independientes para cada región codificante en los que concatenamos las secuencias correspondientes. Una vez teníamos un fichero para cada una de las CDS, se eliminaron las secuencias no variables.

2. Alineamiento de secuencias

Algunas de las herramientas bioinformáticas a las que recurrimos para realizar los análisis posteriores tienen como requerimiento que el alineamiento que se utilice como *input* sea de nucleótidos, por codones y además, debe estar en la pauta de lectura adecuada.

Para cumplir estos requisitos, convertimos nuestros ficheros concatenados a aminoácidos utilizando el software Aliview 1.24 (www.ormbunkar.se/aliview). Seguidamente, utilizamos el software MAFFT v7.310 (<https://mafft.cbrc.jp/alignment/software/>) para llevar a cabo el alineamiento de las secuencias. Finalmente, convertimos el alineamiento resultante a nucleótidos utilizando un script de *R*. Para eliminar los codones stop presentes en el alineamiento se recurrió al script *CleanStopCodons.bf* presente en el software HyPhy (del inglés, *Hypothesis Testing using Phylogenies*) (<http://www.hyphy.org/>).

3. Árbol filogenético

Para observar las relaciones evolutivas entre las secuencias descargadas de adenovirus pertenecientes al género *Mastadenovirus*, se construyó un árbol filogenético. En este caso se trabajó con los ficheros de secuencias de adenovirus de genoma completo.

Utilizando un script de *R*, se eliminaron primero las secuencias codificantes sin función asignada definidas como hipotéticas. Seguidamente, se eliminaron las secuencias no variables, es decir las que no mostraban diferencias. Después, estos genomas únicos se concatenaron en un único fichero y se procedió a realizar el alineamiento.

Para realizar el árbol filogenético utilizamos la herramienta IQ-TREE (<http://www.iqtree.org/>) mediante el método de máxima verosimilitud introduciendo como *input* nuestro alineamiento en formato FASTA. Se recurrió a los comandos de *-m* para la búsqueda del mejor modelo de sustitución que se ajustara a nuestros datos y *-bb* para especificar el número de réplicas *bootstrap*. Los resultados obtenidos fueron visualizados y analizados utilizando FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

4. Recombinación y selección

Previamente al estudio de la presión de selección en cada uno de los genes de adenovirus, se investigó la presencia de recombinación en nuestros alineamientos. Para ello se utilizó la herramienta *Genetic Algorithm for Recombination Detection* (GARD) y el servidor en línea Datamonkey (<https://www.datamonkey.org/>).

METODOLOGÍA

Para calcular los ratios $\frac{dN}{dS}$ (ω) (donde dS corresponde al número de sustituciones sinónimas por sitio no sinónimo y dN a al número de sustituciones no sinónimas por sitio sinónimo) y determinar las presiones de selección posición por posición se recurrió a la herramienta *Single-Likelihood Ancestor Counting* (SLAC), también localizada en el servidor *on line* Datamonkey. Este método utiliza una combinación de aproximaciones de máxima verosimilitud y de conteo para calcular las tasas de sustitución no sinónimas (dN) y sinónimas (dS) en cada posición de un alineamiento dado y su correspondiente filogenia. SLAC asume que la presión de selección en cada una de las posiciones es constante a lo largo de toda filogenia.

Para llevar a cabo este análisis se utilizaron como *input* cada uno de los ficheros de alineamiento en formato FASTA de cada uno de las regiones codificantes por separado. Los alineamientos seguían el formato permitido por la herramienta: alineamiento de nucleótidos, en pauta y en ausencia de codones de stop.

5. Diversidad genética

La diversidad en adenovirus se estudió realizando dos aproximaciones distintas. En la primera de ellas, se estudió la diversidad a lo largo del genoma de las diferentes especies de adenovirus que conforman el género *Mastadenovirus*.

Después de realizar el alineamiento de secuencias, se utilizó el software DNAsp v5.10.01 (<http://www.ub.edu/dnasp/>) para estimar el número promedio de diferencias nucleotídicas por posición entre las secuencias. Para obtener las gráficas, se realizó un análisis por ventana deslizante (*sliding window*) de tamaño 200 pb y un *step size* de 20 pb.

La segunda aproximación consistió en estudiar la diversidad promedio en cada uno de los genes de adenovirus. Para ello, se utilizaron los alineamientos de cada una de las regiones codificantes y se estimó la diversidad promedio tanto a nivel de nucleótidos como de aminoácidos.

Para el cálculo de la diversidad nucleotídica se utilizó el software Variscan versión 2.0.3 (<http://www.ub.edu/softevol/variscan/>) y los alineamientos se convirtieron a formato PHYLIP. Finalmente, para el cálculo del número promedio de sustituciones de aminoácidos por posición entre secuencias, se utilizó el software MEGA versión 10.0.5 (<https://www.megasoftware.net/>) y el modelo de corrección de Poisson.

METODOLOGÍA

6. Células

Las líneas celulares tumorales humanas HeLa (CCL-2) y HeLa H1 (CRL-1958), ambas de adenocarcinoma de cérvix, fueron obtenidas de la *American Type Culture Collection* (ATCC). La línea celular HeLa fue utilizada para la amplificación y la posterior obtención de stocks de adenovirus, mientras que las células HeLa H1, se emplearon para la estimación del título viral mediante el ensayo de titulación en placa. Por otra parte, la línea celular MRC-5 (CCL-171) (fibroblastos embrionarios de pulmón humano), también procedente de la ATCC, fue utilizada como célula hospedadora no tumoral durante la evolución experimental de adenovirus.

Estos tres tipos celulares fueron cultivados en Dulbecco's Modified Eagle Medium (DMEM) suplementado al 10 % de suero bovino fetal inactivado (FBSi), 1% penicilina- estreptomina (P-S) y 1% fungizona (F), en condiciones de humedad relativa (95%), temperatura (37°C) y CO₂ (5%) controladas. La existencia de micoplasma se testó mediante PCR previamente a la realización de los ensayos correspondientes y de forma rutinaria cada vez que se descongelaba una alícuota. Cuando las células alcanzaban la confluencia, se subcultivaban realizando un lavado con tampón fosfato salino (PBS por sus siglas en inglés, de *phosphate buffered saline*) y posteriormente, se añadía 1 ml de tripsina-EDTA 0.25%-0.02%. Tras 5 min de incubación, se procedía a inactivar la tripsina añadiendo 10 ml de DMEM completo.

En el caso de las líneas celulares tumorales, la ratio de subcultivación fue 1:2-1:10 en función de las necesidades y para las células MRC-5 de 1:2-1:6 y únicamente hasta pase 12. La criopreservación se llevó a cabo empleando DMEM completo suplementado al 5% DMSO. Los criotubos se almacenaban a -70°C y al día siguiente en nitrógeno líquido hasta su posterior uso.

7. Virus

El HAdV5 wt (Adwt300) fue amablemente cedido por el Dr. Ramón Alemany del Instituto de Investigación Biomédica de Bellvitge (IDIBELL). El virus fue crecido en células HEK293 y posteriormente purificado mediante centrifugación en gradiente de CsCl.

8. Infección en medio líquido

Las infecciones en medio líquido se realizaron tanto en células MRC-5 como en HeLa, dependiendo del objetivo experimental. La confluencia celular fue establecida en función del tiempo de infección del virus, pues la infección era más tardía o más temprana según la célula huésped utilizada. Las infecciones se realizaban de forma rutinaria en placas de cultivo de 100mm de diámetro. Cuando la monocapa celular mostraba una confluencia aproximada del 70-90%, el medio de cultivo se retiraba y se añadían 5 ml de medio de

METODOLOGÍA

infección (DMEM suplementado al 5% FBSi, 1% P-S, 1% F). Sobre este medio de infección, se agregaba la cantidad de virus necesaria para alcanzar la multiplicidad de infección (MOI) establecida. Las placas se incubaban a 37°C durante 4 h para facilitar la adsorción del virus y posteriormente, se añadían 5 ml de medio de infección. Transcurrido el tiempo de infección deseado y observarse efecto citopático (CPE), se procedía a recoger el sobrenadante y/o la monocapa celular.

9. Stock viral

Para la obtención de stocks de adenovirus, se realizaba una infección en medio líquido, generalmente a MOI alta. Transcurrido el período de infección (normalmente, 72-96 hpi (horas post inoculación)) y observarse CPE, se recogía tanto el sobrenadante como la monocapa celular en un tubo de 15 ml y este se centrifugaba a 500 g durante 5-10 min. El pellet celular obtenido se lavaba con PBS, se resuspendía en 1 ml de DMEM sin suplementar y era sometido a tres ciclos consecutivos de congelación (nitrógeno líquido)-descongelación (baño de agua a 37°C). Después de cada ciclo, la muestra se agitaba vigorosamente durante 30 s. Los restos celulares eran eliminados mediante centrifugación y el sobrenadante resultante se alicuotaba y almacenaba a -70°C para su posterior uso.

10. Titulación en placa

Para la titulación del virus, el día anterior, se sembraban células HeLa H1 en placas de cultivo de seis pocillos. En el momento de la infección, las células debían presentar una confluencia de aproximadamente el 70%. Tras verificar la confluencia bajo el microscopio, el medio de cultivo se aspiraba y se añadía 1 ml de medio de infección a cada uno de los pocillos. Paralelamente, se realizaban diluciones seriadas del stock de virus de título desconocido.

Cada uno de los pocillos se inoculaba con 100 µl de la dilución de virus correspondiente. Transcurrido el periodo de adsorción (4h), el medio de infección se retiraba y se añadían sobre la monocapa 2 ml de medio semisólido (DMEM suplementado al 1% FBSi, 1% P-S, 1% F y Agar Noble 0.8%). Una vez la capa de agar solidificaba, agregábamos 2 ml de medio nutriente (DMEM suplementado al 1% FBSi, 1% P-S y 1% F) y dejábamos incubar durante 5 días a 37°C.

Para observar y contar las calvas formadas por el virus, las células se fijan con formaldehído al 10% durante 30-60 min y se tiñen añadiendo una solución de 2% cristal violeta en formaldehido 2%. Para determinar la concentración de virus, se contaban las calvas que producía una determinada dilución y, posteriormente, se hacían los ajustes pertinentes para llegar a la concentración de virus original.

METODOLOGÍA

Para ello se utilizaba la siguiente fórmula y los títulos se expresaban en unidades formadoras de placas (PFU) /ml.

$$\text{Título stock (PFU/ml)} = \frac{\text{Número de calvas}}{\text{Dilución} \times \text{Volumen inóculo}}$$

11. Dilución límite

Para llevar a cabo el aislamiento de clones únicos de adenovirus, se recurrió al método de dilución límite. Este método es ampliamente utilizado para determinar el título de aquellos virus que no forman calvas. Se basa en realizar diluciones seriadas del virus e inocularlas en réplicas de cultivo de células sembradas en placas de 96 pocillos. Posteriormente, se identifican los pocillos infectados mediante la visualización de CPE bajo el microscopio.

En nuestro caso, realizamos estratégicamente diluciones seriadas del virus, centrándonos en las muestras más diluidas. Al inocularlas en nuestros cultivos, deberemos ser capaces de infectar algunos de los pocillos con una única partícula viral.

Para aislar clones únicos del virus, se realizaron tres ensayos de dilución límite consecutivos. Se sembraron células HeLa en placas de 96 pocillos para que, el día del ensayo, se encontraran aproximadamente al 50% de confluencia. Se prepararon diluciones seriadas del virus en función del título del stock.

Se eliminó el medio de cultivo y cada dilución se inoculó por octuplicado, correspondiendo cada columna de la placa de 96 pocillos a una dilución de virus distinta. Transcurridos 10 días, se identificaron aquellos pocillos que habían sido infectados con una única partícula viral bajo el microscopio. Una vez identificados, se aislaron y se almacenaron a -70°C para su posterior uso.

12. Cálculo del número de partículas virales infecciosas producidas por célula (“*burst size*”)

El número de partículas virales producidas por célula infectada, conocido como *burst size*, nos puede ayudar a estimar el número de ciclos de infección transcurridos durante un determinado proceso experimental, pues la liberación de partículas virales constituye el último paso del ciclo. Para estimar el número de PFUs producidas por célula infectada, se infectaron células HeLa a $\text{MOI} = 10$ en placas de 24 pocillos por duplicado. El sobrenadante se recogió a los 3 y 7 dpi (días post inoculación). Posteriormente, el número de partículas virales se estimó mediante ensayo en placa.

METODOLOGÍA

El *burst size* se calculó como:

$$\text{Burst size (partículas virales/célula)} = \frac{\text{partículas virales a las 72 hpi}}{\text{número de células infectadas}}$$

13. Purificación de DNA viral

Se diseñaron dos protocolos distintos de purificación de DNA viral en función de las necesidades experimentales. Ambos se diseñaron con el objetivo de reducir al máximo las concentraciones de DNA celular y de obtener suficientes cantidades de DNA para NGS sin necesidad de realizar una amplificación de DNA mediante PCR previa. El primero de ellos se basa en utilizar un detergente suave (Tritón X-100) para reducir la lisis de la membrana nuclear durante el proceso de purificación de DNA y evitar la salida del DNA de la célula al medio extracelular. Este método se utilizó cuando el virus se amplificó en células HeLa a alta MOI, pues el número de partículas producidas era muy elevado y la reducción de la lisis de la membrana celular fue suficiente para obtener grandes cantidades de DNA viral.

Sin embargo, cuando el virus se amplificaba en células no tumorales, recurríamos al segundo método de purificación. En este caso, después de la lisis celular realizada mediante ciclos de congelación-descongelación, se añadía a la muestra DNAsa I (FisherScientific) con el objetivo de degradar específicamente todo el DNA celular, pues el DNA viral se encuentra protegido por la cápside.

13.1 Método de purificación utilizando un detergente suave

Después de realizar la infección en medio líquido empleando el tipo celular correspondiente y manteniendo la MOI deseada, se recogió tanto el sobrenadante como la monocapa celular en un tubo de 15 ml.

Las células se recuperaron mediante centrifugación a 500 g durante 10 min y se lavaron con PBS. Esta suspensión celular (aproximadamente 10^7 células) se incubó durante 30 min en hielo añadiéndose 1 ml de tampón de lisis (50 mM Tris-HCl pH 8.5, 150 mM NaCl y 1 % Tritón X-100 (Sigma-Aldrich)). Cada 10 min se agitó vigorosamente la muestra para facilitar la lisis celular. Los restos celulares se eliminaron mediante centrifugación durante 10 min a 4°C y el sobrenadante se trató con 10 mg/ml de RNAsa A durante 1 hora a 37°C.

METODOLOGÍA

Posteriormente, se añadió el tampón de lisis viral (10% SDS, 0.5M EDTA y 10 mg/mL proteinasa K (3115887001 Sigma-Aldrich) y las muestras se incubaron durante 1 hora a 56°C. El DNA viral se purificó añadiendo 2V de UltraPure™ Phenol: Chloroform: Isoamyl Alcohol (25:24:1, v/v) (Sigma-Aldrich) y se precipitó utilizando una mezcla de isopropanol y NaCl. Finalmente, se lavó con EtOH al 70%, se resuspendió en ddH₂O y se almacenó a -20°C. La concentración de DNA viral se determinó empleando el fluorímetro Qubit® 2.0 (Invitrogen) y el Qubit® dsDNA HS Assay Kit (Invitrogen) siguiendo las indicaciones proporcionadas por el fabricante.

13.2 Método de purificación utilizando DNAsa I

Tras la infección realizada en placas de 100 mm, el sobrenadante y la monocapa celular se muestrearon en tubos de 15 ml que se centrifugaron a 500 g durante 5-10 min. El pellet celular se lavó con PBS y se resuspendió en 1 ml de DMEM sin suplementar.

Tres ciclos de congelación-descongelación consecutivos se realizaban para lisar las células y posteriormente, estos restos celulares se eliminaban mediante centrifugación. El sobrenadante resultante se trataba con DNAsa I siguiendo las instrucciones del fabricante durante 30 min a 37°C.

Transcurrido el periodo de incubación, la enzima se inactivó por calor a 65°C durante 10 min en presencia de EDTA y finalmente, se procedió a realizar el tratamiento con RNAsa, la lisis de las cápsides virales y la purificación del DNA viral siguiendo los mismos procedimientos que en el protocolo anterior.

14. Estimación del ratio DNA viral:celular

Para cada muestra, 1 µg de DNA viral se fragmentaba utilizando un sonicador Covaris (Covaris S2 Sonication System) obteniéndose fragmentos de aproximadamente 200 pb. Para la generación de extremos romos se utilizó el NEBNext® End Repair kit (NEB). Después de purificar y concentrar el DNA utilizando el DNA Clean & Concentrator kit (Zymo), se procedió a la ligación de estos fragmentos de DNA en el plásmido pJET 1.2 utilizando el CloneJET PCR Cloning Kit (K1232 ThermoScientific) y la enzima T4 DNA ligasa (Invitrogen).

Seguidamente, el plásmido se transformó mediante choque térmico (42°C 40 s, 2 min en hielo) en células competentes de la cepa DH5α de *Escherichia coli* (NZY) que, finalmente, se sembraron en placas *petri* de LB + Ampicilina (100 µg/ml) (ThermoScientific) ON a 37°C.

METODOLOGÍA

Transcurrido el tiempo de incubación, entre 20-30 colonias fueron sometidas a PCR de colonia empleando una pareja de cebadores que flanquean el sitio de inserción en el plásmido y la Taq DNA polimerasa (VWR). El programa de PCR utilizado fue el siguiente: 95 °C, 3 min; 94 °C, 30 s, 60 °C, 30 s, 72 °C 30 s; 25 ciclos.

El tamaño y la presencia/ausencia de producto de PCR se observó en un gel de agarosa al 0.8% añadiendo el 10% de la muestra en *Loading Buffer* (LB) (Thermo Scientific) y un marcador de peso molecular (Thermo Scientific). Finalmente, se estudió el origen (celular o viral) de las muestras de PCR positivas mediante secuenciación de Sanger. Los cromatogramas (ficheros ab1) generados, se observaron y procesaron utilizando el *Staden Package* (<http://sourceforge.net/projects/staden/>).

El DNA se clasificó mediante la búsqueda de similitud entre nuestras secuencias y las de la *National Center for Biotechnology Information* (NCBI) utilizando la herramienta *Standard Nucleotide BLAST* (BLASTn). Sólo aquellas muestras en las que el porcentaje de DNA viral fue ≥ 70 %, fueron sometidas a NGS.

15. Secuenciación dúplex (DS)

La alta fidelidad de la secuenciación dúplex reside fundamentalmente en la ligación de adaptadores de secuenciación de doble cadena a ambos extremos de nuestro DNA diana fragmentado. Estas etiquetas moleculares, que contienen una secuencia aleatoria pero a su vez complementaria, permiten el rastreo de cada una de las hebras de DNA original y por tanto, la identificación de mutaciones reales. Esto es debido a que seremos capaces de eliminar los artefactos de secuenciación, pues aparecen únicamente en una de las dos hebras de DNA.

Los adaptadores de dúplex se construyeron mediante la hibridación de dos oligonucleótidos, conteniendo uno de ellos la secuencia de cadena simple aleatoria de 12 nucleótidos. Seguidamente, estos cebadores fueron extendidos utilizando el fragmento *Klenow* y digeridos para la obtención de fragmentos cohesivos. Una vez preparados los adaptadores, se procedió a la preparación de librerías tal y como se describió previamente (Kennedy y col. 2014).

El DNA de adenovirus previamente purificado se fragmentó utilizando un sonicador Covaris (Covaris S2 Sonication System) y los fragmentos de DNA resultantes se seleccionaron por tamaño empleando las Ampure XP beads (Beckman Coulter). Las muestras se secuenciaron utilizando el equipo NextSeq (Illumina) obteniéndose lecturas de 150 pb.

METODOLOGÍA

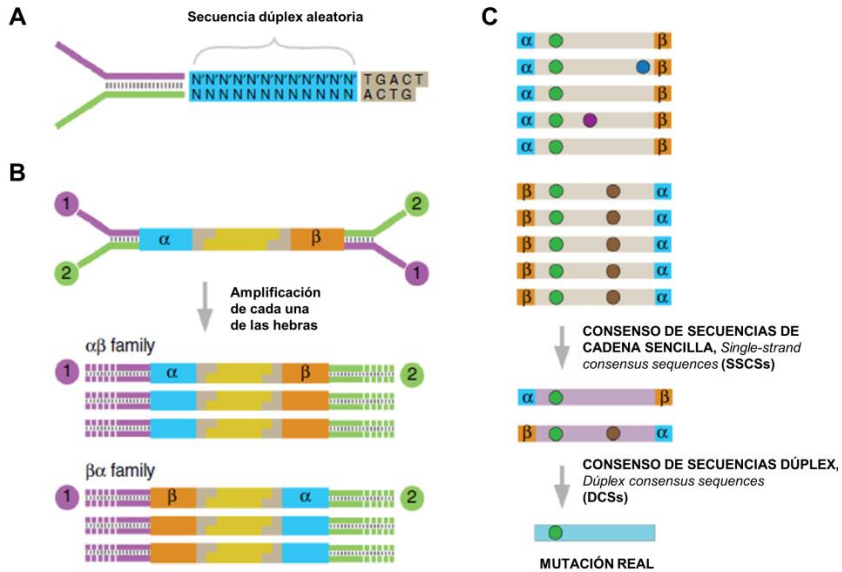


Figura 11. Método de análisis de secuenciación dúplex. **A)** Esquema de los adaptadores de secuenciación dúplex. Estos adaptadores presentan una secuencia de doble cadena complementaria y aleatoria (α y β) que nos permitirá distinguir entre ambas hebras de DNA. **B)** Ligación de los adaptadores de dúplex a ambos extremos de los fragmentos de DNA que se desean secuenciar (color amarillo). Posteriormente a la ligación, se procederá a la amplificación específica de cada una de las hebras de DNA, generándose dos familias de productos de PCR ($\alpha\beta$ y $\beta\alpha$). **C)** Análisis de los datos de secuenciación. El primer paso consiste en obtener los SSCs agrupando aquellas lecturas que pertenecen a las familias $\alpha\beta$ y $\beta\alpha$ y obteniéndose el consenso de las mismas. Finalmente, deben generarse los DCSs comparando los SSCs de las familias $\alpha\beta$ y $\beta\alpha$. Estas secuencias dúplex finales obtenidas nos permitirán identificar las mutaciones reales (color verde) que estarán presentes en ambas hebras de DNA. Este método de análisis de secuencias nos permite eliminar errores de secuenciación (color azul y morado) y errores de PCR introducidos durante la primera ronda de replicación (color verde) (Adaptado de Loeb y col. 2014).

Los archivos FastQ resultantes se procesaron siguiendo la *pipeline* <https://github.com/loeblab/Duplex-Sequencing>, utilizando BWA 0.6.2, Samtools 0.1.19, Picardtools 1.130 y GATK 3.3–0. El análisis computacional consistió en tres pasos fundamentales: emparejamiento de etiquetas y alineamiento inicial, ensamblaje de las secuencias consenso de simple cadena (SSCS, *single stranded consensus sequence*) y, por último, el ensamblaje de secuencias consenso dúplex (DCS, *dúplex consensus sequence*). Finalmente, los datos de DCS procesados se alinearon con la secuencia de referencia correspondiente y se procedió a la identificación de mutaciones.

16. PCR cuantitativa a tiempo real

La presencia de determinadas mutaciones observadas mediante DS se comprobó mediante PCR cuantitativa a tiempo real. Para cada una de las mutaciones, se diseñaron tres parejas de cebadores: uno de ellos localizado aproximadamente 200 pb *upstream* de la posición de la mutación (*Forward*) y otros dos cebadores en los que la base 3' hibrida con la base mutada (*Reverse WT* y *Reverse MUT*); uno de ellos emparejando perfectamente con la base mutante y el otro, con la base WT.

METODOLOGÍA

Las reacciones se llevaron a cabo utilizando el Agilent's Brilliant III Ultra-Fast SYBR Green qPCR Master Mix (600882 Agilent Technologies) en el equipo Agilent AriaMx Real-time PCR (Agilent Technologies). Para cada mutación, se realizaron dos reacciones por separado bajo las mismas condiciones experimentales, una de ellas empleando el primer *Reverse* WT y la otra el *Reverse* MUT (Tabla 2).

Tabla 2. Cebadores utilizados en las reacciones de PCR cuantitativa.

Posición genómica	Orientación cebador	Nombre	Tm	Secuencia
191	F	191_F	65	5'- GGG CGG GTG ACG TAG TAG TG-3'
191	R	191_R	56	5'- TTG TCA CTT CCT GTG TAC AC-3'
191	R	191_R_MUT	54	5'- TTG TCA CTT CCT GTG TAC AT-3'
7295	F	7295_F	58	5'- CAA ACT CTT CGC GGT CTT TC- 3'
7295	R	7295_R	60	5'- GTC ATG GTC AGG GAC ACC TT -3'
7295	R	7295_R_MUT	63	5'- GTC ATG GTC AGG GAC ACC TC -3'
9417	F	7295_F	58	5'-TAG AAG TCC ACG GCG AGG TT- 3'
9417	R	9417_R	71	5'- GCC GTG TCC CCC CTC CCC CA - 3'
9417	R	9417_R_MUT	73	5'- GCC GTG TCC CCC CTC CCC CC - 3'
33565	F	33565_F	60	5'- TAA CTG CAG CAC AGC ACC AC- 3'
33565	R	33565_R	54	5'- GAG GTG GTG AAT TAC AAC AT- 3'
33565	R	33565_R_MUT	56	5'- GAG GTG GTG AAT TAC AAC AC- 3'

Debido a que el emparejamiento de la base 3' es crítico para la fase de extensión y la Taq DNA polimerasa utilizada no presenta actividad 3' exonucleasa, en cada una de las dos reacciones, se amplificara selectivamente el DNA molde mutante o WT. Todas las reacciones se llevaron a cabo por triplicado, utilizando 0.3 ng de DNA de HAdV5 previamente purificado y 200 nM de cada uno de los cebadores.

El programa de qPCR utilizado fue el siguiente: desnaturalización inicial (95°C 3 min), 40 ciclos de amplificación (95°C 5 s, temperatura de *annealing* 10 s, 72°C 10 s) y un ciclo final de *melting* (95°C 30 s, 65°C 30 s, 95°C 30 s). Finalmente, el análisis de los datos obtenidos así como la estimación de los valores de C_t se llevaron a cabo utilizando el software de AriaMx proporcionado por el fabricante del equipo.

17. Análisis de la relación entre tasa de mutación y diversidad *in vivo*

Se calculó la frecuencia (f) de las bases A, T, C y G para cada una de las posiciones genómicas y finalmente, la diversidad nucleotídica se estimó como:

$$H = 1 - f_A^2 - f_T^2 - f_G^2 - f_C^2$$

Para llevar a cabo este análisis de diversidad se utilizó un script de R y los alineamientos obtenidos previamente para una de las regiones codificantes de adenovirus.

METODOLOGÍA

18. Activación e inhibición de la DDR: compuestos químicos

La inhibición farmacológica de la DDR celular se llevó a cabo utilizando tres inhibidores distintos. Cada uno de ellos suprime específicamente la actividad de una de las tres principales rutas que intervienen en la reparación del DNA en células eucariotas.

Previamente a la inoculación, se llevaba a cabo la inhibición de las rutas de reparación de DNA mediante la exposición de las células a estos inhibidores durante 2 h. El compuesto KU-55933 (SML1109 Sigma-Aldrich), se disolvió en DMSO a una concentración de trabajo de 10 mM y se utilizó para la inhibición de la proteína ATM. El compuesto NU7741 (A8315 ApexBio), supresor de DNA-PK, se disolvió en DMSO a una concentración de trabajo de 5 mM. Por último, el compuesto VE-821 se utilizó para inhibir a ATR y se disolvió en DMSO a una concentración final de 15 mM (A2521 ApexBio).

La neocarzinostatina (NCS) (0.5mg/ml N9162 Sigma-Aldrich) se utilizó como inductor de la DDR, pues es capaz de causar daños en el DNA. Este compuesto se utilizó para testar la efectividad de los inhibidores de la DDR en nuestro tipo celular. Las células, previamente tratadas con los inhibidores de la DDR eran expuestas a 1000 ng/ml de NCS durante 30 min a 37°C y posteriormente, la actividad de las rutas de reparación del DNA se testaba realizando ensayos de inmunofluorescencia.

19. Ensayos de citotoxicidad

Como indicador de la viabilidad celular se empleó el compuesto no fluorescente resazurina (R7017 Sigma-Aldrich), de color azul y permeable a las células. En el interior de las células metabólicamente activas, la resazurina es reducida a resorufina, un compuesto de color rojo y fluorescente. Por ello, la resazurina nos permite medir la viabilidad celular en base al estudio del poder reductor de las células vivas en cultivo.

Se sembró el tipo celular correspondiente en placas de 96 pocillos a la confluencia celular deseada. Como blanco, se utilizaron tres pocillos a los que únicamente se les añadió DMEM completo. Al día siguiente, tras inspeccionar las células bajo el microscopio, se aspiró el medio de cultivo y se trató a las células con diferentes concentraciones del compuesto correspondiente por triplicado (100 µl/ pocillo). Tres pocillos se trataron únicamente con DMSO como control negativo del tratamiento. Paralelamente, se preparó un stock de resazurina (100x) disolviendo 2.5 mg de reactivo en 10 ml de ddH₂O. Se mezcló en el vórtex y se filtró utilizando un filtro de 0.2 µM. A partir de este stock, se preparó un stock 2x que se almacenó a 4°C para su posterior uso. Después del tratamiento y la incubación pertinentes, se añadió resazurina a todos los pocillos a una concentración final de 0.1 mM y la placa se incubó durante 4 h a 37°C en oscuridad. Finalmente, la fluorescencia se midió entre 540 nm y 590 nm en el sistema de multi-detección SPARK® (TECAN).

20. Inmunofluorescencia

En función del objetivo experimental y el equipo utilizado para visualizar la fluorescencia, las células se sembraron directamente en placas de seis pocillos (InCell Analyzer 2000) o en cubreobjetos circulares de 12 mm (Thermo Scientific) dentro de placas de 12 pocillos (microscopio de fluorescencia convencional).

Tras el tratamiento de la monocapa celular y la incubación correspondiente, el medio de cultivo se aspiró y la monocapa se lavó con PBS. Tras fijar las células a 4°C durante 15 min añadiendo paraformaldehído 4 %, se procedió a la permeabilización de la membrana celular agregando Triton X-100 0.1% en PBS y dejando incubar 10 min a temperatura ambiente. Después, se realizaron tres lavados de 5 min con PBS en agitación. Para reducir posibles uniones inespecíficas del anticuerpo, las células previamente fijadas se incubaron durante 1 h en solución de bloqueo (1% BSA, 10% goat serum en 0.1% PBS-Tween) y se volvió a lavar con PBS. Seguidamente, se procedió a la incubación con el correspondiente anticuerpo primario diluido en solución de bloqueo a 4°C durante 16 h.

Tabla 3. Anticuerpos utilizados durante los ensayos de inmunofluorescencia

Anticuerpo (Casa comercial)	Antígeno	Origen	Dilución	Conjugado
Anti-phospho-Histone H2A.X (05-636 Merck Millipore)	γH2AX at Ser 139	Ratón	1: 200	NA
Anti-p-Atm (sc-47739 Santa Cruz Biotechnology)	(p)-ATM at Ser 1981	Ratón	1: 200	NA
Anti-DNA PKcs (phospho S2056) (ab18192 Abcam)	(p)-DNA-PK at Ser 2056	Conejo	1: 200	NA
Phospho-Chk1 (Ser345) (2348 Cell Signaling)	(p)-Chk-1 at Ser 345	Conejo	1:50	NA
Goat Anti-Mouse IgG (ab150113 Abcam)	Raton	Cabra	1:1000	Alexa Fluor® 488
Goat Anti-Rabbit IgG (ab150077 Abcam)	Conejo		1:1000	Alexa Fluor® 488

Tras la incubación toda la noche, la monocapa se lavó tres veces con PBS para eliminar el exceso de anticuerpo no unido y, a continuación, se incubó durante 1 h con el anticuerpo secundario diluido en solución de bloqueo y en condiciones de oscuridad.

Posteriormente, los núcleos celulares se tiñeron con DAPI (D9542 Sigma-Aldrich) a una dilución 1:1000 (1 µg /ml) durante 10 min a temperatura ambiente, en agitación y en condiciones de oscuridad. Tras la tinción, se realizaron dos lavados rápidos con ddH₂O y tres lavados con PBS a temperatura ambiente.

METODOLOGÍA

En el caso de utilizar cubre-objetos circular como soporte, este se montó en un porta-objetos de tamaño 25 x 75 mm (Gold Seal Products) sobre dos gotas de FluorSave™ Reagent (Merck Millipore) con la ayuda de unas pinzas y evitándose la aparición de burbujas. Una vez seco, se guardó a 4°C y en oscuridad para su posterior visualización bajo el microscopio de fluorescencia. En este caso las imágenes tomadas fueron procesadas con el software ImageJ, versión 1.5 (National Institutes of Health).

Si se trataba de placas de seis pocillos, tras los lavados con PBS, las células se dejaban en esta solución de lavado a 4°C para ser posteriormente analizadas en el microscopio automatizado InCell Analyzer 2000 (GE Healthcare). La intensidad de fluorescencia fue cuantificada empleando el software proporcionado por el fabricante del equipo.

21. Evolución experimental

Para determinar si existe una posible relación entre el sistema de reparación del DNA de la célula huésped y la tasa de mutación de adenovirus, se llevó a cabo un ensayo de evolución experimental.

Se infectaron células procedentes de tejido sano MRC-5 en las que se asume un correcto funcionamiento de estas rutas de señalización y

células MRC-5 tratadas con inhibidores de las tres quinasas (ATM, ATR y DNA-PK) que modulan las tres principales rutas de la DDR. Tras confirmar que las células MRC-5, sembradas el día anterior en placas de 6 pocillos, mostraban una confluencia del 70%, se eliminó el medio de cultivo y se añadió 1 ml de medio de infección únicamente en aquellos pocillos en los que las células no iban a ser tratadas con los inhibidores químicos de la DDR. En los pocillos restantes, se agregó el correspondiente tratamiento de inhibición de la DDR: KU55933 10 μM , NU7741 1 μM , VE-821 5 μM , en DMEM 5% FBS. Las placas se incubaron durante 2 h a 37°C.

Después de los correspondientes tratamientos, se realizó la infección a MOI baja, 0.1 PFU/célula, y las placas se incubaron durante 4h a 37°C. Tras del periodo de adsorción, el medio se aspiró y se añadió medio de infección fresco en presencia o ausencia de los correspondientes inhibidores químicos de la DDR y las placas se dejaron en el incubador a 37° C.

A las 72 hpi, se recogió tanto la monocapa celular como el sobrenadante de infección y se realizó el procedimiento habitual para la obtención de un stock de adenovirus, cuya concentración se estimó más adelante mediante titulación en placa. Una vez conocido el número de PFUs, se procedió a repetir de nuevo la infección de cultivos frescos hasta pase 10.

22. Curvas de crecimiento

Las curvas de crecimiento nos permiten estudiar el transcurso del ciclo viral y por tanto, aspectos como el número máximo de progenie producida o la velocidad con la que replica el virus. Para realizar las curvas de crecimiento, el tipo celular deseado se infectó a una MOI establecida (generalmente baja) por triplicado en placas multipocillo. La infección se realizó siguiendo el procedimiento de infección en medio líquido. Tras el periodo de adsorción, se eliminó el inóculo, se añadió medio fresco y las placas se incubaron a 37°C.

A diferentes tiempos de post-infección (normalmente cada 12 hpi), se recogió tanto el sobrenadante de infección como la monocapa celular. Cada una de estas muestras se sometió al procedimiento de obtención de stock y se almacenó a -70°C para su posterior uso. El número de PFUs para cada muestra se calculaba mediante el método de titulación en placa. Con estos datos, construíamos una curva de crecimiento que venía representada por el título de virus (PFU/ml) a diferentes tiempos de infección.

23. Identificación de deleciones: métodos no convencionales

Para la identificación de deleciones se utilizaron dos aproximaciones distintas. El tamaño de nuestro inserto en las librerías era de 300 pb y la longitud de nuestras secuencias de 150 pb.

Debido a que no se espera solapamiento entre lecturas pareadas, decidimos estudiar la distancia de mapeo entre lecturas para identificar deleciones en nuestras muestras de secuenciación.

El primer método ha sido recientemente descrito por Chateigner y colegas (Chateigner y col. 2015). Esta aproximación consiste en mapear todas las lecturas pareadas a la secuencia de referencia utilizando Bowtie 2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) sin tener en cuenta la información de pares. Posteriormente, las lecturas que han mapeado se extraen y se calcula la distancia de mapeo entre lecturas pareadas teniendo en cuenta su posición de mapeo y orientación. La obtención de valores de distancia más grandes de lo esperado entre lecturas pareadas constituye un indicador de la presencia de deleciones.

La segunda estrategia consistió en la identificación directamente de discontinuidades entre lecturas consecutivas teniendo en cuenta las posiciones de mapeo después de llevar a cabo el análisis de DS. Nuevamente, la existencia de grandes distancias entre lecturas consecutivas indica la presencia de una deleción.

RESULTADOS Y DESARROLLO ARGUMENTAL

**1. Estudio de la presión de selección y la
diversidad genética en el género
*Mastadenovirus***

1.1 Banco de datos de secuencias y filogenia

Para llevar a cabo el análisis de diversidad y de la presión de selección en adenovirus, se descargaron secuencias de genoma completo pertenecientes al género *Mastadenovirus* de GenBank. Se decidió realizar el estudio empleando secuencias pertenecientes a este género debido a que es el género de la familia *Adenoviridae* para el que existe un mayor número de secuencias disponible.

La restricción a un único género de estudio se basa en que la comparación de secuencias entre organismos cercanos filogenéticamente, facilita la identificación de regiones conservadas y variables durante la evolución y el acceso a la descripción detallada de la función de las diferentes regiones del genoma (Figura 12). Esto nos permitirá distinguir entre aquellos genes o regiones que son esenciales (evolutivamente conservados) y los que no lo son tanto (regiones variables o específicas) dentro del género.

Durante la búsqueda, algunas secuencias fueron excluidas del análisis debido a que las anotaciones eran incompletas o ausentes en algunos casos. También fueron excluidas del análisis aquellas secuencias en las que las coordenadas de las regiones codificantes eran incorrectas o incluían errores de secuenciación. Después de este filtrado, se obtuvieron 621 secuencias de genoma completo a partir de las cuales se construyeron las bases de datos.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

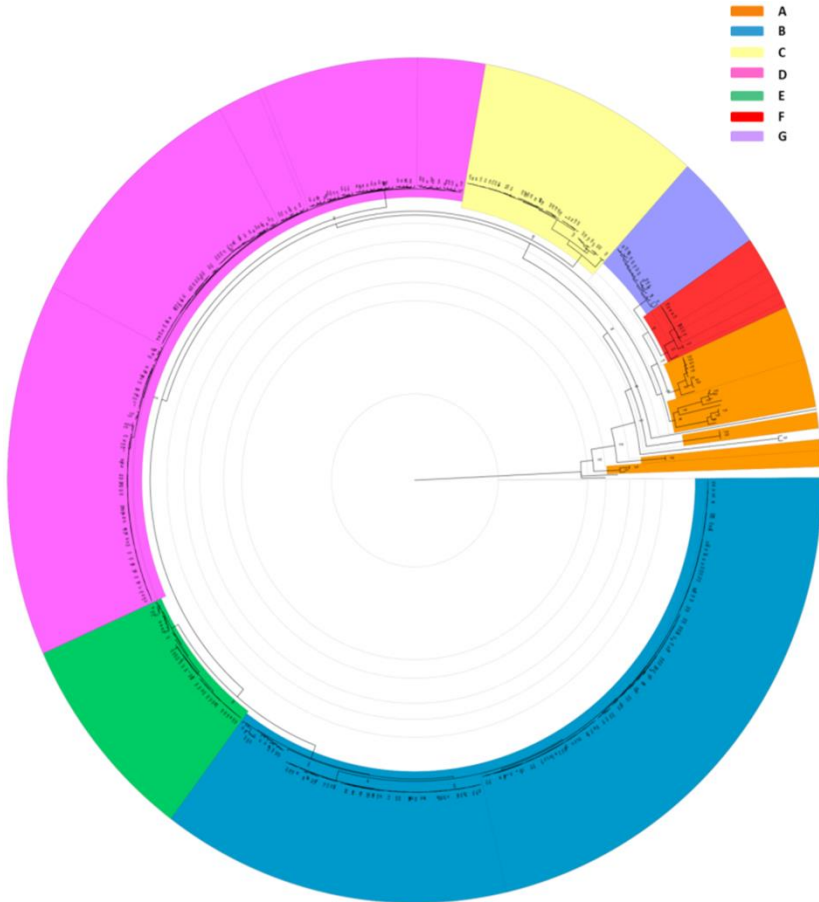


Figura 12. Análisis filogenético de nuestra base de datos de secuencias pertenecientes al género *Mastadenovirus*. El árbol fue construido utilizando IQ-TREE mediante el método de máxima verisimilitud, utilizando 1000 réplicas *bootstrap* y a partir de secuencias de genoma completo.

En el anexo 1 del material suplementario se encuentra la lista de secuencias de GenBank utilizadas para el análisis.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

El género *Mastadenovirus* se caracteriza por infectar a un amplio rango de hospedadores mamíferos. De hecho, en nuestro banco de datos existían secuencias procedentes de virus capaces de infectar a un amplio rango de hospedadores (Figura 13). Sin embargo, la mayoría de secuencias procedían de adenovirus humanos (74.43 %) seguidas por hospedadores primates (15.47%), suponemos por su interés biomédico. La mayoría de especies estaban representadas en nuestro banco de datos (Figura 13). Se observó una mayor representación de adenovirus clasificados dentro de las especies B y D (35.75% y 32.21%, respectivamente), seguidas por las especies C (9.18 %) y E (7.57%). Más detalles sobre la base de datos de secuencias se muestran en el material suplementario (Anexo 1).

Para cada entrada, se descargó tanto la secuencia de genoma completo como de las regiones codificantes asociadas. A partir de estos ficheros se obtuvieron tres bancos de datos distintos para llevar a cabo los análisis posteriores. Primero, una base de datos de genoma completo. Segundo, una base de datos en la que las secuencias de genoma completo se clasificaron por especie. Tercero, a partir del fichero de regiones codificantes, se obtuvo una base de datos para cada unidad de transcripción. Durante este proceso se eliminaron las secuencias hipotéticas y no variables.

Cabe destacar que, para algunas regiones codificantes, el número de secuencias fue más bajo. A pesar de que algunos genes son compartidos dentro de un mismo género, otros no están presentes en algunas especies.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Por ejemplo, genes como *pol*, *pTP*, *DBP* o *hexon* se espera que estén presentes en la mayoría de las especies. Sin embargo, para genes más específicos como los localizados en E3 y E4, se ha demostrado una mayor variabilidad entre géneros y especies (Davison y col. 2003). Además, dentro de los adenovirus humanos existen diferencias entre especies en el gen que codifica la proteína *fiber*. Aunque la mayoría de los serotipos codifican para una única proteína *fiber*, los tipos 40, 41 y 52 codifican para dos, ampliándose el rango de hospedadores infectados por estos virus (Hung-Yueh y col. 1994; Jones y col. 2007).

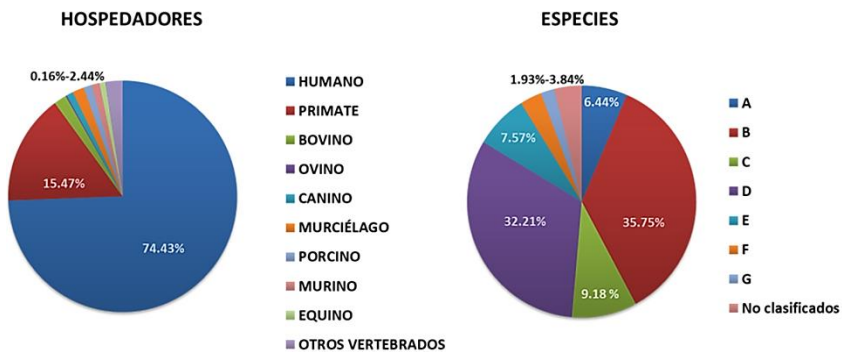


Figura 13. Porcentaje de secuencias correspondientes a cada hospedador y especie en nuestra base de datos. A la izquierda, se muestra el porcentaje de secuencias correspondiente a cada uno de los hospedadores. A la derecha de la figura, se puede observar la representación de la fracción de cada una de las especies de adenovirus presentes en nuestro banco de datos.

1.2 Diversidad en el género *Mastadenovirus*

La diversidad genética dentro del género *Mastadenovirus* se estudió mediante el alineamiento de secuencias correspondientes a cada una de las diferentes especies (A-G) por separado. Seguidamente, se estimó la diversidad nucleotídica (π), entendida como el número promedio de diferencias por posición entre dos secuencias a lo largo del genoma.

Tal como puede observarse en la figura 14, la diversidad varía a lo largo del genoma de las diferentes especies. Para las especies HAdV-A y HAdV-F, se observó un grado de diversidad a lo largo del genoma bastante uniforme, obteniéndose valores más elevados significativamente para HAdV-A que para el resto de especies.

En la mayoría de las especies, se observó un ligero aumento de la diversidad al comienzo y final del genoma y una región central más conservada (exceptuando los elevados picos de diversidad obtenidos para las proteínas *hexon* y *penton* en algunos casos). Esta observación se ajusta perfectamente a la disposición de los genes en el genoma del virus. En la parte central se localizan los genes comunes (replicación de DNA y proteínas de la cápside) y en los extremos se localizan los genes más específicos (interacción con el hospedador).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

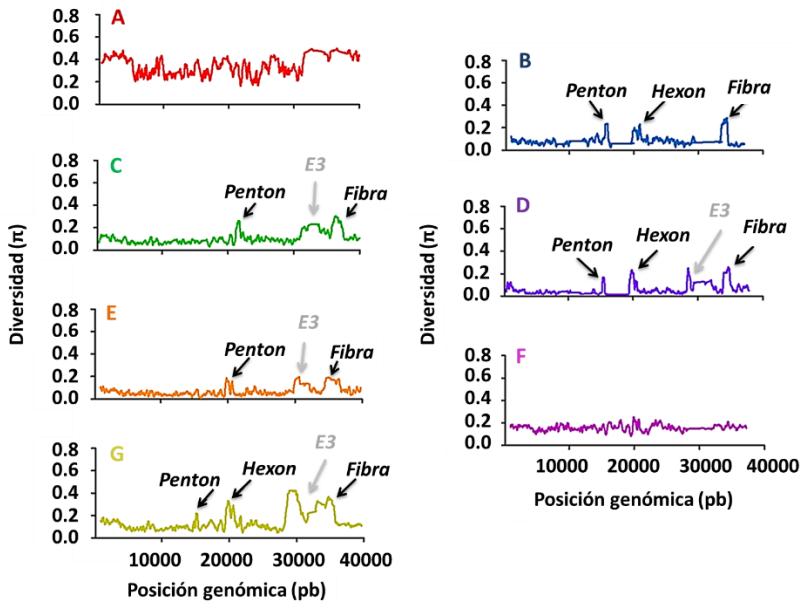


Figura 14. Diversidad nucleotídica estimada para cada una de las especies que conforman el género *Mastadenovirus*. Diversidad nucleotídica promedio por posición para cada una de las especies (A-G) obtenida utilizando el software DNAsp v5. El alineamiento de cada una de las bases de datos de secuencias de genoma completo se llevó a cabo utilizando el software MAFFT. Para construir cada una de las gráficas se empleó la herramienta *sliding window* utilizando un tamaño de ventana de 200 pb y un *step size* de 20 pb. Con una flecha se destacan aquellas zonas del genoma viral para las que se observaron valores de diversidad elevados.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Las regiones para las que se observaron elevados picos de diversidad corresponden a genes que codifican proteínas estructurales (*hexon*, *penton* y *fiber*) y la unidad de transcripción E3. La proteína *hexon* es una de las proteínas estructurales más abundantes de la cápside de adenovirus. Secuencias de ésta y otras proteínas estructurales se utilizan de forma rutinaria para la clasificación de nuevos serotipos de adenovirus, pues contienen la mayoría de los epítomos neutralizantes. La proteína *penton* es el capsómero que sella los 12 vértices de la cápside desde los que se proyecta la proteína *fiber*. Ambas proteínas juegan un papel fundamental en la entrada del virus en la célula hospedadora, el primer paso del ciclo reproductivo viral.

Por último, excepto en la especie F, E3 mostró valores de diversidad elevados. Esta región codifica proteínas virales implicadas en la interacción con el hospedador con el objetivo de modular su respuesta a la infección. Los mecanismos de defensa celular frente a patógenos están en constante evolución y, en respuesta, los virus acumulan diversidad en sus genomas fruto de la selección natural. Uno de los mecanismos desarrollados por la célula hospedadora consiste en el reconocimiento de proteínas presentes en la superficie del virus durante las primeras etapas de la infección. Esto se ajusta a los resultados de diversidad obtenidos, pues adenovirus tiende a acumular diversidad en las proteínas mayoritarias que conforman la cápside.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Por otra parte, se estudió la diversidad genética promedio gen a gen. Una vez se construyó una base de datos para cada una de las regiones codificantes, se procedió a realizar el alineamiento de secuencias y a estimar la diversidad promedio para cada una de las unidades transcripcionales. La diversidad se estimó tanto a nivel de nucleótidos como de aminoácidos (Figura 15). Los resultados obtenidos confirman los análisis de diversidad por especies mostrados anteriormente. Se observó un aumento de la diversidad en ambos extremos (nt: 0.40 ± 0.02 ; aa: 0.86 ± 0.05) y la porción central del genoma (nt: 0.31 ± 0.01 ; aa: 0.46 ± 0.04) mostró un alto grado de conservación, siendo estas diferencias significativas (t-tests: nt: $P = 0.001$; aa: $P < 0.001$).

Seguidamente, se investigó el grado de variabilidad genética presente en los genes que interaccionan con el huésped. Para ello, las regiones codificantes se dividieron en dos grupos en base a la función que desempeñan durante el ciclo de infección.

En el primer grupo, llamado “core”, clasificamos aquellas unidades transcripcionales que poseen un papel fundamental en la replicación del genoma viral y el ensamblaje de nuevas partículas virales, incluyendo las proteínas estructurales. El segundo grupo (“interacción”) acoge a las regiones codificantes que participan en la interacción con el hospedador (Tabla 4).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

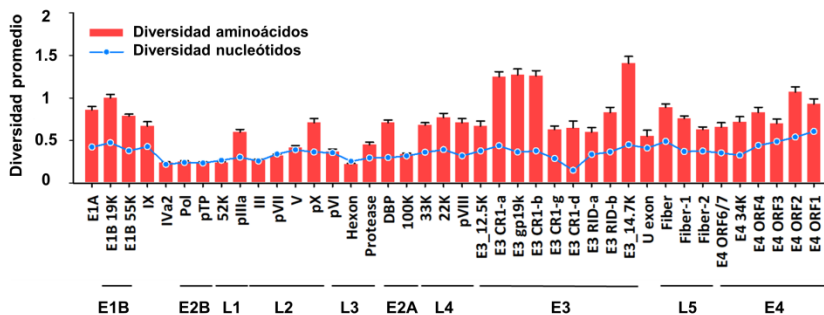


Figura 15. Diversidad promedio obtenida en cada una de las regiones codificantes. La diversidad nucleotídica (azul) y aminoacídica (rojo) promedio para cada una de las regiones codificantes fue obtenida a partir del alineamiento de secuencias del género *Mastadenovirus*. Los alineamientos se realizaron utilizando MAFFT y la diversidad nucleotídica se estimó utilizando los softwares MEGA (aa) y VariScan (nt). La unidad transcripcional a la que pertenece cada uno de los genes se muestra en la parte inferior de la figura.

En el grupo “interacción”, se clasificaron aquellas proteínas que alteran el funcionamiento del ciclo celular, el sistema inmunitario y que hacen posible la entrada del virus en la célula.

Tras el análisis, se observó que las proteínas virales que interaccionan con el hospedador (nt: 0.40 ± 0.02 ; aa: 0.83 ± 0.06) presentan una mayor variabilidad que las proteínas “core” (nt: 0.32 ± 0.02 ; aa: 0.48 ± 0.05), siendo estas diferencias estadísticamente significativas (t-tests: nt: $P < 0.05$; aa: $P < 0.001$).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Tabla 4. Funciones de las proteínas virales divididas en tres grupos: “core” (C), “interacción” (I) e “interacción con la DDR” (I-DDR).

Unidad transcripcional	Proteína	Esencial (E)			Interacción hospedador (I)				Interacción DDR (I-DDR)
		Replicación	Estructural	Ensamblaje	Sistema inmune	Ciclo celular	Entrada	Otros	
E1A	E1A					✓			
E1B	E1B 19K E1B 55K					✓			✓
IX	IX	✓	✓						
Iva2	Iva2			✓					
E2B	Pol pTP	✓ ✓							
L1	52K pIIIa		✓	✓					
L2	III pVII V pX		✓ ✓ ✓						✓
L3	pVI Hexon Protease		✓ ✓	✓					
E2A	DBP 100K	✓		✓				✓	
L4	33K 22K pVIII	✓		✓ ✓					
E3	E3_12.5K E3 CR1-α E3 gp19k E3 CR1-β E3 CR1-γ E3 CR1-δ E3 RID-α E3 RID-β E3_14.7K				✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	✓			
U exon	U exon	✓ ?							
L5	Fiber Fiber-1 Fiber-2						✓ ✓ ✓		
E4	E4orf6/7 E4 34K E4orf4 E4orf3 E4orf2 E4orf1								✓ ✓ ✓ ✓

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Estas diferencias fueron más notables a nivel de aminoácidos que de nucleótidos. En consecuencia, estos resultados sugieren que los genes “core” son preservados en la mayoría de las especies de adenovirus por lo que, en su mayoría desarrollan funciones biológicamente importantes durante el ciclo de infección. No quiere decir que estos genes no hayan mutado, sino que la selección ha actuado sobre ellos a lo largo de la evolución eliminando las variantes menos eficaces o que comprometen la funcionalidad de la proteína. Por el contrario, los genes pertenecientes al grupo “interacción” han sufrido menos restricción durante los períodos de adaptación al hospedador.

Por otra parte, numerosos estudios demuestran la capacidad de los adenovirus para modular la DDR durante la infección. Sin embargo, el posible uso de estas vías de señalización para mantener la estabilidad del genoma viral o por el contrario, su evasión para generar diversidad por parte del virus no se ha investigado todavía. En consecuencia, decidimos estudiar la diversidad en los genes implicados en la modulación de la DDR.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

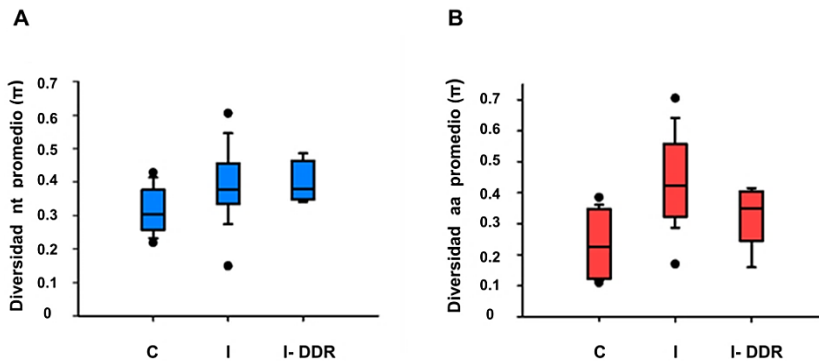


Figura 16. Diversidad nucleotídica (A) y aminoacídica (B) promedio en los tres grupos de estudio en cada uno de los genes de los adenovirus pertenecientes al género de *Mastadenovirus*.

Para ello, el grupo “interacción” se dividió en dos subgrupos y se compararon los niveles de diversidad entre los genes que codifican proteínas virales vinculadas a la DDR (“interacción con la DDR”) y el resto (“interacción”) (Tabla 4).

Sin embargo, no encontramos diferencias significativas entre los grupos “interacción” (nt: 0.39 ± 0.02 ; aa: 0.87 ± 0.06) e “interacción con la DDR” (nt: 0.40 ± 0.02 ; aa: 0.66 ± 0.09) (t-test: nt: $P = 0.840$; aa: $P = 0.087$). A pesar de ello, la obtención de valores de diversidad similares a otras unidades de transcripción como E3 y E1, implicadas en la evasión de la respuesta inmunitaria y la modulación del metabolismo celular para garantizar la replicación del virus, pone de manifiesto su importancia en la biología del virus a pesar de que tenemos bastante poca información sobre ellas (Figura 16).

1.3 Huellas de selección en los genes de *Mastadenovirus*

La presión de selección y las tasas de sustitución sinónima por sitio sinónimo (dS) y no sinónima por sitio no sinónimo (dN) se estimaron para cada una de las regiones codificantes de los adenovirus pertenecientes al género *Mastadenovirus*. A través de estas estimaciones se puede conocer la importancia de la selección en la evolución de las correspondientes secuencias así como el tipo de selección operante.

Una vez se estimaron las tasas de sustitución dN y dS , se calculó el ratio $\frac{dN}{dS}$, conocido como ω . Si el número de sustituciones sinónimas y no sinónimas es similar entonces $\frac{dN}{dS} = 1$ y en promedio no hay selección sobre la secuencia de la proteína. Sin embargo, si el número de sustituciones no sinónimas es mayor que el de sustituciones sinónimas ($\frac{dN}{dS} > 1$), esto indica que la selección positiva favoreció la diversidad en ese codón. El caso contrario ($\frac{dN}{dS} < 1$) refleja la acción de la selección negativa o purificadora para conservar la secuencia y eliminar las variantes deletéreas o de pérdida de función de la proteína.

Una vez obtenida la base de datos de secuencias y realizado el posterior alineamiento para cada una de las regiones codificantes, se procedió a identificar las huellas de selección natural utilizando la herramienta SLAC.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Esta herramienta nos permitió estimar los parámetros dN y dS para cada posición en nuestro alineamiento y de esta manera identificar el tipo de selección operante en cada uno de los codones. Al igual que en el capítulo anterior, las unidades de transcripción se clasificaron en tres grupos distintos en base a la función que desarrollan durante el ciclo de infección (“core”, “interacción” e “interacción con la DDR”) (Tabla 4).

Aunque en un número muy bajo de codones, se identificó selección positiva tanto en genes “core” como en genes implicados en la interacción con el hospedador (Tabla 5). Sin embargo, el porcentaje de codones bajo selección positiva en ambos grupos fue similar (Test de Fisher: $P = 0.77$). Dentro del grupo de “interacción”, se identificaron codones bajo selección positiva en los genes E4orf3, E4orf4, E4orf6/7 y E1B55K, cuya interacción con la DDR ha sido previamente demostrada (Luftig 2014a). Pero nuevamente, las diferencias entre ambos grupos (I e I-DDR) no fueron significativas (Test de Fisher: $P = 0.096$).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

Tabla 5. Análisis de la presión de selección y estimación de los ratios $\frac{dN}{dS}$ promedio para cada uno de los genes de adenovirus pertenecientes al género *Mastadenovirus*.

Unidad de Transcripción	Gen	Secuencias	Selección Neutral	Selección Positiva ($P < 0.05$)	Selección Negativa ($P < 0.05$)	d_s promedio	d_N promedio	ω
E1A	E1A	311	2	0	143	13.36	7.05	0.53
	E1B 19K	158	0	0	80	9.73	5.39	0.55
E1B	E1B 55K	264	0	3	360	18.32	7.31	0.40
IX	IX	163	0	0	87	6.97	3.02	0.43
Iva2	Iva2	209	1	8	314	15.02	4.29	0.29
E2B	Pol	353	1	13	889	18.89	5.16	0.27
	pTP	323	1	0	552	12.85	3.27	0.25
L1	52K	227	2	0	333	14.81	3.52	0.24
	pIIIa	347	0	0	590	22.18	5.22	0.23
	III	269	0	0	468	16.39	4.21	0.26
L2	pVII	216	0	0	178	12.94	3.22	0.25
	V	247	5	0	313	9.63	3.34	0.35
	pX	143	0	0	201	14.39	4.25	0.30
L3	pVI	231	0	0	301	4.01	3.10	0.77
	Hexon	302	0	0	868	27.77	5.83	0.21
	Protease	234	1	0	191	14.61	4.94	0.34
E2A	DBP	301	0	0	415	16.69	5.48	0.33
	100K	221	0	0	415	16.69	5.48	0.33
L4	33K	171	0	2	114	21.24	6.03	0.28
	22K	139	2	0	62	8.91	5.09	0.57
	pVIII	274	0	1	208	25.03	7.34	0.29
E3	E3_12.5K	118	0	0	54	6.22	2.89	0.46
	E3 CR1- α	146	0	0	104	11.18	5.23	0.47
	E3 gp19k	126	0	0	44	5.34	3.28	0.61
	E3 CR1- β	162	1	6	202	19.78	9.54	0.48
	E3 CR1- γ	88	0	0	1	0.52	0.31	0.60
	E3 CR1- δ	11	1	4	89	6.25	2.95	0.47
	E3 RID- α	104	0	0	70	10.83	2.61	0.24
	E3 RID- β	121	0	0	65	10.18	4.59	0.45
U exon	E3_14.7K	113	0	0	117	16.29	6.74	0.41
	U exon	75	0	0	29	3.74	2.36	0.63
	Fiber	205	1	0	441	14.73	5.47	0.37
L5	Fiber-1	54	0	0	203	8.37	2.97	0.35
	Fiber-2	48	1	0	253	5.88	1.95	0.33
E4	E4orf6/7	44	2	1	58	9.85	4.71	0.48
	E4 34K	38	1	0	0	1.17	1.05	0.90
	E4orf4	166	0	1	51	6.85	3.12	0.46
	E4orf3	99	1	1	50	7.35	2.75	0.37
	E4orf2	133	0	0	90	15.04	3.94	0.26
	E4orf1	269	0	0	110	9.97	3.63	0.36

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

En promedio, los ratios $\frac{dN}{dS}$ estimados fueron significativamente más bajos en los genes “core” ($\frac{dN}{dS} = 0.312 \pm 0.018$) que en los implicados en la interacción con el hospedador ($\frac{dN}{dS} = 0.430 \pm 0.023$; t-test: $P < 0.001$).

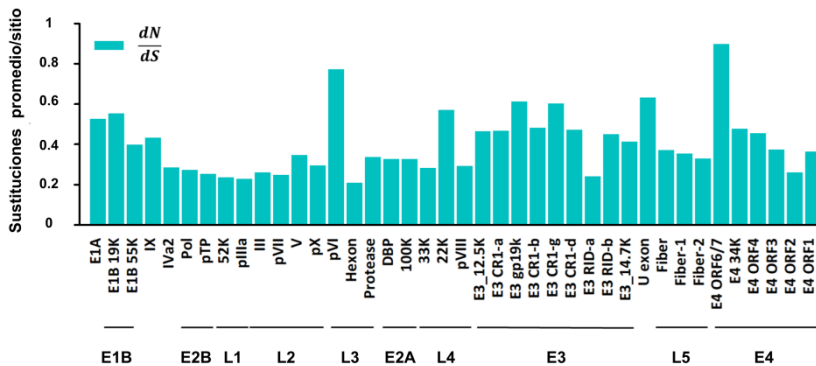


Figura 17. Análisis de las mutaciones sinónimas y no sinónimas en cada una de las regiones codificantes. El número de secuencias utilizadas en el análisis para cada gen se especifica en la Tabla 5. El alineamiento de las secuencias por codones se realizó utilizando MAFFT y la estimación de los valores de dN y dS por codón empleando la herramienta SLAC. La unidad de transcripción a la que pertenece cada gen se muestra en la parte inferior de la figura.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 1

De nuestro análisis de selección se deduce que los genes “core” han estado sujetos a una fuerte selección negativa. Por el contrario, para los genes “interacción” observamos el efecto contrario es decir, un mayor número de mutaciones no sinónimas en nuestras secuencias y por tanto, ratios $\frac{dN}{dS}$ más altos. Entre estas proteínas encontramos las codificadas por las unidades de transcripción E3 y E1A, ambas involucradas principalmente en la evasión de la respuesta inmunitaria y la modulación del metabolismo celular para garantizar la replicación del virus.

Asimismo, se observaron ratios $\frac{dN}{dS}$ altos para los genes implicadas en la interacción con la DDR. Por ejemplo, las unidades de transcripción E1B y E4 ($\frac{dN}{dS} = 0.417 \pm 0.045$). Sin embargo, al igual que en el análisis de diversidad, las diferencias entre ambos grupos (I e I-DDR), no fueron significativas (t-test: $P= 0.766$).

2. Estimación de la tasa de mutación de HAdV5

2.1 Identificación de mutaciones en HAdV5

A diferencia de los virus de RNA, pocas estimaciones de tasas de mutación han sido obtenidas para virus de DNA. Esto se debe principalmente a que los virus de DNA presentan bajas tasas de mutación fruto de la alta tasa de fidelidad en la replicación. Además, las estimaciones han sido obtenidas mediante el uso de métodos indirectos y fenotípicos que generalmente se basan en regiones específicas del genoma o sufren de sesgo por la acción de la selección (Drake 2005; Cuevas y col. 2009; Cheng y col. 2010; Sanjuan y col. 2010; Renzette y col. 2015). Las técnicas de secuenciación se han convertido en una herramienta de gran utilidad para investigar la diversidad genómica. Sin embargo, la tasa a la que aparecen mutaciones en los virus de DNA es similar a la tasa de error de las plataformas de secuenciación. Esta limitación podría ser solventada mediante el uso de nuevas metodologías de secuenciación de alta fidelidad como la que utilizaremos en este estudio, la DS.

En este capítulo, utilizaremos esta nueva metodología de secuenciación para estudiar la aparición espontánea de mutaciones a lo largo de todo el genoma de HAdV5 con el objetivo de determinar su tasa de mutación. El método de dilución límite se empleó para la obtención de tres partículas infecciosas que actuaran como tres réplicas experimentales durante el ensayo.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Seguidamente, cada una de las réplicas experimentales se amplificó a corto plazo para minimizar al máximo la acción de la selección y, al mismo tiempo, obtener grandes cantidades de DNA viral para llevar a cabo la secuenciación sin necesidad de recurrir a la amplificación mediante PCR.

Se realizaron tres ensayos de dilución límite consecutivos en placas de 96 pocillos con el objetivo de eliminar la diversidad preexistente. De esta manera, las mutaciones que identifiquemos mediante secuenciación dúplex serán mutaciones nuevas. Mediante este proceso, se aislaron tres clones de adenovirus (R1, R2 y R3) que fueron utilizados como réplicas biológicas (Figura 18).

Seguidamente, con el objetivo de obtener suficiente concentración de DNA viral, cada uno de los clones se sometió a dos amplificaciones sucesivas en células HeLa. Estas infecciones en medio líquido, se realizaron a alta MOI con el objetivo de reducir al máximo la eficacia de la selección y permitir así la acumulación de mutaciones de manera quasineutra. Durante la extracción de DNA viral, la lisis celular se realizó empleando un detergente suave (Tritón X-100) para impedir la lisis de la membrana nuclear y por consiguiente, la salida del DNA celular. Esto nos permitió obtener cantidades de DNA viral suficientes (500 ng) para llevar a cabo la secuenciación sin necesidad de realizar previamente una amplificación por PCR. El método de secuenciación dúplex nos permite identificar mutaciones presentes en únicamente una de las hebras de DNA, pudiendo así descartar posibles artefactos de secuenciación.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Obviar el paso de amplificación de nuestro DNA molde mediante PCR disminuye aún más la acumulación de mutaciones que no son reales, lo que nos permite ser más rigurosos y precisos a la hora de analizar nuestros datos de secuenciación. El ratio de DNA viral: celular se estimó en cada una de las muestras antes de proceder a la secuenciación y solo aquellas muestras que presentaban un porcentaje de DNA viral $\geq 70\%$ fueron seleccionadas.

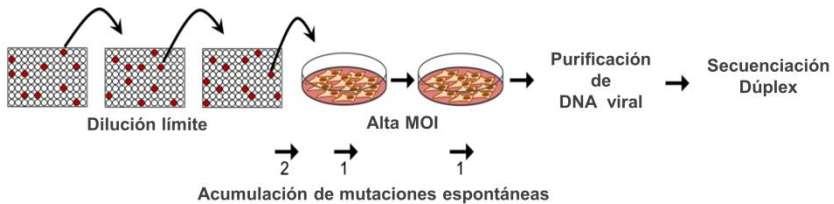


Figura 18. Flujo de trabajo realizado para la estimación de la tasa de mutación de HAdV5. Tres ensayos de dilución límite consecutivos se realizaron con el objetivo de eliminar la diversidad preexistente y aislar tres clones de adenovirus. Seguidamente, estos clones fueron amplificados a alta MOI en células HeLa. El número estimado de ciclos de infección, a lo largo de los cuales tendría lugar la acumulación de mutaciones, viene representado debajo de cada fase. Después de la purificación de DNA, la secuenciación se llevó a cabo directamente sin realizar un paso de amplificación por PCR.

El análisis de secuencias reveló un *coverage* > 2500 para cada una de las réplicas y fuimos capaces de secuenciar $> 99\%$ del genoma de adenovirus (Figura 19).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Se identificaron 68, 78 y 62 polimorfismos en cada una de las réplicas biológicas en 93.2, 115.7 y 123.7 Mb secuenciadas respectivamente (Tabla 6). Con estos datos, fuimos capaces de calcular una frecuencia media de mutación de $(6.4 \pm 0.7) \times 10^{-7}$. Las mutaciones identificadas presentaban frecuencias $< 1 \%$, lo que demostró que no se trataba de sustituciones preexistentes en la población inicial. En caso contrario, los virus aislados presentarían polimorfismos a frecuencias más elevadas debido a que el método de dilución límite supone un cuello de botella importante.

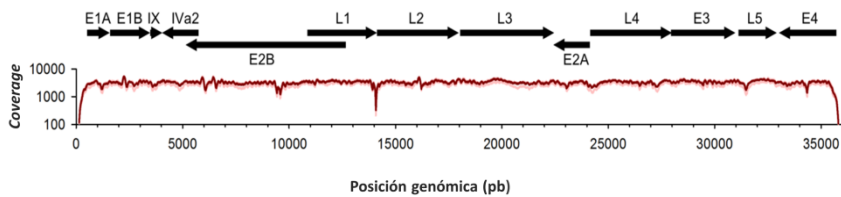


Figura 19. Estructura genómica de adenovirus y cobertura de secuenciación. La línea de color rojo oscuro representa el coverage promedio de las tres réplicas biológicas mientras que en un color rojo más claro, se representan los valores más bajos y altos de cobertura. En la parte superior de la figura, se muestra cada una de las unidades de transcripción y su orientación a lo largo del genoma del virus.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Para evaluar la tasa de error de nuestro método de análisis, también se secuenció el plásmido pUC18 purificado de la bacteria *E. coli*. Esta tasa de error nos permitirá posteriormente distinguir entre mutaciones reales y artefactos de secuenciación. Decidimos utilizar este plásmido como control debido a que la fidelidad de replicación en bacterias es extremadamente alta.

Tabla 6. Mutaciones identificadas mediante DS en HAdV5.

Mutación	Réplica experimental			Total
	1	2	3	
G → A	11	22	13	22.1
C → T	13	13	17	20.7
A → G	9	17	6	15.4
T → C	6	5	7	15.4
G → T	8	2	1	8.7
C → A	8	5	4	5.3
T → G	2	3	3	8.2
A → C	2	2	3	3.8
G → C	2	4	3	3.4
C → G	2	2	1	4.3
A → T	4	1	1	2.4
T → A	1	2	3	2.9
Total	68	78	62	100.0
Mbp secuenciadas	93.2	115.7	123.7	-
Frecuencia de mutación ($\times 10^{-6}$)	0.73	0.67	0.50	-

Una lista más detallada de las mutaciones identificadas puede encontrarse en la sección de material suplementario (Anexo 2).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Tras el análisis, se identificaron únicamente tres mutaciones en pUC18 en 21 Mb secuenciadas. Por tanto, asumiendo ausencia de diversidad en el plásmido, se estimó una tasa de error por base de 1.4×10^{-7} . Por ende, al menos el 80 % de las mutaciones observadas en los tres linajes correspondían a mutaciones reales. Teniendo en cuenta este error de secuenciación, la frecuencia de mutación de las poblaciones experimentales de HAdV5 fue de $(6.4 - 1.4) \times 10^{-7} = 5 \times 10^{-7}$.

Por otro lado, además de las sustituciones nucleotídicas, se identificaron 70 deleciones e inserciones en los tres linajes. Sin embargo, estas indels mostraron frecuencias similares a las observadas en el control de secuenciación y por tanto, no se tuvieron en cuenta para el cálculo de la frecuencia de mutación.

Con el objetivo de verificar la fiabilidad de este método de análisis de secuencias, se testaron cuatro de las mutaciones identificadas mediante secuenciación en cada una de las réplicas biológicas utilizando un método distinto. Basándonos en los datos de DS, las mutaciones seleccionadas fueron las siguientes: 191 (G → A), 7295 (A→G), 9417 (T→G) y 33565 (A→G). Se realizó un ensayo de qPCR en el que se amplificaba específicamente la base wild-type y el alelo mutante. Cada reacción de qPCR se realizó por triplicado y la identificación de cada una de las mutaciones se realizó en las tres réplicas biológicas.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

De hecho, tal y como se muestra en la Tabla 7, deberíamos identificar las mutaciones únicamente en 6 de las 12 muestras analizadas. Tras el ensayo, los valores de C_t para cada una de las muestras fueron analizados para determinar la presencia o ausencia de la sustitución nucleotídica.

Tabla 7. Mutaciones analizadas mediante qPCR. Para cada uno de los polimorfismos identificados se muestra su posición en el genoma y la presencia/ausencia en cada una de las réplicas biológicas.

Mutación	Posición	Réplica experimental		
		1	2	3
G → A	191		✓	
A → G	7295		✓	
T → G	9417	✓	✓	✓
A → G	33565			✓

En el caso de mutaciones detectadas mediante DS, se obtuvieron valores de C_t , para el alelo mutante de 6-12 ciclos más tardíos que los valores obtenidos para la base WT. Sin embargo, para las no detectadas mediante secuenciación dúplex se observaron C_t , para el alelo mutante de 12-20 ciclos más tardíos.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Tal y como se muestra en la figura 20, en el caso de las qPCRs correspondientes a las muestras en las que la mutación había sido previamente identificada mediante DS, se obtuvieron valores de ΔC_t menos negativos que para los linajes en los que la mutación no fue identificada (Test de Mann-Whitney: $P = 0.041$). En una de las muestras, se observó un valor ΔC_t negativo muy alto (se muestra en la gráfica señalado por una flecha). Esto reveló que esta mutación identificada mediante DS era en realidad un artefacto de secuenciación.

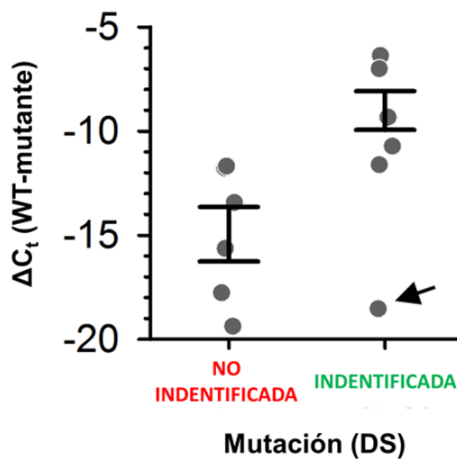


Figura 20. Resultados del análisis de mutaciones por qPCR. En la gráfica se muestran los valores de ΔC_t ($C_{t \text{ cebador WT}} - C_{t \text{ cebador mutante}}$) para cada una de las reacciones de qPCR. Cada punto representa la media de tres replicas experimentales. Las barras de error representan error estándar de la media de las seis reacciones realizadas en cada caso (mutación detectada/no detectada en el análisis de DS).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Tal y como se ha citado en el apartado anterior, teniendo en cuenta la tasa de error por base de 1.4×10^{-7} , al menos el 80 % de las mutaciones observadas en los tres linajes correspondían a mutaciones reales. En este caso, en las seis posiciones estudiadas se ha identificado únicamente un artefacto de secuenciación. Esto se ajusta perfectamente a la observación anterior ya que, nuevamente el 80% de los sitios polimórficos correspondieron a mutaciones reales. Al eliminar esta muestra del análisis, las diferencias entre ambos grupos (mutaciones identificadas y no identificadas) eran aún más significativas ($P = 0.004$).

2.2 Espectro mutacional y tasa de mutación espontánea

Las transiciones fueron dos veces más abundantes que las transversiones. Por su parte, la transición más abundante fue la $G \rightarrow A$ y su complementaria, $C \rightarrow T$ (42%), seguida de $A \rightarrow G / C \rightarrow T$ (24%). La transición $G \rightarrow A / C \rightarrow T$ fue dos veces más abundante que la $A \rightarrow G / C \rightarrow T$, pudiendo esto ser consecuencia de que los motivos CpG presentasen una alta probabilidad de mutación. De hecho, si nos centramos en los 3167 dinucleótidos presentes en el genoma de adenovirus, 50 de ellos presentaban mutaciones $G \rightarrow A / C \rightarrow T$. Sin embargo, de las 13380 bases C y G que no forman parte de motivos CpG en el genoma viral, solo 39 presentaban sustituciones. De estos datos se deduce que la probabilidad de que tenga lugar una sustitución $G \rightarrow A / C \rightarrow T$ en motivos CpG es 2.7 veces más alta (test de Fisher: $P < 0.001$). Sin embargo, se observó que los motivos CpG presentaban solo 1.7 veces mayor probabilidad de mutación ($G \rightarrow A / C \rightarrow T$) que el resto de bases C y G presentes en el genoma viral.

En teoría, el patrón observado no puede explicarse por la acción de la selección en la purga de mutaciones debido a que únicamente se han realizado dos amplificaciones consecutivas del virus. Además, después de las diluciones límite, las infecciones en medio líquido se realizaron a alta MOI, lo que reduce la eficacia de la selección.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Para demostrar que nuestro diseño experimental minimizó la acción purificadora de la selección, se analizaron los cambios sinónimos y no sinónimos. Bajo un modelo neutral y basándonos en el espectro mutacional previamente observado, se espera que el 71.5% de los polimorfismos observados que se localicen en regiones codificantes correspondan a mutaciones no sinónimas. La fracción obtenida fue del 74.5%, un valor que no se desvía significativamente del valor esperado ($\frac{dN}{dS}=1.16$, test Ji-cuadrado: $P = 0.327$).

Con el objetivo de estimar la tasa de mutación de HAdV5 por ciclo de infección celular, se determinó el número de ciclos transcurridos durante el procedimiento experimental.

En ausencia de selección:

$$\text{Tasa de mutación por ciclo de infección} = \frac{\text{frecuencia de mutación}}{\text{número de ciclos de infección}}$$

En primer lugar, se calculó el número de partículas virales infecciosas producidas por célula (*burst size*). Para ello, aproximadamente 5×10^4 células HeLa sembradas en placas de 24 pocillos fueron infectadas por triplicado con HAdV5 a MOI 10. Posteriormente, el número de partículas virales producidas a los 3 y 7 dpi se calculó mediante ensayo en placa.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

El *burst size* se calculó dividiendo el número de partículas virales producidas a las 72 hpi (4.93×10^8 PFUs) por el número de células infectadas. El título a 7 dpi fue ligeramente superior obteniéndose un *burst size* de aproximadamente 2×10^4 PFUs/célula. Gracias a la estimación del *burst size* (10^4 partículas virales/célula), fuimos capaces de saber que aproximadamente transcurrieron 4 ciclos durante el aislamiento y la amplificación de los tres linajes utilizados en este estudio, como se muestra a continuación.

El número de ciclos de infección se calculó utilizando la siguiente ecuación (Sanjuan y col 2010):

$$c = \frac{\log\left(\frac{N_1}{N_0}\right)}{\log B}$$

Donde N_0 es el inóculo, N_1 el título alcanzado, B el *burst size* y c el número de ciclos de infección transcurridos.

Durante la fase de aislamiento utilizando el método de dilución límite, se sembraron 10^4 células por pocillo y la infección se inició a partir de una única partícula viral, por lo que:

$$c = \frac{\log\left(\frac{10^4 \times 10^4}{1}\right)}{\log 10^4} = 2$$

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Cada una de las dos amplificaciones consecutivas a alta MOI supuso 1 ciclo de infección. El número de células sembradas fue de 10^7 y la infección se realizó a una MOI de 10, por lo que:

$$c = \frac{\left(\frac{10^4 \times 10 \times 10^7}{10^8} \right)}{\log 10^4} = 1$$

Conociendo esta información, la tasa de mutación obtenida fue de $5.0 \times 10^{-7} / 4 = 1.3 \times 10^{-7}$ s/n/c.

2.3 Distribución de mutaciones a lo largo del genoma

En estudios previos, se ha observado que la diversidad genética no se distribuye de forma uniforme a lo largo del genoma de adenovirus (Robinson y col. 2011; Robinson y col. 2013). Por este motivo, decidimos estudiar si la tasa de mutación es uniforme o si, por el contrario, existen regiones genómicas en las que la tasa de mutación no es constante.

Aunque se observó que la distribución de mutaciones a lo largo del genoma viral era bastante uniforme, se identificaron dos regiones, de aproximadamente 5 kpb, en las que el número de mutaciones era significativamente más bajo que el esperado (Figura 21).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Una de estas regiones fue la región 18500-23500, en la que se encuentra la unidad transcripcional L3 que codifica para la proteasa viral y dos proteínas de la cápside (VI y hexon). En esta región se identificaron únicamente siete polimorfismos en las tres réplicas biológicas, mientras que el número de sustituciones esperado era de 29.17 (test Binomial: $P < 0.001$). Esta disminución fue significativa en cada una de las tres réplicas biológicas ($P \leq 0.014$).

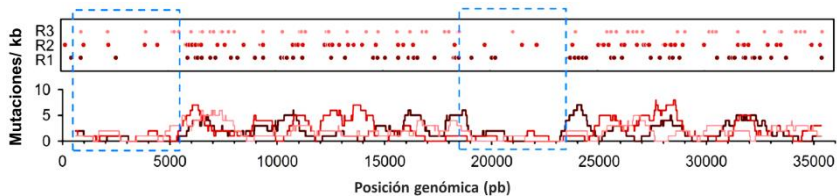


Figura 21. Estructura genómica y mutaciones identificadas mediante DS. Mutaciones obtenidas para cada una de las réplicas biológicas (R1, R2 y R3). Cada uno de los círculos representa una sustitución nucleotídica. En la parte inferior de la figura, para cada réplica biológica, se muestra el número de mutaciones promedio obtenido por *sliding window* usando un tamaño de ventana de 1kb. Las dos regiones genómicas de baja complejidad se indican mediante un cuadro de color azul.

La otra región que presentó un número reducido de mutaciones fue la región 500-5500 (test Binomial: $P < 0.001$). En ella se localizan las unidades de transcripción E1B, IX e IVa2. En este caso, la disminución en el número de mutaciones fue únicamente significativa en dos de las tres réplicas analizadas.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Una posible explicación a la disminución de diversidad observada en estas dos regiones genómica podría ser la acción de la selección purificadora. Sin embargo, el porcentaje de mutaciones no sinónimas en ambas regiones, L3/ E1B-lva2 (76.9%) fue similar al observado en el resto del genoma del virus (74.5%). Tampoco se observaron diferencias significativas en la frecuencia de dinucleótidos CpG y además, la frecuencia de mutación disminuyó cuando las sustituciones $G \rightarrow A$ / $C \rightarrow T$ fueron eliminadas del análisis. Por tanto, no hemos sido capaces de dilucidar los posibles mecanismos a través de los cuales tiene lugar esta variación en la tasa de mutación a lo largo del genoma.

A pesar de lo observado en estas dos regiones, para el resto de unidades de transcripción, el número de mutaciones observadas y esperadas correlacionó perfectamente, asumiéndose por tanto que la tasa de mutación a lo largo del resto del genoma viral es generalmente constante (test Ji-cuadrado: $P= 0.672$; Figura 22A).

Además, para estudiar si existe una distribución uniforme o por el contrario, las mutaciones tienden a acumularse en una zona específica del genoma, se investigó la distribución empírica de la distancia entre mutaciones consecutivas. Asumiendo una tasa de mutación constante, estas distancias deberían seguir una distribución geométrica. Aunque nuestros datos siguieron el patrón de este modelo, se observó un ligero aumento en la proporción de mutaciones que mostraban distancias pequeñas, lo que sugiere un ligero nivel de *clustering* (Figura 22B).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

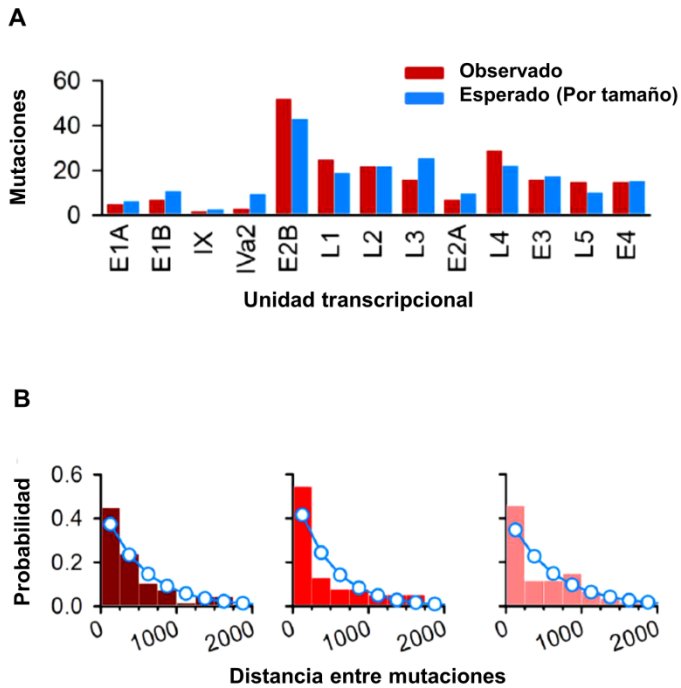


Figura 22. Distribución de mutaciones a lo largo del genoma viral. **A.** Número de sustituciones nucleotídicas identificadas (color rojo) y esperadas (color azul) en cada una de las unidades de transcripción, asumiéndose una tasa de mutación constante a lo largo del genoma. **B.** El histograma muestra la distribución de las distancias entre mutaciones consecutivas observadas mientras que, la línea azul representa la distribución geométrica (modelo nulo). Cada una de las gráficas representa una réplica biológica, siguiéndose el mismo patrón de colores que en la Figura 15.

2.4 Hot spots mutacionales

Se identificaron 11 posiciones genómicas en las que se observó la misma mutación en al menos dos de las tres réplicas biológicas. Como estas mutaciones aparecían de forma recurrente, fueron interpretadas como posibles zonas calientes de mutación. La mayoría de estos cambios fueron inserciones y deleciones (Tabla 8).

En la posición 9417, en la que se encuentra E2B, se observó la sustitución T → G en las tres réplicas a una frecuencia órdenes de magnitud superior a la tasa de mutación obtenida (0.28%, 0.54% y 0.77%, respectivamente).

Una característica que compartían todas estas mutaciones era que se encontraban en un contexto de baja complejidad de secuencia. Estos tramos homopoliméricos reducen la fidelidad de la replicación y además aumentan la tasa de error de NGS. Las secuencias de DNA repetitivas o de baja complejidad son capaces de formar estructuras secundarias que provocan la inhibición de la DNA polimerasa y la disociación del complejo de replicación, favoreciendo el desapareamiento de ambas hebras de DNA. Bajo esta situación, la hebra de DNA naciente puede anillarse con cualquier otra repetición, generándose inserciones y deleciones. Este tipo de error que tiene lugar durante la replicación de DNA se conoce como *replication slippage* o de desplazamiento de hebra. Por este motivo, estos cambios podrían ser considerados como errores de secuenciación.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Para la posición 14073 (inserción), que mapea con el final de L1, se observó una bajada del *coverage* de secuenciación de > 1000x a 200x aproximadamente. Sin embargo, el análisis de la cobertura en otras regiones en las que se identificaron puntos calientes de mutación, como por ejemplo la región 34337-34347 y 1161-1170, no reveló una disminución significativa de cobertura como en el caso anterior.

Tabla 8. Mutaciones recurrentes en las diferentes réplicas biológicas de HAdV5.

Posición genómica	Réplica	Mutación	Contexto de secuencia
1165	1, 2, 3	Inserción	TGGTGTGGTA A TTTTTTTTTT
1216	2, 3	Inserción	TTGTATTGTG A TTTTTTTAAA
7096	1, 2, 3	Inserción	TATCCTGTCC C TTTTTTTTCC
8617	1, 2, 3	Inserción	CCCCGGAGGT A GGGGGGGCTC
9417	1, 2, 3	T→G	CTGGCGCGG T GGGGGAGGGG
11222	2, 3	A→C	CGGGCCCGG C ACTACCTGGAC
14073	1, 2, 3	Inserción	GAGAATGTTTT A AAAAAAAAAA
16602	1, 2	Inserción	GAGATCTATG G CCCCCGAAG
34336	1, 2, 3	Inserción/delección	GAAGAACCAT G TTTTTTTTTT
35122	2, 3	Inserción	AATAAAATA A CAAAAAACAT
35215	1, 2	Inserción	GCGGTGACCG T AAAAAACTG

2.5 Relación entre la tasa de mutación y diversidad de secuencias *in vivo*

Para investigar si existe una relación entre la tasa de mutación y la diversidad genética *in vivo*, se utilizaron las bases de datos obtenidas en el capítulo anterior. Se emplearon los alineamientos en formato FASTA procedentes de cada una de las unidades de transcripción de las secuencias de adenovirus pertenecientes al género *Mastadenovirus*. Decidimos realizar el análisis de diversidad genética por genes en lugar de trabajar con secuencias de genoma completo. Esta decisión facilitó el alineamiento de secuencias, pues adenovirus presenta un tamaño genómico considerable. La diversidad nucleotídica (Nei and Li. 1979) se calculó, utilizando un script de R, como la probabilidad de que dos secuencias de DNA elegidas al azar dentro de una muestra poblacional difieran en una posición nucleotídica.

Las posiciones en las que se identificaron mutaciones mediante secuenciación Dúplex mostraron valores de diversidad significativamente superiores (diversidad promedio= 0.84 ± 0.02) que aquellas en las que no observó mutación (diversidad promedio= 0.39 ± 0.00 ; t-test: $P < 0.001$). Las diferencias observadas en la distribución de diversidad entre sitios mutados y no mutados sugieren un efecto de la tasa de mutación sobre la diversidad *in vivo* (Figura 24AB).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Seguidamente, se analizó la diversidad en aquellas regiones en las que previamente habíamos observado una baja tasa de mutación (L3 y E1B-Iva2). En este caso, se observó que la distribución de los valores de diversidad fue muy similar en estas regiones y el resto del genoma (Figura 24C). Probablemente esto sea debido a que en estas zonas del genoma, existen regiones de alta variabilidad, tal y como se observó en el capítulo 1. Aunque se obtuvieron valores de diversidad para las regiones de baja complejidad significativamente más elevados (diversidad promedio: baja complejidad= 0.43 ± 0.00 , resto del genoma= 0.36 ± 0.00), estas diferencias fueron muy pequeñas como para sacar conclusiones de manera fiable (t-test: $P < 0.05$). En consecuencia, la baja tasa de mutación observada en estas regiones parece no limitar la diversidad *in vivo*.

Por último, estudiamos también si las mutaciones observadas de forma recurrente estaban presentes *in vivo*. En el caso de que fuesen puntos calientes de mutación reales, deberíamos obtener valores de diversidad elevados para esas posiciones en nuestras secuencias de GenBank. Si fuesen artefactos, la mayoría de estas posiciones no deberían ser polimórficas o, por el contrario, deberíamos observar estos cambios sistemáticamente, pues estas secuencias también provienen de reacciones de secuenciación. Los análisis revelaron que aquellas posiciones que mostraban cambios de forma recurrente en nuestro análisis mostraban elevada diversidad *in vivo* (diversidad promedio= 0.80 ± 0.10).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

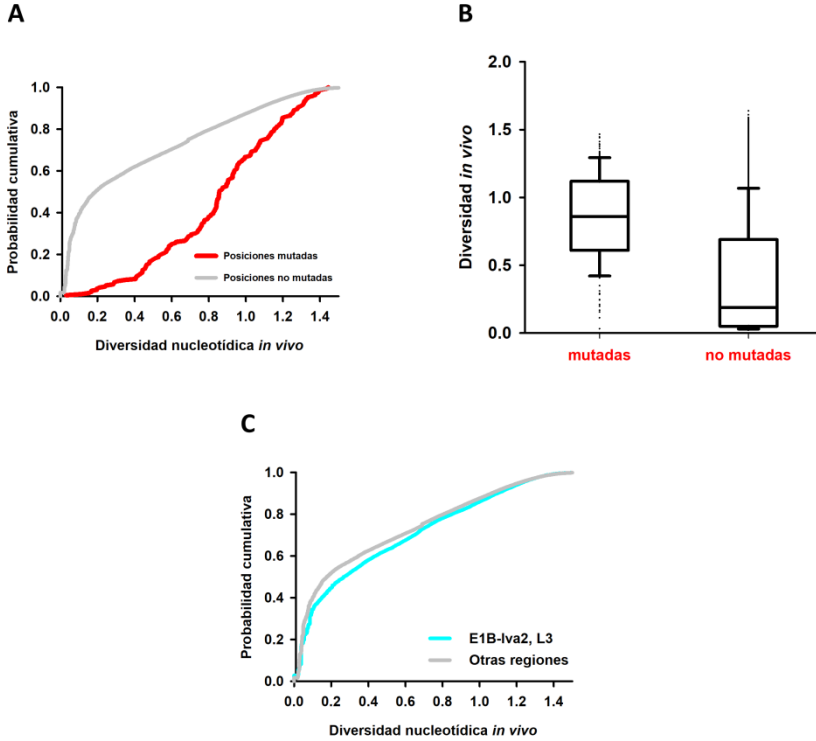


Figura 24. Diversidad nucleotídica obtenida a partir de las secuencias de GenBank. **A)** Se compara la probabilidad acumulativa en las posiciones en las que previamente habíamos identificado mutaciones (en rojo) y en aquellas en las que se había observado la base WT (en gris) **B)** Se compara la probabilidad en las zonas en las que se observa una tasa de mutación baja (en azul) con el resto de regiones genómicas (en gris). **C)** Valores de diversidad *in vivo* obtenidos tanto para las posiciones en las que se identificó una mutación mediante DS como para las que se observó la base WT.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 2

Esto demuestra que estas posiciones son realmente puntos calientes de mutación y no artefactos de DS. Sin embargo, los valores de diversidad en promedio obtenidos para los puntos calientes de mutación fueron similares a los obtenidos para el resto de posiciones mutadas (diversidad promedio= 0.86 ± 0.02 ; t-test: $P = 0.92$).

3. Estudio del posible papel del sistema de reparación del DNA sobre la tasa de mutación de HAdV5

3.1 Inhibición de la DDR y evolución experimental de HAdV5 en células MRC-5

En el capítulo anterior, fuimos capaces de estimar la tasa de mutación del HAdV5 mediante la identificación de mutaciones a lo largo de todo el genoma haciendo uso de una nueva metodología de secuenciación de alta fidelidad. La estimación se realizó bajo condiciones de laboratorio controladas y minimizando la acción de la selección. La tasa de mutación obtenida se encuentra dentro del rango de los virus de DNA. Asimismo, esta tasa de sustitución fue similar o incluso más baja que la de la célula huésped utilizada durante el estudio. Esto podría poner de manifiesto el acceso de HAdV5 al sistema de reparación de DNA celular. Además, durante el primer capítulo, los genes virales involucrados en la interacción con la DDR mostraron valores de diversidad y ratios $\frac{dN}{dS}$ cercanos a otros genes implicados en otros tipos de interacción con el hospedador. Como por ejemplo las regiones codificantes E1 y E3, implicadas en modular la respuesta celular frente a la infección con el objetivo de promover la infección y garantizar la replicación del genoma viral. Esta observación pone de manifiesto la importancia de la funcionalidad de las proteínas virales que interaccionan con la DDR en la biología del virus.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

De aquí nace la hipótesis que consiste en que el acceso a la DDR por parte de HAdV5 podría explicar la baja tasa de mutación estimada anteriormente. Para investigar esta cuestión, decidimos poner a punto un ensayo de evolución experimental en el que sometimos a HAdV5 a diez pases de evolución en células normales MRC-5 tratadas con inhibidores químicos de la DDR. Si nuestra hipótesis es cierta y el virus explota la DDR para reparar los genomas virales nacientes durante la infección, la supresión de estas vías de reparación dará como resultado un aumento en la diversidad. Por tanto, después de realizar el ensayo de evolución experimental, la diversidad genética en cada una de las poblaciones de virus evolucionadas en células tratadas y no tratadas, se estudió mediante DS.

La inhibición de la DDR se llevó a cabo empleando inhibidores químicos previamente utilizados en otros estudios. Se utilizaron los inhibidores químicos KU55933 (inhibidor de ATM) (Hickson y col. 2004) VE-821 (inhibidor de ATR) (Prevo y col. 2012) y NU7741 (inhibidor de DNA-PK) (Leahy y col. 2004) para suprimir cada una de las vías de señalización que modulan la respuesta al daño en el DNA celular. Previamente a la realización del ensayo de evolución experimental, la citotoxicidad de cada uno de estos compuestos químicos fue evaluada en células MRC-5. Asimismo, se testó la toxicidad de la combinación de los tres inhibidores químicos (triple inhibición de la DDR).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

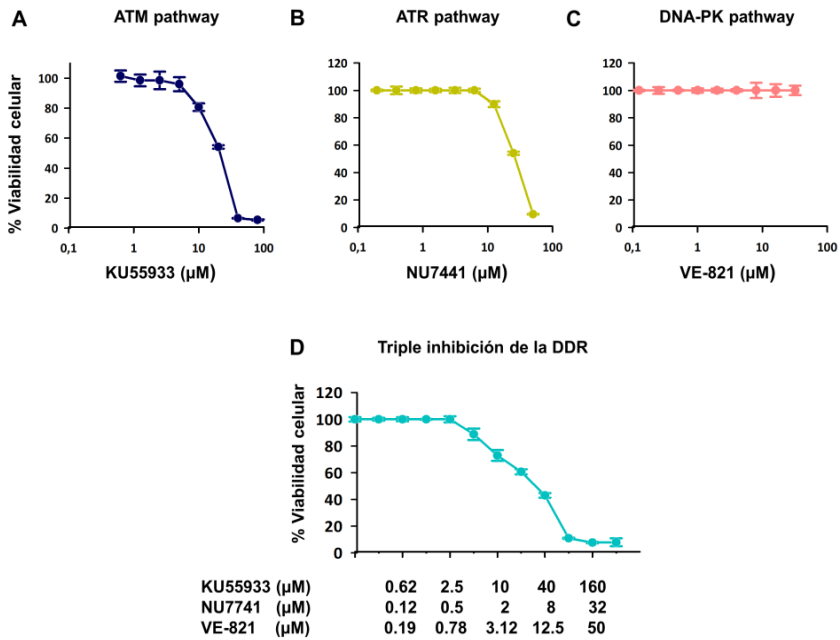


Figura 26. Citotoxicidad de los inhibidores de la DDR en células MRC5. Las células se trataron con diferentes concentraciones del inhibidor de ATM (A), ATR (B) y DNA-PK (C). Además, la inhibición total de la DDR también fue testada mediante la exposición de las células a diferentes combinaciones de los tres inhibidores (D). Las barras de error representan el error estándar de la media de las tres réplicas independientes realizadas en el estudio.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

En placas de 96 pocillos y por triplicado, las células fueron expuestas a diferentes concentraciones de cada uno de los inhibidores durante 72h. El análisis de los resultados reveló que las concentraciones utilizadas en estudios previos no mostraban toxicidad en cultivo (Figura 26). Se seleccionó la combinación de 10 μ M KU55933, 1 μ M NU7741 y 5 μ M VE-821 para la inhibición de la DDR sin comprometer excesivamente la viabilidad celular.

La posible perturbación del ciclo de infección por parte de estos compuestos químicos se evaluó realizando curvas de crecimiento. Previamente a la infección, las células MRC-5 fueron tratadas con cada uno de los inhibidores durante 2 h. Seguidamente, fueron inoculadas con adenovirus a baja MOI (0.1 PFU/célula) en presencia de cada uno de los compuestos químicos. El número de partículas virales infecciosas producido a las 0, 48 y 72 hpi se calculó mediante ensayo en placa.

La inhibición de las rutas de ATR y DNA-PK dio lugar a una ligera disminución significativa del título del virus a las 72 hpi (t-tests: $P < 0.001$) mientras que la supresión de la vía de reparación de ATM no mostró diferencias con respecto a la condición sin tratamiento (t-test: $P = 0.76$). La combinación de los tres tratamientos tampoco afectó significativamente al crecimiento del virus (t-test: $P = 0.14$) (Figura 27).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

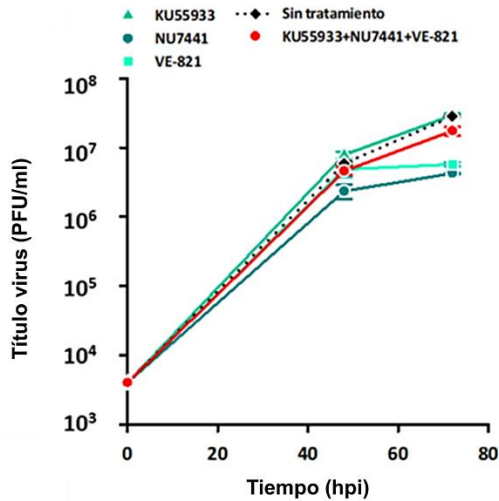


Figura 27. Curva de crecimiento de HAdV5 en presencia de los inhibidores de la DDR. Tras el tratamiento previo de las células con cada uno de los inhibidores, se procedió a la infección de la monocapa celular a baja MOI y el título viral se estimó a diferentes tiempos de infección y a tiempo cero (T0). Cada curva fue ensayada por triplicado y las barras de error representan el error estándar de la media de tres réplicas experimentales independientes.

Seguidamente, decidimos testar la eficacia de cada uno de los inhibidores en nuestro tipo celular. Para estudiar si realmente las vías de reparación están funcionando en presencia de los inhibidores es necesario generar un estrés genotóxico y posteriormente investigar su funcionamiento.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

La NCS es una proteína producida por la bacteria *Streptomyces carzinostaticus* utilizada como droga antitumoral radiomimética. Este compuesto presenta un cromóforo lábil capaz de intercalarse en el DNA que, en presencia de tior, produce especies bi-radicales altamente reactivas que generan roturas en ambas hebras del DNA (DSBs) (Beerman y col. 1983; Kappen y col. 1987). Al mismo tiempo, este agente antitumoral es capaz de generar roturas en una única hebra de DNA (SSBs o *single-strand DNA breaks*) en las bases A y T. En consecuencia, decidimos utilizar este compuesto como inductor de la DDR para testar la efectividad de los inhibidores químicos.

Primeramente, la actividad de este compuesto químico se testó en células MRC-5 mediante ensayos de inmunofluorescencia utilizando como marcador de daño en el DNA la fosforilación de la histona H2AX en la Ser 139 (γ H2AX). La presencia de focos de γ H2AX en el núcleo celular es un marcador de daño en el DNA ampliamente utilizado (Sharma y col. 2012). La fosforilación de la H2AX por las quinasas ATR, ATM y DNA-PK en respuesta a una situación de estrés genotóxico ha sido ampliamente demostrada (Bakkenist y Kastan 2003; Ward y col. 2004; Mukherjee y col. 2006).

Las células fueron expuestas a diferentes concentraciones de NCS (0-1000 ng/ml) durante 30 min. Después del período de incubación, las células se fijaron y se incubaron con el anticuerpo frente a γ H2AX.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

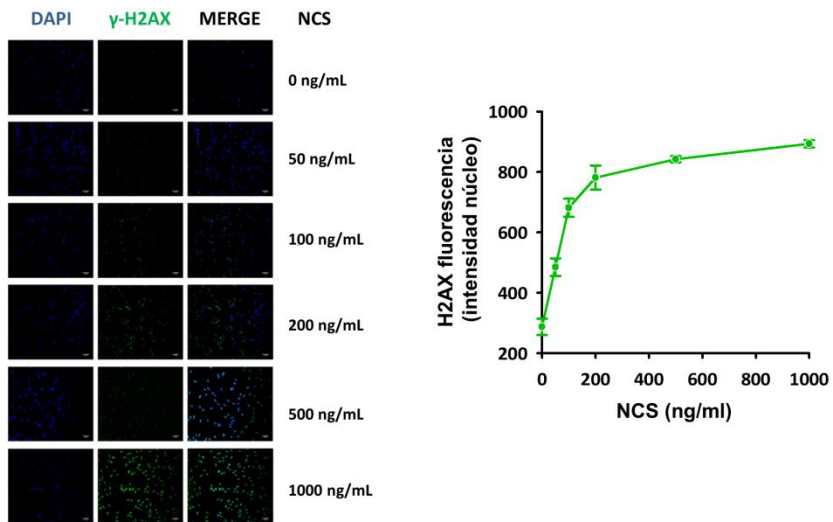


Figura 28. Inducción de la DDR mediante el tratamiento con NCS. Las células MRC-5 fueron expuestas a las concentraciones de NCS indicadas en la figura para inducir daño en el DNA y causar la activación de las vías de reparación del DNA. En la parte izquierda de la figura se representan los focos de γ H2AX y su localización en el núcleo (tinción DAPI) como respuesta al tratamiento con NCS. A la derecha en la imagen se representa la cuantificación de la intensidad de fluorescencia de γ H2AX en respuesta a la exposición a diferentes concentraciones de NCS. Cada condición experimental fue ensayada por triplicado. La intensidad de fluorescencia se midió en 100-300 focos nucleares en cada una de las réplicas experimentales. Las barras de error representan el error estándar de la media.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

La intensidad de fluorescencia en el núcleo celular se observó utilizando un microscopio de fluorescencia convencional y fue cuantificada empleando el software del equipo INCell 2000 (Figura 28). El análisis de los resultados reveló una perfecta correspondencia entre la dosis de NCS y la magnitud del efecto (fosforilación de la H2AX), lo que demuestra que el antitumoral produce daño en el DNA en nuestra línea celular de manera eficaz. La concentración más alta de NCS testada (1000 ng/ml) resultó en una elevada intensidad de fluorescencia. En consecuencia, esta dosis fue la establecida para testar posteriormente la eficacia de los inhibidores del sistema post-replicativo celular.

Una vez establecido el compuesto utilizado para activar la DDR, se procedió a evaluar la efectividad de los inhibidores de las tres quinasas que median las tres principales rutas de la DDR: ATM, ATR y DNA-PK. El ensayo consistió en realizar tres procedimientos consecutivos. En primer lugar, se procedió a inactivar las tres vías de señalización mediante el tratamiento de las células con los correspondientes inhibidores (KU55933, VE-821 y NU7441). Seguidamente, se generaba un estrés genotóxico en la célula a través del tratamiento con NCS a la concentración anteriormente establecida. Por último, la actividad de cada una de las rutas de reparación en respuesta al daño en el DNA generado, se testó realizando ensayos de inmunofluorescencia.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

Cabe destacar que los supresores de la DDR estuvieron presentes en el medio de cultivo durante todo el procedimiento experimental. En el ensayo de inmunofluorescencia se utilizaron anticuerpos primarios específicos frente a la forma fosforilada de la propia proteína transductora (ATM y DNA-PK) o de otras proteínas situadas *downstream* en la ruta correspondiente (Chk1, en el caso de la ruta de ATR).

El tratamiento con NCS provocó un aumento significativo de la actividad de las rutas de la DDR. Tras el tratamiento, se observó un aumento del $67.67 \pm 0.84\%$ (ATM), $44.70 \pm 1.87\%$ (ATR) y $86 \pm 1.18\%$ (DNA-PK), con respecto a los niveles de fluorescencia basales (t-tests: $P < 0.001$ en todos los casos; Figura 29).

Además, el tratamiento con los inhibidores de la DDR disminuyó, de manera muy eficiente y significativa la actividad de las tres rutas diana (t-tests: $P < 0.001$, en todos los casos; Figura 29), alcanzándose niveles de fluorescencia similares a las condiciones basales. Para células inoculadas con HAdV5 se obtuvieron resultados similares ya que no se observaron diferencias significativas entre células infectadas y no infectadas (t-tests: $P > 0.05$).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

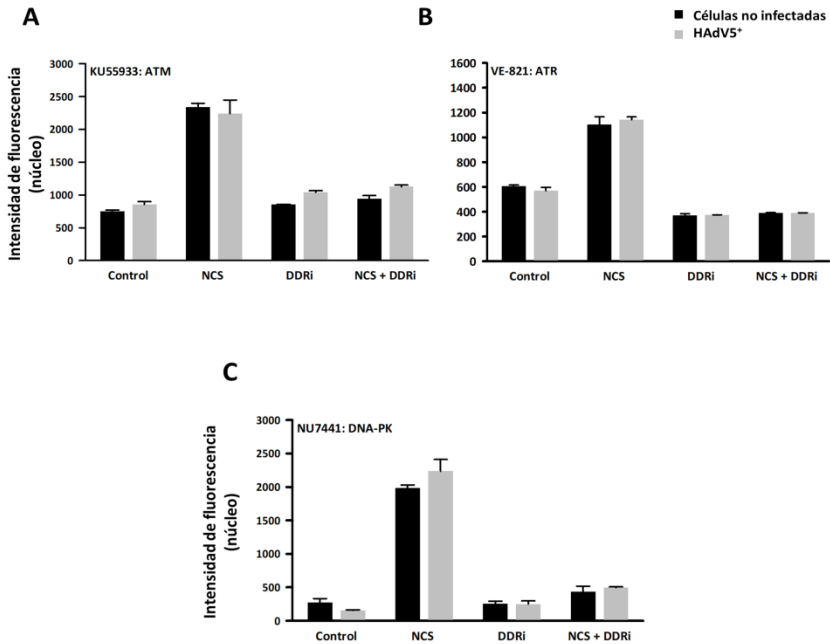


Figura 29. Efectividad de los inhibidores de la DDR. Células MRC-5 no infectadas (negro) e infectadas con HAdV5 (gris) fueron tratadas con el inductor de la DDR (NCS), con cada uno de los inhibidores de la DDR (DDRI) o siguiendo el procedimiento anteriormente explicado (NCS + DDRI). La monocapa celular se trató con DMSO como condición sin tratamiento (control). Los inhibidores utilizados fueron KU55933 (**A**), VE-821 (ATR) (**B**) y NU7441 (DNA-PK) (**C**). El funcionamiento de cada una de las vías de reparación se testó mediante inmunofluorescencia y la posterior cuantificación de la fluorescencia emitida por las proteínas diana como respuesta a cada uno de los tratamientos. Los datos mostrados corresponden a la media de dos réplicas experimentales. La intensidad de fluorescencia se midió en 100-300 focos presentes en el núcleo celular para cada una de las condiciones experimentales. Las barras de error representan el error estándar de la media. DDRI: Inhibidor químico de la DDR.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

El efecto de la inhibición de la DDR sobre la diversidad de HAdV5 se estudió mediante evolución experimental. En el capítulo 1, se estudió la tasa de mutación de nuestro virus modelo en células tumorales, en las que las vías de reparación están comprometidas. Sin embargo, para estudiar las posibles consecuencias de la interacción entre el virus y la DDR sobre la tasa de mutación del virus, necesitamos un hospedador no tumoral, en este caso la línea celular MRC-5, en el que la DDR funciona de manera adecuada. Se realizaron tres ensayos de dilución límite consecutivos con el objetivo de aislar tres clones únicos de adenovirus (virus fundadores) y eliminar la diversidad preexistente en la población. Cada uno de estos virus fundadores (R1, R2 y R3) se sometió a diez pases seriados de evolución en nuestro tipo celular a baja MOI (0.1 PFU/célula).

La evolución experimental se realizó en condiciones normales (sin tratamiento) y células tratadas con la combinación de los tres inhibidores de la DDR (KU55933 + VE-821 + NU7441: triple inhibición de la DDR (DDRi)). Después de cada pase de evolución, se estimaba el título viral alcanzado por cada una de las poblaciones evolucionadas. Los resultados se muestran en la figura 30.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

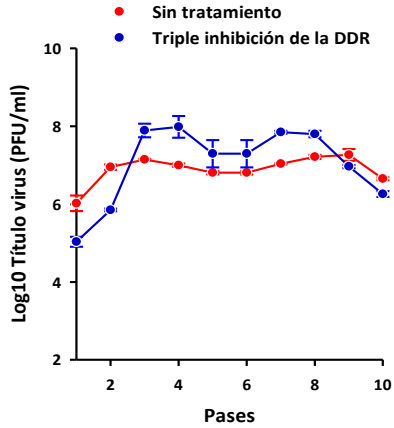


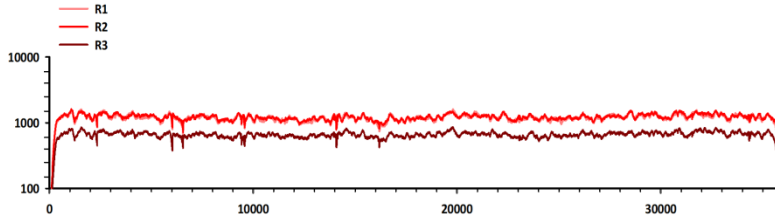
Figura 30. Títulos obtenidos durante los pases seriados de evolución. Los datos mostrados corresponden a la media de tres réplicas experimentales. Las barras de error representan el error estándar de la media.

3.2 Estudio de las poblaciones evolucionadas mediante secuenciación dúplex

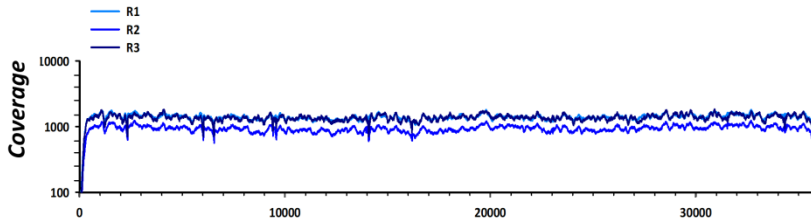
Después de realizar el ensayo de evolución experimental, cada una de las poblaciones evolucionadas se amplificó respetando las condiciones experimentales establecidas durante el ensayo de evolución. Esta amplificación fue realizada con el objetivo de obtener suficientes cantidades de DNA viral para llevar a cabo la secuenciación de DNA sin necesidad de una amplificación previa mediante PCR. Tras observarse CPE, se procedió a la purificación del DNA viral. Antes de llevar a cabo la lisis de las cápsides virales, se realizó un tratamiento con DNAsa I para eliminar el exceso de DNA celular.

Seguidamente, se procedió a la secuenciación dúplex de cada una de las poblaciones evolucionadas y los tres virus fundadores (N= 9) utilizando la plataforma Illumina. Además, al igual que en el capítulo anterior, se secuenció el plásmido bacteriano pUC18 para evaluar la fidelidad de la secuenciación. Tras la secuenciación, se obtuvo un *coverage* promedio de 1191 lecturas por base y el 94.5% de las posiciones genómicas mostraron valores de cobertura superiores a 500. Esto hizo posible el estudio de la diversidad en cada una de las poblaciones virales en profundidad y con elevada precisión (Figura 31).

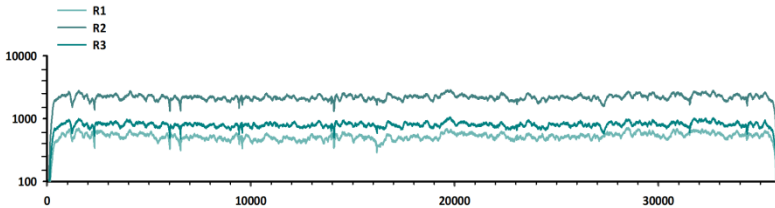
A VIRUS FUNDADOR



B SIN TRATAMIENTO



C TRIPLE INHIBICION



Posición genómica (pb)

Figura 31. Coverage promedio en cada una de las poblaciones evolucionadas y el virus fundador. La cobertura de secuenciación obtenida para el virus fundador (A) y cada una de las condiciones experimentales (B y C).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

En las nueve poblaciones secuenciadas, se identificaron entre 68 y 215 sitios polimórficos (Tabla 9). La mayoría de las mutaciones (68.2%) mostraron frecuencias en el orden de 0.1%, mientras que el 31.8% de las mutaciones identificadas aparecieron en más de una lectura y alcanzaron frecuencias entre 0.1% y 1%. Sin embargo, únicamente 14 de estas mutaciones mostraron frecuencias superiores al 1%. Las frecuencias de mutación calculadas como el ratio entre número total de mutaciones y el número total de bases secuenciadas se muestran en la Tabla 9 y se hallaron dentro del rango de $3.2 - 8.3 \times 10^{-6}$ s/n. Las frecuencias de mutación obtenidas fueron bajas pero al mismo tiempo, más altas que las observadas para el plásmido control pUC18 (1.6×10^{-6} ; t-test: $P < 0.001$). Esto indica que la mayoría de las mutaciones identificadas son reales y no corresponden a artefactos de secuenciación.

Tabla 9. Mutaciones identificadas mediante DS en las poblaciones evolucionadas (sin tratamiento y DDRi) y el virus fundador.

Réplica experimental	1			2			3		
	Fundador	Sin tratamiento	DDRi	Fundador	Sin tratamiento	DDRi	Fundador	Sin tratamiento	DDRi
Sitios polimórficos	122	215	81	171	125	182	80	113	68
Mutaciones totales	159	396	118	272	243	320	96	408	92
Mpb secuenciadas	42.9	43.6	23.6	49.8	33.2	48.9	18.8	78.1	28.5
Frecuencia de mutación ($\times 10^6$)	2.8	4.9	3.4	3.4	3.8	3.7	4.3	1.4	2.4

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

La frecuencia de mutación fue ligeramente más elevada en las poblaciones evolucionadas que en los virus fundadores (t-test: $P = 0.019$), pero no se encontraron diferencias significativas entre los virus evolucionados en células no tratadas (frecuencia promedio = $3.37 \times 10^6 \pm 1.03 \times 10^6$) y tratadas con los inhibidores de la DDR (frecuencia promedio = $3.17 \times 10^6 \pm 3.93 \times 10^5$) (t-test: $P > 0.05$).

Por otro lado, la distribución de mutaciones a lo largo del genoma no fue completamente uniforme. En algunos casos, el número de mutaciones observado por unidad de transcripción se desvió significativamente del número de mutaciones esperado (Figura 32). El número de mutaciones esperado se calculó asumiendo una tasa de mutación constante y teniendo en cuenta el tamaño de cada una de las regiones codificantes. La unidad L1 y las regiones no codificantes mostraron diversidad elevada mientras que IVa2 mostró menos diversidad de la esperada por azar. Desviaciones entre el número de mutaciones observado y el esperado también se observaron en las regiones codificantes IX, L4 y E4 para algunas poblaciones. Sin embargo, cabe destacar que, en general, la distribución de mutaciones fue similar a lo largo del genoma de las dos poblaciones evolucionadas y el virus fundador. Podría ocurrir que, durante la evolución experimental, la selección hubiese actuado eliminando las mutaciones perjudiciales reduciéndose el número de mutaciones observado en algunos genes más que otros.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

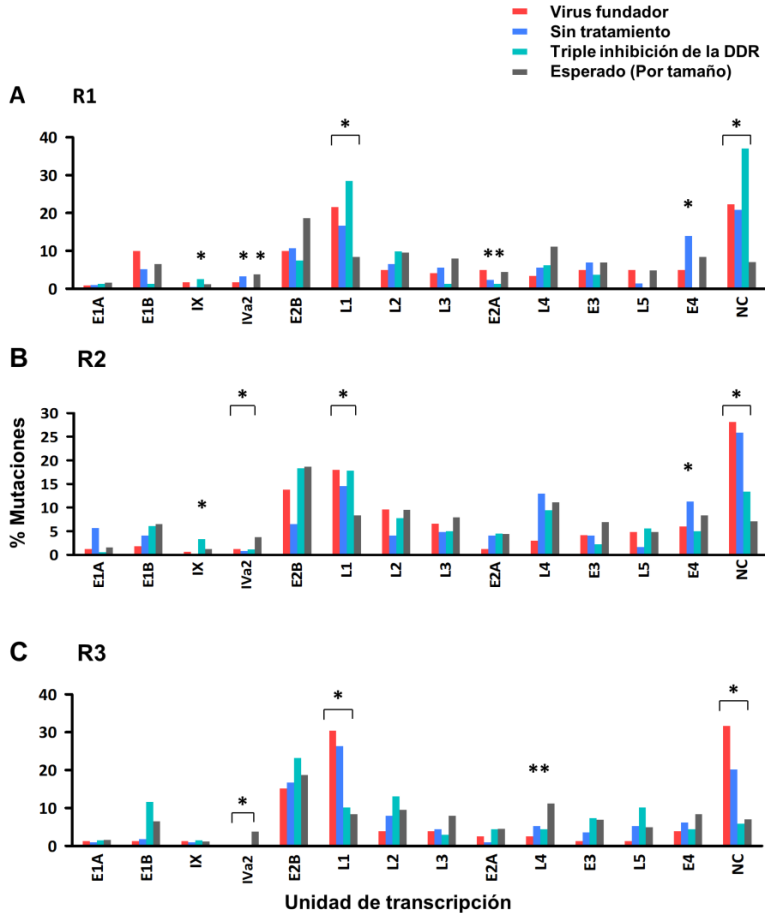


Figura 32. Mutaciones identificadas en cada una de las unidades de transcripción a lo largo del genoma de HAdV5. El número de mutaciones observado y esperado (en base al tamaño de la unidad transcripcional) en cada una de las regiones codificantes. El recuento se realizó para cada una de las réplicas de las condiciones experimentales: sin tratamiento (azul oscuro) y triple inhibición de la DDR (celeste), así como para el virus fundador (rojo). NC: regiones no codificantes.

* Test Ji-cuadrado: $P < 0.05$

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

Finalmente, debido a que las posiciones mutadas en el capítulo 2 presentaban alta diversidad in vivo, se estudió si estas sustituciones tendían también a aparecer en el virus fundador y los linajes evolucionados (en este capítulo 3). De las 257 posiciones variables identificadas en el capítulo anterior, 15 de ellas (5.83%) aparecían en alguna de las réplicas de los virus fundadores. En el caso de las poblaciones evolucionadas en presencia y ausencia de los inhibidores, se observó que 11 y 10 (4.28% y 3.89%) de estos sitios variables estaban también mutados, respectivamente. Tal y como puede observarse en la tabla 10, los sitios en los que se identificaron mutaciones en el capítulo 2 tienen mayor probabilidad de aparecer mutados (en virus fundadores y poblaciones evolucionadas) que el resto de posiciones genómicas en este capítulo (Test de Fisher: $P < 0.001$, en todos los casos).

Tabla 10. Tabla de contingencia en la que se muestra el recuento de mutaciones presentes y ausentes en los capítulos 2 y 3. En la tabla también se muestra el número de mutaciones esperadas en cada categoría (en color gris).

		CAPÍTULO 3					
		VIRUS FUNDADOR		SIN TRATAMIENTO		DDRI	
		PRESENCIA	AUSENCIA	PRESENCIA	AUSENCIA	PRESENCIA	AUSENCIA
CAPÍTULO 2	PRESENCIA	15 (1.97)	242 (255.02)	11 (2.54)	246 (254.46)	10 (1.96)	247 (255.04)
	AUSENCIA	259 (272.03)	35202 (35188.97)	342 (350.46)	35118 (35109.54)	262 (270.04)	35209 (35200.96)

3.3 Identificación de deleciones mediante métodos alternativos

Aunque se detectaron pequeñas inserciones y deleciones mediante secuenciación dúplex, su frecuencia fue similar a la observada en el plásmido control de secuenciación. En consecuencia, no pudimos descartar que estos cambios fuesen errores de secuenciación. Sin embargo, el análisis de secuenciación dúplex no nos permite la detección de grandes deleciones. Por consiguiente, decidimos utilizar dos metodologías alternativas para identificar grandes deleciones en nuestras muestras.

En teoría, el tamaño de inserto en nuestras librerías de Illumina era de 300 pb y el tamaño de las lecturas de 150 pb. Teniendo en cuenta esta información, el primer método consiste en mapear todas las lecturas a la secuencia de referencia. Seguidamente, se calculó la distancia entre lecturas pareadas teniendo en cuenta su posición de mapeo y orientación (Chateigner y col. 2015). La existencia de grandes distancias de mapeo entre lecturas pareadas es un indicador de grandes deleciones. Sin embargo, tal y como puede observarse en la figura 33, no se hallaron grandes diferencias en la frecuencia de las distancias de mapeo entre las poblaciones evolucionadas y el virus fundador.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

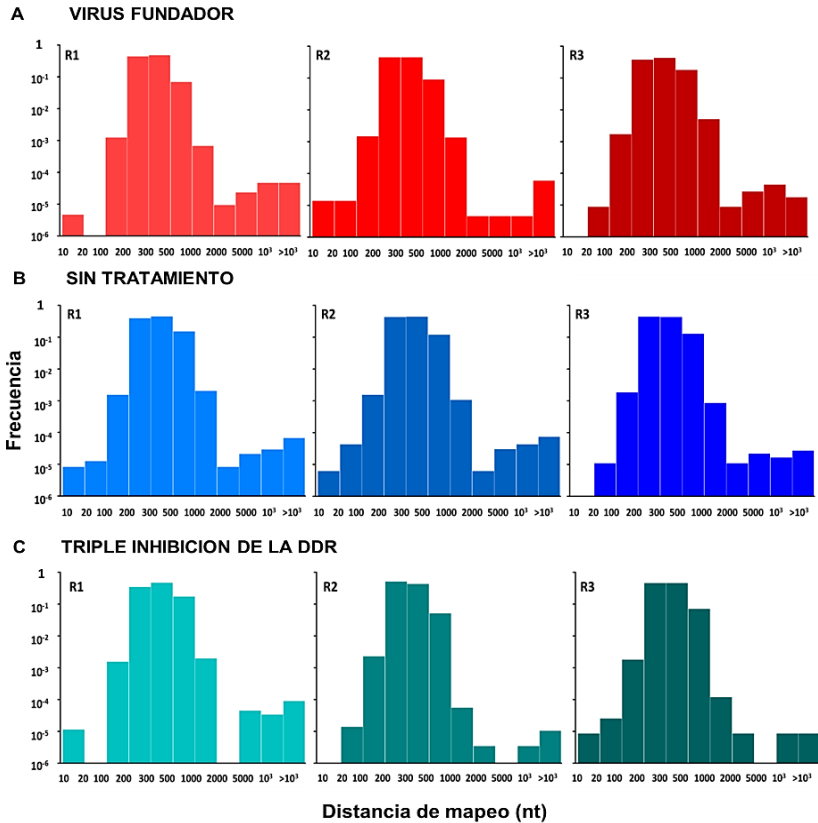


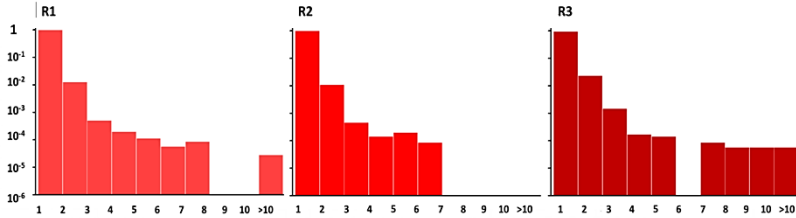
Figura 33. Distribución de los tamaños de las distancias de mapeo entre lecturas pareadas para cada una de las nueve poblaciones virales secuenciadas. La frecuencia de cada distancia de mapeo entre lecturas pareadas se representa para cada virus evolucionado (sin tratamiento en azul; inhibición de la DDR en verde) así como para los virus fundadores (rojo). Los valores de frecuencia fueron obtenidos dividiendo el número de *gaps* existentes dentro de cada intervalo de tamaño por el número de secuencias pareadas analizadas.

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

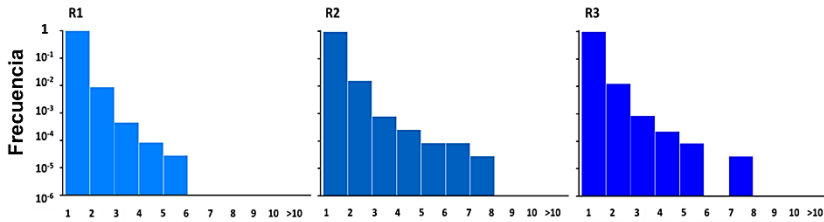
La segunda estrategia consistía en estudiar la existencia de discontinuidades entre las posiciones de mapeo de las lecturas. Después de llevar a cabo el análisis de DS, se extrajeron las posiciones de mapeo de cada una de las lecturas. De nuevo, la existencia de discontinuidades más grandes de lo esperado indicaría la existencia de una deleción. Sin embargo, al igual que en el caso anterior, no se observaron diferencias entre la frecuencia y el tipo de deleción entre las poblaciones evolucionadas en células no tratadas y tratadas con los inhibidores de la DDR y el virus fundador (Figura 34).

RESULTADOS Y DESARROLLO ARGUMENTAL: CAPÍTULO 3

A VIRUS FUNDADOR



B SIN TRATAMIENTO



C TRIPLE INHIBICION DE LA DDR

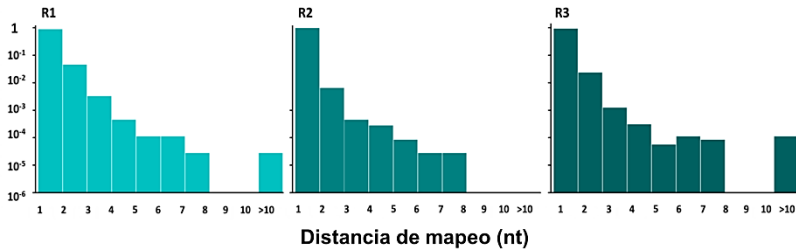


Figura 34. Distribución de los tamaños de las distancias de mapeo entre lecturas consecutivas para cada una de las nueve poblaciones virales secuenciadas. La frecuencia de cada distancia de mapeo entre lecturas pareadas se representa para cada virus evolucionado (sin tratamiento en azul; inhibición de la DDR en verde) así como para los virus fundadores (rojo). Los valores de frecuencia fueron obtenidos dividiendo el número de *gaps* existentes dentro de cada intervalo de tamaño por el número de secuencias pareadas analizadas.

4. DISCUSIÓN GENERAL

4.1 La diversidad a lo largo del genoma de adenovirus no es uniforme

Nuestros análisis de diversidad realizados utilizando secuencias pertenecientes al género *Mastadenovirus* revelaron que la diversidad a lo largo del genoma viral no es uniforme. Excepciones a esta observación son las especies A y F para las que los valores de diversidad fueron aparentemente constantes a lo largo del genoma viral.

En general, se observó hipervariabilidad en los genes que codifican las proteínas mayoritarias que conforman la cápside (*hexon*, *penton* y *fiber*). Ebner y col. compararon secuencias de *hexon* procedentes de 51 serotipos humanos observándose tres zonas variables, las cuales mostraban una alta divergencia entre serotipos y especies (Ebner y col. 2005). Esta divergencia era tanto a nivel de longitud como de secuencia de aminoácidos. Hasta la fecha, se han descrito hasta nueve regiones hipervariables (Crawford-Miksza y Schnurr 1996; Rux y col. 2003) localizadas en la parte de la superficie de la proteína que se encuentra en contacto con el medio (Roberts y col. 2006). Además, el intercambio de las regiones hipervariables de un serotipo a otro dio lugar a la neutralización del virus por la respuesta humoral.

DISCUSIÓN GENERAL

Es por esta razón que estas regiones hipervariables se utilizan en la clasificación de especies de adenovirus mediante la realización de pruebas serológicas (Roy y col. 1998).

Las proteínas *penton* y *fiber* presentan dominios implicados en la interacción virus-hospedador, jugando un papel fundamental en la unión e internalización del virus en la célula. De hecho, un estudio reveló que dodecaedros recombinantes formados por ambas proteínas virales eran capaces de unirse a la célula y entrar en su interior hasta alcanzar el núcleo (Fender y col. 1997). La existencia de diversidad entre especies en la proteína *fiber* ha sido previamente demostrada. En el caso de adenovirus humanos, las especies A-E presentan una única proteína *fiber* mientras que las especies F y G, presentan dos *fibers* (Hung-Yueh y col. 1994; Kovács y col. 2005; Jones y col. 2007). Otra muestra de variabilidad en la proteína *fiber* es la excepción encontrada en la especie D, cuya proteína *fiber* es capaz de unirse al receptor CAR pero su entrada en la célula solo tiene lugar cuando su *fiber* entra en contacto con el ácido siálico (Arnberg 2009).

Por otra parte, se ha descrito previamente que la variabilidad de secuencia en la proteína *penton* está restringida al *loop* variable y el motivo RGD, que se encuentra expuesto al medio e interacciona con las integrinas de la superficie celular (Chiu y col. 1999). En adenovirus humanos, este motivo únicamente está presente en las especies A, B, C, D, E y G.

DISCUSIÓN GENERAL

En la especie F, *penton* presenta los motivos (RGAD e IGDD) en los serotipos 40 y 41, respectivamente (Albinsson y Kidd 1999). Una de las razones por las que dentro del género *Mastadenovirus* se observa diversidad en ambas proteínas (*fiber* y *penton*) podría ser el amplio rango de hospedadores que son capaces de infectar los virus pertenecientes a este género. La única especie que no mostró alta variabilidad fue la especie F. La razón podría ser el número reducido de secuencias disponible y utilizado para llevar a cabo el análisis.

Otra de las regiones que mostró diversidad fue la unidad de transcripción E3. La región E3 codifica proteínas que modulan la respuesta antiviral en la que participan las células B, T y NK (Windheim y col. 2004). Esta unidad de transcripción está únicamente presente en los géneros *Mastadenovirus* y *Siadenovirus* (Davison y col. 2003). En adenovirus humanos, se ha confirmado que el número de ORFs localizados en esta región es variable entre especies. De ahí que se haya observado variabilidad en la longitud de E3 entre distintos adenovirus (Singh y col. 2013). Por ejemplo, en HAdV B se han identificado hasta nueve regiones codificantes mientras que en HAdV F únicamente cinco. Algo similar ocurre en E4 y se cree que es debido a eventos de duplicación y delección de genes a lo largo de la evolución de los adenovirus (Davison y col. 2003). Ambos mecanismos evolutivos son llevados a cabo por virus de DNA grandes como por ejemplo HSV para la generación de nuevos genes (Shackelton y Holmes 2004).

DISCUSIÓN GENERAL

A pesar de su relevancia en la biología del virus, algunas de las funciones de las proteínas codificadas por la unidad E3 son todavía desconocidas (Robinson y col. 2011). Uno de los ORFs más estudiado y para el que hemos obtenido valores elevados de diversidad, es el que codifica la proteína gp19k. Esta proteína viral está implicada en la inhibición de presentación de antígenos de la superficie celular al retener el complejo principal de histocompatibilidad I (MCH I) (Schowalter y col. 1997). Otras proteínas codificadas por esta unidad de transcripción juegan un papel fundamental en la inhibición de la entrada en apoptosis de las células infectadas (como por ejemplo 14.7K) o en el bloqueo de las respuestas inmunes innata y adaptativa, garantizando la replicación del virus (Benedict y col. 2001; Hilgendorf y col. 2003; Windheim y col. 2004).

Los extremos del genoma de adenovirus también mostraron valores de diversidad elevados. En ellos se localizan los genes más específicos del género/hospedador, como por ejemplo E1A, E1B y E4. Las unidades de transcripción E1A y E1B son estrictamente necesarias para que tenga lugar la replicación y la expresión génica viral de manera eficiente. Al igual que la región E3, las proteínas codificadas por estos genes también interaccionan con numerosas dianas celulares. Durante los primeros estadios de la infección, la proteína E1A interacciona con proteínas del hospedador para garantizar la expresión génica viral (Bandara y La Thangue 1991; Arroyo y Raychaudhuri 1992; Parreño y col. 2001).

DISCUSIÓN GENERAL

Por su parte, E1B además de desregular el ciclo celular promoviendo la entrada de la célula en fase S, interacciona con la proteína viral E4orf6 formando un complejo capaz de inhibir determinadas proteínas celulares, entre las que se encuentran algunos componentes de la DDR (Endter y Dobner 2004; Berk 2005; Weitzman y col. 2010).

La región genómica en la que se observó un mayor grado de conservación fue aquella en la que se localizan los genes *core* (IVa2, E2B, L1, L2, L3, E2A y L4). Estas unidades de transcripción se encuentran en la porción central del genoma y están implicadas en funciones generales del ciclo de infección. Las proteínas codificadas por estos genes participan en la replicación del genoma viral, el ensamblaje de los genomas nacientes y en la formación de la estructura del virión (Leppard 2014). Al mismo tiempo, en estas regiones genómicas altamente conservadas aparecen picos de diversidad específicos correspondientes a zonas hipervariables. Concretamente en los genes *penton* y *hexon*, tal y como se ha discutido previamente.

Esta localización de genes específicos del género/huésped en ambos extremos del genoma y una zona central más conservada no es exclusiva de adenovirus. También la encontramos en otros virus de dsDNA como las familias *Herpesviridae* y *Poxviridae*, donde la mayoría de genes localizados en los extremos tienen un papel fundamental en la interacción con el hospedador (McGeoch y Davison 1999; Upton y col. 2003).

DISCUSIÓN GENERAL

Por tanto, parece que la evolución de los adenovirus tiene lugar mediante la conservación de regiones específicas del genoma (situadas en la región central) y la acumulación de diversidad en regiones codificantes implicadas en la interacción con el hospedador (situadas en ambos extremos).

Debido a que los virus dependen de sus hospedadores para su replicación, la evolución de su genoma también está fuertemente influenciada por el huésped. La célula huésped utiliza una gran variedad de estrategias para tratar de eliminar al virus y garantizar su supervivencia (tenOever 2016). En consecuencia, los patógenos evolucionan para superar estas respuestas celulares. Esto explica la hipervariabilidad observada en las proteínas que conforman la cápside de adenovirus. En este sentido, se ha observado que los picornavirus, myxovirus y retrovirus cambian constantemente sus glicoproteínas de la envuelta para tratar de evadir su reconocimiento por el sistema inmune (Vossen y col. 2002). Otros virus de DNA con mayor capacidad de codificar proteínas como los poxvirus o herpesvirus han desarrollado mecanismos más sofisticados para evadir las respuestas celulares (Crow y col. 2016).

Este es el caso también de los adenovirus, pues presentan unidades de transcripción como E1A, E1B, E3 y E4 que codifican proteínas capaces de antagonizar ciertos mediadores del sistema inmunitario como MHC-I, inhibir la entrada de la célula en apoptosis mediada por TNF o modular la DDR garantizando la replicación del virus (Weitzman y Ornelles 2005).

DISCUSIÓN GENERAL

En consecuencia, la alta diversidad observada en estas regiones del genoma del virus involucradas en la interacción con componentes celulares refleja el proceso de adaptación al hospedador.

Al mismo tiempo, estudiamos la presión de selección gen a gen y los resultados obtenidos se ajustaron perfectamente a los análisis de diversidad previos. Los ratios $\frac{dN}{dS}$ obtenidos fueron más elevados en aquellas regiones implicadas en la interacción con el huésped que en los genes “*core*”. Esto implica la acción de la selección purificadora en aquellas regiones más conservadas del genoma del virus. Debido a que estas proteínas participan en funciones básicas para la replicación del virus, la aparición de variantes que modifiquen o que comprometan su funcionalidad, deberán ser eficientemente eliminadas por la selección natural.

DISCUSIÓN GENERAL

4.2 Tasa de mutación de HAdV5

La tasa de mutación espontánea en virus determina tanto su capacidad evolutiva como el potencial de infectar nuevos hospedadores, generar nuevas resistencias frente a medicamentos o evadir el sistema inmune. A diferencia de los virus de RNA, para los que se han realizado numerosas estimaciones de tasas de mutación, muy pocas han sido realizadas en virus de DNA. Los virus de DNA evolucionan más lentamente que los de RNA y en consecuencia, presentan tasas de mutación más bajas y cercanas a los límites de detección de la secuenciación de Sanger y NGS (Sanjuán y col. 2010; Geoghegan y Holmes 2018; Peck y Lauring 2018).

Estas limitaciones técnicas son probablemente la causa de que hoy en día dispongamos de muy poca información sobre tasas de mutación en virus DNA. Además, muchas de las estimaciones realizadas se basan en regiones específicas del genoma o sufren de sesgo por la acción de la selección (Drake 2005; Cuevas y col. 2009; Cheng y col. 2010; Sanjuan y col. 2010; Renzette y col. 2015). En este sentido, hemos sido capaces de estimar por primera vez de forma fiable utilizando DS (Schmitt y col. 2012; Kennedy y col. 2014), la tasa de mutación espontánea de un virus de dsDNA a nivel de genoma completo. Esta estimación fue realizada bajo condiciones experimentales de laboratorio controladas y minimizando al máximo la acción de la selección.

DISCUSIÓN GENERAL

La tasa de mutación inferida fue de 1.7×10^{-7} s/n/c. Esta tasa de mutación fue similar a la observada en otros virus de DNA, pero órdenes de magnitud inferior a la observada en virus de ssDNA (Duffy y col. 2008).

Las tasas de mutación por ciclo de infección son más informativas desde el punto de vista evolutivo que las tasas por ronda de replicación debido a que un ciclo de infección celular equivale a una generación viral. De hecho, incorporan procesos no necesariamente asociados a la replicación, como por ejemplo el daño al DNA, la edición del material genético o la reparación. Aparentemente, la estimación de la tasa de mutación por ronda de replicación da información más concreta acerca de la fidelidad de replicación del virus. Sin embargo, a partir de datos como los obtenidos en este trabajo no parece adecuado inferir estimaciones de tasa de mutación por ronda de replicación. Si aproximadamente 10^4 partículas infecciosas de HAdV5 son producidas por célula, deberían haber ocurrido por lo menos $\log_2(10^4) = 13.3$ rondas de replicación viral semi-conservativa por célula. Por tanto, la tasa de mutación por ronda de replicación para HAdV5 sería de $1.3 \times 10^{-7} / 13.3 = 9.8 \times 10^{-9}$. Cabe destacar que esta estimación está sujeta a errores durante el cálculo del *burst size*. Por ejemplo, en el caso de que el *burst size* fuera de 10^3 partículas infecciosas por célula, esto arrojaría una tasa de mutación por ronda de replicación ligeramente distinta de $1.3 \times 10^{-7} / \log_2(10^3) = 13.0 \times 10^{-9}$.

DISCUSIÓN GENERAL

El valor obtenido es muy bajo y por tanto, esta estimación carece de sentido pues parece obvio que no nos da información sobre la fidelidad de la enzima de replicación.

Por otro lado, según nuestros resultados obtenidos y los estudios realizados por otros investigadores (Robinson y col. 2011, 2013; Dhingra y col. 2019), la diversidad no es uniforme a lo largo del genoma de adenovirus. Esta observación nos lleva a plantearnos si las tasas de mutación varía de acuerdo a ello. En relación a esto, hemos observado que la tasa de mutación a lo largo del genoma de adenovirus es bastante uniforme. Únicamente se hallaron dos regiones genómicas de aproximadamente 5 kpb en las que la tasa de mutación fue más baja de lo esperado. En una de estas dos regiones se localiza la unidad de transcripción L3, que codifica para la proteasa viral y dos proteínas de la cápside, VI y *hexon*. En un estudio reciente se investigó la tasa de mutación a lo largo del genoma viral de HIV (Geller y col. 2015). El análisis reveló una disminución significativa de la tasa de mutación en aquellas regiones del genoma en las que se localizan los dominios más externos de la envoltura viral. Al igual que ocurre con la proteína gp120 de HIV, este punto frío mutacional es inesperado, pues la proteína *hexon* es uno de los genes más variables del genoma de adenovirus. Estos resultados no van en concordancia con los mecanismos de hipermutación desarrollados en respuesta al sistema inmune por otros patógenos que también mutan lentamente (Arambula y col. 2013; Moxon y col. 2006; Rando y Verstrepen 2007; Teng y Papavasiliou 2007).

DISCUSIÓN GENERAL

Por ejemplo, el virus vaccinia (VACV) utiliza los conocidos acordeones genéticos para aumentar el número de copias del gen K3L. Por medio de este mecanismo consigue aumentar de manera específica el número de mutaciones en este locus y evadir el sistema inmune (Elde y col. 2012).

Por otro lado, se ha observado una tasa de mutación alta en las regiones del genoma de bacterias que codifican para proteínas de la superficie (Moxon y col. 2006) o en *loci* implicados en el reconocimiento de las bacterias en algunos bacteriófagos (Medhekar y Miller 2007). En principio, los adenovirus también deberían beneficiarse de la acumulación de mutaciones en determinadas regiones de su genoma, sin embargo los puntos calientes de mutación que fuimos capaces de identificar se restringían a unos pocos nucleótidos y se localizaban en regiones genómicas de baja complejidad. Uno de estos puntos calientes de mutación, localizado en la unidad de transcripción L1, fue identificado en otro estudio como punto caliente de recombinación (Robinson y col. 2013). Este *hot spot* mutacional está ubicado entre una región conservada y otra más variable. Podría ocurrir que, al tratarse de una zona de baja complejidad genómica, los errores cometidos durante la replicación fueran muy frecuentes y esto causara el reclutamiento de proteínas del sistema post-replicativo celular, aumentando la probabilidad de recombinación.

DISCUSIÓN GENERAL

Por último, los niveles de variación genética tienen un papel importante en la evolución y conservación de las especies. Estos niveles parecen ser muy variables entre especies y poblaciones naturales. Sin embargo, los determinantes de estas diferencias son todavía un misterio. En esta tesis doctoral, se observó una perfecta correlación entre tasa de mutación y diversidad *in vivo* en nuestro virus modelo. La tasa de mutación y el tamaño efectivo poblacional (N_e) determinan conjuntamente la diversidad genética intraespecífica. Sin embargo, a menudo el papel de la tasa de mutación es ignorado (Corbett-Detig y col. 2015; Ellegren y Galtier 2016b; Romiguier y col. 2014). Generalmente se ha pensado que la fuente de diversidad genética en virus consiste fundamentalmente en los errores cometidos durante la replicación. Sin embargo, numerosos estudios han mostrado que la diversidad genética en las poblaciones virales depende de múltiples procesos tanto dependientes del virus como del hospedador (Sanjuán y Domingo-Calap 2016a).

4.3 Probabilidad de mutación alta en dinucleótidos CpG

En general, el porcentaje de dinucleótidos CpG en el genoma de vertebrados es menor que el esperado por azar. Se piensa que esto es debido a que la metilación tiene lugar casi exclusivamente en citosinas seguidas de una guanina en la misma cadena (dinucleótidos CpG) (Hodgkinson y Eyre-Walker 2011). La metilación es una de las modificaciones epigenéticas más estudiadas que consiste en la adición de un grupo metilo (-CH₃) al carbono 5 del nucleótido citosina.

Las citosinas metiladas tienden a convertirse en timinas por medio de desaminación enzimática o espontánea, observándose elevadas tasas de sustitución C → T. Sin embargo, el conocimiento sobre metilación o tasas de sustitución elevadas en dinucleótidos CG es muy pobre (Pontecorvo y col. 2000).

En el genoma de HAdV5, los dinucleótidos CpG mostraron ser 2.7 veces más propensos a mutación que el resto de bases C o G en el genoma (Risso-Ballester y col. 2016). Una explicación a esta observación podría ser la metilación de estos residuos durante la infección. El impacto de la metilación de citosinas en dinucleótidos CpG sobre el ciclo de vida ha sido ampliamente estudiado en virus capaces de establecer infecciones latentes o de reactivarse. También ha sido investigada en virus capaces de integrarse en el genoma hospedador como *Retroviridae* o *Papillomaviridae* (Takács y col. 2001; Shalginskikh y col. 2013).

DISCUSIÓN GENERAL

Sin embargo, existe muy poca información acerca de la frecuencia y el papel de la metilación en virus que se replican de manera activa. Bonvicini y col. demostraron por primera vez metilación de CpG en un virus de ssDNA, concretamente el parvovirus B19 (Bonvicini y col. 2012). Estos y otros estudios han mostrado que los virus latentes a menudo muestran metilación en sus genomas mientras que, los virus de replicación activa no muestran signos de metilación o están hipometilados (Hoelzer y col. 2008). En este sentido, adenovirus no es una excepción ya que, varios estudios han demostrado que el genoma de HAdV5 está muy poco metilado (Gunthert y col. 1976; Hoelzer y col. 2008). Por tanto, en principio, la metilación parece no ser la causa de la alta probabilidad de mutación observada en los dinucleótidos CpG en HAdV5.

4.4 La DDR y la tasa de mutación de HAdV5

Según la conocida regla de Drake (Drake 1991a) existe una correlación inversa entre el tamaño del genoma y la tasa de mutación en organismos que poseen DNA como material genético. Entre estos microorganismos se encuentran las bacterias, los eucariotas unicelulares y los virus. Esta regla fue consecuencia de la observación de una tasa de mutación constante de aproximadamente 0.003 sustituciones por genoma por ronda de replicación (s/g/r).

Teniendo en cuenta que HAdV5 tiene un tamaño de genoma de 35.9 kpb, esto se traduce en 0.0046 s/g/r, una estimación que se ajusta perfectamente a la regla de Drake. Sin embargo, si comparamos la tasa de mutación obtenida para HAdV5 con la obtenida para otros virus de DNA de tamaños genómicos distintos, esta regla podría ser cuestionable. Por ejemplo, para el citomegalovirus murino (MCMV), se obtuvo una tasa similar a la estimada para HAdV5 en este trabajo, pero por día en lugar de por ciclo de infección (1.4×10^{-7}) (Cheng y col. 2010). La tasa de mutación de HSV se estimó mediante el estudio de la aparición de mutaciones deletéreas en el gen *tk* utilizando ganciclovir (Drake 2005). Por medio de este ensayo, se obtuvo una tasa de 5.9×10^{-8} s/n/c. Si se tienen en cuenta los tamaños genómicos de HAdV5, HSV y CMV (35.9, 150 y 230-236 kpb, respectivamente), se obtienen tasas de mutación que difieren aproximadamente en un orden de magnitud y además, estas son más bajas de lo esperado para los virus de DNA

DISCUSIÓN GENERAL

según la regla de Drake (0.003 s/g/r). Cabe destacar que los procedimientos utilizados para estimar las tasas de mutación de HSV y MCMV sufren de limitaciones técnicas. El hecho de que estas estimaciones se basen en únicamente una pequeña porción del genoma o sufran del sesgo introducido por la acción de la selección natural hace que no podamos aventurarnos en sacar conclusiones. Sin embargo, esto sugiere que otros factores, además del tamaño genómico podrían determinar las tasas de mutación en virus de DNA. Por ejemplo, el hecho de que el virus utilice la DNA polimerasa huésped o codifique la suya propia podría ser uno de ellos (Sanjuán y Domingo-Calap, 2016b). Otro factor importante que podría determinar la tasa de mutación en virus de DNA podría ser su estructura genómica. El ssDNA es más susceptible de ser modificado por la célula hospedadora o de ser dañado por factores tanto exógenos como endógenos (Chan y col. 2012).

La tasa de mutación en células humanas se encuentra en el orden de 10^{-9} s/n/r (Lynch 2010). Sin embargo, las células tumorales han mostrado tasas de mutación en el rango de 10^{-7} - 10^{-6} (Bielas y col. 2006; Williams y col. 2016). Esta diferencia de hasta tres órdenes de magnitud podría ser resultado del mal funcionamiento de los *checkpoints* durante el ciclo celular en las células tumorales (Curtin 2012). En el capítulo 2, la tasa de mutación de HAdV5 se estimó utilizando células tumorales HeLa. Por tanto, la tasa de mutación obtenida fue similar o incluso más baja que la de su célula hospedadora.

Esta observación podría ser debida al acceso del virus a la DDR. Sin embargo, las células HeLa no son un buen modelo para investigar esta posibilidad, puesto que por lo general las células tumorales presentan las rutas de la DDR alteradas. Por ello hemos elegido células MRC-5 en su lugar para realizar los experimentos llevados a cabo en el capítulo 3.

Las proteínas virales implicadas en la interacción con la DDR mostraron valores de diversidad y de ω más altos que los observados para los genes “core” pero al mismo tiempo, similares a los observados en otras proteínas que interaccionan con el hospedador, como por ejemplo las codificadas por E3. Además, el análisis de selección reveló selección positiva en dos de ellos (E4orf3 y E4orf6), lo que refleja la importancia de estas proteínas en la biología del virus.

Sin embargo, se desconoce si la explotación de las vías de reparación por parte de los virus de DNA podría tener alguna consecuencia sobre la estabilidad de su genoma. Este aprovechamiento podría explicar la baja tasa de mutación obtenida tanto para HAdV5 como para otros virus de DNA (Sanjuán y col. 2016). Puesto que los virus de RNA no pueden acceder a la DDR, esta interacción también podría explicar las diferencias halladas entre los virus de RNA y DNA.

DISCUSIÓN GENERAL

Recurrimos a la DS de alta fidelidad para testar el efecto de la inhibición de las tres principales rutas que modulan la DDR en la generación de diversidad en HAdV5.

Se realizó un ensayo de evolución experimental en el que HAdV5 se sometió a diez pases seriados de evolución en células no tumorales (MRC-5) tratadas con inhibidores químicos de la DDR. Las frecuencias de mutación, a diferencia de las tasas de mutación, no miden el número de sustituciones producidas de *novo* por ciclo de infección, sino el número de mutaciones producidas en una población en un momento dado. Las frecuencias de mutación observadas son dependientes de la tasa de mutación pero también de otros factores como el número de generaciones transcurridas o la selección, entre otros mecanismos evolutivos (Uil y col. 2011). Las poblaciones evolucionadas en ausencia de inhibición de la DDR revelaron frecuencias de mutación promedio de $(9.3 \pm 0.7) \times 10^{-6}$ s/n. Teniendo en cuenta que aproximadamente transcurrió un ciclo de infección en cada uno de los pases de evolución e ignorando la acción de la selección, se obtiene una tasa de mutación de $3.4 \times 10^{-6} / 10 = 3.4 \times 10^{-7}$ s/n/c, un valor similar al obtenido previamente en el capítulo 2.

No obstante, debido a la esperable acción de la selección y otros factores, por lo general no es adecuado transformar frecuencias observadas de mutación en tasas de mutación de esta manera. De hecho, las frecuencias de mutación aumentaron muy poco en las poblaciones evolucionadas en comparación con el virus fundador.

Esto sugiere que tras la evolución, el número de mutaciones acumuladas en el genoma del virus fue bajo.

Podría ocurrir que la tasa de HAdV5 en células MRC-5 fuera extremadamente baja y que la mayoría de sustituciones identificadas fueran errores de secuenciación. Esto es poco probable debido a que el plásmido control utilizado durante la secuenciación mostró una frecuencia de mutación más baja. También existe la posibilidad de que la mayoría de las mutaciones identificadas en las poblaciones evolucionadas en respuesta a la inhibición de la DDR fueran altamente perjudiciales o letales para el virus. Esto haría que la selección actuara sobre estos genotipos y por consiguiente, tendría lugar la eliminación de estas variantes en la población. Esta hipótesis también es poco probable ya que, a pesar de la selección, las frecuencias de mutación obtenidas en este ensayo fueron alrededor de 10 veces más altas que las obtenidas en el estudio anterior (capítulo 2). La deriva genética tampoco parece ser una explicación a la no acumulación de diversidad tras la evolución experimental, pues se utilizó una MOI de 0.1 PFU/célula, lo que implica que cada pase se inició inoculando 7×10^4 unidades infecciosas.

Tampoco observamos diferencias significativas en las frecuencias de mutación entre las poblaciones evolucionadas en células tratadas y no tratadas con los inhibidores químicos de la DDR. Por tanto, nuestros resultados sugieren que, aunque numerosos estudios han demostrado previamente la interacción entre adenovirus y la DDR, esta interacción virus-huésped parece no tener un efecto sobre la

DISCUSIÓN GENERAL

tasa de mutación del virus (Cvijović y col. 2018). En consecuencia, la baja tasa de mutación previamente observada en adenovirus podría ser fruto únicamente de la actividad 3' exonucleasa de su DNA polimerasa (King y col. 1997; Brenkman y col. 2002)

Como se ha discutido previamente, nuestros datos pueden estar sujetos a interpretaciones alternativas. Además de la posible purga de mutaciones a través de la acción rápida de la selección o de la deriva genética fruto de la generación de posibles cuellos de botella entre pases, otros detalles técnicos podrían explicar la baja diversidad observada.

Existe la posibilidad de que el ruido de la secuenciación enmascarase los efectos de la inhibición de la DDR o que la supresión química de estas rutas al ser incompleta, haya permitido la actividad residual de las vías de señalización. Estos problemas podrían ser abordados en futuros trabajos mediante la utilización de metodologías como RNAi o CRISPR. Sin embargo, en experimentos preliminares, no hemos sido capaces de crear una línea celular de MRC-5 deficiente para ATM mediante CRISPR.

En conclusión, la secuenciación profunda de alta fidelidad no reveló ningún efecto de la inhibición de DDR sobre las frecuencias de mutación de adenovirus. Esto sugiere que los adenovirus no utilizan la DDR para corregir los errores de replicación. A pesar de ello, nuestros resultados deben validarse utilizando métodos alternativos antes de descartar de manera concluyente este rol para la DDR.

CONCLUSIONES FINALES

CONCLUSIONES FINALES

- La diversidad a lo largo del genoma de adenovirus no es uniforme, a excepción de las especies A y F.
- Los genes que codifican las proteínas mayoritarias que conforman la cápside (*hexon*, *penton* y *fiber*) así como la unidad de transcripción E3, implicada en la modulación del sistema inmune mostraron hipervariabilidad.
- El genoma de adenovirus presenta una zona central conservada, en la que se localizan los genes *core* y tiende a acumular diversidad en ambos extremos, donde se localizan los genes involucrados en la interacción con el hospedador.
- Los genes *core* muestran mayor presión de selección negativa y valores de diversidad promedio más bajos que los genes implicados en la interacción con el huésped.
- Los genes que codifican proteínas que interaccionan con la DDR presentan un grado de diversidad y presiones de selección similares a los obtenidos para otras unidades de transcripción involucradas en la modulación de la respuesta celular.

CONCLUSIONES FINALES

- El desarrollo de nuevas metodologías de secuenciación de alta fidelidad nos ha permitido la estimación de la tasa de mutación de virus de DNA. En concreto, hemos sido capaces de estimar la tasa de mutación de HAdV5 mediante DS bajo condiciones mínimas de selección. La tasa de mutación por ciclo de infección obtenida fue de 1.3×10^{-7} , una tasa de mutación similar a la estimada para otros virus de dsDNA.
- La PCR a tiempo real es una herramienta útil para la identificación de variantes de manera específica utilizando cebadores reversos cuya base 3' hibrida con la base *wildtype* o mutante y el posterior análisis de los valores de C_t obtenidos.
- En HAdV5, nuestros resultados revelan una alta probabilidad de mutación en dinucleótidos CpG. Sin embargo, debido a que otros estudios han demostrado que el genoma del adenovirus esta hipometilado, los mecanismos que subyacen esta observación son desconocidos.
- Las regiones de baja complejidad en el genoma de HAdV5 parecen constituir puntos calientes mutacionales debido a que han mostrado la acumulación de mutaciones de manera recurrente.

CONCLUSIONES FINALES

- A pesar de que, en general, se observó una tasa de mutación bastante uniforme a lo largo del genoma del virus, se identificaron dos regiones de aproximadamente 5kpb en las que el número de mutaciones fue bajo. En estas regiones se localizan las unidades de transcripción E1B, IX, IVa2 y L3.
- Las posiciones para las que se identificaron mutaciones mediante DS, muestran valores de diversidad *in vivo* significativamente más elevados que los observados en el resto del genoma del virus. En consecuencia, nuestros análisis revelan una correlación entre tasa de mutación y diversidad *in vivo* en HAdV5. Además, estas posiciones también mostraron mayor probabilidad de mutación que el resto de posiciones, tanto en los virus fundadores como en las poblaciones evolucionadas del capítulo 3.
- La evolución experimental realizada en HAdV5 en células procedentes de tejido sano tratadas con inhibidores de la DDR reveló que dicho tratamiento no tuvo aparentemente ningún efecto sobre las frecuencias de mutación del virus.
- La interacción entre adenovirus y la DDR parece no determinar la tasa de mutación del virus. La alta fidelidad de la polimerasa viral quizás sea la explicación a su baja tasa de mutación.

BIBLIOGRAFÍA

Aaskov, J. 2006. Long-term transmission of defective RNA viruses in humans and aedes mosquitoes. *Science* 311: 236–38.

Acevedo, A., L. Brodsky, and R. Andino. 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505: 686–90.

Adeyemi, R.O., S. Landry, M.E. Davis, M.D. Weitzman, and D.J. Pintel. 2010. Parvovirus minute virus of mice induces a DNA damage response that facilitates viral replication. *PLoS Pathogens* 6: e1001141.

Ahi, Y.S., S. V. Vemula, A.O. Hassan, G. Costakes, C. Stauffacher, and S.K. Mittal. 2015. Adenoviral L4 33K forms ring-like oligomers and stimulates ATPase activity of IVa2: implications in viral genome packaging. *Frontiers In Microbiology* 6: 318.

Alba, R., D. Cots, P. Ostapchuk, A. Bosch, P. Hearing, and M. Chillon. 2011. Altering the Ad5 packaging domain affects the maturation of the Ad particles. *PLoS One* 6: e19564.

Albinsson, B., and A.H. Kidd. 1999. Adenovirus type 41 lacks an RGD alpha(v)-integrin binding motif on the penton base and undergoes delayed uptake in A549 cells. *Virus Research* 64: 125–36.

BIBLIOGRAFÍA

Amstutz, B., M. Gastaldelli, S. Kälin, N. Imelli, K. Boucke, E. Wandeler, J. Mercer, S. Hemmi, and U.F. Greber. 2008. Subversion of CtBP1-controlled macropinocytosis by human adenovirus serotype 3. *The EMBO Journal* 27: 956–69.

Anderson, C.W., M.E. Young, and S.J. Flint. 1989. Characterization of the adenovirus 2 virion protein, mu. *Virology* 172: 506–12.

Anderson, E.J. 2010. Prevention and treatment of viral diarrhea in pediatrics. *Expert Review Of Anti-Infective Therapy* 8: 205–17.

Arambula, D., W. Wong, B.A. Medhekar, H. Guo, M. Gingery, E. Czornyj, M. Liu, S. Dey, P. Ghosh, and J.F. Miller. 2013. Surface display of a massively variable lipoprotein by a Legionella diversity-generating retroelement. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 110: 8212–17.

Arnberg, N. 2009. Adenovirus receptors: implications for tropism, treatment and targeting. *Reviews In Medical Virology* 19: 165–78.

Arnold, J.J., M. Vignuzzi, J.K. Stone, R. Andino, and C.E. Cameron. 2005. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *The Journal Of Biological Chemistry* 280: 25706–16.

BIBLIOGRAFÍA

Arroyo, M., and P. Raychaudhuri. 1992. Retinoblastoma-repression of E2F-dependent transcription depends on the ability of the retinoblastoma protein to interact with E2F and is abrogated by the adenovirus E1A oncoprotein. *Nucleic Acids Research* 20: 5947–54.

Avgousti, D.C., A.N. Della Fera, C.J. Otter, C. Herrmann, N.J. Pancholi, and M.D. Weitzman. 2017. Adenovirus core protein VII downregulates the dna damage response on the host genome. *Journal Of Virology* 91.

Baker, A., K.J. Rohleder, L.A. Hanakahi, and G. Ketner. 2007. Adenovirus E4 34k and E1b 55k oncoproteins target host DNA ligase IV for proteasomal degradation. *Journal Of Virology* 81: 7034–40.

Bakkenist, C.J., and M.B. Kastan. 2003. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* 421: 499–506.

Bandara, L.R., and N.B. La Thangue. 1991. Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature* 351: 494–97.

Beerman, T.A., G. Mueller, and H. Grimmond. 1983. Effects of neocarzinostatin on chromatin in HeLa S3 nuclei. *Molecular Pharmacology* 23: 493–99.

BIBLIOGRAFÍA

Belgnaoui, S.M., K.A. Fryrear, J.O. Nyalwidhe, X. Guo, and O.J. Semmes. 2010. The viral oncoprotein tax sequesters DNA damage response factors by tethering MDC1 to chromatin. *The Journal Of Biological Chemistry* 285: 32897–905.

Benedict, C.A., P.S. Norris, T.I. Prigozy, J.-L. Bodmer, J.A. Mahr, C.T. Garnett, F. Martinon, J. Tschopp, L.R. Gooding, and C.F. Ware. 2001. Three adenovirus E3 proteins cooperate to evade apoptosis by tumor necrosis factor-related apoptosis-inducing ligand receptor-1 and -2. *Journal Of Biological Chemistry* 276: 3270–78.

Benkő, M., and B. Harrach. 2003. Molecular evolution of adenoviruses. *Current Topics In Microbiology And Immunology* 272: 3–35.

Berk, A.J. 2005. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene* 24: 7673–85.

Bielas, J.H., K.R. Loeb, B.P. Rubin, L.D. True, and L.A. Loeb. 2006. Human cancers express a mutator phenotype. *Proceedings Of The National Academy Of Sciences* 103: 18238–42.

Blackford, A. N., R.N. Patel, N.A. Forrester, K. Theil, P. Groitl, G.S. Stewart, A.M.R. Taylor, I.M. Morgan, T. Dobner, R.J.A. Grand, and A.S. Turnell. 2010. Adenovirus 12 E4orf6 inhibits ATR activation by promoting TOPBP1 degradation. *Proceedings Of The National Academy Of Sciences* 107: 12251–56.

BIBLIOGRAFÍA

Blackford, Andrew N, R.K. Bruton, O. Dirlik, G.S. Stewart, A.M.R. Taylor, T. Dobner, R.J.A. Grand, and A.S. Turnell. 2008. A role for E1B-AP5 in ATR signaling pathways during adenovirus infection. *Journal Of Virology* 82: 7640–52.

Bonvicini, F., E. Manaresi, F. Di Furio, L. De Falco, and G. Gallinella. 2012. Parvovirus B19 DNA CpG Dinucleotide methylation and epigenetic regulation of viral expression. *PLoS ONE* 7: e33316.

Brenkman, A.B., E.C. Breure, and P.C. van der Vliet. 2002. Molecular architecture of adenovirus dna polymerase and location of the protein primer. *Journal Of Virology* 76: 8200–8207.

van Breukelen, B., A.B. Brenkman, P.E. Holthuisen, and P.C. van der Vliet. 2003. Adenovirus type 5 DNA binding protein stimulates binding of DNA polymerase to the replication origin. *Journal Of Virology* 77: 915–22.

Bull, J.J., M.R. Badgett, H.A. Wichman, J.P. Huelsenbeck, D.M. Hillis, A. Gulati, C. Ho, and I.J. Molineux. 1997. Exceptional convergent evolution in a virus. *Genetics* 147: 1497–1507.

Burckhardt, C.J., M. Suomalainen, P. Schoenenberger, K. Boucke, S. Hemmi, and U.F. Greber. 2011. Drifting motions of the adenovirus receptor CAR and immobile integrins initiate virus uncoating and membrane lytic protein exposure. *Cell Host & Microbe* 10: 105–17.

BIBLIOGRAFÍA

Burma, S., B.P. Chen, M. Murphy, A. Kurimasa, and D.J. Chen. 2001. ATM Phosphorylates Histone H2AX in response to dna double-strand breaks. *Journal Of Biological Chemistry* 276: 42462–67.

Burma, S., and D.J. Chen. 2004. Role of DNA–PK in the cellular response to DNA double-strand breaks. *DNA Repair* 3: 909–18.

Butcher, D. 1995. Muller's ratchet, epistasis and mutation effects. *Genetics* 141: 431–37.

Carson, C.T., R.A. Schwartz, T.H. Stracker, C.E. Lilley, D. V Lee, and M.D. Weitzman. 2003. The Mre11 complex is required for ATM activation and the G2/M checkpoint. *The EMBO Journal* 22: 6610–20.

Černý, J., B. Černá Bolfíková, J.J. Valdés, L. Grubhoffer, and D. Růžek. 2014. Evolution of tertiary structure of viral rna dependent polymerases. *PLoS ONE* 9: e96070.

Cervelli, T., J.A. Palacios, L. Zentilin, M. Mano, R.A. Schwartz, M.D. Weitzman, and M. Giacca. 2008. Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. *Journal Of Cell Science* 121: 349–57.

Chan, K., J.F. Sterling, S.A. Roberts, A.S. Bhagwat, M.A. Resnick, and D.A. Gordenin. 2012. Base damage within single-strand DNA underlies in vivo hypermutability induced by a ubiquitous environmental agent. *PLoS Genetics* 8: e1003149.

Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348: 454–55.

Chao, Lin, C.U. Rang, and L.E. Wong. 2002. Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage phi6. *Journal Of Virology* 76: 3276–81.

Charlesworth, B., M.T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134: 1289–1303.

Charrel, R.N., P. De Micco, and X. de Lamballerie. 1999. Phylogenetic analysis of GB viruses A and C: evidence for cospeciation between virus isolates and their primate hosts. *Journal Of General Virology* 80: 2329–35.

Chateigner, A., A. Bézier, C. Labrousse, D. Jiolle, V. Barbe, and E. Herniou. 2015. Ultra Deep Sequencing of a Baculovirus Population Reveals Widespread Genomic Variations. *Viruses* 7: 3625–46.

Chaurushiya, M.S., C.E. Lilley, A. Aslanian, J. Meisenhelder, D.C. Scott, S. Landry, S. Ticau, C. Boutell, J.R. Yates, B.A. Schulman, T. Hunter, M.D. Weitzman, and M.D. Weitzman. 2012. Viral E3 ubiquitin ligase-mediated degradation of a cellular E3: viral mimicry of a cellular phosphorylation mark targets the RNF8 FHA domain. *Molecular Cell* 46: 79–90.

BIBLIOGRAFÍA

Chaurushiya, M.S., and M.D. Weitzman. 2009. Viral manipulation of DNA repair and cell cycle checkpoints. *DNA Repair* 8: 1166–76.

Chen, P.H., D.A. Ornelles, and T. Shenk. 1993. The adenovirus L3 23-kilodalton proteinase cleaves the amino-terminal head domain from cytokeratin 18 and disrupts the cytokeratin network of HeLa cells. *Journal Of Virology* 67: 3507–14.

Cheng, T.P., M.C. Valentine, J. Gao, J.T. Pingel, and W.M. Yokoyama. 2010. Stability of murine cytomegalovirus genome after in vitro and in vivo passage. *Journal Of Virology* 84: 2623–28.

Chiu, C.Y., P. Mathias, G.R. Nemerow, and P.L. Stewart. 1999. Structure of adenovirus complexed with its internalization receptor, alphavbeta5 integrin. *Journal Of Virology* 73: 6759–68.

Chu, W., and D. Pavan-Langston. 1979. Ocular surface manifestations of the major viruses. *International Ophthalmology Clinics* 19: 135–67.

Ciccia, A., and S.J. Elledge. 2010. The DNA damage response: making it safe to play with knives. *Molecular Cell* 40: 179–204.

Collaco, R.F., J.M. Bevington, V. Bhriugu, V. Kalman-Maltese, and J.P. Trempe. 2009. Adeno-associated virus and adenovirus coinfection induces a cellular DNA damage and repair response via redundant phosphatidylinositol 3-like kinase pathways. *Virology* 392: 24–33.

BIBLIOGRAFÍA

Coppola, N., L. Onorato, C. Minichini, G. Di Caprio, M. Starace, C. Sagnelli, and E. Sagnelli. 2015. Clinical significance of hepatitis B surface antigen mutants. *World Journal Of Hepatology* 7: 2729.

Corbett-Detig, R.B., D.L. Hartl, and T.B. Sackton. 2015. Natural selection constrains neutral diversity across a wide range of species. *PLOS Biology* 13: e1002112.

Crawford-Miksza, L., and D.P. Schnurr. 1996. Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. *Journal Of Virology* 70: 1836–44.

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

Crow, M.S., K.K. Lum, X. Sheng, B. Song, and I.M. Cristea. 2016. Diverse mechanisms evolved by DNA viruses to inhibit early host defenses. *Critical Reviews In Biochemistry And Molecular Biology* 51: 452–81.

Cuesta, R., Q. Xi, and R.J. Schneider. 2001. Preferential translation of adenovirus mRNAs in infected cells. *Cold Spring Harbor Symposia On Quantitative Biology* 66: 259–67.

Cuevas, J.M., R. Geller, R. Garijo, J. López-Aldeguer, and R. Sanjuán. 2015. Extremely high mutation rate of hiv-1 in vivo. *PLOS Biology* 13: e1002251.

BIBLIOGRAFÍA

Cuevas, J.M., S. Duffy, and R. Sanjuán. 2009. Point mutation rate of bacteriophage PhiX174. *Genetics* 183: 747–49.

Cuevas, J.M., F. González-Candelas, A. Moya, and R. Sanjuán. 2009. Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo. *Journal Of Virology* 83: 5760–64.

Curtin, N.J. 2012. DNA repair dysregulation from cancer driver to therapeutic target. *Nature Reviews Cancer* 12: 801–17.

Cvijović, I., A.N. Nguyen Ba, and M.M. Desai. 2018. Experimental studies of evolutionary dynamics in microbes. *Trends In Genetics* 34: 693–703.

Dahl, J., J. You, and T.L. Benjamin. 2005. Induction and utilization of an atm signaling pathway by polyomavirus. *Journal Of Virology* 79: 13007–17.

Davis, A.J., B.P.C. Chen, and D.J. Chen. 2014. DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. *DNA Repair* 17: 21–29.

Davison, A.J., M. Benko, and B. Harrach. 2003. Genetic content and evolution of adenoviruses. *Journal Of General Virology* 84: 2895–2908.

Dhingra, A., E. Hage, T. Ganzenmueller, S. Böttcher, J. Hofmann, K. Hamprecht, P. Obermeier, B. Rath, F. Hausmann, T. Dobner, and A. Heim. 2019. Molecular evolution of Human Adenovirus (HAdV)

species *C*. *Scientific Reports* 9: 1039.

Domingo-Calap, P., and R. Sanjuán. 2011. Experimental evolution of RNA versus DNA viruses. *Evolution* 65: 2987–94.

Domingo, E. 2006. *Quasispecies: Concept and implications for virology*. Springer.

Doronin, K., K. Toth, M. Kuppaswamy, P. Krajcsi, A.E. Tollefson, and W.S.M. Wold. 2003. Overexpression of the ADP (E3-11.6K) protein increases cell lysis and spread of adenovirus. *Virology* 305: 378–87.

Drake, J.W. 1991a. A constant rate of spontaneous mutation in DNA-based microbes. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 88: 7160–64.

Drake, J.W. 1991b. Spontaneous mutation. *Annual Review Of Genetics* 25: 125–46.

Drake, J.W. 2005. On the mutation rate of herpes simplex virus type 1. *Genetics* 170: 969–70.

Drake, J.W., B. Charlesworth, D. Charlesworth, and J.F. Crow. 1998. Rates of spontaneous mutation. *Genetics* 148: 1667–86.

Duffy, S., and E.C. Holmes. 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. *Journal Of Virology* 82: 957–65.

BIBLIOGRAFÍA

Duffy, S., L.A. Shackelton, and E.C. Holmes. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews. Genetics* 9: 267–76.

E, X., and T. Kowalik. 2014. The DNA damage response induced by infection with human cytomegalovirus and other viruses. *Viruses* 6: 2155–85.

Ebner, K., W. Pinsker, and T. Lion. 2005. Comparative sequence analysis of the hexon gene in the entire spectrum of human adenovirus serotypes: phylogenetic, taxonomic, and clinical implications. *Journal Of Virology* 79: 12635–42.

Elde, N.C., S.J. Child, M.T. Eickbush, J.O. Kitzman, K.S. Rogers, J. Shendure, A.P. Geballe, and H.S. Malik. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* 150: 831–41.

Elena, S.F., F. González-Candelas, I.S. Novella, E.A. Duarte, D.K. Clarke, E. Domingo, J.J. Holland, and A. Moya. 1996. Evolution of fitness in experimental populations of vesicular stomatitis virus. *Genetics* 142: 673–79.

Ellegren, H., and N. Galtier. 2016. Determinants of genetic diversity. *Nature Reviews Genetics* 17: 422–33.

BIBLIOGRAFÍA

Endter, C., and T. Dobner. 2004. Cell transformation by human adenoviruses. *Current Topics In Microbiology And Immunology* 273: 163–214.

Ewing, G.B., and J.D. Jensen. 2016. The consequences of not accounting for background selection in demographic inference. *Molecular Ecology* 25: 135–41.

Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78: 737–56.

Fender, P., R.W. Ruigrok, E. Gout, S. Buffet, and J. Chroboczek. 1997. Adenovirus dodecahedron, a new vector for human gene transfer. *Nature Biotechnology* 15: 52–56.

Firth, C., A. Kitchen, B. Shapiro, M.A. Suchard, E.C. Holmes, and A. Rambaut. 2010. Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Molecular Biology And Evolution* 27: 2038–51.

Flint, S. J., Enquist, L. W., Racaniello, V.R.&, and A.M. Skalka. 2004. *Principles of Virology: Molecular Biology, Pathogenesis and Control of Animal Viruses*. Washington DC: ASM PRESS.

Fukui, K. 2010. DNA mismatch repair in eukaryotes and bacteria. *Journal Of Nucleic Acids* 2010: 1–16.

BIBLIOGRAFÍA

Gaggar, A., D.M. Shayakhmetov, and A. Lieber. 2003. CD46 is a cellular receptor for group B adenoviruses. *Nature Medicine* 9: 1408–12.

Gago, S., S.F. Elena, R. Flores, and R. Sanjuan. 2009. Extremely high mutation rate of a hammerhead viroid. *Science* 323: 1308–1308.

García-Arenal, F., A. Fraile, and J.M. Malpica. 2001. Variability and genetic structure of plant virus populations. *Annual Review Of Phytopathology* 39: 157–86.

Garcia-Diaz, M., and K. Bebenek. 2007. Multiple functions of DNA polymerases. *Critical Reviews In Plant Sciences* 26: 105–22.

Gastaldelli, M., N. Imelli, K. Boucke, B. Amstutz, O. Meier, and U.F. Greber. 2008. Infectious adenovirus type 2 transport through early but not late endosomes. *Traffic* 9: 2265–78.

Ge, L., J. Zhang, X. Zhou, and H. Li. 2007. Genetic structure and population variability of tomato yellow leaf curl China virus. *Journal Of Virology* 81: 5902–7.

Geller, R., P. Domingo-Calap, J.M. Cuevas, P. Rossolillo, M. Negroni, and R. Sanjuán. 2015. The external domains of the HIV-1 envelope are a mutational cold spot. *Nature Communications* 6: 8571.

Geoghegan, J.L., and E.C. Holmes. 2018. Evolutionary virology at 40. *Genetics* 210: 1151–62.

BIBLIOGRAFÍA

Gerrish, P.J., and R.E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. *Genetics* 102–103: 127–44.

Gillespie, K.A., K.P. Mehta, L.A. Laimins, and C.A. Moody. 2012. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. *Journal Of Virology* 86: 9520–26.

Girard, M., J.-P. Bouché, L. Marty, B. Revet, and N. Berthelot. 1977. Circular adenovirus DNA-protein complexes from infected HeLa cell nuclei. *Virology* 83: 34–55.

Good, B.H., A.M. Walczak, R.A. Neher, and M.M. Desai. 2014. Genetic diversity in the interference selection limit *PLoS Genetics* 10: e1004222.

Gräble, M., and P. Hearing. 1990. Adenovirus type 5 packaging domain is composed of a repeated element that is functionally redundant. *Journal Of Virology* 64: 2047–56.

Gruhne, B., R. Sompallae, D. Marescotti, S.A. Kamranvar, S. Gastaldello, and M.G. Masucci. 2009. The Epstein-Barr virus nuclear antigen-1 promotes genomic instability via induction of reactive oxygen species. *Proceedings Of The National Academy Of Sciences* 106: 2313–18.

BIBLIOGRAFÍA

Gunthert, U., M. Schweiger, M. Stupp, and W. Doerfler. 1976. DNA methylation in adenovirus, adenovirus-transformed cells, and host cells. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 73: 3923–27.

Gutierrez, B., M. Escalera-Zamudio, and O.G. Pybus. 2019. Parallel molecular evolution and adaptation in viruses. *Current Opinion In Virology* 34: 90–96.

Gutiérrez, S., Y. Michalakis, and S. Blanc. 2012. Virus population bottlenecks during within-host progression and host-to-host transmission. *Current Opinion In Virology* 2: 546–55.

Hearing, P., R.J. Samulski, W.L. Wishart, T. Shenk, and M.J. Imperiale. 1987. Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome. *Journal Of Virology* 61: 2555–58.

Hein, J., S. Boichuk, J. Wu, Y. Cheng, R. Freire, P.S. Jat, T.M. Roberts, and O. V. Gjoerup. 2009. Simian Virus 40 Large T Antigen disrupts genome integrity and activates a dna damage response via Bub1 binding. *Journal Of Virology* 83: 117–27.

Hickson, I., Y. Zhao, C.J. Richardson, S.J. Green, N.M.B. Martin, A.I. Orr, P.M. Reaper, S.P. Jackson, N.J. Curtin, and G.C.M. Smith. 2004. Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. *Cancer Research* 64: 9152–59.

BIBLIOGRAFÍA

Hilgendorf, A., J. Lindberg, Z. Ruzsics, S. Höning, A. Elsing, M. Löfqvist, H. Engelmann, and H.-G. Burgert. 2003. Two distinct transport motifs in the adenovirus E3/10.4-14.5 proteins act in concert to down-modulate apoptosis receptors and the epidermal growth factor receptor. *The Journal Of Biological Chemistry* 278: 51872–84.

Hilleman, M.R., and J.H. Werner. 1954. Recovery of new agent from patients with acute respiratory illness. *Experimental Biology And Medicine* 85: 183–88.

Hodgkinson, A., and A. Eyre-Walker. 2011. Variation in the mutation rate across mammalian genomes. *Nature Reviews Genetics* 12: 756–66.

Hoeben, R.C., and T.G. Uil. 2013. Adenovirus DNA replication. *Cold Spring Harbor Perspectives In Biology* 5: a013003.

Hoelzer, K., L.A. Shackelton, and C.R. Parrish. 2008. Presence and role of cytosine methylation in DNA viruses of animals. *Nucleic Acids Research* 36: 2825–37.

Hong, S.S., E. Szolajska, G. Schoehn, L. Franqueville, S. Myhre, L. Lindholm, R.W.H. Ruigrok, P. Boulanger, and J. Chroboczek. 2005. the 100k-chaperone protein from adenovirus serotype 2 (subgroup c) assists in trimerization and nuclear localization of hexons from subgroups c and b adenoviruses. *Journal Of Molecular Biology* 352: 125–38.

BIBLIOGRAFÍA

Hong, Y. Bin, Y. Choi, and G. Jung. 2004. Increased DNA polymerase fidelity of the Lamivudine resistant variants of human hepatitis B virus DNA polymerase. *Journal Of Biochemistry And Molecular Biology* 37: 167–76.

Hosokawa, K., and M.T. Sung. 1976. Isolation and characterization of an extremely basic protein from adenovirus type 5. *Journal Of Virology* 17: 924–34.

Hughes, A.L., and R. Friedman. 2000. Evolutionary diversification of protein-coding genes of hantaviruses. *Molecular Biology And Evolution* 17: 1558–68.

Hung-Yueh, Y., N. Pieniazek, D. Pieniazek, H. Gelderblom, and R.B. Luftig. 1994. Human adenovirus type 41 contains two fibers. *Virus Research* 33: 179–98.

Jee, J., A. Rasouly, I. Shamovsky, Y. Akivis, S.R. Steinman, B. Mishra, and E. Nudler. 2016. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature* 534: 693–96.

Jeong, S.-Y., A. Kumagai, J. Lee, and W.G. Dunphy. 2003. Phosphorylated claspin interacts with a phosphate-binding site in the kinase domain of Chk1 during ATR-mediated activation. *The Journal Of Biological Chemistry* 278: 46782–88.

Jiricny, J. 2013. Postreplicative mismatch repair. *Cold Spring Harbor Perspectives In Biology* 5: a012633–a012633.

BIBLIOGRAFÍA

Jones, M. S., B. Harrach, R.D. Ganac, M.M.A. Gozum, W.P. dela Cruz, B. Riedel, C. Pan, E.L. Delwart, and D.P. Schnurr. 2007. New adenovirus species found in a patient presenting with gastroenteritis. *Journal Of Virology* 81: 5978–84.

Jones, Morris Saffold, II, B. Harrach, R.D. Ganac, M.M.A. Gozum, W.P. dela Cruz, B. Riedel, C. Pan, E.L. Delwart, and D.P. Schnurr. 2007. New adenovirus species found in a patient presenting with gastroenteritis. *Journal Of Virology* 81: 5978.

Kalin, S., B. Amstutz, M. Gastaldelli, N. Wolfrum, K. Boucke, M. Havenga, F. DiGennaro, N. Liska, S. Hemmi, and U.F. Greber. 2010. macropinocytotic uptake and infection of human epithelial cells with species b2 adenovirus type 35. *Journal Of Virology* 84: 5336–50.

Kanagawa, T. 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal Of Bioscience And Bioengineering* 96: 317–23.

Kaplan, N.L., R.R. Hudson, and C.H. Langley. 1989. The hitchhiking effect revisited. *Genetics* 123: 887–99.

Kappen, L.S., T.E. Ellenberger, and I.H. Goldberg. 1987. Mechanism and base specificity of DNA Breakage in intact cells by neocarzinostatin. *Biochemistry* 26: 384–90.

BIBLIOGRAFÍA

Karen, K.A., and P. Hearing. 2011. Adenovirus core protein VII protects the viral genome from a DNA damage response at early times after infection. *Journal Of Virology* 85: 4135–42.

Kennedy, S.R., M.W. Schmitt, E.J. Fox, B.F. Kohn, J.J. Salk, E.H. Ahn, M.J. Prindle, K.J. Kuong, J.-C. Shen, R.-A. Risques, and L.A. Loeb. 2014. Detecting ultralow-frequency mutations by Duplex Sequencing. *Nature Protocols* 9: 2586–2606.

King, A.J., W.R. Teertstra, L. Blanco, M. Salas, and P.C. van der Vliet. 1997. Processive proofreading by the adenovirus DNA polymerase. Association with the priming protein reduces exonucleolytic degradation. *Nucleic Acids Research* 25: 1745–52.

Kinjo, T., J. Ham-Terhune, J.-M. Peloponese, and K.-T. Jeang. 2010. Induction of reactive oxygen species by human T-cell leukemia virus type 1 tax correlates with DNA damage and expression of cellular senescence marker. *Journal Of Virology* 84: 5431–37.

Kojaoghlanian, T., P. Flomenberg, and M.S. Horwitz. The impact of adenovirus infection on the immunocompromised host. *Reviews In Medical Virology* 13: 155–71.

Kovács, G.M., B. Harrach, A.N. Zakhartchouk, and A.J. Davison. 2005. Complete genome sequence of simian adenovirus 1: an Old World monkey adenovirus with two fiber genes. *The Journal Of General Virology* 86: 1681–86.

Kovesdi, I., R. Reichel, and J.R. Nevins. 1987. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proceedings Of The National Academy Of Sciences* 84: 2180–84.

Kumagai, A., and W.G. Dunphy. 2003. Repeated phosphopeptide motifs in Claspin mediate the regulated binding of Chk1. *Nature Cell Biology* 5: 161–65.

Kumagai, A., J. Lee, H.Y. Yoo, and W.G. Dunphy. 2006. TopBP1 Activates the ATR-ATRIP complex. *Cell* 124: 943–55.

de la Peña, M., S.F. Elena, and A. Moya. 2000. Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. *Evolution; International Journal Of Organic Evolution* 54: 686–91.

Lang, G.I., and A.W. Murray. 2008. Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 178: 67–82.

Leahy, J.J.J., B.T. Golding, R.J. Griffin, I.R. Hardcastle, C. Richardson, L. Rigoreau, and G.C.M. Smith. 2004. Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries. *Bioorganic & Medicinal Chemistry Letters* 14: 6083–87.

Lee, J.-H., and T.T. Paull. 2005. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* 308: 551–54.

BIBLIOGRAFÍA

Lemey, P., O.G. Pybus, S. Van Dooren, and A.-M. Vandamme. 2005. A Bayesian statistical analysis of human T-cell lymphotropic virus evolutionary rates. *Infection, Genetics And Evolution: Journal Of Molecular Epidemiology And Evolutionary Genetics In Infectious Diseases* 5: 291–98.

Leopold, P.L., G. Kreitzer, N. Miyazawa, S. Rempel, K.K. Pfister, E. Rodriguez-Boulan, and R.G. Crystal. 2000. Dynein- and microtubule-mediated translocation of adenovirus serotype 5 occurs after endosomal lysis. *Human Gene Therapy* 11: 151–65.

Leppard, K.N. 2014. Adenoviruses: molecular biology. *Reference Module In Biomedical Sciences*.

Li, G.-M. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Research* 18: 85–98.

Liang, X., M.T. Pickering, N.-H. Cho, H. Chang, M.R. Volkert, T.F. Kowalik, and J.U. Jung. 2006. Deregulation of DNA damage signal transduction by herpesvirus latency-associated M2. *Journal Of Virology* 80: 5862–74.

Lilley, C.E., C.T. Carson, A.R. Muotri, F.H. Gage, and M.D. Weitzman. 2005. DNA repair proteins affect the lifecycle of herpes simplex virus 1. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 102: 5844–49.

BIBLIOGRAFÍA

Lilley, C.E., R.A. Schwartz, and M.D. Weitzman. 2007. Using or abusing: viruses and the cellular DNA damage response. *Trends In Microbiology* 15: 119–26.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature* 362: 709–15.

Lindert, S., M. Silvestry, T.-M. Mullen, G.R. Nemerow, and P.L. Stewart. 2009. Cryo-electron microscopy structure of an adenovirus-integrin complex indicates conformational changes in both penton base and integrin. *Journal Of Virology* 83: 11491–501.

Liu, H., L. Jin, S.B.S. Koh, I. Atanasov, S. Schein, L. Wu, and Z.H. Zhou. 2010. Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science (New York, N.Y.)* 329: 1038–43.

Luftig, M.A. 2014. Viruses and the DNA damage response: Activation and antagonism. *Annual Rev Virol* 1: 605–25.

Luo, Y., and J. Qiu. 2013. Parvovirus infection-induced DNA damage response. *Future Virology* 8: 245–57.

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

Lynch, M. 2010. Evolution of the mutation rate. *Trends In Genetics* 26: 345–52.

BIBLIOGRAFÍA

Ma, H.-C., and P. Hearing. 2011. Adenovirus structural protein iiiia is involved in the serotype specificity of viral DNA packaging. *Journal Of Virology* 85: 7849–55.

Mabit, H., M.Y. Nakano, U. Prank, B. Saam, K. Dohner, B. Sodeik, and U.F. Greber. 2002. Intact microtubules support adenovirus and herpes simplex virus infections. *Journal Of Virology* 76: 9962–71.

Machida, K., G. McNamara, K.T.-H. Cheng, J. Huang, C.-H. Wang, L. Comai, J.-H.J. Ou, and M.M.C. Lai. 2010. Hepatitis C virus inhibits DNA damage repair through reactive oxygen and nitrogen species and by interfering with the ATM-NBS1/Mre11/Rad50 DNA repair pathway in monocytes and hepatocytes. *The Journal Of Immunology* 185: 6985–98.

Mahaney, B.L., K. Meek, and S.P. Lees-Miller. 2009. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochemical Journal* 417: 639–50.

Maier, O., D.L. Galan, H. Wodrich, and C.M. Wiethoff. 2010. An N-terminal domain of adenovirus protein VI fragments membranes by inducing positive membrane curvature. *Virology* 402: 11–19.

Maier, O., and C.M. Wiethoff. 2010. N-terminal α -helix-independent membrane interactions facilitate adenovirus protein VI induction of membrane tubule formation. *Virology* 408: 31–38.

Mangel, W.F., and C. San Martín. 2014. Structure, function and dynamics in adenovirus maturation. *Viruses* 6: 4536–70.

Matsuoka, S., B.A. Ballif, A. Smogorzewska, E.R. McDonald, K.E. Hurov, J. Luo, C.E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S.P. Gygi, and S.J. Elledge. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316: 1160–66.

McGeoch, D., and A.J. Davison. 1999. The molecular evolutionary history of the herpesviruses. *Origin And Evolution Of Viruses*: 441–65.

Medhekar, B., and J.F. Miller. 2007. Diversity-generating retroelements. *Current Opinion In Microbiology* 10: 388–95.

Meek, K., V. Dang, and S.P. Lees-Miller. 2008. DNA-PK: the means to justify the ends? *Advances In Immunology* 99: 33–58.

Meier, O., and U.F. Greber. 2004. Adenovirus endocytosis. *The Journal Of Gene Medicine* 6: S152–63.

Menéndez-Arias, L. 2009. Mutation rates and intrinsic fidelity of retroviral reverse transcriptases. *Viruses* 1: 1137–65.

Mesters, J.R., J. Tan, and R. Hilgenfeld. 2006. Viral enzymes. *Current Opinion In Structural Biology* 16: 776–86.

BIBLIOGRAFÍA

Minskaia, E., T. Hertzig, A.E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr. 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 103: 5108–13.

Mohni, K.N., C.M. Livingston, D. Cortez, and S.K. Weller. 2010. ATR and ATRIP are recruited to herpes simplex virus type 1 replication compartments even though atr signaling is disabled. *Journal Of Virology* 84: 12152–64.

Morris, S.J., G.E. Scott, and K.N. Leppard. 2010. Adenovirus Late-Phase Infection Is Controlled by a Novel L4 Promoter. *Journal Of Virology* 84: 7096–7104.

Moxon, R., C. Bayliss, and D. Hood. 2006. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annual Review Of Genetics* 40: 307–33.

Moya, A., E.C. Holmes, and F. González-Candelas. 2004. The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews Microbiology* 2: 279–88.

Mukherjee, B., C. Kessinger, J. Kobayashi, B.P.C. Chen, D.J. Chen, A. Chatterjee, and S. Burma. 2006. DNA-PK phosphorylates histone H2AX during apoptotic DNA fragmentation in mammalian cells. *DNA Repair* 5: 575–90.

BIBLIOGRAFÍA

Muller, H.J. 1964. The relation of recombination to mutational advance. *Mutation Research/Fundamental And Molecular Mechanisms Of Mutagenesis* 1: 2–9.

Murali, V.K., D.A. Ornelles, L.R. Gooding, H.T. Wilms, W. Huang, A.E. Tollefson, W.S.M. Wold, and C. Garnett-Benson. 2014. Adenovirus Death Protein (ADP) is required for lytic infection of human lymphocytes. *Journal Of Virology* 88: 903–12.

Mysiak, M.E., C. Wyman, P.E. Holthuisen, and P.C. van der Vliet. 2004. NFI and Oct-1 bend the Ad5 origin in the same direction leading to optimal DNA replication. *Nucleic Acids Research* 32: 6218–25.

Nakano, M.Y., K. Boucke, M. Suomalainen, R.P. Stidwill, and U.F. Greber. 2000. The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol. *Journal Of Virology* 74: 7085–95.

Nemerow, G.R., L. Pache, V. Reddy, and P.L. Stewart. 2009. Insights into adenovirus host cell interactions from structural studies. *Virology* 384: 380–88.

Nichol, D., J. Rutter, C. Bryant, A.M. Hujer, S. Lek, M.D. Adams, P. Jeavons, A.R. Anderson, R.A. Bonomo, and J.G. Scott. 2018. Collateral sensitivity is contingent on the repeatability of evolution. *BioRxiv*: 185892.

BIBLIOGRAFÍA

Novella, I.S., E.A. Duarte, S.F. Elena, A. Moya, E. Domingo, and J.J. Holland. 1995. Exponential increases of RNA virus fitness during large population transmissions. *Proceedings Of The National Academy Of Sciences* 92: 5841–44.

Ostapchuk, P., M.E. Anderson, S. Chandrasekhar, and P. Hearing. 2006. The L4 22-Kilodalton protein plays a role in packaging of the adenovirus genome. *Journal Of Virology* 80: 6973–81.

Park, S.-C., and J. Krug. 2007. Clonal interference in large populations. *Proceedings Of The National Academy Of Sciences* 104: 18135–40.

Parkinson, J., S.P. Lees-Miller, and R.D. Everett. 1999. Herpes simplex virus type 1 immediate-early protein vmw110 induces the proteasome-dependent degradation of the catalytic subunit of DNA-dependent protein kinase. *Journal Of Virology* 73: 650–57.

Parreño, M., J. Garriga, A. Limón, J.H. Albrecht, and X. Graña. 2001. E1A modulates phosphorylation of p130 and p107 by differentially regulating the activity of G1/S cyclin/CDK complexes. *Oncogene* 20: 4793–4806.

Peck, K.M., and A.S. Luring. 2018. Complexities of viral mutation rates *Journal Of Virology* 92.

BIBLIOGRAFÍA

Pennella, M.A., Y. Liu, J.L. Woo, C.A. Kim, and A.J. Berk. 2010. Adenovirus E1B 55-kilodalton protein is a p53-SUMO1 E3 ligase that represses p53 and stimulates its nuclear export through interactions with promyelocytic leukemia nuclear bodies. *Journal Of Virology* 84: 12210–25.

Pereira-Gómez, M., and R. Sanjuán. 2015. Effect of mismatch repair on the mutation rate of bacteriophage ϕ X174. *Virus Evolution* 1: vev010.

Perelson, A.S. 2002. Modelling viral and immune system dynamics. *Nature Reviews Immunology* 2: 28–36.

Plyusnin, A., and S.P. Morzunov. 2001. Virus evolution and genetic diversity of hantaviruses and their rodent hosts. *Current Topics In Microbiology And Immunology* 256: 47–75.

Pontecorvo, G., B. De Felice, and M. Carfagna. 2000. Novel methylation at GpC dinucleotide in the fish *Sparus aurata* genome. *Molecular Biology Reports* 27: 225–30.

Prevo, R., E. Fokas, P.M. Reaper, P.A. Charlton, J.R. Pollard, W.G. McKenna, R.J. Muschel, and T.B. Brunner. 2012. The novel ATR inhibitor VE-821 increases sensitivity of pancreatic cancer cells to radiation and chemotherapy. *Cancer Biology & Therapy* 13: 1072–81.

BIBLIOGRAFÍA

Querido, E., P. Blanchette, Q. Yan, T. Kamura, M. Morrison, D. Boivin, W.G. Kaelin, R.C. Conaway, J.W. Conaway, and P.E. Branton. 2001. Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes & Development* 15: 3104–17.

Quiñones-Mateu, M.E., S. Avila, G. Reyes-Teran, and M.A. Martinez. 2014. Deep sequencing: Becoming a critical tool in clinical virology. *Journal Of Clinical Virology* 61: 9–19.

Rando, O.J., and K.J. Verstrepen. 2007. Timescales of genetic and epigenetic inheritance. *Cell* 128: 655–68.

Reddy, V.S., S.K. Natchiar, P.L. Stewart, and G.R. Nemerow. 2010. Crystal structure of human adenovirus at 3.5 Å resolution. *Science* 329: 1071–75.

Rekosh, D.M., W.C. Russell, A.J. Bellet, and A.J. Robinson. 1977. Identification of a protein linked to the ends of adenovirus DNA. *Cell* 11: 283–95.

Renzette, N., C. Pokalyuk, L. Gibson, B. Bhattacharjee, M.R. Schleiss, K. Hamprecht, A.Y. Yamamoto, M.M. Mussi-Pinhata, W.J. Britt, J.D. Jensen, and T.F. Kowalik. 2015. Limits and patterns of cytomegalovirus genomic diversity in humans. *Proceedings Of The National Academy Of Sciences* 112: E4120–28.

BIBLIOGRAFÍA

Ribeiro, R.M., H. Li, S. Wang, M.B. Stoddard, G.H. Learn, B.T. Korber, T. Bhattacharya, J. Guedj, E.H. Parrish, B.H. Hahn, G.M. Shaw, and A.S. Perelson. 2012. Quantifying the diversification of hepatitis c virus (hcv) during primary infection: estimates of the in vivo mutation rate *PLoS Pathogens* 8: e1002881.

Risso-Ballester, J., J.M. Cuevas, and R. Sanjuán. 2016. Genome-wide estimation of the spontaneous mutation rate of human adenovirus 5 by high-fidelity deep sequencing *PLoS Pathogens* 12: e1006013.

Roberts, D.M. et al. 2006. Hexon-chimaeric adenovirus serotype 5 vectors circumvent pre-existing anti-vector immunity. *Nature* 441: 239–43. Robinson, C.M., D. Seto, M.S. Jones, D.W. Dyer, and J. Chodosh. 2011. Molecular evolution of human species D adenoviruses. *Infection, Genetics And Evolution* 11: 1208–17.

Robinson, C.M., J. Rajaiya, X. Zhou, G. Singh, D.W. Dyer, and J. Chodosh. 2011. The E3 CR1-gamma gene in human adenoviruses associated with epidemic keratoconjunctivitis. *Virus Research* 160: 120–27.

Robinson, C.M., G. Singh, J.Y. Lee, S. Dehghan, J. Rajaiya, E.B. Liu, M.A. Yousuf, R.A. Betensky, M.S. Jones, D.W. Dyer, D. Seto, and J. Chodosh. 2013. Molecular evolution of human adenoviruses. *Scientific Reports* 3: 1812.

BIBLIOGRAFÍA

Romiguier, J. et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515: 261–63.

Ross, P.J., M.A. Kennedy, C. Christou, M. Risco Quiroz, K.L. Poulin, and R.J. Parks. 2011. Assembly of helper-dependent adenovirus dna into chromatin promotes efficient gene expression. *Journal Of Virology* 85: 3950–58.

Rowe, W.P., R.J. Huebner, L.K. Gilmore, R.H. Parrott, and T.G. Ward. 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings Of The Society For Experimental Biology And Medicine. Society For Experimental Biology And Medicine* 84: 570–73.

Roy, A., A. Bhardwaj, P. Datta, G.C. Lander, and G. Cingolani. 2012. Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure* 20: 1403–13.

Roy, S., P.S. Shirley, A. McClelland, and M. Kaleko. 1998. Circumvention of immunity to the adenovirus major coat protein hexon. *Journal Of Virology* 72: 6875–79.

Russell, W.C. 2009. Adenoviruses: update on structure and function. *Journal Of General Virology* 90: 1–20.

Russell, W.C., Laver W.G., and P.J. Sanderson. 1968. Internal components of adenovirus. *Nature* 219: 1127–30.

Rux, J.J., P.R. Kuser, and R.M. Burnett. 2003. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *Journal Of Virology* 77: 9553–66.

Sanjuan, R., M.R. Nebot, N. Chirico, L.M. Mansky, and R. Belshaw. 2010. Viral mutation rates. *Journal Of Virology* 84: 9733–48.

Sanjuán, R., M. Pereira-Gómez, and J. Risso. 2016. Genome instability in DNA viruses. In *Genome Stability*, Elsevier, 37–47.

Sanjuán, R. 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *PLoS Pathogens* 8: e1002685.

Sanjuán, R. 2018. Collective properties of viral infectivity. *Current Opinion In Virology* 33: 1–6.

Sanjuán, R., and P. Domingo-Calap. 2016. Mechanisms of viral mutation. *Cellular And Molecular Life Sciences : CMLS* 73: 4433–48.

Sanjuán, R., E. Lázaro, and M. Vignuzzi. 2012. Biomedical implications of viral mutation and evolution. *Future Virology* 7: 391–402.

BIBLIOGRAFÍA

Sato, Y., T. Kamura, N. Shirata, T. Murata, A. Kudoh, S. Iwahori, S. Nakayama, H. Isomura, Y. Nishiyama, and T. Tsurumi. 2009. Degradation of phosphorylated p53 by viral protein-ECS E3 ligase complex. *PLoS Pathogens* 5: e1000530.

Scheffner, M., B.A. Werness, J.M. Huibregtse, A.J. Levine, and P.M. Howley. 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 63: 1129–36.

Schmitt, M.W., S.R. Kennedy, J.J. Salk, E.J. Fox, J.B. Hiatt, and L.A. Loeb. 2012. Detection of ultra-rare mutations by next-generation sequencing. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 109: 14508–13.

Schotsaert, M., and A. García-Sastre. 2014. Influenza Vaccines: A moving interdisciplinary field. *Viruses* 6: 3809–26.

Schowalter, D., J. Tubb, M. Liu, C. Wilson, and M. Kay. 1997. Heterologous expression of adenovirus E3-gp19K in an E1a-deleted adenovirus vector inhibits MHC I expression in vitro, but does not prolong transgene expression in vivo. *Gene Therapy* 4: 351–60.

Schröfelbauer, B., Y. Hakata, and N.R. Landau. 2007. HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 104: 4130–35.

BIBLIOGRAFÍA

Schwartz, R.A., C.T. Carson, C. Schuberth, and M.D. Weitzman. 2009. Adeno-Associated Virus replication induces a DNA damage response coordinated by DNA-Dependent Protein Kinase. *Journal Of Virology* 83: 6269–78.

Shackelton, L. A., and E.C. Holmes. 2006. Phylogenetic evidence for the rapid evolution of Human B19 Erythrovirus. *Journal Of Virology* 80: 3666–69. Shackelton, Laura A., and E.C. Holmes. 2004. The evolution of large DNA viruses: combining genomic information of viruses and their hosts. *Trends In Microbiology* 12: 458–65.

Shackelton, Laura A, C.R. Parrish, U. Truyen, and E.C. Holmes. 2005. High rate of viral evolution associated with the emergence of carnivore parvovirus. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 102: 379–84.

Shalginskikh, N., A. Poleshko, A.M. Skalka, and R.A. Katz. 2013. Retroviral DNA methylation and epigenetic repression are mediated by the antiviral host protein Daxx. *Journal Of Virology* 87: 2137–50.

Sharma, A., K. Singh, and A. Almasan. 2012. Histone H2AX phosphorylation: a marker for dna damage. In Humana Press, Totowa, NJ, 613–26.

Shendure, J., and H. Ji. 2008. Next-generation DNA sequencing. *Nature Biotechnology* 26: 1135–45.

BIBLIOGRAFÍA

Shirata, N., A. Kudoh, T. Daikoku, Y. Tatsumi, M. Fujita, T. Kiyono, Y. Sugaya, H. Isomura, K. Ishizaki, and T. Tsurumi. 2005. Activation of ataxia telangiectasia-mutated DNA damage checkpoint signal transduction elicited by herpes simplex virus infection. *The Journal Of Biological Chemistry* 280: 30336–41.

Singh, G., C.M. Robinson, S. Dehghan, M.S. Jones, D.W. Dyer, D. Seto, and J. Chodosh. 2013. Homologous recombination in E3 genes of human adenovirus species D. *Journal Of Virology* 87: 12481–88.

Singhal, G., E. Leo, S.K.G. Setty, Y. Pommier, and B. Thimmapaya. 2013. Adenovirus E1A oncogene induces rereplication of cellular dna and alters DNA replication dynamics. *Journal Of Virology* 87: 8767–78.

Smith, E.C., H. Blanc, M. Vignuzzi, M.R. Denison, and M.R. Denison. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics *PLoS Pathogens* 9: e1003565.

Smith, E.C., N.R. Sexton, and M.R. Denison. 2014. Thinking outside the triangle: replication fidelity of the largest RNA viruses. *Annual Review Of Virology* 1: 111–32.

Smith, J., L.M. Tho, N. Xu, and D.A. Gillespie. 2010. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Advances In Cancer Research* 108: 73–112.

Smith, J.G., C.M. Wiethoff, P.L. Stewart, and G.R. Nemerow. 2010. Adenovirus. In *Current Topics in Microbiology and Immunology*, , 195–224.

Soria, C., F.E. Estermann, K.C. Espantman, and C.C. O'Shea. 2010. Heterochromatin silencing of p53 target genes by a small viral protein. *Nature* 466: 1076–81.

Squarzanti, D.F., R. Sorrentino, M.M. Landini, A. Chiesa, S. Pinato, F. Rocchio, M. Mattii, L. Penengo, and B. Azzimonti. 2018. Human papillomavirus type 16 E6 and E7 oncoproteins interact with the nuclear p53-binding protein 1 in an in vitro reconstructed 3D epithelium: new insights for the virus-induced DNA damage response. *Virology Journal* 15: 176.

Steinhauer, D.A., E. Domingo, and J.J. Holland. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122: 281–88.

Stewart, G.S., B. Wang, C.R. Bignell, A.M.R. Taylor, and S.J. Elledge. 2003. MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature* 421: 961–66.

BIBLIOGRAFÍA

Stracker, T.H., C.T. Carson, and M.D. Weitzman. 2002. Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. *Nature* 418: 348–52.

Stracker, T.H., T. Usui, and J.H.J. Petrini. 2009. Taking the time to make important decisions: the checkpoint effector kinases Chk1 and Chk2 and the DNA damage response. *DNA Repair* 8: 1047–54.

Strunze, S., M.F. Engelke, I.-H. Wang, D. Puntener, K. Boucke, S. Schleich, M. Way, P. Schoenenberger, C.J. Burckhardt, and U.F. Greber. 2011. Kinesin-1-mediated capsid disassembly and disruption of the nuclear pore complex promote virus infection. *Cell Host & Microbe* 10: 210–23.

Subramanian, T., S. Vijayalingam, and G. Chinnadurai. 2006. Genetic identification of Adenovirus type 5 genes That Influence Viral Spread. *Journal Of Virology* 80: 2000.

Sun, S., S. Gao, K. Kondabagil, Y. Xiang, M.G. Rossmann, and V.B. Rao. 2012. Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. *Proceedings Of The National Academy Of Sciences* 109: 817–22.

Suzuki, Y., K. Katayama, S. Fukushi, T. Kageyama, A. Oya, H. Okamura, Y. Tanaka, M. Mizokami, and T. Gojobori. 1999. Slow evolutionary rate of GB virus C/hepatitis G virus. *Journal Of Molecular Evolution* 48: 383–89.

BIBLIOGRAFÍA

Takács, M., J. Segesdi, K. Balog, M. Mezei, G. Tóth, and J. Minárovits. 2001. Relative deficiency in CpG dinucleotides is a widespread but not unique feature of Gammaherpesvirinae genomes. *Acta Microbiologica Et Immunologica Hungarica* 48: 349–57.

Taylor, T.J., and D.M. Knipe. 2004. Proteomics of herpes simplex virus replication compartments: association of cellular DNA replication, repair, recombination, and chromatin remodeling proteins with ICP8. *Journal Of Virology* 78: 5856–66.

Teng, G., and F.N. Papavasiliou. 2007. Immunoglobulin somatic hypermutation. *Annual Review Of Genetics* 41: 107–20.

tenOever, B.R. 2016. The evolution of antiviral defense systems. *Cell Host & Microbe* 19: 142–49.

Tollefson, A.E., B. Ying, K. Doronin, P.D. Sidor, and W.S.M. Wold. 2007. Identification of a new human adenovirus protein encoded by a novel late l-strand transcription unit. *Journal Of Virology* 81: 12918–26.

Tribouley, C., P. Lutz, A. Staub, and C. Keding. 1994. The product of the adenovirus intermediate gene IVa2 is a transcriptional activator of the major late promoter. *Journal Of Virology* 68: 4450–57.

BIBLIOGRAFÍA

Turnell, A.S., and R.J. Grand. 2012. DNA viruses and the cellular DNA-damage response. *The Journal Of General Virology* 93: 2076–97.

Tyler, R.E., S.G. Ewing, and M.J. Imperiale. 2007. formation of a multiple protein complex on the adenovirus packaging sequence by the Iva2 protein. *Journal Of Virology* 81: 3447–54.

Uil, T.G., J. Vellinga, J. de Vrij, S.K. van den Hengel, M.J.W.E. Rabelink, S.J. Cramer, J.J.M. Eekels, Y. Ariyurek, M. van Galen, and R.C. Hoeben. 2011. Directed adenovirus evolution using engineered mutator viral polymerases. *Nucleic Acids Research* 39: e30–e30.

Umemura, T., Y. Tanaka, K. Kiyosawa, H.J. Alter, and J.W.-K. Shih. 2002. Observation of positive selection within hypervariable regions of a newly identified DNA virus (SEN virus). *FEBS Letters* 510: 171–74.

Upton, C., S. Slack, A.L. Hunter, A. Ehlers, and R.L. Roper. 2003. Poxvirus orthologous clusters: toward defining the minimum essential poxvirus genome. *Journal Of Virology* 77: 7590–7600.

Vossen, M., E. Westerhout, C. Söderberg-Nauclér, and E. Wiertz. 2002. Viral immune evasion: a masterpiece of evolution. *Immunogenetics* 54: 527–42.

Walkiewicz, M.P., N. Morral, and D.A. Engel. 2009. Accurate single-day titration of adenovirus vectors based on equivalence of protein VII nuclear dots and infectious particles. *Journal Of Virological Methods* 159: 251–58.

Ward, I.M., K. Minn, and J. Chen. 2004. UV-induced Ataxia-telangiectasia-mutated and Rad3-related (ATR) activation requires replication stress. *Journal Of Biological Chemistry* 279: 9677–80.

Weber, J.M. 2007. Synthesis and assay of recombinant adenovirus protease. *Methods In Molecular Medicine* 131: 251–55.

Weiden, M.D., and H.S. Ginsberg. 1994. Deletion of the E4 region of the genome produces adenovirus DNA concatemers. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 91: 153–57.

Weigand, H., and F. Leese. 2018. Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal Of The Linnean Society* 184: 528–83.

Weitzman, M.D., C.E. Lilley, and M.S. Chaurushiya. 2010. Genomes in conflict: maintaining genome integrity during virus infection. *Annual Review Of Microbiology* 64: 61–81.

Weitzman, M.D., and D.A. Ornelles. 2005. Inactivating intracellular antiviral responses during adenovirus infection. *Oncogene* 24: 7686–96.

BIBLIOGRAFÍA

Wensing, A.M., V. Calvez, H.F. Günthard, V.A. Johnson, R. Paredes, D. Pillay, R.W. Shafer, and D.D. Richman. 2016. 2017 Update of the drug resistance mutations in HIV-1. *Topics In Antiviral Medicine* 24: 132–33.

Wickham, T.J. 1994. Integrin alpha v beta 5 selectively promotes adenovirus mediated cell membrane permeabilization. *The Journal Of Cell Biology* 127: 257–64.

Wickham, T.J., P. Mathias, D.A. Cheresh, and G.R. Nemerow. 1993. Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell* 73: 309–19.

Wiethoff, C. M., H. Wodrich, L. Gerace, and G.R. Nemerow. 2005. adenovirus protein VI mediates membrane disruption following capsid disassembly. *Journal Of Virology* 79: 1992–2000.

Wilkinson, D.E., and S.K. Weller. 2004. Recruitment of cellular recombination and repair proteins to sites of herpes simplex virus type 1 DNA replication is dependent on the composition of viral proteins within prereplicative sites and correlates with the induction of the DNA damage response. *Journal Of Virology* 78: 4783–96.

Wilkinson, D.E., and S.K. Weller. 2006. Herpes simplex virus type I disrupts the ATR-dependent DNA-damage response during lytic infection. *Journal Of Cell Science* 119: 2695–2703.

BIBLIOGRAFÍA

Williams, M.J., B. Werner, C.P. Barnes, T.A. Graham, and A. Sottoriva. 2016. Identification of neutral tumor evolution across cancer types. *Nature Genetics* 48: 238–44.

Windheim, M., A. Hilgendorf, and H.G. Burgert. 2004. Immune evasion by adenovirus E3 proteins: exploitation of intracellular trafficking pathways. *Current Topics In Microbiology And Immunology* 273: 29–85.

Wohl, B.P., and P. Hearing. 2008. Role for the L1-52/55K protein in the serotype specificity of adenovirus DNA packaging. *Journal Of Virology* 82: 5089–92.

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

Wu, K., D. Orozco, and P. Hearing. 2012. The adenovirus L4-22K protein is multifunctional and is an integral component of crucial aspects of infection. *Journal Of Virology* 86: 10474–83.

Ying, B., A.E. Tollefson, and W.S.M. Wold. 2010. Identification of a previously unrecognized promoter that drives expression of the UXP transcription unit in the human adenovirus type 5 genome. *Journal Of Virology* 84: 11470–78.

BIBLIOGRAFÍA

Young, C.S.H. 2003. The structure and function of the adenovirus major late promoter. *Current Topics In Microbiology And Immunology* 272: 213–49.

Yuste, E., S. Sánchez-Palomino, C. Casado, E. Domingo, and C. López-Galíndez. 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *Journal Of Virology* 73: 2745–51.

Zhang, W., and M.J. Imperiale. 2000. Interaction of the adenovirus Iva2 protein with viral packaging sequences. *Journal Of Virology* 74: 2687–93.

Zhang, W., J.A. Low, J.B. Christensen, and M.J. Imperiale. 2001. Role for the adenovirus Iva2 protein in packaging of viral DNA. *Journal Of Virology* 75: 10446–54.

Zhao, H., M. Chen, and U. Pettersson. 2014. A new look at adenovirus splicing. *Virology* 456–457: 329–41.

Zijderveld, D.C., and P.C. van der Vliet. 1994. Helix-destabilizing properties of the adenovirus DNA-binding protein. *Journal Of Virology* 68: 1158–64.

Zou, L., and S.J. Elledge. 2003. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 300: 1542–48.

INFORMACIÓN SUPLEMENTARIA

Anexo 1

**SECUENCIAS UTILIZADAS PARA LLEVAR A CABO EL ANÁLISIS
DE DIVERSIDAD Y DE HUELLAS DE SELECCIÓN EN EL GÉNERO
*Mastadenovirus***

INFORMACIÓN SUPLEMENTARIA

Anexo 2

MUTACIONES IDENTIFICADAS MEDIANTE DS

Lista de las mutaciones identificadas mediante DS en cada uno de las tres replicas experimentales, incluyendo inserciones y deleciones. Se proporciona información sobre: posición genómica en la secuencia de referencia AY601635, base mutante, cobertura de secuenciación, numero de lecturas que presentan la mutación, unidad transcripcional , gen, posición en aa, codón mutado y sustitución de aa.

Enlace: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5100877/>

