



La emergencia viral como consecuencia de la interacción entre las variabilidades genéticas del virus y huésped

Julia Hillung

Programa de Postgrado en Biotecnología

**Directores: Prof. Santiago F. Elena Fito
Dr. José M. Cuevas Torrijos**



MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

INSTITUTO DE BIOLOGIA MOLECULAR Y CELULAR DE PLANTAS

Santiago F. Elena Fito, Doctor en Ciencias Biológicas y Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Molecular y Celular de Plantas (IBMCP), centro mixto del CSIC y de la Universidad Politécnica de Valencia y José M. Cuevas Torrijos, Investigador Doctor de la Universitat de València en el Instituto Cavanilles de Biodiversidad y Biología Evolutiva (ICBIBE),

CERTIFICAN:

que Doña Julia Hillung, Licenciada en Biología por la Universidad de *Gottfried Wilhelm Leibniz Universität Hannover*, ha realizado bajo su supervisión la Tesis Doctoral titulada "*La emergencia viral como consecuencia de la interacción entre las variabilidades genéticas del virus y huésped*".

Para que así conste, en cumplimiento de la legislación vigente, firman el presente certificado en Valencia a 17 de Junio de 2015.

Fdo. Santiago F. Elena Fito

Fdo. José M. Cuevas Torrijos

Agradecimientos / Danksagung

Un trabajo científico nunca es la labor de una única persona, por lo tanto ahora es el momento de agradecer a todos los que han hecho posible realizar el camino que me llevó a defender la Tesis.

En primer lugar, a las personas que me acogieron amablemente en el laboratorio, me proporcionaron un tema muy interesante, me ofrecieron su ayuda de manera constante y ayudaron al proceso de desarrollo del proyecto - mis directores Santiago Elena y José Cuevas. Vuestra inestimable experiencia, crítica constructiva, pero también vuestra comprensión y paciencia, contribuyeron considerablemente a mi positiva experiencia de postgrado.

Utilizo esta oportunidad para expresar mi gratitud también a todos los compañeros de laboratorio con los que me cruce a lo largo de estos años de proyecto: Susana, Mark, Stéphanie, Nicolas, Guillaume, Anouk, José Luis, Guillermo R., Guillermo B., Héctor, Fernando, Beilei, Rebeca y Silvia. Estoy sinceramente agradecida por tener la oportunidad de conocer vuestros puntos de vista sobre cuestiones relacionadas con el proyecto y otros, por la inspiración y por el simpático asesoramiento que me habéis proporcionado durante mi estancia en el laboratorio. Particularmente, quiero dar las gracias a Paqui, Àngels y Paula; sin vuestra ayuda, algunas partes técnicas del proyecto no habrían sido realizables. Habéis proporcionado las condiciones idóneas de trabajo tanto en laboratorio como en el invernadero. Agradezco también a Francisco García por el asesoramiento y el extenso análisis estadístico de los datos de micromatrices.

Quiero expresar mi agradecimiento a Antonio Mas, cuyas decisiones y esfuerzos, a nivel tanto académico como burocrático, me han facilitado venir a España y cuyo apoyo durante los estudios de Master me ayudó a iniciar mi vida laboral como investigadora.

Llegados a este punto, me gustaría dar las gracias a mi familia, por apoyarme incondicionalmente durante toda mi vida y por respetar mis decisiones.

Mil gracias a Toni, por acompañarme y apoyarme en la fase más difícil de la Tesis. A mis suegros - gracias por las comidas reconstituyentes de los domingos.

Mi agradecimiento también a mis amigos, todos habéis sido una fuente de inspiración y ánimos.

Finalmente, la realización de este trabajo no habría sido posible sin la financiación del Ministerio de Economía y Competitividad.

Eine wissenschaftliche Arbeit ist nie das Werk einer einzigen Person, somit ist es an der Zeit mich bei all denen zu bedanken, die mir die Realisierung der Doktorarbeit ermöglicht haben.

An erster Stelle bei den Menschen, die mich herzlich im Labor aufgenommen und mir ein interessantes Forschungsthema bereitgestellt haben - meinen Direktoren Santiago Elena und José Cuevas. Ihr habt mir im Entwicklungsprozess des Projekts konstanterweise eure Hilfe angeboten. Eure unschätzbare Erfahrung, konstruktive Kritik, aber auch euer Verständnis und eure Geduld haben zu meiner positiven Erfahrung als Doktorandin eine Menge beigetragen.

Darüber hinaus möchte diese Gelegenheit nutzen um meine Dankbarkeit auch gegenüber den Arbeitskollegen auszusprechen, die ich im Laufe des Jahres kennengelernt habe: Susana, Mark, Stephanié, Nicolás, Guillaume, Anouk, José Luis, Guillermo R., Guillermo B. Héctor, Fernando, Beilei, Rebeca und Silvia. Für die Gelegenheit eure Ansichten zu projektrelevanten und projektunrelevanten Fragen erfahren zu haben sowie für eure Inspiration und eure freundliche Beratung, die Ihr mir während meiner Zeit im Labor gegenüber gebracht habt, bin ich euch aufrichtig dankbar. Insbesondere möchte ich Paqui, Àngels und Paula danken: Ohne euch wären einige technische Abschnitte dieses Projektes nicht möglich gewesen. Ihr habt euch um die besten Arbeitsbedingungen im Labor und im Gewächshaus ausgezeichnet gekümmert. Mein Dank gilt auch Francisco Garcia für seine freundliche Beratung und umfangreiche statistische Datenanalyse von Microarrays.

Mein Einstieg in das Berufsleben als Forscherin in Spanien wäre nicht so reibungslos geworden ohne sowohl akademischen als auch bürokratischen Bemühungen von Antonio Mas. Ich danke dir dafür und auch für die Unterstützung während des Masterstudiums.

An dieser Stelle möchte ich mich bei meiner Familie für ihre bedingungslose Unterstützung und für das Respektieren meiner Entscheidungen bedanken.

Vielen Dank an Toni, der mich in der schwierigsten Phase der Dissertation stets begleitet und unterstützt hat. An meine Schwiegereltern - danke für die rekonstruktive Sonntagsmahlzeiten.

Ein herzliches Dankeschön auch an meine Freunde. Für mich wart ihr eine Quelle der Inspiration und Motivation.

Zum Schluss, die Realisierung dieser Arbeit wäre ohne die Finanzierung vom Spanischem Ministerium für Wirtschaft und Wettbewerbsfähigkeit nicht möglich gewesen, dafür auch vielen Dank.

ÍNDICE

AGRADECIMIENTOS / DANKSAGUNG	3
PRÓLOGO	9
RESUMEN.....	11
I INTRODUCCIÓN	13
1.1 Diversidad estructural y genética en virus	13
1.1.1 Potyvirus y estructura de la cápsida	14
1.1.2 Tamaño y organización del genoma de <i>Potyvirus</i>	14
1.1.3 Ciclo de vida de <i>Potyvirus</i>	17
1.1.4 Características particulares de TEV	23
1.1.5 Mecanismos de resistencia del huésped frente al virus	24
1.2 Herramientas de la evolución viral	26
1.2.1 Mutaciones adaptativas	27
1.2.2 Tiempo de generación	28
1.2.3 Tamaño de la población viral.....	29
1.2.5 Arquitectura de genomas	29
1.2.6 Estructuras secundarias de ARN	29
1.2.7 Factores ambientales.....	30
1.2.8 Generalistas y especialistas en la evolución viral.....	30
1.2.9 Eficacia viral	31
1.3 La evolución experimental.....	32
1.4 <i>Arabidopsis thaliana</i>.....	33
1.5 Micromatrices de ARN	33
OBJETIVOS	35
II MATERIALES Y MÉTODOS.....	37
2.1 Evolución experimental	37
2.1.1 Vector plasmídico	37
2.1.2 Población viral ancestral.....	37
2.1.3 Pases seriados en ecotipos de <i>A. thaliana</i>	38
2.2 Análisis fenotípico de los linajes virales evolucionados.....	38
2.2.1 Infección cruzada.	38
2.2.2 Extracción de ARN para detección de infección.....	39
2.2.3 Detección de ARN de TEV en plantas inoculadas	40

2.2.4 Cuantificación de la acumulación viral.....	40
2.2.5 Preparación de muestras para la secuenciación del genoma viral.....	41
2.2.5 Transcripción <i>in vitro</i>	42
2.2.6 Clonación por mutagénesis dirigida.....	43
2.2.7 Inoculación de plantas de cinco ecotipos de <i>A. thaliana</i> con los clones mutantes derivados de TEV-At17b.....	44
2.2.8 Medición de infectividad, virulencia, eficacia y epistasia.....	45
2.2.9 Análisis estadísticos.....	46
2.2.10 Estadísticas de la red de infección.....	47
2.3 Análisis transcriptómico de plantas infectadas con TEV.....	48
2.3.1 Elección de muestras para hibridación de micromatrices.....	48
2.3.2 Extracción de ARN para el ensayo de micromatrices.....	50
2.3.3 Etiquetado de ARN e hibridación de micromatrices.....	50
2.3.4 Análisis de datos de Micromatrices.....	51
III RESULTADOS Y DISCUSIÓN.....	53
3.1 Heterogeneidad en el perfil de interacciones virus-huésped.....	53
3.1.1 Estima de síntomas de infección, carga viral e infectividad de TEV- At17b.....	54
3.1.2 Comparación de los patrones de expresión genética entre los ecotipos infectados.....	57
3.1.3 Determinación de genes con expresión significativamente alterada.....	57
3.1.4 Diferencias y similitudes entre ecotipos en patrones de expresión.....	58
3.1.5 Análisis funcional de genes alterados.....	61
3.1.6 Discusión.....	65
3.2 Evolución experimental de TEV en huéspedes genéticamente heterogéneos.....	68
3.2.1 Medida de la adaptación a cada huésped local.....	69
3.2.2 La especificidad de la adaptación.....	72
3.2.3 Evolución de infectividad y virulencia.....	74
3.2.4 Análisis de redes de infección.....	75
3.2.5 Evolución genómica.....	76
3.2.6 Asociación entre diversidad molecular y variación genética para la eficacia.....	79
3.2.7 Selección para la eficiencia traduccional.....	80
3.2.8 Discusión.....	83
3.3 Evaluación de efectos de las mutaciones convergentes fijadas durante la adaptación del virus a diferentes ecotipos de huésped.....	90
3.3.1 Exploración de datos de eficacia dentro del huésped.....	91
3.3.2 El efecto de la eficacia dentro del huésped es mayor en huéspedes simpátricos que en alopátricos.....	94
3.3.3 Patrones de epistasia y robustez de los paisajes adaptativos.....	96
3.3.4 Discusión.....	99

3.4 Transcriptoma de interacción gen a gen del virus con sus huéspedes nuevos	101
3.4.1 Análisis de la adaptación local: diferencias en la expresión de genes entre linajes virales evolucionados en diferentes ecotipos	102
3.4.2 Esclarecimiento de dianas específicas de ecotipo y generales de la adaptación viral	108
3.4.3 Efecto de la evolución en los ecotipos alternativos en el rendimiento en el ecotipo original Ler-0	119
3.4.4 Evaluación de las diferencias en la expresión génica de linajes generalistas y especialistas	124
3.4.5 Discusión	134
IV DISCUSIÓN GENERAL	137
V CONCLUSIONES FINALES.....	145
REFERENCIAS	147

Prólogo

Los virus emergentes están reconocidos como una amenaza no sólo para la salud humana, sino también para actividades como la agricultura, ganadería e incluso para las especies en peligro de extinción. Los virus tienen una historia de millones de años, pero aun así sabemos poco acerca de ellos. Principalmente estamos al tanto de las enfermedades que causan, pero queda un campo de investigación enorme para entender la compleja interacción virus-huésped y poder predecir la emergencia viral, entendida ésta como la aparición de nuevos virus o de nuevas variantes de virus conocidos con propiedades distintas. Asimismo, una perspectiva interesante y necesaria para estudiar los virus es a través de donde prosperan, a través de su huésped. La diversidad genética del huésped para la susceptibilidad a la infección asume un papel clave en la aparición, propagación y prevalencia de las enfermedades infecciosas. También hay que destacar que la diversidad biológica de los virus es más grande que la de cualquier otro tipo de patógeno conocido, lo que constituye posiblemente una de las causas más importantes de la existencia de los virus en todos los ambientes celulares. En particular, los virus de ARN son los que presentan una mayor diversidad genética a nivel poblacional. Esta característica, junto con tiempos de generación cortos y grandes tamaños poblacionales, ha convertido a estos patógenos en excelentes modelos para estudiar la evolución. Además, los experimentos con virus de ARN de plantas permiten estudiar varios aspectos de la compleja interacción virus-huésped en condiciones muy cercanas a las naturales, donde se dispone de múltiples organismos completos como huéspedes potenciales y aun así estando sujeto a mínimas consideraciones éticas y de bioseguridad, tan limitantes en estudios con virus animales.

Los resultados de este trabajo se dividen en cuatro capítulos. La primera parte del trabajo presentará la dependencia de la susceptibilidad a la infección por el aislado del *Virus del grabado del tabaco At-17b* (TEV-At17b) en diferentes ecotipos de *Arabidopsis thaliana*. En la segunda parte, se presentarán los resultados de la evolución experimental de TEV en cinco ecotipos genéticamente heterogéneos de *A. thaliana*. En el tercer capítulo se caracterizarán las mutaciones virales putativamente adaptativas resultantes de la evolución experimental. Por último, el cuarto capítulo se centrará en el análisis de las interacciones que pueden producirse durante la adaptación del virus a un nuevo huésped y definir el destino de la evolución. En general, este trabajo aborda la cuestión de cómo puede condicionar la heterogeneidad genética del huésped el destino evolutivo del virus.

Resumen

La variabilidad genética en las poblaciones de huéspedes para la susceptibilidad al patógeno es esencial para la propagación de un virus emergente. Los modelos matemáticos predicen que la tasa de difusión del patógeno se frena a medida que aumenta la frecuencia y la diversidad de alelos de resistencia en la población huésped^{148,230}. Sin embargo, las pruebas experimentales de esta hipótesis son escasas. Por ese motivo, este trabajo aborda la cuestión de cuánta heterogeneidad genética del huésped es necesaria para el cambio del destino evolutivo del virus. El modelo de patosistema experimental utilizado en este trabajo está compuesto por el *Virus del grabado del tabaco* (TEV), aislado At17b (TEV-At17b), previamente adaptado por Agudelo-Romero *et al.* (2008)⁵ al ecotipo Ler-0 de *A. thaliana*, y por diferentes ecotipos de esta planta.

En la primera parte de la Tesis, aplicando el método de hibridación de micromatrices de ARN, se caracterizó el transcriptoma de seis ecotipos de *A. thaliana* (Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1) tras ser infectados con el aislado TEV-At17b. El resultado de la infección se caracterizó fenotípicamente y también se determinó el patrón de expresión de genes alterados por la infección, así como las diferencias entre los distintos ecotipos. La respuesta era heterogénea, si bien los ecotipos se pudieron agrupar en dos grupos según su perfil transcriptómico.

El propósito de la segunda parte del proyecto era explorar el efecto de la variabilidad genética intraespecífica del huésped en la dinámica de adaptación del virus. Con este fin, se realizó por triplicado un experimento de evolución con TEV-At17b en cinco ecotipos diferentes de *A. thaliana* (Di-2, Ei-2, Ler-0, St-0 y Wt-1). Después de la fase de evolución, se inocularon todos los ecotipos con todos los linajes virales evolucionados en un experimento de infección cruzada y se evaluaron la infectividad, la eficacia y la virulencia de cada combinación virus-ecotipo. Los resultados obtenidos se utilizaron para construir una red de infecciones. Las poblaciones de TEV, evolucionadas en cada uno de los cinco ecotipos de *A. thaliana*, aumentaron su eficacia, virulencia e infectividad en cada huésped de una manera compatible con un modelo gen-a-gen de interacciones parásito-huésped: los ecotipos difíciles de infectar fueron infectados por linajes generalistas, mientras que los ecotipos más susceptibles fueron infectados por cualquier linaje evolucionado.

A continuación se caracterizaron los genomas completos de los virus evolucionados y se encontraron siete casos de mutaciones convergentes. En esta parte del proyecto, para obtener un mejor conocimiento de la base molecular de la interacción gen-a-gen, se generaron todas estas mutaciones de forma individual, así como combinaciones

específicas, y se probaron sus efectos en plantas de los cinco ecotipos. Además, se evaluó la topografía del paisaje adaptativo subyacente.

Usando el método de micromatrices de ARN, en el último capítulo de la Tesis se identificó la respuesta transcriptómica de cada ecotipo a la infección del virus evolucionado en simpatria. Los ecotipos del huésped se clasificaron en grupos de acuerdo a las similitudes en sus perfiles de expresión génica. Se evaluaron las categorías funcionales de los genes afectados por la infección y se determinaron las intersecciones y diferencias entre los ecotipos. A continuación, se caracterizó la respuesta transcriptómica del huésped ancestral *Ler-0* a la infección con cada uno de los linajes evolucionados, cada uno adaptado a un ecotipo local diferente. Los cambios en las interacciones virus-huésped a nivel de expresión génica en *Ler-0* eran homogéneos, pero mostraban cierta heterogeneidad en las categorías funcionales enriquecidas. Por último, se analizó la respuesta transcriptómica inducida por la cepa viral más especializada y la más generalista en los cinco ecotipos. Se encontró que el virus especialista inducía una respuesta más fuerte y más heterogénea en todos los ecotipos, mientras que el virus generalista alteraba la expresión génica de manera similar en todos los ecotipos del huésped. Por tanto, la estrategia adaptativa de un virus especialista parecía ser específica para su ecotipo simpátrico, pagando un coste de eficacia en su interacción con los huéspedes alopátricos. Por el contrario, la estrategia adaptativa del virus generalista consistía en la alteración de vías comunes para todos los ecotipos.

I Introducción

1.1 Diversidad estructural y genética en virus

Los virus son unidades infecciosas obligatoriamente parasíticas únicamente capaces de replicar utilizando la maquinaria celular de un organismo y desarrollando la mayor parte de su ciclo vital dentro de la célula huésped. Los genomas virales carecen de una maquinaria propia para biosíntesis de proteínas, ni tampoco codifican procesos metabólicos de producción de energía.

Existen patógenos virales para todo tipo de organismos, bacterias, arqueas, y células de hongos, plantas y animales. No es habitual que un mismo virus infecte huéspedes en diferentes reinos biológicos, siendo la excepción algunos virus de plantas que replican en los insectos vectores mientras que son transmitidos a otras plantas^{83,146}.

En su composición más sencilla, las partículas virales, o viriones, son material genético con una envoltura proteica que protege el material genético y puede servir también de vehículo de transmisión. La envoltura proteica, también llamada cápsida, puede tener diferente complejidad. En algunos virus, se compone de unidades multiméricas de un solo polipéptido, mientras que en otros las cápsidas están compuestas por complejos heteroméricos. Las proteínas de cápsida pueden formar subunidades discretas o componentes estructurales morfológicamente distinguibles, denominadas capsómeros.

La composición y estructura de los genomas virales presenta una enorme variabilidad⁹. El material genético que lleva un virus puede ser ADN o ARN, monocatenario o bicatenario, lineal, circular o segmentado. Los genomas virales de cadena simple pueden ser de polaridad positiva (es decir, de la misma polaridad que el ARN mensajero (ARNm)), de polaridad negativa, o de ambas polaridades (una mezcla de las dos; ambisentidos)⁹.

La gran mayoría de los virus conocidos actualmente presentan ARN como material genético. Debido a su enorme dispersión, los virus de ARN son considerados los patógenos más importantes, tanto en animales como en plantas. El tipo de genoma que lleva un virus es determinante para sus procesos de expresión de genes y para la replicación del genoma. Esta Introducción se va a centrar en los virus de ARN monocatenario de polaridad positiva, dado que es la estructura correspondiente a nuestro modelo experimental.

Existen tres órdenes y 33 familias distintas conocidas de virus de ARN monocatenario de polaridad positiva, sin tener en cuenta las familias y especies todavía no clasificadas. Lo que todos tienen en común es que usan su genoma como ARNm para sintetizar una o varias proteínas, que en el caso de tener la forma de poliproteínas serán procesadas por

proteasas virales o celulares en componentes únicos. Este tipo de virus codifican en su genoma para una polimerasa de ARN dependiente de ARN, enzima implicada en la síntesis de las cadenas negativas que aparecen como productos intermediarios y en la transcripción de éstas que da lugar a nuevas moléculas de ARN genómico⁹.

1.1.1 Potyvirus y estructura de la cápsida

Los potyvirus pertenecen al género *Potyvirus*, que según el Comité Internacional de Taxonomía de Virus (ICTV), consta de 146 especies conocidas y representa el género más grande dentro de la familia *Potyviridae*. El género *Potyvirus* es también el segundo más extenso de virus de plantas. Muchos potyvirus son importantes patógenos de cultivos agrícolas, ya que infectan una amplia gama de especies de plantas mono- y dicotiledóneas y se han encontrado en todas las partes del mundo. Las infecciones por potyvirus son difíciles de controlar debido a su escasa persistencia, pero eficaz transmisión a través de más de 25 especies de áfidos^{98,121,203}. Además, algunos de las especies también se transmiten por semillas²²².

Una de las características fenotípicas distintivas de los potyvirus es la inducción de formación de cuerpos de inclusión^{165,223,250}. Estas estructuras a menudo tienen forma de molinillos con aspas o de espirales que se forman en el citoplasma de las células durante del ciclo de infección, pero muchos potyvirus también inducen cuerpos de inclusión citoplasmáticos amorfos y algunos forman inclusiones nucleares²⁰³.

Los potyvirus constituyen un grupo de virus de plantas caracterizados por presentar habitualmente partículas sin envoltura filamentosas y flexibles, cuyo tamaño varía entre 680-900 nm de largo y entre 11-15 nm de ancho⁶⁹. Cann propone que la flexibilidad es probablemente un atributo importante, porque la capacidad de plegamiento reduce la probabilidad de rotura o daño del virión⁹.

El virión está formado por unas 2000 unidades de proteína que rodean una molécula de ARN de aproximadamente 10 kb de tamaño. Fraenkel-Conrat *et al.*⁸⁸ demostraron en 1957 que se podían formar espontáneamente partículas virales a partir de subunidades purificadas de proteína de cubierta y de ARN sin ninguna influencia adicional, lo cual se explicaría asumiendo que la partícula estuviera en un estado de mínima energía libre, y por lo tanto, constituyera una estructura favorecida a partir de los componentes^{88,89}. Esta estabilidad es una característica importante de la partícula viral en los potyvirus.

1.1.2 Tamaño y organización del genoma de *Potyvirus*

El genoma viral, compuesto por ARN, está caracterizado por su tamaño extremadamente pequeño. En todos los virus de ARN conocidos su rango de tamaño oscila entre las 2,3 kb que presentan miembros de la familia *Narnaviridae* hasta las 30 kb de los miembros

de la familia *Coronaviridae*, si bien el tamaño medio oscila alrededor de 10 kb⁷⁵, mientras que los genomas de virus de ADN pueden ser hasta 1000 veces más grandes. Varios factores se han sugerido como limitantes para el tamaño del genoma. Las cadenas largas de ARN son frágiles, lo que explicaría el menor tamaño de los virus que poseen este tipo de genoma⁹. En el caso particular de los virus de ARN de plantas también existe la hipótesis de que los plasmodesmos, a través de los cuales el virus se mueve de célula a célula, pueden limitar el tamaño del genoma^{9,208}. Además, un genoma pequeño se replica más rápido que un genoma largo, lo que tiene importantes repercusiones en el tamaño poblacional del virus. Además, el tamaño del genoma viral puede estar condicionado por el hecho de que los virus de ARN presentan una replicación intrínsecamente propensa a errores. En este sentido, genomas de mayor tamaño presentarían mayor carga mutacional, donde la mayoría de las mutaciones tendrían un carácter deletéreo, lo cual supondría un impacto potencialmente negativo en la evolución del virus¹¹⁵. En comparación, los virus de ADN bicatenario se replican con mayor fidelidad y tienen tasas de mutación mucho más bajas, lo que les habría permitido alcanzar mayores tamaños genómicos. Sin embargo, existe al menos una excepción dentro de los virus de ARN, que viene del mundo de los virus de animales y tiene el genoma más grande conocido (26-32 kb); es el caso de los nidovirales. Estos virus tienen un conjunto de enzimas vinculadas a su replicasa y que se encuentran implicadas en mecanismos de reparación, lo que sugiere que estos virus son capaces de reducir la tasa de mutación y así evitar el mayor inconveniente que supone presentar genomas de mayor tamaño^{100,166}. En cualquier caso, a pesar del desafío que le supone al virus la reducción de su tamaño genómico, los virus han conseguido maximizar su capacidad genética mediante la codificación de genes de forma solapada. El ARN genómico de los potyvirus contiene una sola pauta abierta de lectura (ORF) que codifica para una poliproteína, la cual es procesada proteolíticamente mediante proteasas codificadas por el propio virus (Figura 1). La región N-terminal de la proteína codifica las proteínas virales maduras P1 y HC-Pro, y es variable entre distintos potyvirus y otros miembros de la familia *Potyviridae*. Las regiones centrales y C-terminales de la poliproteína tienen una arquitectura conservada y codifican las proteínas virales maduras P3, 6K1, CI, 6K2, VPg, NIaPro, NIb y CP. El procesamiento de la poliproteína precursora de las proteínas individuales se produce por medio de las proteasas virales NIaPro, HC-Pro y P1², como indican las flechas en la Figura 1. Además de la ORF larga, en los potyvirus existe una ORF corta, que codifica para la proteína P3N-PIPO, solapada dentro de la región que codifica P3, pero que se traduce en una pauta de lectura +2 en relación a P3. La expresión de P3N-PIPO parece tener lugar como un producto de fusión de la parte N-terminal de P3 (ca. 25 kDa) con un motivo altamente conservado en el extremo 5' de PIPO. Esta fusión parece surgir a través del desplazamiento de la pauta de lectura

ribosomal o por deslizamiento transcripcional⁴⁴. La longitud de PIPO es bastante variable entre las diferentes especies de potyvirus, variando entre 60 y 115 aminoácidos⁴⁴.

Los extremos 5' terminal y 3' terminal del genoma de potyvirus están flanqueados por regiones no codificantes. Dentro del extremo 5' no codificante existe una estructura secundaria de ARN que funciona como sitio interno de entrada del ribosoma (IRES)³⁴. Además el genoma de los potyvirus está asociado con la proteína VPg en el extremo 5' terminal²²⁴ y con una cola de poliadenina de tamaño variable en el extremo 3' terminal^{136,229}.

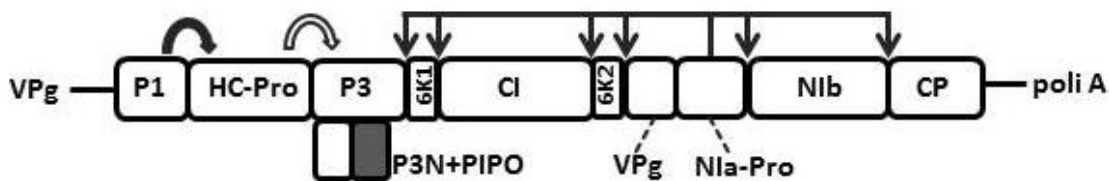


Figura 1. Esquema de la organización genómica de TEV, como representante del género *Potyvirus*. Los cuadros representan proteínas maduras individuales que se obtienen tras el procesado de la poliproteína por las proteinasas virales N1aPro, P1 y HC-Pro, tal y como indican las flechas. Las cajas inferiores a la pauta de lectura grande muestran la pauta de lectura alternativa que da lugar a la proteína P3N-PIPO. (Figura adaptada de Revers *et al.*⁹¹).

Las proteínas de los potyvirus son multifuncionales. Una de las proteínas con más funciones conocidas es HC-Pro, que además de ser la proteinasa encargada de la auto-escisión del extremo N-terminal³⁵, también está involucrada en un número de procesos infecciosos tales como la transmisión por áfidos¹⁰¹, movimientos de célula a célula²⁰⁷ y a larga distancia²¹⁰, supresión del silenciamiento génico¹⁴⁹ y desarrollo de síntomas¹⁹⁷. La proteína de la cubierta de los potyvirus (CP) también está implicada en multitud de interacciones con factores virales, con el huésped y con el vector de transmisión²⁹.

La Tabla 1 resume las funciones descritas para cada una de las proteínas virales, si bien las funciones más relevantes se describirán con detalle en los siguientes apartados de esta Introducción.

Tabla 1. Funciones de las proteínas de potyvirus.

Proteína	Funciones
P1	<ul style="list-style-type: none"> ▪ Serina proteinasa ▪ Factor accesorio para la amplificación del virus ▪ Adaptación al huésped
HC-Pro	<ul style="list-style-type: none"> ▪ Cisteína proteinasa ▪ Factor ayudante para la transmisión por áfidos ▪ Supresión del silenciamiento del ARN ▪ Aumento de cantidad de partículas virales
P3	<ul style="list-style-type: none"> ▪ Amplificación del virus ▪ Adaptación al huésped
P3N-PIPO	<ul style="list-style-type: none"> ▪ Movimiento célula a célula del virus
6K1	<ul style="list-style-type: none"> ▪ Posible modulación de la actividad de P3
CI	<ul style="list-style-type: none"> ▪ Helicasa del ARN ▪ Replicación del ARN ▪ Formación de estructuras de molinillo con aspas ▪ Movimiento del virus célula a célula
6K2	<ul style="list-style-type: none"> ▪ Proliferación de vesículas de membrana ▪ Transporte hacia la membrana
VPg	<ul style="list-style-type: none"> ▪ Proteína ligada al extremo 5' del genoma viral ▪ Cebador de replicación viral ▪ Traducción del ARN ▪ Movimiento del virus
NlaPro	<ul style="list-style-type: none"> ▪ Cisteína proteinasa ▪ ADNasa
Nlb	<ul style="list-style-type: none"> ▪ Polimerasa de ARN dependiente de molde de ARN (replicasa viral)
CP	<ul style="list-style-type: none"> ▪ Protección del ARN genómico ▪ Movimiento del virus ▪ Transmisión por áfidos

1.1.3 Ciclo de vida de *Potyvirus*

1.1.3.1 Transmisión viral

Al contrario de los animales, las plantas no disponen de receptores en la superficie de la célula que permitan la entrada del virus, ni tienen una membrana celular descubierta que permita la fusión de los virus para lograr la inclusión en la célula huésped. Un desafío especial a la hora de la transmisión para un virus de plantas es la gruesa y sólida pared celular vegetal. Éste sería el motivo por el cual los virus de plantas han evolucionado hacia el uso de vectores de transmisión, que perforan para ellos la pared celular y además les sirven de vehículo. Los vectores más comunes son artrópodos, pero nematodos, hongos e incluso bacterias también pueden transmitir virus²⁴⁷. En el caso de

los potyvirus, sus vectores son artrópodos que durante la alimentación de la planta infectada adquieren de forma pasiva el virus y lo transfieren a una planta sana, su próxima fuente de alimentación. La transmisión es un proceso específico para cada virus, es decir, un virus en particular puede ser transmitido por una sola especie o un género particular de vector y no por otros. Además, la interacción entre un virus y su vector específico es muy variable: algunos virus tan solo se enganchan en las partes bucales del vector, mientras que otros llegan a multiplicarse en las células de sus vectores^{83,146}. Hay unas pocas excepciones, como el caso de los tobamovirus, cuyo modo de transmisión consiste en el contacto mecánico pasivo de planta a planta o por semillas, sin implicar ningún vector en especial. A pesar de que algunos potyvirus se transmiten por el polen o por las semillas, éste no es el caso de TEV, que sirve de modelo en este trabajo.

1.1.3.2 Multiplicación viral

El primer paso de la replicación viral es la entrada del virión en una célula del huésped. Una vez el virión penetra en el interior, enseguida puede empezar la traducción del ARN en la poliproteína. Tras el procesado de la poliproteína en proteínas maduras, se inicia el proceso de multiplicación: primero NIb genera la cadena negativa a partir del ARN genómico y después estos anti-genomas se usan como molde para que NIb sintetice la cadena positiva, la que a su vez puede utilizarse como ARNm, implicarse en nuevos pasos de replicación o bien ser encapsidada formando nuevos viriones.

En cuanto a la localización de la multiplicación viral, muchos trabajos demuestran que, al igual que en el caso de los virus animales^{60,104,137}, los potyvirus también se replican en vesículas producidas por remodelación de endomembranas de la célula huésped²¹⁸. Estudios con varios potyvirus confirman que la proteína 6K2, que contiene un dominio hidrófobo central²⁰⁰, se asocia con VPg y NIaPro en las membranas derivadas del retículo endoplasmático^{144,219} e induce la formación de vesículas citoplasmáticas que se distribuyen por todo el sistema de membrana del retículo endoplasmático (RE) cortical y perinuclear^{17,18}. Estas vesículas constituyen los sitios de replicación viral⁵³. Las vesículas se forman acumulando la proteína 6K2 en los sitios de salida del RE (ERES) de manera COPI y COPII dependiente, lo que sugiere que la biogénesis de vesículas depende del transporte retrogrado y anterógrado entre el retículo endoplasmático y el aparato de Golgi²⁵². La colocación de 6K2 con los cuerpos de Golgi¹⁴⁵ apoya esta teoría. De esta manera, las vesículas inducidas por 6K2 se mueven por el citoplasma usando los microfilamentos de actina sin que los microtúbulos estén implicados en el proceso^{53,252}, dirigiéndose a los cloroplastos, donde se fusionan con la membrana e inducen invaginaciones en la misma²⁵¹. La presencia de ARN y de intermediarios replicativos de ARN en estas vesículas asociadas a cloroplastos^{96,108,160} indica que son sitios de

replicación viral, mientras que el RE es el sitio donde los potyvirus inician la traducción del genoma y donde se forman las vesículas 6K2²⁵¹. Se ha demostrado también que los cloroplastos asociados a las vesículas inducidas por 6K2 están presentes en una larga estructura globular perinuclear, que contiene RE compacto, aparato de Golgi y coaptámeros COPII¹⁰³. Para resumir las conclusiones de estos estudios, la replicación parece tener lugar en cloroplastos incluidos dentro de la estructura globular perinuclear y después las vesículas 6K2 brotan de los ERES y se mueven hacia la membrana plasmática y plasmodesmos para la transmisión del virus a las células vecinas¹⁰³. Sin embargo, no se descarta la posibilidad de que la replicación también tenga lugar en la estructura perinuclear fuera de los cloroplastos o en los cloroplastos fuera de la estructura perinuclear.

No sólo las proteínas 6K2, VPg y NIaPro están implicadas de forma esencial en la multiplicación de los potyvirus, sino que otras proteínas virales también lo están. Se ha demostrado que las proteínas 6K2, NIa, NIaPro y NIb también se localizan en las vesículas replicativas^{17,53,56,73,161}. La replicasa NIb, forma el núcleo del complejo replicativo¹¹⁶. Además, la helicasa CI, y las proteínas VPg y NIa están ligadas al extremo 5' del ARN viral interactuando con NIb^{82,171}, mientras que la proteína 6K2 sirve como anclaje a la membrana^{200,218}. Las proteínas P3 y 6K1 también son imprescindibles en la replicación viral^{128,132,161}, aunque su papel no está muy bien determinado. La proteína P3 se encuentra asociada tanto con RE como con las vesículas replicativas^{56,76}, y al igual que 6K1, interacciona con varias proteínas virales²⁶⁹. La proteína HC-Pro tiene la función de silenciamiento de la actividad supresora de ARN de la planta. De esta forma, HC-Pro protege los intermediarios replicativos y las hebras genómicas de ARN viral cuando no llevan cápsida^{27,126}. Sin embargo, hay indicios que provienen de los estudios de interacciones de HC-Pro con otras proteínas virales y de planta que demuestran que HC-Pro puede estar involucrada de forma directa en la multiplicación de potyvirus⁸. P1 no es esencial para la multiplicación de potyvirus, pero actúa como un factor accesorio para la amplificación del genoma²⁴⁵. CP y P3N-PIPO no parecen desempeñar un papel en este proceso¹⁵⁵, a pesar de lo cual se ha demostrado en TEV que para completar el proceso de multiplicación es necesaria la traducción hasta el codón 138 de CP y la colocación de un elemento de ARN que actúa en *cis* entre los codones 211 y 246¹⁵⁵.

Aparte de los factores virales y celulares involucrados en la replicación del ARN, en las vesículas replicativas también están presentes los factores celulares necesarios para la biosíntesis e incluso para el procesamiento de proteínas, lo que indica que la traducción del ARN viral y su replicación pueden estar acoplados^{53,111,120}. Por otro lado, es posible que los factores de traducción de plantas incluidos en vesículas, aparte de su papel en la traducción del ARN viral, estén directamente involucrados en la síntesis del ARN genómico viral, ya que algunos de ellos interactúan con la replicasa viral NIb.

Uno de los factores celulares que interacciona con proteínas virales y participa de forma directa en la multiplicación del virus es el factor de iniciación de la traducción eIF4E y sus isoformas. Este factor celular interacciona con VPg^{41,143,144} y con 6K2, VPg y NIaPro¹⁷. Varios estudios *in vitro* demostraron que la polimerasa viral NIb interacciona con la proteína de unión a poli(A) 2 (PABP2, otro factor de traducción), con la proteína de choque térmico 70 (Hsc/HSP70-3, una chaperona celular) y con el factor de elongación de traducción eEF1A^{17,73,239}. Las plantas de *Nicotiana benthamiana* y de *A. thaliana* que presentan niveles alterados de expresión de HSP70-15 han sido menos susceptibles a la infección con varios potyvirus, debido a la inhibición de la traducción del ARN viral^{111,124}. Experimentos *in vitro* también demostraron que tanto VPg, NIaPro como NIb interaccionan con la proteína PABP8, que se encuentra en *A. thaliana*; además, VPg y NIaPro interaccionan también con PABP4⁷³. En estudios con mutantes se confirmó que PABP es un factor importante para la acumulación del virus.

Huang *et al.*¹¹⁷ demostraron que una proteína parecida a la ARN helicasa de melocotonero y de *Arabidopsis* (AtRH8), que está relacionada con eIF4A, interacciona con la proteína VPg del virus de la sharka (PPV) en vesículas 6K2. Los mutantes que no expresaban esta proteína eran resistentes por completo a la infección con PPV y TuMV, demostrando que esta ARN helicasa tiene un papel crucial en la multiplicación de potyvirus¹¹⁷. Recientemente, se demostró en TuMV que las proteínas de la familia SNARE Syp 71 y la proteína similar a SNARE Vap27-1 son reclutadas a las estructuras tubulares inducidas por 6K2²⁵³, donde la proteína 6K2 interacciona con Vap27-1 e induce su unión a Syp71, posiblemente sirviendo de enlazador entre 6K2 y Syp71⁹¹.

Algunos factores relevantes para la infección con potyvirus, como DBP1³¹, DIP2³⁷ y eIF4E^{107,164}, no parecen estar directamente involucrados en la replicación, pero sí tener un papel importante en la regulación de la interacción del virus con el huésped.

Para la regulación de procesos de traducción y replicación que parecen estar ligados funcionalmente y localmente, se propusieron varios factores implicados. Análisis transcriptómicos revelaron que la expresión de las proteínas ribosomales RPS2, RPS6, RPL7, RPL13 y RPL19 está afectada durante las infecciones con potyvirus^{57,266} y que estas proteínas son requeridas para la traducción de potyvirus. También parece importante la proteína P1, que se une específicamente a la subunidad ribosomal 60S y parece estimular la traducción de las proteínas virales¹⁵⁸. Al contrario, la expresión de la co-chaperona CPIP (proteína que interactúa con CP)¹¹⁴, asociada a la chaperona HSC70, induce la degradación de CP y así mejora la traducción viral a costa de la replicación. También se describió por Khan *et al.*¹³⁰ que la proteína viral VPg mejora la eficiencia de la traducción del ARN e inhibe la transcripción del ARN destinado a ser encapsidado. Sin embargo, el caso de regulación por la proteína VPg es más complejo y requiere varios factores celulares adicionales a los nombrados arriba. La mejora de la

traducción mediada por VPg es dependiente de la presencia de los IRES y de las interacciones o bien VPg/eIF4F o VPg/eIF(iso)4F¹³⁰. Se ha descrito que la proteína VPg está unida covalentemente al extremo 5' del ARN viral, imitando la estructura presente en el extremo 5' del ARN mensajero celular e interactuando con varios factores celulares implicados en la traducción²⁰². Sin embargo, la mejora de la traducción promovida por VPg no parece requerir la unión covalente al ARN viral, sino que esta unión únicamente sería necesaria para la ronda inicial de traducción de ARN⁹¹. Otra posibilidad que encuentra apoyo en varios estudios, en los que se demostró la interacción de VPg con la replicasa viral Nib^{82,269}, es que la proteína VPg covalentemente unida esté implicada en la síntesis de ARN y no en su traducción, ejerciendo de cebador^{11,172,193}.

1.1.3.3 Movimiento de *Potyvirus*

En el transporte del virus se distingue entre el transporte de célula a célula y el transporte a larga distancia. Para una propagación eficaz del virus en la planta, las partículas virales amplificadas en las vesículas de membrana primero tienen que moverse por plasmodesmos que representan la unión directa de los citoplasmas celulares. Después de cruzar las células del mesófilo, el haz vascular de las hojas, las células vasculares del parénquima y las células acompañantes, el virión finalmente entra en el tubo criboso. Una vez dentro, es cuando se inicia el transporte a larga distancia. A partir de aquí, los viriones son transportados por la savia del floema a lugares distantes, donde saldrán del haz vascular para iniciar nuevos sitios de infección y difundir de forma sistémica en toda la planta.

La movilidad no está restringida únicamente a los viriones, sino que también está presente en los complejos replicativos, los cuales se mueven a lo largo de microfilamentos de actina con la participación de la miosina XI-K^{53,56,103,172,252} y luego pasan entre distintas células a través del haz vascular¹⁰³⁻¹⁰⁵.

Las proteínas virales claves de movimiento son CP y CI^{36,62,63}. Rojas *et al.*²⁰⁷ demostraron que también HC-Pro puede estar implicada en el movimiento intracelular, aumentando el límite de tamaño de partículas que pueden pasar por los plasmodesmos. Del mismo modo, también se vio implicada la proteína VPg, que parece unirse a HC-Pro¹⁰⁹. Esta hipótesis encontró apoyo en varias pruebas de interacciones de las proteínas tanto *in vitro*²³⁶ como en levaduras¹⁴⁷ y plantas²⁶⁹.

Más recientemente, también se descubrió el papel crucial de la proteína P3N-PIPO en el movimiento del virus. Esta proteína interactúa de forma directa con polímeros de CI y los dirige hacia los plasmodesmos utilizando la ruta secretoria. El complejo CI/P3N-PIPO, formado en las membranas del retículo endoplasmático de la estructura perinuclear inducida por el virus, se desplaza hacia los plasmodesmos a través del sistema secretor y se ancla a ellos gracias a la interacción del dominio PIPO con PCaP1²⁴⁶. PCaP1e es

una proteína de planta de unión catiónica e hidrófila que está anclada a la membrana plasmática a través de la miristoilación de un residuo de glicina¹⁷⁴. A continuación, más moléculas de CI se unen al complejo CI/P3N-PIPO/PCaP1 por auto-interacción y se forman estructuras cónicas. La proteína CP se desplaza, en conjunción con moléculas recién sintetizadas de ARN viral liberadas de las vesículas de replicación, a lo largo de microfilamentos de actina hacia el plasmodesmo mientras se forman viriones o nuevos complejos replicativos, uniéndose a estructuras cónicas de CI para moverse a través de los plasmodesmos a la célula vecina.

Además de PCaP1, otros dos factores del huésped que podrían tener un papel en el movimiento de célula a célula han sido identificados. Los dos interactúan con VPg. El primero es PVIP, que es parte de una pequeña familia de proteínas que contiene un dominio de regulación de expresión génica a través de modificación de histonas⁴⁹. En *Arabidopsis*, los genes PVIP2 y PVIP1 actúan como reguladores centrales en el desarrollo mediado por auxina²⁴⁰. Sin embargo, los factores celulares PVIP se localizan en el núcleo, y por lo tanto, es más probable que la interacción VPg/PVIP regule el transporte modificando la expresión de otros genes sin estar directamente involucrada en el transporte viral. El segundo factor que está putativamente implicado en el transporte, aunque su papel aún está por dilucidar, es la proteína celular eIF4E y su isoforma eIF(iso)4E. Su interacción con VPg está bien caracterizada para la multiplicación del virus y hay indicios de que, al menos en TEV, esta proteína está implicada en el transporte de complejos replicativos del virus de célula a célula⁴⁸.

El mecanismo de transporte vascular o de larga distancia es poco conocido. El movimiento a larga distancia implica varios pasos: (i) la entrada del virus en las células del floema (ii) la entrada en el haz vascular (iii) el transporte a lo largo del haz vascular y (iv) la salida del haz vascular uniéndose al transporte de los hidratos de carbono. Cada uno de estos pasos puede representar un cuello de botella para la propagación del virus, ya que el virus tiene que traspasar varias barreras para llegar a su destino. Se ha demostrado que el tamaño de la población viral que invade la parte de la planta hasta entonces no infectada depende de la concentración de virus en la savia o/y de las barreras que impone el huésped¹¹⁰.

No se sabe si el movimiento del virus a larga distancia se produce en forma de viriones o como complejo replicativo, ni tampoco se conocen factores celulares específicamente involucrados en el proceso, pero algunos factores necesarios para la propagación sistémica han sido caracterizados, así como los factores que la delimitan⁹¹. Las proteínas virales con alguna implicación en el movimiento a larga distancia son CP, VPg y 6K2. A pesar de la poca información disponible de factores celulares que promueven la propagación del virus, se han caracterizado varios factores de restricción del transporte a larga distancia. Los más descritos son los genes de restricción de

movimiento de TEV (RTM) identificados en *A. thaliana*¹⁵⁵. Se han descrito cinco genes dominantes de esta clase, denominados *RTM1*, *RTM2*, *RTM3*, *RTM4* y *RTM5*. Todos ellos están involucrados en la resistencia de la planta frente a la infección viral^{50,155,258}, si bien únicamente se han caracterizado en profundidad los tres primeros. La presencia de alelos dominantes en los tres loci caracterizados es necesaria para el bloqueo del movimiento sistémico de TEV, mientras que mutaciones recesivas homocigotas en cualquiera de los tres loci permite una infección sistémica^{42,43,258}. El papel de los otros dos loci *RTM4* y *RTM5* aún no ha sido comprobado experimentalmente.

RTM1 codifica una proteína lectina jacalina con similitud de secuencia a varias proteínas de unión a mirosinasa que participan en la respuesta de defensa⁴². *RTM2* codifica una proteína cuya región N-terminal es similar a pequeñas proteínas de choque térmico de plantas, mientras que la región C-terminal tiene un dominio transmembrana²⁵⁵. A su vez, *RTM3* codifica una proteína con una homología al dominio meprina y TRAF (MATH) en su región N-terminal y un dominio de hélice súperenrollado en su extremo C-terminal⁵¹. *RTM4* y *RTM5* solo han sido caracterizados genéticamente⁵⁰, pero aún no han sido clonados para realizar ensayos que apoyen su papel en la resistencia de la planta. Las proteínas *RTM1* y *RTM3* forman un complejo multimérico⁵². La interacción directa de este complejo con CP no se pudo demostrar, pero una mutación en CP aumenta la susceptibilidad de la planta⁵⁸.

Los mecanismos por los cuales estas proteínas restringen el movimiento de TEV a larga distancia aún no están claros. Se han propuesto varias hipótesis: (i) las partículas del virus o los complejos replicativos podrían ser secuestrados por el complejo RTM en el proceso de movimiento cuando están en el haz vascular, (ii) el complejo RTM podría reducir la accesibilidad del virus a factores o estructuras celulares requeridos para el movimiento a larga distancia o (iii) el complejo RTM podría activar una respuesta inmunológica en la planta como consecuencia de la infección por el virus que provocara restricción del movimiento.

1.1.4 Características particulares de TEV

TEV pertenece al género *Potyvirus* de la familia *Potyviridae* y a la superfamilia de los picornavirus. Su material genético consiste en ARN de polaridad positiva. El virus es transmisible de forma no persistente por más de 10 especies de pulgón¹²⁹, entre otros *Myzus persicae*, *Macrosiphum euphorbiae* y *Aphis fabae*¹²⁵. Su gama de huéspedes consta de más de 120 especies en 19 familias diferentes de plantas, entre los cuales sus huéspedes principales pertenecen a la familia de las *Solanaceae*. Los síntomas sistémicos típicos inducidos por TEV en sus huéspedes primarios son nervaduras más claras, grabado necrótico en las hojas, tallo corto y malformaciones de las hojas²²², pero la gravedad y extensión de los síntomas dependen de la cepa de virus, así como de las

especies de huéspedes infectados. La Figura 2 muestra síntomas de infección con TEV en tres huéspedes diferentes pertenecientes a la familia *Solanaceae* (*Nicotiana tabacum*, *N. benthamiana* y *Capsicum annuum*).

La existencia de TEV en el campo es muy común en América del Norte y del Sur. Se han detectado casos en Canadá, Estados Unidos (inclusive Hawái), México, Puerto Rico y Venezuela.

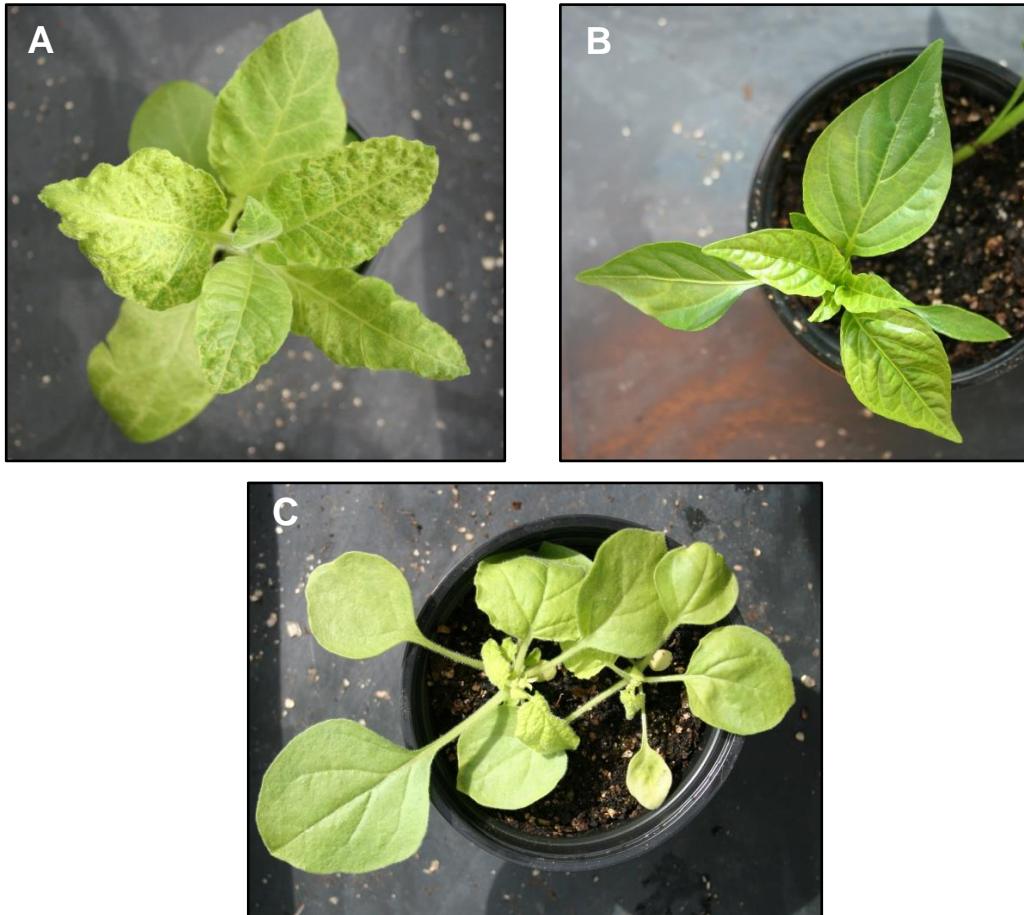


Figura 2. Síntomas en diferentes plantas huéspedes de TEV. **(A)** *N. tabacum* **(B)** *C. annuum* y **(C)** *N. benthamiana*

Lo que se sabe de la replicación de TEV es común a otros potyvirus y ha sido descrito en detalle anteriormente en el apartado 1.1.3.2. Como otros potyvirus, también TEV se ensambla espontáneamente formando una nucleocápsida filamentosa y flexible con una longitud de alrededor de 800 nm y una anchura de aproximadamente 13 nm. La capa proteica, formada por varios monómeros de la proteína CP y polímeros de la proteína CI, también es parte de la partícula viral^{95,183,194}.

1.1.5 Mecanismos de resistencia del huésped frente al virus

Los mecanismos moleculares por los cuales un virus puede dar lugar a progenie infecciosa dentro de una célula dada del huésped, donde el virus utiliza el metabolismo celular en su propio beneficio, a menudo implican un daño a la célula y, ocasionalmente,

incluso la muerte del organismo. En el contexto de la adaptación viral a su huésped, se está promoviendo una coevolución de virus y células, en la cual el huésped, tanto animal como vegetal, ha desarrollado mecanismos de defensa que ayudan a mantener el equilibrio en las interacciones virus-patógeno.

Existen dos modelos generales de defensa del huésped. El primero está especialmente adaptado para la resistencia en plantas. En este modelo, un genotipo de parásito puede infectar a todos los genotipos de huésped y un genotipo del huésped es universalmente susceptible⁸⁴. La resistencia ocurre cuando los agentes patógenos producen ciertas proteínas conocidas como factores de avirulencia (*Avr*). Estos factores de avirulencia son únicos para cada patógeno. Los productos de los genes *Avr* son identificados en la célula huésped por factores que son productos de genes de resistencia (*R*). Los genes *R* de la planta tienen que coincidir con los genes *Avr* del parásito para que la planta sea inmune al parásito intracelular²⁴³. En el caso de que no exista esta coincidencia, el patógeno puede colonizar el huésped y evitar su respuesta inmunitaria. Una respuesta típica de resistencia inducida en plantas es la reacción de hipersensibilidad (HR), que consiste en la muerte celular localizada y la consecuente detención de la propagación viral en el tejido vegetal. Flor fue el primero en demostrar que la herencia de ambos, de la resistencia del huésped y la capacidad del parásito para causar la enfermedad, está controlada por sus correspondientes parejas de genes, lo que acuñó como modelo de interacción gen-a-gen⁸⁴. El segundo modelo de resistencia es el de alelos coincidentes (MA). Este modelo es más conocido por los zoólogos de invertebrados y está basado en el modelo de respuesta inmunitaria al reconocimiento de lo propio y de lo extraño. La infección no es posible a menos que el parásito posea todos los alelos que coinciden con los del huésped⁹⁰. En definitiva, los dos modelos representan los extremos opuestos del mismo continuo, es decir, se requiere una correspondencia genética exacta entre el patógeno y el huésped para determinar el destino de la infección.

Los genes *R* de numerosos huéspedes han sido descritos para muchos virus de plantas¹⁵⁷. Sin embargo, no se ha descubierto todavía ningún gen específico de resistencia en los huéspedes naturales de TEV, si bien está muy bien documentado que la defensa antiviral está mediada por silenciamiento de ARN⁶¹. Estos mecanismos están basados en el reconocimiento específico de secuencias por pequeñas moléculas de ARN y operan para que los eucariotas puedan protegerse de los patógenos. Además de ser un componente crucial del sistema inmune innato contra los patógenos bacterianos y virales^{176,182}, el silenciamiento por ARN en plantas es uno de los mecanismos más importantes de regulación de genes. Muchos virus producen, en alguna fase de su ciclo de infección, intermediarios de ARN de doble cadena que inducen el silenciamiento de ARN. El ARN de doble cadena es reconocido y digerido por la enzima Dicer, que posee

dominios de ARNasa de tipo III capaces de cortar los ARNs de doble cadena para formar moléculas pequeñas de 21-24 nucleótidos de longitud. Estas pequeñas moléculas de ARN se denominan ARNs de interferencia o siARN. En la siguiente etapa, el siARN se une a un complejo con actividad nucleasa para formar el complejo RISC (complejo de silenciamiento inducido por ARN). La actividad helicasa de RISC separa las dos hebras del siARN, una de las cuales permanece unida al complejo y sirve como molde para reconocer otras hebras de ARN a degradar. El núcleo del complejo RISC está formado por proteínas pertenecientes a la familia argonauta (AGO), que promueven la degradación de ARNs virales, limitando consecuentemente la acumulación y propagación del virus dentro del organismo huésped^{233,237}.

Los virus han desarrollado estrategias para suprimir los efectos defensivos de silenciamiento de ARN en un huésped. Por ejemplo, la proteína HC-Pro interfiere en la propagación de la señal de silenciamiento a través de la planta²²⁷. Además, los virus interactúan con funciones celulares reguladas por pequeños ARNs endógenos, redirigiendo estas funciones de varias maneras^{40,74,151}.

En el caso particular de *A. thaliana*, empleada como modelo experimental en esta Tesis, cabe destacar que sus ecotipos varían en la susceptibilidad a la infección por TEV¹⁵⁵. Algunos ecotipos permiten el movimiento a larga distancia desde la hoja de roseta, en la que el virus se inocula, a los tejidos de la inflorescencia no inoculada. Por el contrario, otros ecotipos soportan la replicación en las hojas inoculadas, pero no permiten el movimiento sistémico¹⁵⁵. Los determinantes genéticos de la susceptibilidad de *A. thaliana* a la infección viral han sido estudiados a fondo³⁰. En concreto, el sistema de resistencia de *A. thaliana* mejor descrito, aún sin una completa comprensión de sus mecanismos, es el complejo multigénico RTM. Este complejo está compuesto por los loci RTM, como se ha descrito en detalle en la sección 1.1.3.3 de esta Introducción. La susceptibilidad de ciertos ecotipos a TEV es debida a mutaciones en cualquiera de los tres loci RTM²⁵⁸.

1.2 Herramientas de la evolución viral

El estudio de los virus de ARN presenta un enorme interés desde el punto de vista de la patología, pero al mismo tiempo, la capacidad evolutiva de los virus de ARN los convierte en un sistema idóneo para los estudios de evolución. Un virus de ARN de cadena positiva, una vez entra en el citoplasma del huésped, puede apropiarse de la maquinaria de la célula para iniciar la biosíntesis de sus proteínas y la replicación, sin tener la necesidad de transportar su genoma al núcleo de la célula huésped, y por tanto, sin correr el riesgo de ser degradado en el proceso por los mecanismos de defensa de la planta. Estos agentes infecciosos son únicos en su capacidad de almacenar la información genética de forma eficiente y productiva en una o varias moléculas de ARN.

Además, los virus de ARN tienen características distintivas que subyacen especialmente en este tipo de genoma y les permiten ser capaces de generar gran variabilidad genética, saltar las fronteras de las especies y propagarse en nuevos huéspedes. Estas características son los mecanismos básicos del cambio evolutivo: altas tasas de mutación y recombinación, grandes tamaños poblacionales y cortos tiempos de generación^{3,65,68,190}.

1.2.1 Mutaciones adaptativas

Las altas tasas de mutación de los virus de ARN dependen en gran parte de la polimerasa de ARN, que es muy propensa a errores. Por poner un ejemplo, la estima de la tasa de mutación de TEV es de $1,2 - 3 \times 10^{-5}$ sustituciones por nucleótido y por generación, una tasa representativa para la mayoría de los virus de ARN^{212,215,241}. En promedio, cada nueva secuencia contiene una mutación respecto a su ancestro. Además, la tasa de mutación viral también es dependiente del huésped en el que tiene lugar la replicación¹⁸⁹. La elevada capacidad de generación de mutantes de los virus de ARN y su amplia extensión entre diferentes huéspedes es especialmente importante para la supervivencia de virus cuyo huésped a menudo son plantas anuales. De hecho, la mayoría de los virus de plantas tienen este tipo de genoma.

Las mutaciones se pueden calificar comúnmente en relación a la eficacia como neutrales, beneficiosas o deletéreas. Las mutaciones neutrales se acumulan por deriva genética a velocidad constante y este proceso depende sólo del tamaño de la población y de la tasa de mutación¹³¹. Por otra parte, la fijación de mutaciones beneficiosas y la eliminación de mutaciones deletéreas dependen de la selección, y por lo tanto, del medio ambiente, además del tamaño de la población y de su heterogeneidad^{122,180}. En la mayoría de los casos, la transferencia del virus entre diferentes tipos de huéspedes no lleva directamente a una infección epidémica, sino que resulta en una infección puntual y limitada. Cuando el salto se produce a un huésped parcialmente accesible, que permite la replicación del virus de forma mínimamente eficiente, la adaptación del virus al nuevo huésped va a estar probablemente determinada en gran medida por procesos epistáticos y pleiotrópicos^{77,80,97,140}. La pleiotropía es un fenómeno en el cual genes únicos contribuyen a varias características fenotípicas y, en consecuencia, una mutación en un gen pleiotrópico puede tener un efecto sobre algunos o todos los rasgos de forma simultánea. En la evolución se habla de pleiotropía antagonista cuando la selección de un rasgo favorece la eficacia, mientras que la selección de otro rasgo codificado por el mismo alelo tiene el efecto contrario.

El efecto de las mutaciones sobre la eficacia será dependiente del contexto genético y ambiental donde se producen las mutaciones. Bajo condiciones de epistasia, el efecto de mutaciones dobles y múltiples no es la suma de las mutaciones únicas (efectos aditivos),

sino que en estos casos existe una interacción significativa entre las mutaciones, que puede ser antagonista (efecto menor a la asunción de aditividad) o sinergista (efecto mayor a la asunción de aditividad).

La selección que interviene sobre la evolución puede ser positiva o negativa. La selección positiva favorece la presencia de mutaciones adaptativas, mientras que la selección negativa elimina o tiende a la eliminación de las mutaciones deletéreas. La frecuencia de una mutación dada en la población viral también puede estar determinada por la deriva genética, que resulta de los cambios en el tamaño de la población que pueden conducir a la fijación estocástica de mutaciones, independientemente de su valor selectivo. El principal resultado de la deriva genética es una nueva composición aleatoria de los genomas en una población¹⁵⁶.

Si el virus no es capaz de replicar en el huésped nuevo, su proliferación va a depender de la preexistencia de mutantes virales que sean capaces de iniciar la replicación y recuperar la eficacia de la población viral. Este conjunto de mutantes, que forman un grupo de virus con genomas estrechamente relacionados y que compiten en un entorno altamente mutagénico, ha sido definido clásicamente con el término de cuasiespecies⁶⁴. Uno de los aspectos clave de esta definición es que la unidad de selección no es el virus individual, sino el conjunto de genotipos interconectados^{39,46,167,213}. Este enjambre de genotipos constituye una reserva de posibilidades evolutivas, exploradas previamente por el virus, y que pueden ser utilizadas de nuevo si un cambio ambiental lo requiere¹⁵⁶. La capacidad de tener diferentes variantes presentes en la población incrementa la probabilidad de que al menos una variante sea eficaz en las condiciones ambientales impuestas por el nuevo huésped. La propiedad de los virus de mantener diferentes variantes de mutantes relacionados en la población sin que afecten a su fenotipo se denomina robustez mutacional o plasticidad fenotípica. La robustez mutacional es considerada una característica esencial en la capacidad adaptativa de los virus, porque les permite hacer frente a las variaciones frecuentes en su entorno⁷⁸.

1.2.2 Tiempo de generación

Los tiempos de generación en virus de ARN están determinados por el tamaño pequeño de su genoma y por la velocidad de la replicasa. Un genoma pequeño se replica más rápido que un genoma grande y una polimerasa de baja fidelidad, como es el caso de la polimerasa del ARN dependiente de ARN, es más rápida que las polimerasas del ARN dependientes de ADN. Evidentemente, estas propiedades tienen importantes repercusiones en la tasa de mutación y en los tamaños poblacionales de virus. La baja fidelidad de la polimerasa de ARN debe de ser un rasgo adaptativo y resultado de un compromiso con la velocidad de replicación^{79,93}.

1.2.3 Tamaño de la población viral

Otro factor relevante en la evolución es el tamaño de la población. En virología, este parámetro se refiere a la población de individuos altamente relacionados que constituye la progenie de un genotipo inicial. Cuando las condiciones ambientales imponen tamaños poblacionales bajos, las mutaciones presentes en los genomas se transmiten a la mayoría de los virus de la progenie, lo que resulta en una mayor fijación de mutaciones en la secuencia consenso. La reducción en el tamaño de la población implica que los procesos de competencia y selección de los genotipos mejor adaptados tienen lugar entre un número reducido de genomas. Además, estas condiciones favorecen la acción de la deriva, y dado que la mayoría de las mutaciones tienen efectos deletéreos sobre la eficacia, el resultado de las reducciones repetidas en el tamaño poblacional es por lo general una disminución en la eficacia de la población. Por el contrario, grandes tamaños poblacionales permiten la competencia entre un gran número de genomas, aumentando la probabilidad de fijación de mutaciones ventajosas, lo que resulta en el aumento de eficacia¹⁵⁶.

1.2.5 Arquitectura de genomas

La arquitectura de los genomas virales de ARN es otra de sus características sustanciales. En el pequeño genoma de los potyvirus existen pautas de lectura solapantes⁴⁴. Los mecanismos que permiten la superposición de genes incluyen el desplazamiento de la pauta de lectura ribosomal, el uso de codones de inicio no AUG y el corte y empalme del ARN²¹. Las proteínas creadas por solapamiento de genes suelen ser proteínas accesorias que juegan un papel en la patogenicidad viral o en la propagación¹⁹⁶. Estas superposiciones permiten al virus la compresión del genoma, aumentando el repertorio de proteínas del virus sin aumentar la longitud del genoma^{15,44,221}.

1.2.6 Estructuras secundarias de ARN

Las estructuras secundarias de ARN presentan un papel fundamental en muchos procesos celulares diferentes, incluyendo transcripción, traducción, localización, corte y empalme del ARN, transporte, estabilidad y actividad catalítica. Los virus necesitan interactuar con la maquinaria de síntesis del huésped para completar su ciclo de replicación. En este contexto, los genomas virales contienen estructuras secundarias de ARN cruciales para las diversas etapas del ciclo de vida viral y la evasión de las defensas del huésped. Estas estructuras de ARN incluyen IRES²⁶³, señales de empaquetamiento y de corte y empalme de ARN, pseudonudos, estructuras secundarias parecidas a ARN de transferencia, motivos de pauta de lectura ribosomal y elementos reguladores cis⁹.

Teniendo en cuenta las estructuras secundarias del ARN, la selección de mutaciones sinónimas ya no puede considerarse como neutra^{191,226}. Los cambios sinónimos podrían dar lugar a una progenie no viable debido a potenciales cambios conformacionales dentro de las estructuras secundarias de ARN y que pueden influir en la patogénesis del virus, y en definitiva, en su viabilidad^{47,225}.

1.2.7 Factores ambientales

A pesar de que las características genómicas de los virus de ARN son determinantes en su evolución, hay que destacar que los virus emergentes son la consecuencia de la compleja interacción de las dinámicas genéticas del patógeno y del huésped^{77,187}, donde numerosos factores a menudo desconocidos son requisitos previos para el salto del virus al nuevo huésped. Algunos de estos factores pueden ser cambios en el medio ambiente y la posterior selección de los genotipos con mayor eficacia en las nuevas condiciones^{187,264}. En general, un aumento en la disponibilidad de posibles huéspedes, tanto en el tiempo como en el espacio, facilita enormemente la evolución y propagación de patógenos. La disponibilidad de gran cantidad de huéspedes, altamente relacionados filogenéticamente y presentes durante todo el año, aumenta la probabilidad de la emergencia de un virus^{59,150,187,264}. Sin embargo, también es posible que el salto a un huésped poco relacionado con el original no impida al virus una replicación eficaz, de tal manera que el virus pueda llegar a provocar una epidemia en la población una vez el virus consiga la extensión local en la nueva generación de huéspedes¹⁸⁷.

1.2.8 Generalistas y especialistas en la evolución viral

La adaptación viral a un nuevo huésped es un complejo proceso de cambio de interacciones entre el huésped y el patógeno. El huésped constituye el entorno de evolución del parásito intracelular. Éste es un entorno complejo y muy variable en el espacio y en el tiempo. La heterogeneidad del huésped es un factor determinante en la adaptación del virus y representa un factor de selección. Los virus varían en su rango de huéspedes y pueden clasificarse en base a este criterio en (*i*) especialistas, aquellos que infectan sólo una o unas pocas especies relacionadas de huéspedes y (*ii*) generalistas, que son aquellos que muestran una amplia gama de huéspedes donde son capaces de replicar⁷⁷. Uno de los objetivos más recurrentes en los estudios de evolución es el de determinar por qué algunos virus muestran fuertes tropismos de células y tejidos, siendo capaces de replicarse únicamente en tipos celulares muy limitados, mientras que otros evolucionan para poder infectar y replicarse en varios tipos de células diferentes.

La especialización sobre un solo huésped podría permitir a los virus una reducción de la competencia interespecífica y así conseguir altos valores de eficacia en este huésped. Sin embargo, si llegado el caso estos genotipos altamente especializados tuvieran que

cambiar de huésped, el coste en la eficacia que tendrían que pagar podría ser muy alto⁹⁴. Por otro lado, las ventajas del generalismo son más evidentes: un virus generalista sería capaz de beneficiarse de varios huéspedes, potenciando así su eficacia. A pesar de ello, se han descrito pocos casos de virus generalistas, lo que puede significar que el generalismo conlleva un coste potencialmente alto como consecuencia de la incapacidad para utilizar los recursos de un huésped dado de forma óptima²⁶⁰. Debido a esto, cabe pensar que la evolución tiende a favorecer estrategias especialistas, más si tenemos en cuenta que la evolución actúa más rápidamente en nichos estrechos^{260,265}.

A nivel molecular, nos encontramos con dos fenómenos que afectan al grado de adaptación a un huésped. El más directo y sencillo es la pleiotropía antagonista, definida como la situación en la que las mutaciones que son favorecidas por la selección por ser beneficiosas en un huésped resultan perjudiciales en otro⁹². El segundo mecanismo es el relativo a la acumulación de mutaciones neutrales o casi neutrales por efecto de la deriva en un huésped dado, pero que resultan claramente deletéreas en uno alternativo¹²⁷. Ambos mecanismos implican diferencias en los efectos de eficacia de mutaciones en los distintos huéspedes, pero no se trata de fenómenos equivalentes. Mientras que la selección natural es la única razón para el primer mecanismo, la deriva genética lo es para el segundo.

Como se ha comentado anteriormente, la existencia de variación genética en las poblaciones es esencial para la propagación de un virus emergente. De hecho, la homogeneidad en la población de huéspedes resulta por lo general en la especialización⁷¹ y conduce a la adaptación local del virus²⁶⁵.

1.2.9 Eficacia viral

La eficacia viral es un concepto complejo que originalmente se definió como la capacidad de un virus para producir una progenie infecciosa en un entorno dado⁶⁷. Más ampliamente, tal y como se ha utilizado también en este trabajo, esta definición se usa para referirse a la eficacia replicativa⁶⁶. La eficacia replicativa representa los efectos combinados de todas las demás propiedades fenotípicas necesarias para la supervivencia y reproducción de un determinado genotipo en un determinado ambiente¹⁴¹ e incluye múltiples componentes en el caso de virus de plantas, tales como la entrada celular, replicación, traducción, ensamblaje, encapsidación, liberación del virión de la célula huésped y capacidad de movimiento a corta y larga distancia, así como la resistencia a respuestas antivirales e incluso las tasas de transmisión. La mayoría de estos componentes de la eficacia implica la interacción con factores celulares del huésped⁷.

La eficacia viral replicativa no es un valor constante, sino que depende del entorno y de las perturbaciones externas que pueden modificar completamente el valor cuantitativo de la eficacia de un genotipo o de una población entera^{7,156}. Como se ha descrito anteriormente, una población viral es un enjambre dinámico y heterogéneo de genomas relacionados.

1.3 La evolución experimental

Las alteraciones más comunes para los virus que conllevan consecuencias evolutivas son las perturbaciones ambientales, los cambios en el tamaño de la población y la disponibilidad de huéspedes. En condiciones de laboratorio, estos factores se pueden aplicar de forma controlada a la evolución de las poblaciones virales. Al finalizar un experimento de evolución, se pueden determinar las características de los virus resultantes de la adaptación a determinadas condiciones, como por ejemplo la eficacia viral y las variantes moleculares adquiridas por los distintos linajes virales.

Cuando se pretende optimizar la eficacia en un nuevo huésped, se llevan a cabo pases seriados con un gran número de partículas virales de partida en cada pase. Los pases iniciados a partir de tamaños poblacionales grandes pueden implicar que muchas partículas virales infecten cada célula individual. Cuando esto sucede, los procesos de selección incluyen la interferencia y la competencia entre los genotipos virales dentro o fuera de la célula. La interferencia extracelular tiene más importancia para los virus de animales, que tienen que competir por los receptores celulares para poder iniciar la infección. En plantas no hay receptores específicos para la entrada del virus, y además, en condiciones experimentales tampoco existen interacciones selectivas con el vector de transmisión, ya que se emplea la inoculación mecánica. Por tanto, las interferencias extracelulares se limitarían a la selección de los genomas más estables fuera del huésped.

La interferencia intracelular de los genomas virales es causada por la competencia por las vías implicadas en la maquinaria de biosíntesis. Debido a la heterogeneidad de la población viral, se presentarán genotipos con distintos grados de viabilidad en su interacción con el huésped e incluso una fracción de la población será inevitablemente defectiva. En muchos casos, los virus defectivos pueden replicar con éxito dentro de la célula si existen formas capaces de suministrar las proteínas intactas y complementar las funciones ausentes en estos genotipos potencialmente inviables. Por lo general, la cantidad de genomas defectivos mantiene un equilibrio con la cantidad de genomas totalmente viables. Sin embargo, hay situaciones en las que este equilibrio se rompe, como por ejemplo cuando hay un aumento suficientemente grande en la tasa de mutación¹⁰². En esta situación, la cantidad de genomas defectivos puede exceder en gran medida a la de genomas capaces de codificar las proteínas correctas y provocar lo

que se conoce como catástrofe de error, conducente en último término a la extinción de la población viral¹⁰².

En condiciones experimentales, sólo unas pocas partículas virales consiguen entrar en la célula e iniciar una infección. Esta parte de la población se denomina población efectiva. Dado que la población efectiva suele ser considerablemente menor que la población potencialmente infecciosa, es habitual la observación de fenómenos de cuello de botella. Estas condiciones suelen conducir a una baja heterogeneidad inicial en la población viral, donde sólo los genotipos efectivos pueden sobrevivir, ya que las partículas defectivas aisladas no se verán favorecidas por procesos de complementación a nivel intracelular y serán eliminadas necesariamente²¹⁷.

1.4 *Arabidopsis thaliana*

A. thaliana es una planta anual, rara vez bianual, con un ciclo de vida inferior a cinco meses. Es una planta herbácea, que alcanza alturas de hasta 30 centímetros, donde las hojas basales forman una roseta y su tallo es redondo. Esta planta presenta flores hermafroditas blancas, de dos a cuatro milímetros de diámetro, principalmente en el período de abril a mayo. La planta también tiene silicuas con frutos, que pueden ser de 10 a 20 milímetros de largo.

En 1907, F. Laibach descubrió el número de cromosomas de *A. thaliana* ($2n = 10$) y destacó el potencial de esta planta para la experimentación genética, entre otras razones por la brevedad de su ciclo vital. En 1996, más de doscientos científicos crearon el proyecto de investigación AGI (Iniciativa del Genoma de *Arabidopsis*). Cuatro años más tarde, se presentó el primer mapa genético de la planta, con 25498 genes identificados que codifican 11000 familias de proteínas^{155,155,238,238}.

A. thaliana tiene muchas ventajas como planta modelo para la identificación de factores del huésped implicados en la infección viral. Una de ellas es la variabilidad genética, que se refleja en la existencia de diferentes ecotipos. Ecotipo es un término utilizado principalmente en botánica para designar una variante genotípica de una especie que se asocia con un tipo particular de hábitat⁹⁴. A pesar de no ser el huésped natural de TEV, las plantas de *Arabidopsis* disponen de un sistema de resistencia específico de TEV, denominado RTM (descrito en la sección 1.1.3.3), que es en parte responsable de la heterogeneidad en la susceptibilidad de los ecotipos de *A. thaliana* a la infección por TEV.

1.5 Micromatrices de ARN

Los virus de plantas son parásitos intracelulares obligados, que alteran normalmente la fisiología del huésped, desviando el uso normal de metabolitos y dirigiéndolo hacia la producción de componentes específicos del virus. Para ello, los componentes virales,

tanto proteínas como el ácido nucleico, deben establecer múltiples y complejas interacciones entre ellos y con un gran número de factores del huésped^{13,257}. Como respuesta a la infección viral, las células tratan de compensar los efectos causados por el virus regulando positiva o negativamente la expresión de algunos genes. Una forma de acceder a estudiar estas complejas interacciones entre el huésped y el patógeno es identificar el cambio en los patrones de expresión de genes en el huésped como consecuencia directa o indirecta de la infección viral. La llegada de las tecnologías de micromatrices de alto rendimiento ha hecho posible el estudio exhaustivo de las redes de expresión génica^{209,256,259,262}.

Objetivos

En esta Tesis Doctoral se plantearon los siguientes objetivos:

1. Caracterizar el transcriptoma de siete ecotipos de *A. thaliana* (Col-0, Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1) tras ser infectados con el aislado TEV-At17b.
2. Evolucionar experimentalmente el aislado del virus TEV-At17b en siete ecotipos diferentes de *A. thaliana*.
3. Caracterizar el fenotipo de infección con los virus evolucionados en los diferentes ecotipos.
4. Caracterizar los genomas de los linajes virales evolucionados en cada ecotipo.
5. Identificar la respuesta transcriptómica de cada ecotipo a la infección del virus evolucionado en simpatria.
6. Explorar el efecto de la variabilidad genética intraespecífica del huésped en la dinámica de adaptación del virus.
7. Analizar la respuesta transcriptómica inducida por la cepa viral evolucionada como más especializada y como más generalista en todos los ecotipos.

II Materiales y métodos

2.1 Evolución experimental

2.1.1 Vector plasmídico

Relevante para experimentos 3.1 –3.3

A lo largo de la presente Tesis Doctoral se ha utilizado el plásmido pMTEV²⁰, que contiene el ADN complementario de TEV (número de acceso en GenBank DQ986288) y donde se han introducido diversas mutaciones puntuales en función de las necesidades de cada experimento.

2.1.2 Población viral ancestral

Relevante para experimentos 3.1 y 3.2

En un experimento anterior realizado en nuestro laboratorio, se llevó a cabo una dinámica de adaptación del TEV al nuevo huésped *A. thaliana* ecotipo Ler-0⁵. Una planta infectada procedente del último pase de evolución de este experimento (pase 16) había sido homogeneizada en nitrógeno líquido con mortero y alicuotada convenientemente. Este material infeccioso fue utilizado como punto de partida en el primer experimento de esta Tesis (ver apartado 3.1) y para el inicio de la evolución experimental (ver apartado 3.2).

El material infeccioso procedente del experimento anterior⁵ fue amplificado con objeto de obtener el suficiente material para poder iniciar el primer experimento de esta Tesis. Para ello, se siguió el siguiente procedimiento:

1. Se resuspendió una alícuota de tejido infeccioso en tampón C (50 mM de ácido bórico, pH = 8,0, EDTA 5 mM; proporción 100 µL de tampón C/100 mg de tejido).
2. Tras centrifugación corta, que sedimentó el material sólido, se pasó el sobrenadante, que contenía los viriones, a otro tubo. A dicho tubo se le añadió un 10% de carborundum (100 mg/mL) para facilitar la entrada de los viriones durante el proceso de inoculación mecánica.
3. Se inocularon 50 plantas de *A. thaliana* ecotipo Ler-0 de 21 días de edad (4 µL de inóculo/hoja, 2 hojas/planta), para lo cual se frotaban las hojas con una varilla de vidrio ligeramente curvada.
4. Las plantas fueron mantenidas en un invernadero BSL-2 con ciclos de 16:8 horas de luz:oscuridad y 24:20 °C de temperatura día:noche hasta la recogida de muestras a los 21 días después de la infección (dpi).
5. Las plantas que presentaron síntomas fueron recogidas. Las hojas inoculadas y la raíz fueron eliminadas y el resto fue homogeneizado con mortero en nitrógeno

líquido. El tejido infeccioso obtenido fue alicuotado y almacenado a -80 °C hasta su uso.

La secuencia consenso completa del nuevo aislado fue obtenida y analizada como se describe abajo en el apartado de secuenciación (2.2.5). Esta secuencia era idéntica a la descrita anteriormente para TEV-At17⁵ con la única diferencia de una mutación adicional no sinónima (G6816A; M2224I) en el gen NIaPro. Debido a esta diferencia, el aislado obtenido se denominará en adelante TEV-At17b.

2.1.3 Pases seriados en ecotipos de *A. thaliana*

Relevante para experimento 3.2

A partir de la población ancestral (obtenida como se ha descrito en el anterior apartado II.2.1.2) se procedió a inocular cada uno de los siete ecotipos de *A. thaliana* (Col-0, Di-2, Ei-2, Ler-0, Oy-0, St-0 y WT-1). Para ello, se resuspendieron 2,5 g de material infeccioso en 5 mL de tampón C, añadiendo a continuación un 10 % de carborundum (100 mg/mL). Se inoculó un total de 50 plantas de 21 días de edad para cada uno de los siete ecotipos. Tal y como se describió en el apartado anterior 2.1.2, se inocularon dos hojas de cada planta (4 µL/hoja) por frotamiento con varilla de vidrio. Las plantas se mantuvieron en las condiciones descritas anteriormente. La infección de las plantas se confirmaba a partir de los 14 dpi mediante comprobación visual de los síntomas o bien utilizando la técnica de RT-PCR, como se describe en el apartado 2.2.3. Las plantas con infección confirmada se recogieron 21 dpi, se pulverizaron individualmente en mortero con nitrógeno líquido y se almacenaron a -80 °C. A partir de este primer pase, se eligieron al azar tres plantas infectadas (es decir 3 réplicas experimentales) de cada ecotipo para preparar nuevas mezclas infecciosas con las que continuar el siguiente pase de evolución. Para cada uno de los tres linajes/réplicas de cada ecotipo se inoculaban al menos diez plantas para tener mayor certeza de que alguna resultara infectada. Sin embargo, los pases se hacían a partir de una única planta infectada por linaje, es decir, sin juntar varias plantas para la infección del siguiente pase. Esta planta era elegida al azar y la infección confirmada por visualización de síntomas o por RT-PCR. De esta manera, la evolución experimental se continuó durante 15 pases. Al término de dicha evolución, se escogieron al azar cinco plantas infectadas de cada linaje, las cuales fueron homogeneizadas conjuntamente y almacenadas a -80 °C. Los pases seriados en Col-0 y Oy-0 no han sido posibles por la baja susceptibilidad de estos ecotipos al virus.

2.2 Análisis fenotípico de los linajes virales evolucionados

2.2.1 Infección cruzada.

Relevante para experimento 3.2

Las 15 cepas virales evolucionadas (3 réplicas biológicas × 5 ecotipos) se inocularon en su ecotipo de huésped local, así como en los cuatro ecotipos alternativos (Figura 3). Para este fin, las muestras de tejido infeccioso resultantes de la evolución experimental fueron cuantificadas y diluidas a la misma concentración, la más alta posible, para todos los linajes. A continuación, se inoculó una hoja de cada planta de *A. thaliana* de 21 días de edad. Para cada combinación ecotipo-linaje se inocularon al menos diez plantas. Las plantas control se inocularon sólo con tampón C y se mantuvieron en las mismas condiciones que las plantas inoculadas con viriones.

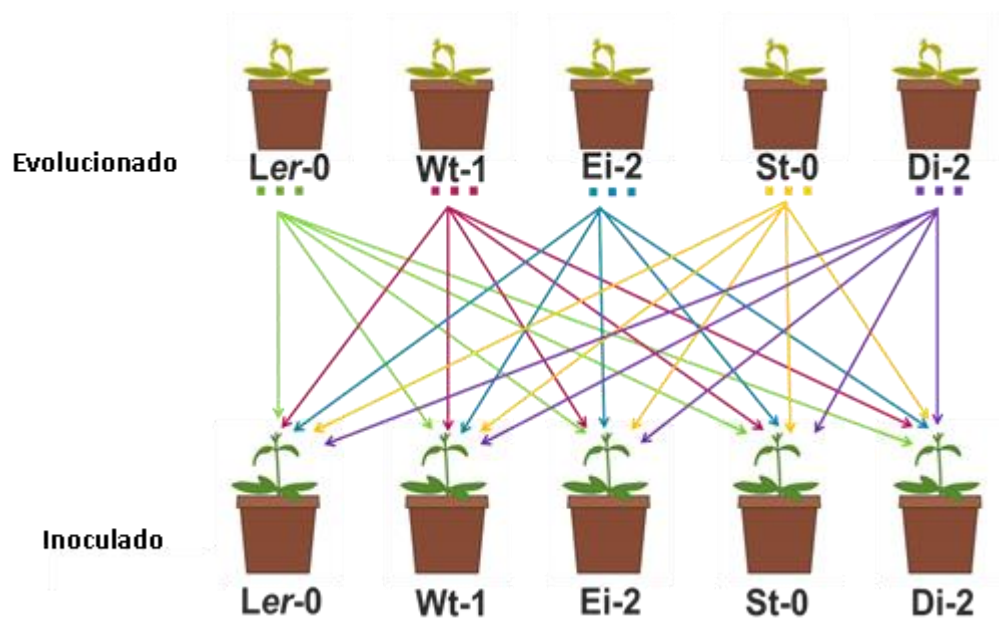


Figura 3. Diseño experimental de infección cruzada. Cada una de las tres réplicas biológicas obtenidas para cada ecotipo tras la evolución experimental se utilizó para infectar cada uno de los cinco ecotipos implicados en el experimento.

La infección sistémica de las distintas combinaciones se comprobó mediante la observación de síntomas o por RT-PCR, tras lo cual se cosecharon todas las plantas infectadas, enteras pero sin raíz. Se pesaron tanto las plantas control, inoculadas sólo con tampón, como las plantas infectadas. De cada combinación virus-huésped, se trituraron un mínimo de tres plantas con ayuda de un mortero y se guardaron a $-80\text{ }^{\circ}\text{C}$.

2.2.2 Extracción de ARN para detección de infección

Relevante para experimentos 3.1 –3.4

En aquellos casos donde los síntomas no eran visibles en las plantas inoculadas, la infección necesitaba ser confirmada mediante RT-PCR. Para ello, a los 14 dpi se arrancaba una hoja de cada planta y se introducía en un tubo de 1 mL que contenía una bolita de acero. A continuación, los tubos se sumergían en nitrógeno líquido para que el tejido de planta se congelara, tras lo cual el tejido congelado se homogeneizaba

mecánicamente mediante agitación a una frecuencia de 3 Hz durante 20 segundos en un molino mezclador MM400 (Retsch). El tejido homogeneizado se resuspendía en un volumen igual de tampón TCES (0,2 M Tris, pH = 8,0, 0,2 M NaCl, 50 mM EDTA, 2% SDS). A partir de este momento, se seguían los pasos descritos a continuación:

1. Se añadía a la mezcla de tejido resuspendido en tampón TCES un volumen igual de fenol saturado.
2. Tras agitación vigorosa, se centrifugaba durante 5 minutos a 14500 rpm.
3. Se transfería la fase acuosa a un tubo nuevo y se añadía LiCl 7,5 M y EDTA 50 mM en proporción 1:0,5 (fase acuosa: LiCl + EDTA). El proceso de precipitación de los ácidos nucleicos se prolongaba durante dos horas en hielo.
4. Se centrifugaba durante 10 minutos a velocidad máxima, tras lo cual se eliminaba el sobrenadante con micropipeta.
5. Se realizaba un lavado con etanol 70% frío (14500 rpm, 10 minutos).
6. Se eliminaba el sobrenadante y se secaba el precipitado al aire, tras lo cual se resuspendía en 100 µL de agua tratada con dietilpirocarbonato (DEPC) 0,1%.
7. Se cuantificaba la concentración en las muestras usando un espectrofotómetro ND-1000 NanoDrop (Thermo Scientific).

2.2.3 Detección de ARN de TEV en plantas inoculadas

Relevante para experimentos 3.1 – 3.4

Como se comentó anteriormente, en aquellos casos donde los síntomas no eran visibles en las plantas inoculadas, la infección necesitaba ser confirmada mediante RT-PCR. Para ello, una vez obtenido el ARN total de las plantas problema (protocolo descrito en el apartado 2.2.2), éste era utilizado para la potencial detección de la presencia de virus. En concreto, se utilizó el *kit* OneStep RT-PCR (Takara) con un volumen de reacción de 10 µL por muestra, donde se añadían 100 ng de ARN total, siguiendo las instrucciones del fabricante. El cebador directo, PC90-95F 5'-GCTGTATTGAAAGTGCGAC-3' y el cebador inverso PC86-91R 5'-AGGCCCAACTCTCCGAAAG-3' amplifican 334 nucleótidos de una región conservada del gen Nlb³². El perfil de amplificación fue el siguiente: 5 min a 42 °C, 10 s a 95 °C, seguidos por 40 ciclos de 5 s a 95 °C y 20 s a 60 °C. Los productos de PCR fueron visualizados en un gel de electroforesis de 1% de agarosa.

2.2.4 Cuantificación de la acumulación viral

Relevante para experimentos 3.1 – 3.3

La cuantificación se llevó a cabo mediante PCR a tiempo real aplicando el protocolo y los reactivos del *kit* OneStep SYBR[®] PrimeScript™ RT-PCR Kit II (Perfect Real Time). Este

kit permite realizar las reacciones de retrotranscripción y PCR en un solo paso. Los cebadores para la PCR cuantitativa fueron diseñados usando PRIMEREXPRESS v.2 (Applied Biosystems). La pareja de cebadores TEV7DACP689F 5'-TTGGTCTTGATGGCAACGTG-3' y TEV7D A739R 5'-TGTGCCGTTTCAGTGTCTTCCT-3' amplifica un fragmento de 30 nucleótidos en el gen de la proteína CP. La mezcla de reacción fue preparada según las instrucciones del manual del *kit*. La concentración de todas las muestras fue ajustada a 50 ng/μL y se pusieron 2 μL de muestra por reacción. Las amplificaciones fueron realizadas usando un Applied Biosystems® 7500 Real-Time PCR System, de acuerdo con el siguiente perfil : 5 min a 42 °C , 10 s a 95 °C y 40 ciclos de 5 s a 95 °C y 34 s a 60 °C. Todas las muestras se cuantificaron por triplicado. Los resultados de la cuantificación fueron examinados usando el programa SDS7500 v.2.2.2 (Applied Biosystems). La carga viral fue determinada en pg de ARN viral en 100 ng de ARN total de planta. La cuantificación absoluta del ARN viral se llevó a cabo usando una curva patrón de ARN de TEV con concentración conocida preparada a partir de seis puntos de dilución seriada. Para obtener el ARN de TEV, el plásmido pMTEV fue transcrito *in vitro* (ver apartado 2.2.5) y el ARN cuantificado con un espectrofotómetro ND-1000 NanoDrop (Thermo Scientific). Para igualar los efectos inhibitorios que pudieran provenir de la planta, así como para tener la misma cinética de amplificación en muestras y en la curva estándar¹⁶², el ARN de TEV para la curva patrón fue diluido en ARN de planta sana. Este ARN fue extraído de una planta crecida en las mismas condiciones que las plantas a partir de las cuales se extrajeron las muestras a cuantificar.

2.2.5 Preparación de muestras para la secuenciación del genoma viral

Relevante para experimentos 3.1 y 3.2

El ARN total de las plantas infectadas ha sido extraído usando el *kit* Invitrap® Spin Universal RNA Mini Kit (Stratec) según protocolo del manual del *kit*. La secuencia completa del genoma viral podía obtenerse mediante la amplificación de tres fragmentos de PCR de tamaño parecido (Tabla 2). En primer lugar, se obtenía el ADN complementario (ADNc) usando la transcriptasa inversa del virus M-MuLV (Thermo Scientific) y un cebador específico para cada uno de los tres fragmentos a amplificar en la PCR. A continuación, el ADNc se usó para amplificación por PCR con la polimerasa Phusion (Thermo Scientific) de los tres fragmentos que comprendían el genoma completo del virus. La secuenciación se llevó a cabo por la empresa GenoScreen (Lille, Francia) usando el método Sanger. En total, se utilizaron 16 cebadores (Tabla 2) para obtener la secuencia consenso del genoma completo, cuyo análisis se realizó mediante el programa Lasergene (DNASStar Inc. Madison WI).

Tabla 2. Cebadores empleados para la retrotranscripción (*), PCR (**) y secuenciación (***) del genoma completo del virus.

Fragmento	Cebador	Secuencia	Localización
1	PC40-45seqr(*, **)	ATCCAACAGCACCTCTCAC	3876-3894
	PC2-10seqf (**, ***)	GCAATCAAGCATTCTACTTC	46-66
	PC11-18seqf (***)	CAATTGTTTCGCAAGTGTGC	762-781
	PC19-26seqf (***)	ACACGTACTGGCTGTCAGCG	1454-1474
	PC27-34seqf (***)	GGACTTCACCAAGTTTATAAGG	2144-2166
	PC35-44seqf (***)	GGCTATGCTGTGACCTCTG	2853-2872
	PC45-48seqf (***)	TTGACGCTGAGCGGAGTGATGG	3541-3563
2	PC67-77seqr (*, **)	AATGCTTCCAGAATATGCC	6700-6681
	PC45-48seqf (**)	TTGACGCTGAGCGGAGTGATGG	3541-3563
	PC49-53seqf (***)	CCCGTGAAACTCAAGATAG	4281-4300
	PC40-45seqr (***)	ATCCAACAGCACCTCTCAC	3875-3894
	PC54-63seqf (***)	TTAAGCTACACACTTGTGAGAC	4972-4994
	PC64-72seqf (***)	CAATGAACCAAGTCTATTTC	5669-5689
	PC73-80seqf (***)	TCATTACAAACAAGCACTTG	6376-6396
3	PC97-101seqr (*, **)	CGCACTACATAGGAGAATTAG	9471-9492
	PC73-80seqf (**, ***)	TCATTACAAACAAGCACTTG	6376-6396
	PC81-89seqf (***)	CACAAAGCATGTGGTTAAAG	7058-7078
	PC90-95seqf (***)	GCTGTATTGAAAGTGCAC	7767-7786
	PC96-97seqf (***)	TACGATATTCCAACGACTG	8481-8500
	PC98-101seqf (***)	AGTACGAAACAGTGGAAGTAG	9179-9200

2.2.5 Transcripción *in vitro*

Relevante para experimento 3.3

Previamente a la transcripción *in vitro*, los plásmidos han sido linearizados con la enzima de restricción *Bgl*II, según el protocolo recomendado por el fabricante de la enzima.

Los plásmidos linearizados fueron purificados usando el método de fenol:cloroformo:alcohol isoamílico (25:25:1), como se describe a continuación:

1. Se añadió un volumen de fenol (pH = 4,2): cloroformo:alcohol isoamílico por volumen de solución de ADN y se mezcló vigorosamente usando un agitador
2. La mezcla se centrifugó a 14500 rpm durante 5 minutos, lo que dio lugar a la formación de dos fases, transfiriéndose la fase acuosa (fase superior) a un tubo nuevo y desechándose el resto.

3. Se añadió un volumen de acetato sódico por cada diez volúmenes de fase acuosa, un volumen de glicógeno por cada 100 volúmenes de fase acuosa y dos volúmenes de etanol absoluto refrigerado por volumen de fase acuosa, tras lo cual se dejó precipitar durante 15 minutos en hielo.
4. A continuación, se centrifugó a velocidad máxima durante 5 minutos y se eliminó el sobrenadante con una micropipeta.
5. Se añadió 1 mL de etanol 70% refrigerado y se centrifugó durante 5 minutos a 14500 rpm.
6. Se quitó el sobrenadante con una micropipeta. Los tubos se secaron al aire y el precipitado se resuspendió en agua tratada con DEPC 0,1%.

La transcripción *in vitro* fue efectuada usando el *kit* mMESSAGE mMACHINE[®] SP6 (Ambion[®]) según protocolo del manual del *kit*. El ARN resultante de la reacción se precipitó usando LiCl del *kit* mMESSAGE mMACHINE[®] SP6 según las recomendaciones del manual.

2.2.6 Clonación por mutagénesis dirigida

Relevante para experimento 3.3

Como paso previo, sobre el plásmido correspondiente al genotipo TEV-At17, con excepción de la mutación C6666U, que no pudo ser añadida por problemas técnicos, fue introducida la mutación adicional G6816A detectada en el aislado inicial de la evolución experimental. En adelante, este plásmido será conocido como pMTEV-At17b. De esta forma, se conseguía reproducir en gran medida el fondo genético correspondiente a la población inicial utilizada en nuestros experimentos.

A continuación, partiendo de este fondo genético, tratamos de reproducir las distintas combinaciones de mutaciones convergentes observadas en los distintos linajes del experimento de adaptación (Tabla 3), lo cual se llevó a cabo mediante mutagénesis dirigida utilizando como molde el plásmido pMTEV-At17b. Los cebadores para la mutagénesis han sido diseñados usando el programa PRIMEREXPRESS v.2 (Applied Biosystems) (Tabla 3). La mezcla de PCR se preparó en un volumen de 20 μ L donde se añadía 4 μ L de Tampón HF 5x, 200 μ M de dNTPs, 0,2 μ M de cada uno de los dos cebadores directo e inverso, 5% de DMSO, 0,02 U/ μ L de polimerasa Phusion (Finnzymes) y 100 ng de plásmido molde. El perfil del termociclador para la PCR de mutagénesis consistió en 30 s de desnaturalización a 98 °C, seguidos por 30 ciclos de 10 s a 98 °C, 30 s a 60 °C y 3 min a 72 °C, terminando con 10 min de elongación a 72 °C. A continuación, los productos de PCR de mutagénesis se incubaron con la enzima de restricción *DpnI* (Fermentas) durante 3 h a 37 °C con el fin de digerir el molde de ADN metilado. La transformación se realizó utilizando células electrocompetentes de *Escherichia coli* DH5 α a las que se añadieron 2 μ L de producto de PCR digerido con

DpnI, tras lo cual se sembraron en placas de LB agar suplementado con 50 µg/mL de ampicilina. Las colonias de bacterias se inocularon en 5 mL de medio líquido LB suplementado con 100 µg/mL de ampicilina y se incubaron durante 16 h a 37 °C agitando a 250 rpm. La extracción de ADN plasmídico se realizó utilizando el *kit Wizard*[®] Plus SV Minipreps DNA Purification System (Promega) siguiendo el manual del usuario. La incorporación de la mutación se confirmó por secuenciación.

Tabla 3. Cebadores de mutagénesis dirigida utilizados para reproducir las mutaciones convergentes detectadas en los distintos linajes tras el experimento de adaptación. El nucleótido correspondiente al cambio que se pretende introducir está subrayado en la secuencia del cebador.

Nucleótido	Aminoácido	Proteína	Cebador directo	Cebador inverso
G6816A	M2224I	Nla-Pro	5'- GCGTGCCGAAAACTTC AT <u>G</u> AATTGTTGACAAA TCAGG-3'	5'- CCTGATTTGTCAACAAT <u>I</u> CTA TGAAGTTTTTCGGCACGC-3'
C795U	G217G	P1	5' CGTCACATGTATGG <u>T</u> G GAGGAAAAGGGTG 3'	5' CACCCTTTTCTCTC <u>A</u> CCATA CATGTGACG 3'
G1272U	sinónima	Hc-Pro	5' GTTGAGGGTGAGT <u>C</u> IGG AGACAGTGTAC 3'	5' GTGACACTGTCTCC <u>A</u> GACTCA CCCTCAAC 3'
C2116U	sinónima	Hc-Pro	5' CATTTTCTTTGCTCTA <u>T</u> TA GTGAATGTCAAGG 3'	5' CCTTGACATTCACTA <u>A</u> TAGAG CAAAGAAAATG 3'
G3639A	sinónima	Cl	5' CATCTACATGCAGAGTTT <u>A</u> GATGATTACGTTACAA CC 3'	5' GGTTGTAACGTAATCATC <u>T</u> AA ACTCTGCATGTAGATG 3'
G6420A	sinónima	Nla-Pro	5' GAAGAAATAATGGAACA CT <u>A</u> TGGTCCAATCACTA CATG 3'	5' CATGTAGTGATTGGACCA <u>A</u> T AGTGTCCATTATTCTTC 3'
C8636U	S2831L	CP	5' CTGGA <u>A</u> CTTCAGGAACA TTCT <u>I</u> AGTTCCACGAATA AATGC 3'	5' GCATTTATTCGTGGA <u>A</u> CTAA GAATGTTCCCTGAAGTTCCAG 3'
A9240G	sinónima	CP	5' GTACTGCAGAGG <u>A</u> GGAC ACTGAACGG 3'	5' CCGTTCA <u>G</u> TGTC <u>C</u> CTCTGC AGTAC 3'

2.2.7 Inoculación de plantas de cinco ecotipos de *A. thaliana* con los clones mutantes derivados de TEV-A17b.

Relevante para experimento 3.3

A. thaliana no se infecta fácilmente con ARN desnudo, como sí pasa con otros huéspedes, como por ejemplo *N. benthamiana*. Por esa razón, todos los clones obtenidos por mutagénesis dirigida fueron transcritos *in vitro* a ARN como está descrito arriba (2.2.5) y los productos de transcripción se utilizaron como inóculo para plantas de *N. benthamiana*. Las hojas de *N. benthamiana* que mostraron síntomas de infección se recogieron a los 10 dpi, se homogeneizaron y se cuantificó su carga viral. El tejido infeccioso de todos los genotipos de virus se resuspendió en tampón C (50 mM de ácido bórico, pH = 8,0, EDTA 5 mM) y se diluyó a la misma concentración para todos los genotipos. Después de una corta centrifugación que permitió sedimentar los trozos de tejido de planta, se cogieron 10 µL del sobrenadante para inocular hojas de *A. thaliana*. Para mejorar la entrada del virus a las células, se aplicó carborundum al 10%. Las hojas se frotaron finalmente con una varilla de vidrio. Las plantas que se usaron para la infección tenían 21 días después de la siembra y se recogieron para valorar sus características de infección a los 21 dpi. La inoculación de ecotipos se realizó en tres bloques temporales manteniendo condiciones constantes de crecimiento de plantas y las condiciones del inóculo.

2.2.8 Medición de infectividad, virulencia, eficacia y epistasia

2.2.8.1 Medición de infectividad

Relevante para experimentos 3.1 - 3.3

Para comparar la infectividad, virulencia y eficacia de los linajes evolucionados en todos los ecotipos de *A. thaliana*, los cinco ecotipos fueron inoculados con cada una de las 15 poblaciones virales evolucionadas en un experimento de infección cruzada (ver apartado 2.2.1). La infección se verificó mediante RT-PCR y la infectividad, I , se determinó como el porcentaje de plantas infectadas del total de plantas inoculadas.

2.2.8.2 Medición de virulencia

Relevante para experimentos 3.2 y 3.3

La parte aérea de las plantas infectadas y control se pesó con una precisión de 10 mg a los 21 dpi. La virulencia, V , se definió como la reducción en el peso debido a la infección y se calcula como $V = 1 - P/\bar{P}$, donde P es el peso de una planta infectada y \bar{P} es el peso promedio de las plantas no infectadas del mismo ecotipo.

2.2.8.3 Medición de eficacia relativa

Relevante para experimentos 3.2 y 3.3

La tasa de crecimiento malthusiano por día de un solo bloque de muestras (Experimento 3.2) se calculó como $m = \frac{1}{t} \log(C_t/C_0)$, donde C_t es el número de genomas de TEV/100 ng de ARN total cuantificado a t dpi. La eficacia relativa se calculó como $W_{x,h} = \exp(m_{x,h} - m_{\text{TEV-At17b},h})$, donde m_x y $m_{\text{TEV-At17b}}$ son las tasas de crecimiento malthusiano del aislado viral x y del aislado ancestral TEV-At17b, respectivamente, evaluados en el mismo huésped h .

2.2.8.4 Medición de epistasia

Relevante para experimento 3.3

La epistasia entre pares de mutaciones x e y (Experimento 3.3) se evaluó como $\varepsilon_{xy,h} = W_{00,h}W_{xy,h} - W_{x0,h}W_{0y,h}$, donde $W_{00,h}$, $W_{xy,h}$, $W_{x0,h}$, y $W_{0y,h}$ son, respectivamente, la eficacia del aislado ancestral TEV-At17b, del doble mutante y de cada mutante simple¹³³, todos evaluados en el huésped con ecotipo h . En aquellos casos en los que la epistasia resultó ser significativa, se procedió además a determinar su tipo: magnitud, signo o señal de reciprocidad. Para ello, hemos empleado las condiciones de desigualdad desarrollados por Poelwijk *et al.*¹⁹²

2.2.9 Análisis estadísticos

Relevante para experimento 3.2

Los datos de eficacia relativa fueron analizados utilizando GLM (distribución normal y función de enlace identidad). El modelo tiene cuatro factores aleatorios, el huésped local (LH), el linaje evolutivo del virus (L), el huésped alternativo de prueba (TH) y la réplica experimental biológica (R , es decir, una planta individual del huésped de prueba inoculada con el virus de un linaje evolucionado en un determinado huésped local). LH y TH son factores ortogonales, L se anida dentro de LH (representado como $L(LH)$) y R se anida dentro de la interacción $TH \times L(LH)$. La ecuación del modelo es:

$$W_{ijklm} = \mu + LH_i + L(LH)_{ij} + TH_k + (LH \times TH)_{ik} + (TH \times L(LH))_{ijk} + R(TH \times L(LH))_{ijkl} + \varepsilon_{ijklm}, \quad (1)$$

donde μ es el gran valor medio y ε_{ijklm} es el error asociado con la medida individual m .

Los datos de infectividad fueron analizados mediante regresión logística binaria GLM (respuesta binomial y función de enlace logit):

$$\log[I_{ijkl}/(1 - I_{ijkl})] = \mu + LH_i + L(LH)_{ij} + TH_k + (LH \times TH)_{ik} + (TH \times L(LH))_{ijk} + \varepsilon_{ijkl}, \quad (2)$$

donde ε_{ijkl} es el error asociado con la medida individual l .

Los datos de virulencia fueron analizados mediante GLM (distribución normal y función de enlace de identidad):

$$V_{ijkl} = \mu + LH_i + L(LH)_{ij} + TH_k + (LH \times TH)_{ik} + (TH \times L(LH))_{ijk} + \varepsilon_{ijkl}. \quad (3)$$

La magnitud de los diferentes efectos incluidos en los modelos fue evaluada usando el estadístico de eta-cuadrado parcial (η_p^2), que representa la proporción de la variabilidad total atribuible a un factor dado. Por convención, los efectos son considerados pequeños para valores $\eta_p^2 < 0.05$, medios para valores $0.05 \leq \eta_p^2 < 0.15$ y grandes para valores $\eta_p^2 \geq 0.15$.

La importancia relativa de la divergencia genética frente al paralelismo fenotípico durante la evolución experimental viene dada por la proporción $I_w = \sigma_G(W)/|\overline{\Delta W}|$ ²⁴⁴, donde $\sigma_G(W)$ es la desviación estándar genética de la eficacia entre linajes y $|\overline{\Delta W}|$ el valor absoluto del cambio promedio de la eficacia comparando con el ancestro común. I_w proporciona una medida de la diferencia genética promedio entre los linajes relativa al cambio evolutivo promedio desde el estado ancestral. Bajo la hipótesis nula de que todo cambio fenotípico está asociado a un cambio genético, $I_w = 1$.

Todos los análisis estadísticos se realizaron con IBM SPSS Statistics v. 21.

2.2.10 Estadísticas de la red de infección

Relevante para experimento 3.2

Una red de infección se considera bipartita cuando contiene dos agentes que interaccionan, en este caso aislados del virus y ecotipos de plantas. Esta red puede representarse como una matriz booleana de tamaño $m \times n$ con entradas asignadas a 1 cuando hay una infección del par (aislado de virus, ecotipo planta), o a 0 cuando ocurre lo contrario. Aquí m es el número de aislados virales y n es el número de ecotipos distintos. La matriz de eficacias W se transformó primero en una matriz booleana que indica si un aislado viral dado actúa significativamente mejor (1) o peor (0) que el ancestral TEV-At17b en un genotipo del huésped dado.

El anidamiento de la matriz de infección se calculó utilizando el algoritmo de cálculo de temperatura de anidamiento implementado en BINMATNEST²⁰⁵. La temperatura T de una matriz de interacción se estima reordenando el orden de las filas de ecotipos y el orden de las columnas de virus de tal manera que muchas de las interacciones ocurran en la porción superior izquierda de la matriz. T cuantifica el grado en que una matriz está perfectamente anidada ($T = 0$) o si la matriz carece de cualquier orden y los elementos se distribuyen al azar ($T = 100$).

Las redes bipartitas se pueden descomponer en componentes inconexos de tal manera que no se observen infecciones cruzadas entre los componentes. La modularidad de las redes de infección bipartitas se computó utilizando el algoritmo estándar de módulos bipartitos recursivamente inducidos¹⁴, que utiliza una búsqueda heurística local para maximizar la modularidad bipartita Q . Q representa con qué frecuencia un ordenamiento particular de virus y ecotipos en módulos se corresponde a

interacciones que están primariamente dentro de un módulo ($Q = 1$ o modular), fuera de los módulos ($Q = -1$ o antimodular) o en algún punto intermedio ($-1 < Q < 1$).

La significación estadística de T y Q se evaluó usando el modelo nulo general propuesto por Bascombe *et al.*¹⁶. En este modelo, la probabilidad de que cada celda de la matriz esté ocupada es el promedio de las probabilidades de ocupación de su fila y columna. Biológicamente, esto significa que la probabilidad de obtener una interacción es proporcional al nivel de generalización (grado) tanto del aislado de virus como del ecotipo de la planta.

2.3 Análisis transcriptómico de plantas infectadas con TEV

2.3.1 Elección de muestras para hibridación de micromatrices

Relevante para experimento 3.1 y 3.4

Los experimentos de hibridación de micromatrices están divididos en dos grandes bloques. El primer bloque de plantas se corresponde con las muestras obtenidas en el primer pase de la evolución experimental (Experimento 3.1), de donde se utilizaron seis ecotipos de *A. thaliana* (Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1) inoculados con el linaje ancestral TEV-At17b (Figura 4A). Para este primer bloque, se procesaron 5 réplicas de cada ecotipo infectado y controles negativos de planta no infectada (datos analizados en apartado 3.1 de esta Tesis).

En el segundo bloque del experimento, se utilizaron plantas procedentes del experimento de infección cruzada. Del total de combinaciones analizadas, se optó por los ecotipos de huéspedes infectados con el virus localmente adaptado, por todos los linajes evolucionados del virus en el ecotipo original Ler-0 y además se analizó el transcriptoma de las plantas donde se había obtenido el linaje más especialista (St-0/3) y el más generalista (Ler-0/1). La Figura 4B muestra el esquema de las muestras elegidas para este segundo bloque (Experimento 3.4). En este caso, se procesaron tres réplicas biológicas de cada categoría de muestra y tres réplicas de planta no infectada por cada ecotipo.

A

Linaje viral (ancestral)	Ecotipo de huésped				
	Ler-0	St-0	Di-2	Wt-1	Ei-2
TEV-At17b	5	5	5	5	5

B

Linaje viral (evolucionado)	Ecotipo de huésped				
	Ler-0	St-0	Di-2	Wt-1	Ei-2
Ler-0/1	3	3	3	3	3
Ler-0/2	3				
Ler-0/3	3				
St-0/1	-	3			
St-0/2	3	3			
St-0/3	3	3	3	3	3
Di-2/1	3		3		
Di-2/2	3		3		
Di-2/3	3		3		
Wt-1/1	3			3	
Wt-1/2	3			3	
Wt-1/3	3			3	
Ei-2/1	3				3
Ei-2/2	3				3
Ei-2/3	3				3

Figura 4. Diseño experimental. **A** y **B**, Representación de muestras usadas para caracterización transcriptómica con micromatrices de ARN. **(A)** Muestras de plantas infectadas con el virus ancestral TEV-At17b. **(B)** Muestras inoculadas con virus evolucionados. Los números en la tabla representan el número de plantas individualmente analizadas para cada categoría. Los números sombreados son combinaciones de virus que infectan su huésped local.

2.3.2 Extracción de ARN para el ensayo de micromatrices

Relevante para experimento 3.1 y 3.4

El ARN que se usó para los ensayos de micromatrices de transcriptómica había sido extraído usando kits comerciales. Para el primer bloque de ensayo (Experimento 3.1) se usó el *kit* RNeasy Plant Mini Kit (Qiagen) y para el segundo bloque (Experimento 3.4) se usó Invitrap® Spin Universal RNA Mini *kit* (Strattec), comprobando con un Agilent 2100 Bioanalyzer (Agilent Technologies) que no había variación en la integridad de ARN obtenido entre los *kits*. En ambos casos, se usaron los protocolos de los *kits* siguiendo las instrucciones de los fabricantes.

2.3.3 Etiquetado de ARN e hibridación de micromatrices

Relevante para experimento 3.1 y 3.4

Para este proyecto se utilizaron plantas individuales como réplicas biológicas. El ARN total fue extraído del tejido homogeneizado de las plantas control o infectadas como se ha descrito en el apartado de extracción de ácidos nucleicos. La integridad del ARN se verificó con un Agilent 2100 Bioanalyzer (Agilent Technologies). Las muestras de ARN para este análisis se prepararon utilizando el *kit* Agilent RNA 6000 Nano (Agilent Technologies) siguiendo las instrucciones del fabricante, tras lo cual se cuantificaron por espectrofotómetro ND-1000 NanoDrop (Thermo Scientific). A continuación, se utilizaron 500 ng de ARN para la amplificación y el etiquetado de la reacción utilizando el *kit* Quick Amp Labeling Kit One-Color (Agilent Technologies) y siguiendo las instrucciones del fabricante. Todas las muestras fueron amplificadas y marcadas con Cy3 y posteriormente verificadas en un Bioanalyzer 2100 como se describió previamente. Como control positivo de amplificación, etiquetado e hibridación, se añadieron a las muestras una serie de ARNs sintéticos del *kit* RNA Spike-In Kit, One-Color (Agilent Technologies). La solución de ARNs sintéticos había sido diluida siguiendo el protocolo del fabricante. El ARN marcado se utilizó para hibridación con un cristal de micromatrices de sondas de *A. thaliana* Col-0 (Agilent Technologies) siguiendo el protocolo de hibridación estándar del Gene Expression Hybridization Kit (Agilent Technologies). Después de la hibridación y el lavado, los cristales se escanearon a 532 nm con un escáner GenePix 4000B (Axon Molecular Devices), usando 10 μ m de resolución y el 100% de la potencia del láser. Los voltajes del tubo fotomultiplicador se ajustaron para igualar la intensidad de la señal global para cada canal, así como para aumentar la relación señal - ruido y reducir el número de puntos de señal de fluorescencia saturada. Las intensidades de cada uno de los puntos fueron cuantificadas usando el programa GenePix Pro 4.1 (Axon Molecular Devices).

Los datos crudos de micromatrices fueron depositadas en el NCBI GEO con números de acceso GSE37269 (Experimento 3.1) y GSE66020 (Experimento 3.4).

2.3.4 Análisis de datos de Micromatrices

2.3.4.1 Primer bloque de muestras

Relevante para experimento 3.1

Los datos de micromatrices fueron normalizados simultáneamente con el programa Babelomics v. 4.2 (<http://babelomics.bioinfo.cipf.es>). Las mediciones de todas las micromatrices fueron normalizadas a una única distribución final en la que las señales de las diferentes muestras estaban calibradas unas respecto a las otras. Los niveles de expresión de todas las sondas que coincidían con un gen dado fueron promediados, lo que reflejaba la expresión del gen. Los puntos de chequeo diseñados como control de calidad del ensayo, la señal de fondo y las estimaciones de hibridación cruzada fueron eliminados en el proceso de normalización de los datos.

Los datos normalizados fueron utilizados en un análisis de expresión diferencial de genes para cada ecotipo de *A. thaliana*, para lo cual se utilizó la herramienta LIMMA de Babelomics. A continuación, se realizaron los análisis de ontología de genes (GO) alterados por la infección utilizando la herramienta FatiGO de Babelomics. La lista de genes significativamente sobre-expresados o suprimidos fue comparada con la lista del genoma completo.

Para el análisis del solapamiento en las listas de genes con expresión alterada entre los ecotipos, se calcularon las matrices de similitud entre listas de genes para todas las comparaciones en parejas de ecotipos. El elemento (i, j) de esta matriz se calculó de acuerdo con $S_{ij} = 2n_{ij}/(n_i + n_j)$, donde n_i y n_j son los genes encontrados en los ecotipos i y j , mientras n_{ij} es el número de genes compartidos entre estos dos ecotipos. A continuación, esta matriz de similitud se utilizó para construir un dendograma usando el método *Neighbor-Joining* con el programa PHYLIP v. 3.69 (<http://evolution.genetics.washington.edu/phylip.html>). Estos análisis se realizaron por separado para genes reprimidos y sobre-expresados.

2.3.4.2 Segundo bloque de muestras

Relevante para experimento 3.4

Los datos del segundo bloque de muestras se normalizaron junto con el primer bloque de muestras, pero diferenciando estadísticamente que se trata de dos bloques de ensayos. La normalización conjunta es necesaria para poder realizar análisis comparando los dos bloques.

La señal de las muestras se estandarizó usando la normalización de cuantiles²⁴. La expresión génica diferencial se llevó a cabo usando el paquete Limma²²⁸ de BIOCONDUCTOR. Los ajustes de los p -valores para el sistema de múltiples pruebas se realizaron de acuerdo con la metodología de Benjamini y Hochberg²².

Dado que muchos términos funcionales se prueban simultáneamente, los resultados de la prueba se corrigen para múltiples ensayos para obtener un p -valor ajustado. El análisis de conjuntos de genes devuelve los p -valores ajustados según el método de tasas de falso descubrimiento (FDR)^{22,23}. Las anotaciones GO para los genes de micromatrices se tomaron de la base de datos de ENSEMBL 73 (<http://www.ensembl.org>).

III Resultados y discusión

3.1 Heterogeneidad en el perfil de interacciones virus-huésped.

Los virus de plantas interactúan con su huésped a nivel molecular y normalmente alteran su fisiología de tal manera que desvían el uso habitual de metabolitos dirigiéndolo hacia la producción de componentes específicos del virus. La consecuencia de esta interacción puede ser tanto que las plantas controlen la propagación de la infección viral como que el virus supere las defensas y establezca una infección productiva. El efecto es muy variable incluso para un determinado par planta-virus, ya que depende de los genotipos de ambos, así como de factores ambientales. Centrándonos sobre todo en el aspecto genético de dichas interacciones, el presente capítulo trata de evaluar cómo las variabilidades genéticas entre huéspedes genéticamente heterogéneos, pero pertenecientes a la misma especie (*i.e.* *A. thaliana*), pueden condicionar la susceptibilidad a la infección con TEV-At17b. Con este objetivo, los ecotipos de *A. thaliana* Col-0, Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1, heterogéneos en la combinación alélica de loci RTM (ver apartado 1.1.3 y la Tabla 4), fueron elegidos como huéspedes e inoculados con la cepa TEV-At17b. Las plantas infectadas fueron recogidas a los 21 dpi y en cada ecotipo se determinó la acumulación del virus, la infectividad y la gravedad de los síntomas. Finalmente se estudió cómo estos elementos fenotípicos se reflejan en el transcriptoma de la planta.

Tabla 4. Lista de ecotipos de *A. thaliana*, origen geográfico y combinación alélica de loci RTM. Tabla extraída de Lalić *et al.* (2010)¹³⁸, con modificaciones.

Ecotipo	Origen	Genotipo
Di-2	Francia	<i>RTM1/RTM1 RTM2/RTM2 RTM3/RTM3</i>
Ei-2	Alemania	<i>rtm1/rtm1 RTM2/RTM2 RTM3/RTM3</i>
Ler-0	Alemania	<i>rtm1/rtm1 RTM2/RTM2 RTM3/RTM3</i>
St-0	Suecia	<i>RTM1/RTM1 RTM2/RTM2 rtm3/rtm3</i>
Wt-1	Alemania	<i>RTM1/RTM1 RTM2/RTM2 RTM3/RTM3</i>
Col-0	EEUU	<i>rtm1/rtm1 RTM2/RTM2 RTM3/RTM3</i>
Oy-0	Noruega	<i>RTM1/RTM1 RTM2/RTM2 rtm3/rtm3</i>

3.1.1 Estima de síntomas de infección, carga viral e infectividad de TEV- At17b

Los síntomas inducidos por TEV-At17b fueron muy variados entre los distintos ecotipos, extendiéndose desde leves (Col-0 y Oy-0) a moderados (St-0, Di-2) y muy graves (Ei-2, Ler-0 y Wt-1) (Figura 5). Los síntomas visibles de la infección han consistido por lo general en el retraso del crecimiento, el grabado de nervaduras pronunciadas y la malformación de la hoja⁵.

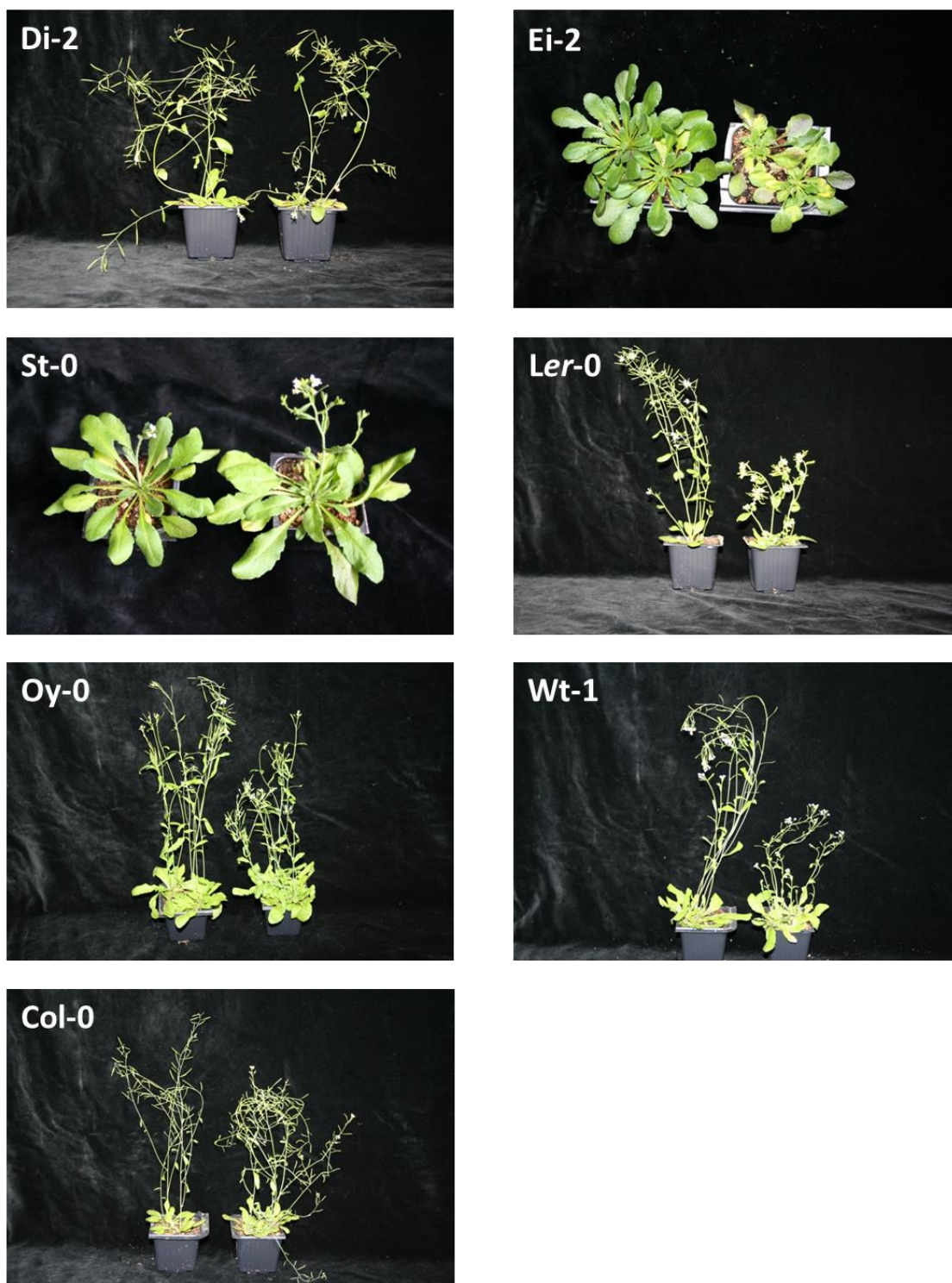


Figura 5. Síntomas inducidos por la infección con TEV-At17b en diferentes ecotipos de *A. thaliana* a los 21 dpi. En cada foto, se muestra a la izquierda la planta sana y a la derecha la planta infectada del correspondiente ecotipo.

Del mismo modo que los síntomas, la infectividad de TEV-At17b también fue heterogénea entre los distintos ecotipos (Figura 6A).

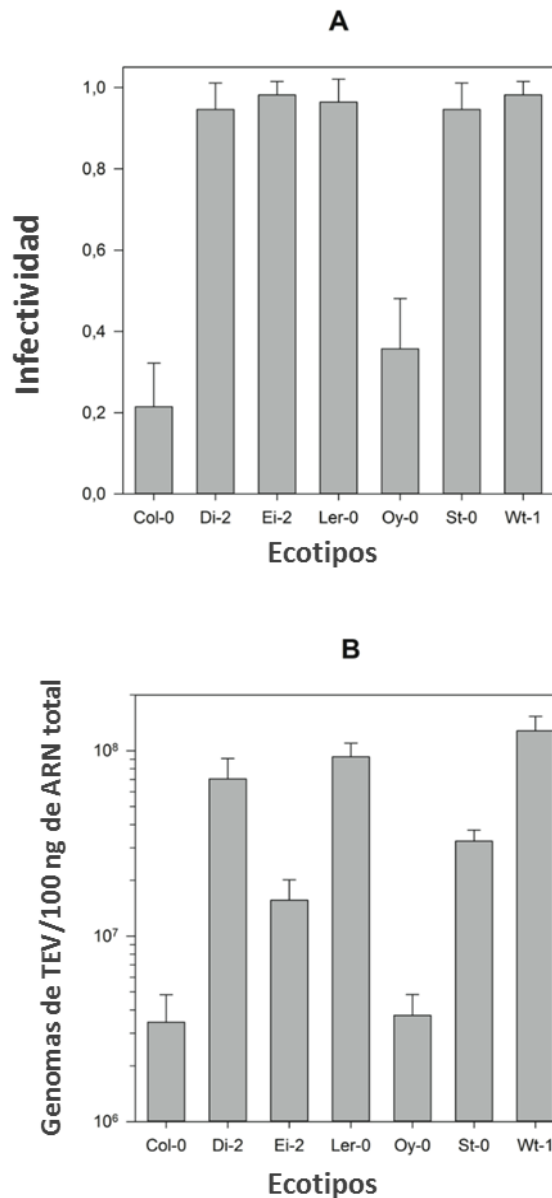


Figura 6. Análisis fenotípico de infección viral en diferentes ecotipos de *A. thaliana*. **(A)** Infectividad de TEV-At17b en cada ecotipo. Las barras de error representan el intervalo de confianza (95%). **(B)** Carga viral de TEV-At17b en cada ecotipo, representado como el número de moléculas de ARN viral en 100 ng de ARN total. Las barras de error representan el error estándar de la media.

En todos los ecotipos, la infectividad de TEV-At17b ha sido significativamente mayor que cero, siendo la más baja en Col-0 y la más alta en Ei-2 y Wt-1. La variación de infectividad es significativa entre los ecotipos ($\chi^2 = 226,815$, 6 g.l., $P < 0,001$), que se pueden clasificar en dos grupos (prueba de Tukey *post hoc*, $P < 0,058$). El primer grupo está formado por los ecotipos Col-0 y Oy-0, con infectividades $\leq 35,71\%$, mientras el segundo grupo está formado por los cinco ecotipos restantes, que presentan infectividades $\geq 96,43\%$.

La Figura 6B muestra la carga viral de los ecotipos a 21 dpi. Dado que los datos no apoyan la homocedasticidad de las varianzas entre grupos (prueba de Levene, $F_{6,51} =$

6,456, $P < 0,001$), para evaluar las diferencias entre los grupos se optó por utilizar la prueba no paramétrica de Kruskal-Wallis. Esta prueba encontró una diferencia altamente significativa ($\chi^2 = 44,113$, 6 g.l., $P < 0,001$) debida a la existencia de dos grupos que no se solapaban (prueba de Tukey, $P \geq 0,171$). El primer grupo estaba formado por Col-0 y Oy-0, ecotipos en los que TEV-At17b ha mostrado la acumulación más baja (carga viral media = $7,26 \times 10^5$ genomas/100 ng de ARN total). El segundo grupo estaba constituido por los cinco ecotipos donde el virus ha llegado a acumulaciones superiores (carga viral media = $4,74 \times 10^7$ genomas/100 ng ARN total, 65 veces mayor que en el primer grupo).

A consecuencia de la baja infectividad y acumulación de TEV-At17b en Col-0, no ha sido posible recoger el material vegetal infectado necesario para su caracterización transcriptómica. Por lo tanto, en las siguientes secciones sólo se utilizaron los seis ecotipos restantes.

3.1.2 Comparación de los patrones de expresión genética entre los ecotipos infectados

Después de evaluar las diferencias fenotípicas debidas a la infección viral a través de los distintos ecotipos de *A. thaliana*, el siguiente objetivo consistió en explorar si estas diferencias de acumulación, infectividad y gravedad de síntomas tenían una correlación a nivel transcriptómico. Para este fin, utilizamos el ensayo de hibridación de ARN marcado de plantas con micromatrices de ARN de un color, donde se representan casi todos los genes de *A. thaliana* Col-0. Para cada ecotipo, se analizó el ARN de cinco plantas independientes infectadas con TEV-At17b, salvo para Oy-0, donde sólo se utilizaron tres. Como controles negativos, se utilizaron cuatro réplicas de plantas de cada ecotipo inoculadas únicamente con tampón.

3.1.3 Determinación de genes con expresión significativamente alterada

Después de la normalización y el análisis estadístico de los datos de micromatrices, se pudieron identificar los genes con expresión alterada entre plantas infectadas con TEV-At17b y plantas control. La Tabla 5 muestra, para los distintos ecotipos, el número de genes cuya expresión estaba significativamente alterada (sobre-expresados y reprimidos) por la infección viral.

Tabla 5. Genes diferencialmente expresados en los distintos ecotipos como respuesta a la infección con TEV-At17b.

Ecotipo	Genes sobre-expresados	Genes reprimidos
Di-2	2662	2733
Ei-2	2070	1971
Ler-0	3406	3405
Oy-0	24	16
St-0	3135	2972
Wt-1	3666	3518

Todos los ecotipos, excepto Oy-0, muestran un gran número de genes diferencialmente expresados en respuesta a la infección con TEV-At17b, donde el valor mínimo lo presenta Ei-2 (4041 genes) y el máximo Wt-1 (7184 genes). En el caso particular de Oy-0, sólo 40 genes habían alterado significativamente su expresión. Globalmente, se ha observado una correlación positiva entre el número de genes alterados y la carga viral ($r_S = 0,943$, 4 g.l., $P = 0,005$). Esta correlación sigue siendo significativa incluso después de la eliminación de los datos de Oy-0 ($r_S = 0,900$, 3 g.l., $P = 0,037$), lo que descarta la posibilidad de que la correlación observada fuera debida en su totalidad a la gran diferencia que existe entre la acumulación en Oy-0 y en el resto de los ecotipos.

3.1.4 Diferencias y similitudes entre ecotipos en patrones de expresión

Para comparar el efecto de la infección de TEV-At17b entre los distintos ecotipos, se compararon las listas de genes alterados, calculando el índice de similitud entre pares de listas (Tabla 6). Los genes regulados positiva o negativamente se observaron por separado.

Tabla 6. Hemimatrices de proporciones de similitudes entre pares de ecotipos. **(A)** Genes sobre-expresados. **(B)** Genes reprimidos.

A

	Di-2	Ei-2	Ler-0	Oy-0	St-0	Wt-1
Di-2	1.0000	0.2303	0.4542	0.0104	0.3288	0.2765
Ei-2		1.0000	0.2765	0.0096	0.2501	0.2273
Ler-0			1.0000	0.0105	0.3397	0.2045
Oy-0				1.0000	0.0108	0.0043
St-0					1.0000	0.2232
Wt-1						1.0000

B

	Di-2	Ei-2	Ler-0	Oy-0	St-0	Wt-1
Di-2	1.0000	0.1747	0.4263	0.0058	0.3222	0.2898
Ei-2		1.0000	0.2236	0.0020	0.1416	0.2463
Ler-0			1.0000	0.0035	0.2531	0.2328
Oy-0				1.0000	0.0074	0.0057
St-0					1.0000	0.2403
Wt-1						1.0000

Las matrices de similitudes se utilizaron para construir dendogramas mediante el método de unión de vecinos, según la similitud en las respuestas de las parejas de ecotipos a la infección por TEV-At17b. La Figura 7 muestra los dendogramas basados en listas de genes sobre-expresados (A) y reprimidos (B).

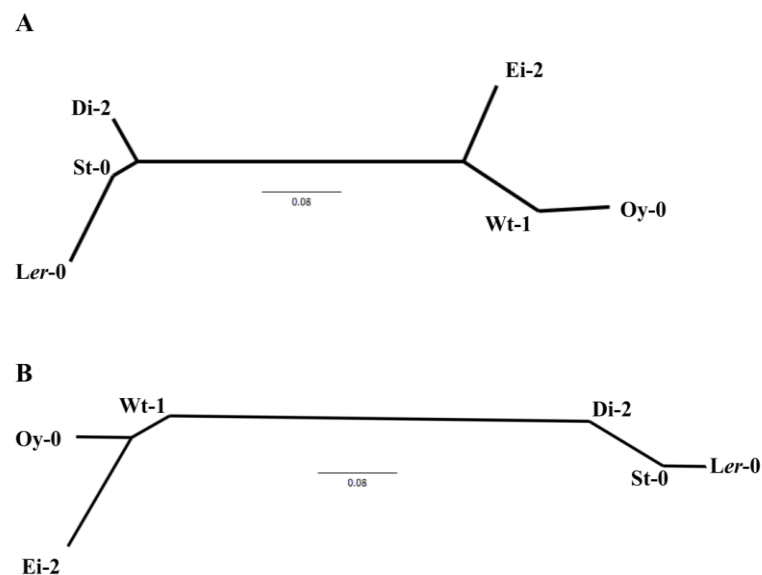


Figura 7. Dendogramas representativos de la similitud en la respuesta transcriptómica de los distintos ecotipos a la infección con TEV-At17b (método *Neighbor-joining*). **(A)** Genes sobre-expresados. **(B)** Genes reprimidos.

Los dendogramas muestran dos grupos diferenciados, tanto para genes sobre-expresados como reprimidos. Un grupo está formado por los ecotipos Di-2, Ler-0 y St-0, al que en adelante nos referiremos como Grupo 1. Del mismo modo, los ecotipos Ei-2, Oy-0 y Wt-1 forman el que llamaremos Grupo 2. La agrupación de ecotipos no se correlaciona ni con síntomas ni con carga viral o infectividad. El Grupo 1 está formado por ecotipos que presentan diversidad en la gravedad de síntomas (moderados para St-0 y Di-2 y graves para Ler-0), al igual que ocurre con los miembros del Grupo 2 (síntomas leves para Oy-0 y graves para Ei-2 y WT-1).

En cuanto a la carga viral, todos los miembros del Grupo 2 muestran una alta acumulación (Figura 7B), mientras que los miembros del Grupo 1 presentan más variabilidad: Ei-2 y Oy-0 tenían valores de acumulación intermedios, mientras Wt-1 presentó la mayor acumulación observada.

Respecto al número de genes alterados a nivel de grupos (Tabla 7), el Grupo 1 tiene 688 genes sobre-expresados y 521 genes reprimidos comunes. Por el contrario, el Grupo 2 comparte muy pocos genes, debido a que la lista de genes alterados en Oy-0 es muy corta (Tabla 5). De hecho, al obviar Oy-0 del análisis, se ve que Ei-2 y Wt-1 en el Grupo 2 todavía comparten 652 genes sobre-expresados y 676 genes reprimidos (Tabla 7). A nivel global, los seis ecotipos sólo comparten cinco genes sobre-expresados (*At1g21520*, *At3g50090*, *At5g03350*, *At5g40990*, *At5g61890*) y uno reprimido (*At5g15240*). Este resultado está obviamente condicionado por el bajo número de genes alterados en Oy-0, ya que al eliminar este ecotipo del análisis se obtiene 157 genes comunes sobre-expresados y 58 genes reprimidos entre el resto de ecotipos. Sin embargo, no hay enriquecimiento significativo de categorías funcionales entre estos genes, posiblemente debido al relativamente bajo número de genes.

Tabla 7. Número de genes comunes sobre-expresados y reprimidos para grupos de ecotipos.

	Grupo de ecotipos	Número de genes comunes
Sobre-expresados	Ler-0 St-2 Di-2	688
	Ei-2 Wt-1 Oy-0	5
	Ei-2 Wt-1	652
	Ler-0 St-2 Di-2 Ei-2 Wt-1 Oy-0	5
	Ler-0 St-2 Di-2 Ei-2 Wt-1	157
Reprimidos	Ler-0 St-2 Di-2	521
	Ei-2 Wt-1 Oy-0	1
	Ei-2 Wt-1	676
	Ler-0 St-2 Di-2 Ei-2 Wt-1 Oy-0	1
	Ler-0 St-2 Di-2 Ei-2 Wt-1	58

Analicemos ahora las funciones proteicas de los genes alterados comunes entre los seis ecotipos. El gen *At1g21520* codifica para una proteína de función desconocida que se localiza en el retículo endoplasmático y se asocia con la respuesta al estrés oxidativo¹⁵². El gen *At3g50090* codifica una exonucleasa con dominios de ribonucleasa similar a H y ribonucleasa T/DNA polimerasa III y se ha descrito en asociación con la degradación de miARNs¹⁹⁵. El gen *At5g03350* codifica para una lectina implicada en la unión de hidratos de carbono que se localiza en el apoplasto, en la pared celular y en cloroplastos²⁵. El gen *At5g40990* codifica para la proteína GDSL lipasa 1 (GLIP1), que está involucrada en la resistencia de plantas. GLIP1 es un componente importante de las dos respuestas de resistencia local y sistémica a través de la vía dependiente de etileno¹³⁴. El gen *At5g61890* codifica para un miembro de la subfamilia del factor de respuesta de etileno B-4 de la familia del factor de transcripción ERF/AP2. Este factor está implicado en las respuestas de defensa bióticas¹¹⁸. Por último, el gen reprimido común *At5g15240* codifica un transportador de aminoácidos transmembrana que puede estar implicado en el estrés osmótico⁹⁹.

3.1.5 Análisis funcional de genes alterados

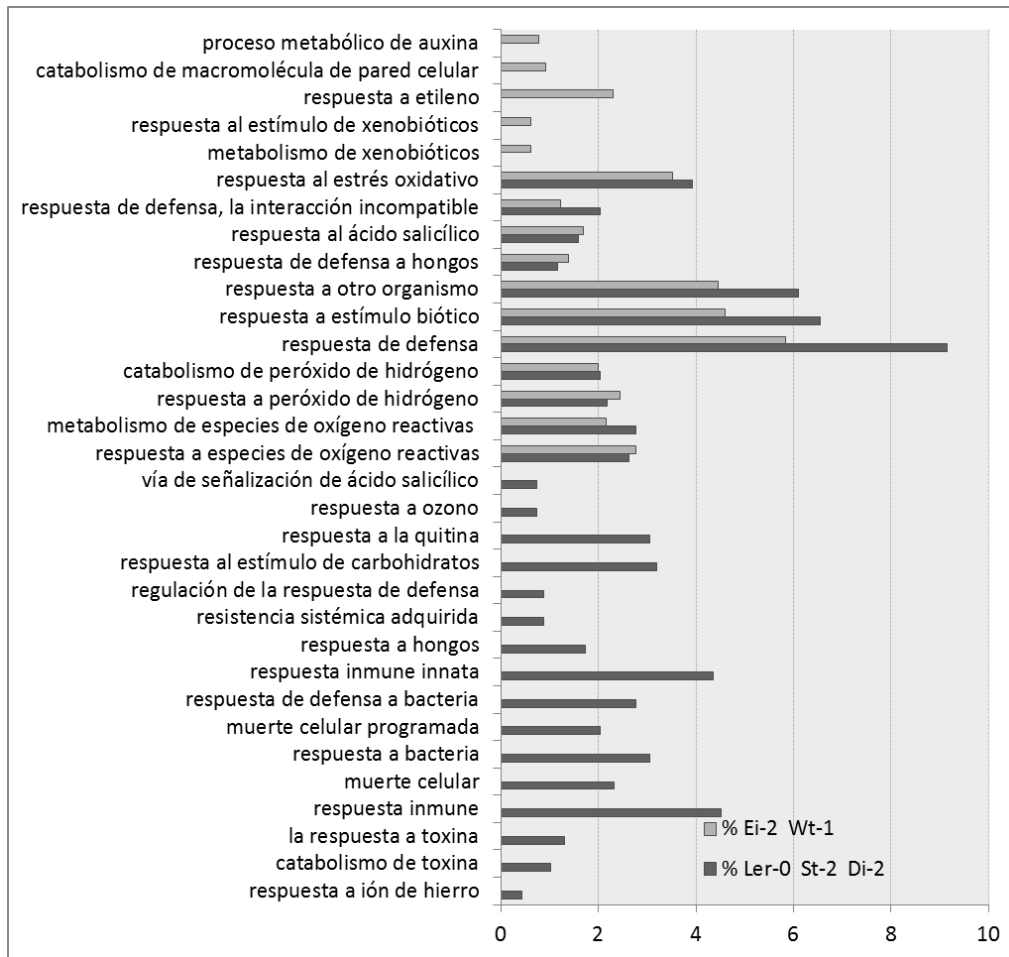
A continuación, se exploraron los procesos biológicos afectados por la infección de TEV-*At17b* en cada ecotipo. Se compararon las categorías funcionales enriquecidas dentro del conjunto de genes significativamente alterados en su expresión con las categorías funcionales distribuidas por el genoma completo. Se seleccionaron las categorías funcionales que están más representadas o enriquecidas en la lista con genes afectados que en el genoma completo. En lugar de describir ecotipos individuales, como en la

sección anterior, se exploraron los aspectos comunes y las diferencias entre los grupos de ecotipos 1 y 2. Para aumentar la sensibilidad del análisis FatiGO e identificar enriquecimientos significativos en los procesos biológicos, excluimos el ecotipo Oy-0 del análisis.

La Figura 8 muestra procesos biológicos significativamente enriquecidos en genes sobre-expresados y reprimidos. Centrándonos en primer lugar en los procesos biológicos enriquecidos para los genes sobre-expresados (Figura 8A), se encontró que ambos grupos de ecotipos compartieron 11 categorías funcionales, incluyendo cinco relacionadas con estrés oxidativo y seis con las respuestas de defensa a las infecciones. No es sorprendente que se haya detectado en todos los ecotipos un importante enriquecimiento en los genes implicados en respuestas mediadas por el ácido salicílico.

Si nos fijamos ahora en las diferencias, observamos dieciséis procesos significativamente enriquecidos en el Grupo 1, pero no en el Grupo 2. Por el contrario, se encontró cinco categorías significativamente enriquecidas en el Grupo 2, pero no en el Grupo 1 (Figura 8A). En concreto, el Grupo 2 presentó un enriquecimiento en los procesos catabólicos de la pared celular, procesos metabólicos de auxina, respuesta al etileno y a estímulos xenobióticos. En cambio, en el Grupo 1 los procesos de defensa eran relativamente más importantes: la señalización mediada por ácido salicílico se potenció aún más que en el Grupo 2, así como la vía de SAR, la respuesta inmune innata y la activación de los genes de apoptosis. Algunas respuestas adicionales a estrés abiótico también se activaron en el Grupo 1. La activación de las respuestas de defensa adicionales en el grupo 1 se correlaciona con mayor acumulación de TEV-A τ 17b en estos ecotipos.

A



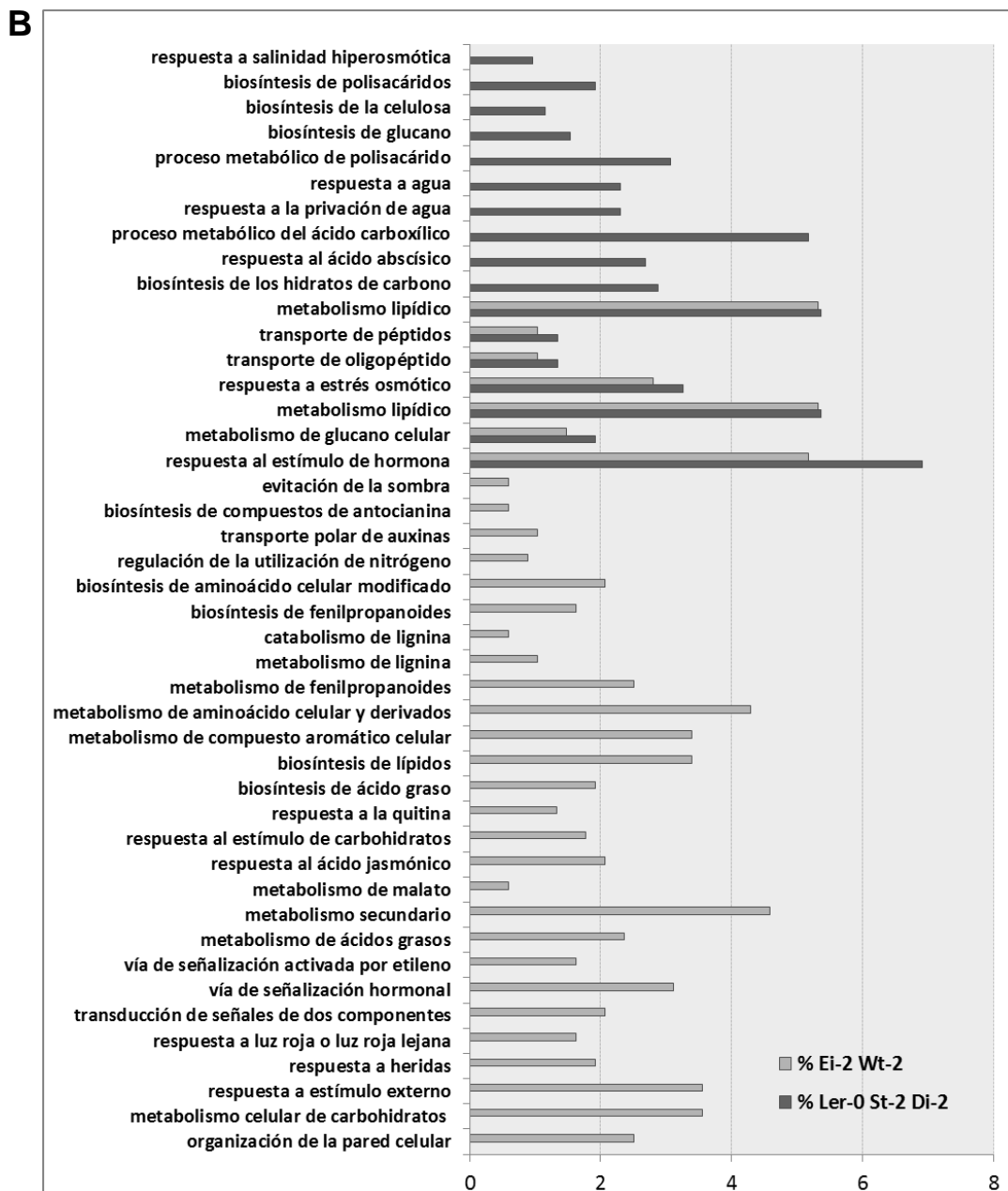


Figura 8. Categorías funcionales significativas identificadas para el grupo 1 (*Ler-0 St-2* y *Di-2*, barras oscuras) y grupo 2 (*Ei-2* y *Wt-2*, barras claras). La longitud de la barra representa el porcentaje de enriquecimiento de cada proceso biológico para genes sobre-expresados (A) y para genes reprimidos (B).

En cuanto a los procesos biológicos enriquecidos entre los genes reprimidos (Figura 8B), se encontró un número reducido de categorías significativamente enriquecidas en ambos ecotipos: transporte de péptidos, metabolismo de lípidos, metabolismo de glucano celular, respuesta a estrés osmótico y respuesta al estímulo hormonal. Diez procesos biológicos fueron enriquecidos entre los genes sobre-expresados en el Grupo 1. Estos genes están implicados en la salinidad hiperosmótica y respuesta a estrés hídrico, procesos implicados en polisacáridos, como por ejemplo celulosa y glucano, metabolismo de los ácidos carboxílicos y respuesta a estímulo de ácido abscísico, una fitohormona con importantes funciones dentro de la fisiología de la planta que participa

en procesos del desarrollo y crecimiento, así como en la respuesta adaptativa a estrés tanto de tipo biótico como abiótico. Del mismo modo, veintiséis categorías fueron significativamente enriquecidas para los genes regulados en el Grupo 2. Estas categorías incluyen evitación de sombra y respuesta a la intensidad de la luz roja, regulación de la utilización del nitrógeno, catabolismo de la lignina y organización de la pared celular, varios procesos del metabolismo secundario (por ejemplo, fenilpropanoides, compuestos aromáticos, malato y ácidos grasos) y varias vías de señalización mediadas por el ácido jasmónico, auxinas y etileno.

3.1.6 Discusión

En este primer capítulo se propuso valorar cómo la variabilidad genética del huésped afecta a la susceptibilidad al virus TEV-*At17b*. Los resultados confirmaron que la adaptación previa del virus ancestral TEV a *A. thaliana* ecotipo *Ler-0*⁵ no era específica de ecotipo. Ha sido posible infectar con éxito otros ecotipos y los efectos que el virus ha producido en estos ecotipos, tanto fenotípicos como transcriptómicos, han sido heterogéneos. Los ecotipos se pudieron clasificar en dos grupos en base a las diferencias en susceptibilidad y carga viral. Los ecotipos Col-0 y Oy-0 han mostrado menor susceptibilidad y carga viral, mientras que los otros cinco ecotipos eran altamente susceptibles a la infección y presentaban altos niveles de acumulación viral. La carga viral además está correlacionada significativamente con el número de genes alterados por la infección, lo que apoya la hipótesis de que cuanto más replicación, y por lo tanto más acumulación viral, más perturbación sobre el metabolismo celular y por lo tanto más genes alterados en sus niveles de expresión.

Los dendogramas basados en similitudes de genes alterados entre los ecotipos, tanto sobre-expresados como reprimidos, tienen topologías casi idénticas y apoyan la existencia de dos respuestas claramente diferenciadas a la infección por TEV-*At17b*. El primer grupo está formado por los ecotipos Di-2, *Ler-0* y St-0. Curiosamente, *Ler-0* y St-0 fueron susceptibles a la cepa ancestral de TEV por sus alelos recesivos presentes en los loci RTM. El complejo multigénico RTM es uno de los sistemas de resistencia de TEV, que incluye loci RTM1 (*At1g05760*), RTM2 (*At5g04890*) y RTM3 (*At3g58350*)^{42,43,155,255,258}. La presencia de alelos dominantes en los tres loci es necesaria para el bloqueo del movimiento sistémico de TEV, mientras que las mutaciones recesivas homocigotas en cualquiera de los tres loci tienden a resultar en una infección sistémica^{42,43}. Por tanto, es comprensible pensar que los alelos RTM marcan gran diferencia en el tipo de respuesta de los grupos de ecotipos a la infección con TEV-*At17b*.

Hemos encontrado que algunas categorías de genes implicados en la respuesta a la infección por TEV-At17b eran específicos de ecotipo, mientras que otras fueron compartidas por diferentes ecotipos. Un conjunto básico de seis genes fue alterado en todos los ecotipos. Curiosamente, cuatro de estos genes estaban involucrados en las respuestas al estrés y uno en el metabolismo de miARNs. Dos de los genes relacionados con el estrés estaban involucrados en la resistencia sistémica dependiente de etileno, lo que puede indicar la relevancia de esta vía en la respuesta a las infecciones virales. Los dos grupos de ecotipos de *A. thaliana* difieren significativamente en la forma en que perciben y responden a la infección con TEV-At17b. Por un lado, los miembros del Grupo 1 tienden a sobre-expresar los genes de defensa y reprimir aquellos implicados en la producción de componentes de la pared celular. Por otro lado, los miembros del Grupo 2 tienden a sobre-expresar los genes implicados en el estrés abiótico y la construcción de la pared celular, mientras se presenta una tendencia a reprimir los genes implicados en el metabolismo secundario y algunas vías reguladas por hormonas. La alteración positiva de la expresión de genes implicados en la construcción celular en miembros del Grupo 2 (*i.e.*, Ei-2 y Wt-1) podría ser un reflejo de síntomas de infección fuertes y alta acumulación del virus en estos ecotipos.

Los huéspedes han desarrollado una variedad de mecanismos para compensar el impacto de la infección viral. Por ejemplo, la teoría de la evolución de caracteres de historia de vida postula que los organismos tienen que distribuir recursos limitados como la energía, los alimentos y el tiempo entre varios procesos vitales, como son el crecimiento, la salud y la reproducción. Cualquier inversión en uno de los procesos reduce los recursos disponibles para otros procesos. La modificación de rasgos de en la historia de la vida y la reasignación de recursos puede permitir que las plantas sean más tolerantes a la infección^{163,188,231}. Para los ecotipos donde la infección es más grave, produciéndose más acumulación de virus, el conjunto de genes sobre-expresados tiende a estar involucrado en las defensas, desviando recursos desde la construcción de nuevos tejidos. Por el contrario, ecotipos más tolerantes, donde el virus se acumula en menor medida pero provoca síntomas graves, tienden a sobre-expresar los genes implicados en el estrés abiótico y genes implicados en la construcción de nuevos tejidos. Por lo general, las plantas donde están activados los genes de defensa presentan escasez en los recursos para el mantenimiento de las vías vitales de la planta a medida que se alteran diversos procesos metabólicos.

Se ha demostrado que las respuestas tempranas de *A. thaliana* a la infección viral, como son los cambios en la expresión de genes relacionados con la defensa, la señalización celular, metabolismo primario y secundario, la transcripción y los transportadores, ya están presentes pocas horas después de la inoculación^{12,119,256}. Por otra parte, también se ha visto que diferentes células y tejidos pueden responder de

forma ligeramente diferente a la infección²⁶⁶ y por lo tanto se podría perder la respuesta específica mezclando diferentes tejidos para el análisis. En cualquier caso, el tiempo de recogida de las plantas se definió a los 21 dpi por tres razones: (i) los síntomas en este punto de tiempo tan avanzado de la infección acumulan efectos de las alteraciones que tienen lugar en diferentes momentos y se esperaría encontrar información más extendida (ii) un trabajo anterior⁵ permitió detectar diferencias significativas entre cepas de TEV en este punto temporal y (iii) este período incluye el efecto del movimiento sistémico y no sólo el movimiento de célula a célula.

Los resultados obtenidos en el presente capítulo nos llevan a pensar que la interacción del aislado TEV-At17b con los diferentes ecotipos es heterogénea, lo cual explica las diferencias observadas en susceptibilidad, acumulación viral y gravedad de los síntomas. Esta heterogeneidad en el perfil fenotípico tiene su reflejo a nivel transcriptómico en diferentes ecotipos de huésped activando/suprimiendo la expresión de diferentes conjuntos de genes.

3.2 Evolución experimental de TEV en huéspedes genéticamente heterogéneos

En el primer capítulo se observó una respuesta heterogénea de diferentes ecotipos de huéspedes de la misma especie a la infección con el aislado TEV-At17b previamente adaptado al ecotipo Ler-0. En el presente capítulo, vamos a tratar de evaluar en qué medida las diferencias genéticas entre los huéspedes de la misma especie condicionan el destino evolutivo de un virus de ARN de plantas. Con este fin, se realizó un experimento de evolución del aislado TEV-At17b en cinco ecotipos genéticamente heterogéneos de *A. thaliana* (tres réplicas experimentales por ecotipo). La evolución experimental consistió en pases seriados del virus en los ecotipos Di-2, Ei-2, Ler-0, St-0 y Wt-1.

Como ya hemos comentado, los ecotipos de *A. thaliana* varían en susceptibilidad al TEV¹⁵⁵; algunos permiten el movimiento a larga distancia desde la hoja inoculada a las hojas no inoculadas, mientras que otros permiten la replicación en las hojas inoculadas, pero no el movimiento sistémico¹⁵⁵. La susceptibilidad depende del sistema multigénico RTM, que está compuesto de los loci *RTM1*, *RTM2* y *RTM3*^{42,43,51,52,155,255,258}. La presencia de alelos dominantes en los tres loci es necesaria para la resistencia, mientras que mutaciones recesivas homocigotas en cualquiera de los tres loci permiten la infección sistémica^{42,43}. La evolución experimental fue también iniciada en los ecotipos Oy-0 y Col-0, pero los niveles marcadamente bajos de infectividad impidieron finalmente llevar a cabo la evolución en estos ecotipos.

Después de 15 pases de 21 días de evolución experimental, se encontró que los virus evolucionados mostraban un aumento de eficacia, infectividad y virulencia en los ecotipos donde habían sido evolucionados. La secuenciación del genoma de los linajes evolucionados permitió identificar una serie de mutaciones potencialmente adaptativas. En este sentido, cabe destacar la detección de algunas mutaciones convergentes en varios linajes de la evolución experimental. A continuación, se trató de comprobar la especificidad de la adaptación del virus, para lo cual se infectaron con las 15 poblaciones de virus evolucionado cada uno de los cinco ecotipos. El resultado de esta infección cruzada fue caracterizado fenotípicamente en términos de infectividad, virulencia y eficacia relativa al linaje viral adaptado a cada huésped. En el primer capítulo de este trabajo, se comparó el efecto de la infección de TEV-At17b en el transcriptoma de los ecotipos (Tabla 3), lo que permitió observar que los ecotipos diferían en la forma de percibir y responder a la infección. En el presente capítulo, se pretendía explorar si la evolución de TEV-At17b en cada ecotipo de *A. thaliana* daría lugar a la especialización o, por el contrario, los nuevos virus evolucionados conservarían la capacidad de infectar

todos los ecotipos, tal y como ocurría en el caso del virus ancestral TEV-At17b. En particular, se evaluó la potencial aparición de una interacción significativa entre el genotipo del huésped y la población viral evolucionada. Asimismo, se aplicaron análisis de red para evaluar el anidamiento y la modularidad de la matriz de infección. Finalmente, se estudiaron las mutaciones convergentes de los linajes virales resultantes de la evolución experimental, su significado a nivel molecular y su potencial papel en la adaptación al nuevo huésped.

3.2.1 Medida de la adaptación a cada huésped local.

En primer lugar, hemos tratado de determinar si cada linaje viral evolucionado ha incrementado la eficacia en su correspondiente ecotipo local de huésped en comparación con el linaje ancestral TEV-At17b. La medida de la eficacia (W) viene definida como acumulación viral, tal y como se describe en los apartados 2.2.8 y 2.2.9. De este modo, se cuantificó W para todos los linajes evolucionados en los distintos ecotipos (Tabla 8).

Tabla 8. Eficacia relativa de cada combinación virus-huésped. Las celdas sombreadas indican los valores de W para un linaje dado obtenidos en el mismo ecotipo donde se evolucionó. Los valores son el promedio de varias réplicas biológicas (entre 5 y nueve) y los errores corresponden a ± 1 SEM. Los asteriscos indican valores significativamente diferentes a los obtenidos para la cepa ancestral TEV-At17b (prueba t para una única muestra, $P < 0,05$; niveles de significación corregidos por el método FDR).

Hospedador local	Linaje	Huésped de prueba				
		Di-2	Ei-2	Ler-0	St-0	Wt-1
Di-2	1	0.990±0.000*	0.984±0.002*	0.998±0.021	1.044±0.015*	1.061±0.036
	2	0.995±0.012	0.988±0.004*	1.002±0.000*	1.024±0.028	1.021±0.004*
	3	1.009±0.004*	0.987±0.015	1.000±0.021	1.041±0.010*	1.054±0.027
Ei-2	1	1.018±0.002*	1.006±0.004	0.954±0.024	1.024±0.015	0.953±0.032
	2	1.001±0.014	1.004±0.003	0.991±0.005	1.027±0.009*	1.015±0.011
	3	1.006±0.002*	0.984±0.011	0.992±0.006	1.031±0.008*	1.003±0.008
Ler-0	1	1.003±0.004	0.980±0.007*	0.953±0.021	1.084±0.038	1.029±0.014
	2	1.001±0.005	1.012±0.004*	1.004±0.024	1.013±0.010	1.014±0.008
	3	0.986±0.004*	0.987±0.051	1.034±0.037	1.028±0.009*	0.978±0.040
St-0	1	0.981±0.005*	0.975±0.016	0.980±0.028	1.027±0.013*	0.995±0.001*
	2	0.998±0.003	0.996±0.008	0.982±0.026	1.057±0.003*	1.021±0.004*
	3	0.992±0.008	0.995±0.007	0.994±0.011	1.136±0.006*	0.991±0.005
Wt-1	1	1.013±0.004*	0.967±0.037	0.997±0.009	0.991±0.036	1.011±0.004*
	2	0.994±0.005	0.999±0.023	0.950±0.014*	1.023±0.010	1.011±0.003*
	3	0.995±0.002*	0.993±0.004	1.004±0.028	1.085±0.019*	1.022±0.005*

Los resultados fueron muy variables entre ecotipos (prueba de Friedman, $\chi^2 = 43,320$, 1 g.l., $P < 0,001$). Ninguno de los linajes evolucionados en Ei-2 o Ler-0 mostró aumento significativos en W (prueba t para una única muestra, 1 cola, $P \geq 0,058$), mientras que únicamente el linaje 3 del ecotipo Di-2 mostró un 0,85% de aumento significativo (prueba t de una muestra, 1 cola, $P = 0,002$). Por el contrario, todos los linajes evolucionados en St-0 (2,74%, 5,65% y 13,62%, respectivamente, prueba t de una muestra, 1 cola $P \leq 0,031$) y en Wt-1 (1,13%, 1,13% y 2,24%, respectivamente; prueba t de una muestra, 1 cola $P \leq 0,005$) mostraron claros incrementos significativos.

A continuación, se trató de explorar si los cambios observados en W reflejaban sucesos concretos de mutación y selección o si podían ser explicados por evolución fenotípica paralela sin una clara base genética.

La relación entre la divergencia genética y el paralelismo fenotípico durante la evolución experimental viene dada por la proporción $I_W = \sigma_G(W)/|\overline{\Delta W}|^{244}$, donde $\sigma_G(W)$ es la desviación genética estándar de W entre linajes y $|\overline{\Delta W}|$ el valor absoluto del cambio promedio de W comparando con el ancestro común. I_W proporciona una medida de la diferencia genética promedio entre los linajes relativa al cambio de eficacia promedio desde el estado ancestral. Bajo la hipótesis nula de que todo cambio fenotípico está asociado a un cambio genético, $I_W = 1$.

Tal y como observamos en la Tabla 8, los linajes evolucionados en Di-2, Ei-2 y Ler-0 muestran muy diversos resultados, mientras que los linajes evolucionados en St-0 y Wt-1 presentan tendencias crecientes similares en su eficacia relativa.

La Tabla 9 muestra las estimaciones de máxima verosimilitud de $\sigma_G(W)$, $|\overline{\Delta W}|$ e I_W para cada huésped local. $I_W > 1$ para los linajes evolucionados en Di-2, Ei-2 y Ler-0, lo que indica la evolución fenotípica en paralelo, o dicho de otro modo, las diferencias entre los linajes independientes son pequeñas en relación con la variación media de la eficacia de la cepa ancestral TEV-A17b. Sin embargo, no se pudo rechazar la hipótesis nula de una contribución igual de paralelismo y divergencia ($z \leq 1,277$, $P \geq 0,101$) debido a la gran varianza asociada a las estimaciones de I_W (Tabla 8). Por el contrario, $I_W < 1$ para St-0 y Wt-1, ($z \geq 2,196$, $P \leq 0,014$), lo que indica que la contribución de las diferencias genéticas entre los linajes era más importante que los cambios promedios en la eficacia relativa.

La infectividad I mostró una respuesta más débil a la evolución en los distintos ecotipos que la observada para W . Un menor número de linajes mostraron una variación significativa, teniendo también una tendencia variable en signo. Por un lado, el linaje St-0/1 redujo su I en un 83,34% (GLM; $\chi^2 = 9,515$, 1 g.l., $P = 0,002$). Por otro lado, los linajes Di-2/1 (118,55%), Di-2/3 (122,61%), Wt-1/2 (55,55%), y Wt-1/3 (113,88%) (GLMs; $\chi^2 \geq 3,552$, 1 g.l., $P \leq 0,030$) mostraron un aumento significativo en I .

Los resultados de estos análisis apuntan a que la respuesta evolutiva de TEV-A17b depende del genotipo del huésped. Los linajes evolucionados en St-0 y Wt-1 aumentaron la eficacia en sus huéspedes locales como consecuencia de potenciales cambios genéticos durante la evolución. Por el contrario, la respuesta adaptativa viral en el ecotipo Di-2 fue más débil, e incluso nula para Ei-2 y Ler-0, donde además se mostró un fuerte paralelismo fenotípico.

Tabla 9. Divergencia contra paralelismo en eficacia relativa para los linajes evolucionados independientemente en los cinco ecotipos de *A. thaliana*.

Huésped local	$ \overline{\Delta W} $	$\text{Var} \overline{\Delta W} $	$\sigma_G(W)$	$\text{Var}(\sigma_G(W))$	$I_W (\pm 1 \text{ SEM})$
Di-2	$2,533 \times 10^{-3}$	$9,402 \times 10^{-5}$	$1,198 \times 10^{-4}$	$2,913 \times 10^{-5}$	$4,321 \pm 56,183$
Ei-2	$2,300 \times 10^{-3}$	$1,502 \times 10^{-4}$	$4,657 \times 10^{-5}$	$7,016 \times 10^{-9}$	$2,967 \pm 7,121$
Ler-0	$3,067 \times 10^{-3}$	$1,675 \times 10^{-3}$	$3,105 \times 10^{-4}$	$9,338 \times 10^{-7}$	$5,746 \pm 80,002$
St-0	$6,957 \times 10^{-2}$	$3,173 \times 10^{-3}$	$2,054 \times 10^{-3}$	$2,951 \times 10^{-6}$	$0,652 \pm 0,076$
Wt-1	$1,390 \times 10^{-2}$	$4,107 \times 10^{-5}$	$1,106 \times 10^{-5}$	$4,923 \times 10^{-10}$	$0,239 \pm 0,058$

3.2.2 La especificidad de la adaptación

A continuación se exploró si la adaptación de TEV-A17b a los distintos ecotipos que diferían en su susceptibilidad a la infección supondría un coste de eficacia en el huésped ancestral Ler-0. Para ello, se utilizaron pruebas *t* emparejadas para comparar la eficacia de los linajes en su nuevo huésped con las obtenidas en Ler-0. Los linajes evolucionados en Di-2, Ei-2 y Wt-1 no mostraron diferencias significativas entre los dos huéspedes, es decir, los linajes virales evolucionados en estos ecotipos no tuvieron un coste de eficacia en Ler-0 ($t_3 \leq 2,058$, 1 cola $P \geq 0,075$). Por el contrario, los linajes evolucionados en St-0 mostraron un coste significativo de eficacia en Ler-0 del 8,93% ($t_3 = 3,129$, 1 cola $P = 0,028$).

Posteriormente, se estudió la posibilidad de que la adaptación del virus a un nuevo ecotipo estuviera asociada con la eficacia en un ecotipo alternativo. Para ello, se compararon los valores de eficacia de los linajes evolucionados en sus nuevos huéspedes locales con los valores estimados en cada uno de los tres nuevos huéspedes alternativos. Los resultados variaban ampliamente entre los huéspedes locales (Tabla 8). En promedio, los aislados evolucionados en Di-2 presentaban mayor eficacia en St-0 y Wt-1 ($t_3 \geq 3,656$, 1 cola $P \leq 0,034$) que la población ancestral, pero no en Ei-2. Las poblaciones virales evolucionados en Ei-2 eran más eficaces que el ancestro en St-0 y Di-2 ($t_3 = 3,346$, 1 cola $P = 0,039$), pero sin cambio de eficacia en Wt-1. Los linajes evolucionados en Wt-1 y St-0 eran en promedio igualmente eficientes en todos los huéspedes alternativos. La Tabla 10 muestra los resultados del análisis GLM descrito por la ecuación (1), definida en la sección de Materiales y Métodos.

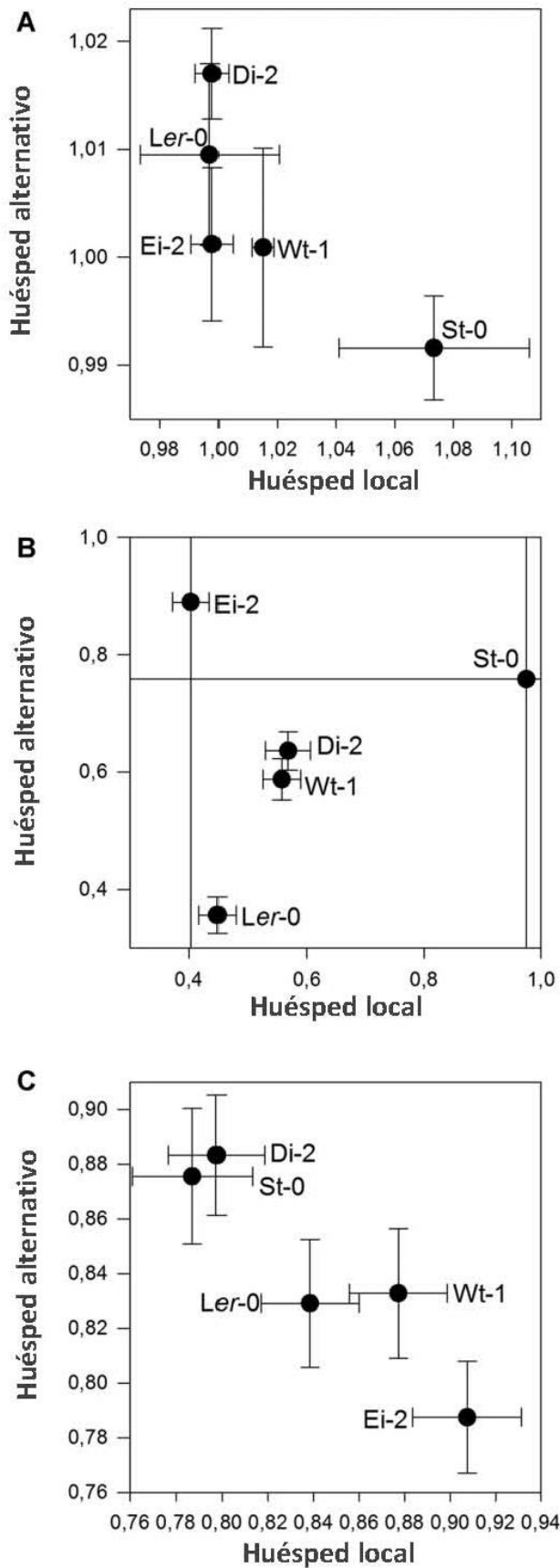


Figura 9. (A) Eficacia relativa promedio en el huésped local (eje X) y en huéspedes alternativos (eje Y). **(B)** Infectividad promedio en el huésped local (eje X) y en huéspedes alternativos (eje Y). **(C)** Virulencia promedio en el huésped local (eje X) y en huéspedes alternativos (eje Y).

Cada punto representa el promedio de los linajes evolucionados. En concreto, para cada huésped local se disponía de las tres estimas de los linajes evolucionados, mientras que para los huéspedes alternativos se promediaron 12 estimas (todas salvo las de los linajes evolucionados en un huésped local dado). Las barras de error representan ± 1 SEM. Los intervalos de error para la infectividad de Ei-2 en huéspedes alternativos, y para St-0 en el huésped local y en alternativos, superan el intervalo [0, 1].

Tabla 10. Análisis GLM de la varianza respecto a las tres características medidas para todos los linajes evolucionados en los cuatro nuevos huéspedes (Di-2, Ei-2, St-0 y Wt-1).

Fuente de variación	Eficacia relativa				Infectividad				Virulencia			
	χ^2 de Wald	d.f.	P	η_p^2	χ^2 de Wald	d.f.	P	η_p^2	χ^2 de Wald	d.f.	P	η_p^2
Intersección μ	366741 75,7	1	<0,001	1	0	1	1,000	0,95	6743,1	1	<0,001	1
Huésped local <i>LH</i>	555,6	4	<0,001	0,32	17,6	4	0,002	0,08	19,1	4	0,001	0,14
Linaje <i>L(LH)</i>	2667,07	10	<0,001	0,18	81,8	10	<0,001	0,67	89,8	10	<0,001	0,38
Huésped alternativo <i>TH</i>	14629,1 2	4	<0,001	0,76	61,6	4	<0,001	0,55	12,8	4	0,012	0,14
<i>LH</i> × <i>TH</i>	4655,74	16	<0,001	0,28	39,4	16	0,001	0,38	82,5	16	<0,001	0,36
<i>TH</i> × <i>L(LH)</i>	12124,3	40	<0,001	0,19	86,9	39	<0,001	1	149,9	40	<0,001	0,16
Replica biológica <i>R(TH</i> × <i>L(LH))</i>	53712,9	295	<0,001	0,98								

Tanto huésped local como huésped alternativo tienen claros efectos altamente significativos en W ($\eta_p^2 > 0,15$ en todos los casos). Cabe destacar que la interacción entre los huéspedes locales y alternativos ($LH \times TH$) también tiene un efecto claro y altamente significativo, lo que sugiere que las restricciones selectivas impuestas por un huésped local dado afectan al rendimiento posterior en los huéspedes seleccionados. De acuerdo con esto, los linajes también son heterogéneos en la respuesta a su huésped local ($L(LH)$), así como en su rendimiento a través de los huéspedes ($TH \times L(LH)$).

En resumen, se puede decir que todos los linajes han evolucionado hacia un carácter generalista, ya que la adaptación a un ecotipo dado siempre se asocia a un aumento de eficacia en al menos uno de los huéspedes alternativos. Sin embargo, no todos los ecotipos han seleccionado poblaciones virales igualmente generalistas. Existe una correlación negativa significativa entre la eficacia en huéspedes locales y alternativos ($r_s = -0,900$, 3 g.l., 1 cola $P = 0,019$). Los huéspedes más permisivos, como St-0, han seleccionado virus más especializados, mientras que los huéspedes menos permisivos, como es el caso de Ei-2, han seleccionado virus más generalistas.

3.2.3 Evolución de infectividad y virulencia

El análisis de los datos de infectividad (I) muestra unos resultados similares a los obtenidos con las estimas de eficacia. Los datos se ajustaron a un modelo de regresión logística dado por la ecuación (2), definida en la sección de Materiales y Métodos.

Se detectó un efecto muy significativo para LH (aunque de tamaño medio, $\eta_p^2 < 0.15$), así como para TH y la interacción de ambos (en estos dos casos, de gran efecto, Tabla 10), sugiriendo que las restricciones selectivas impuestas por el huésped local afectan posteriormente a la infectividad en los huéspedes alternativos. En general, los linajes evolucionados en $St-0$ muestran los valores más altos de infectividad en sus huéspedes locales, mientras que los linajes evolucionados en $Ei-2$ muestran el porcentaje más bajo de infectividad en su ecotipo local (Figura 9B). Respecto a los ecotipos alternativos, en promedio, los niveles más altos de infectividad se encontraron en $Ei-2$ y los más bajos en $Ler-0$.

Los datos de virulencia (V) se ajustaron a la ecuación (3), que viene definida en la sección de Materiales y Métodos. Se detectó un efecto altamente significativo para LH , TH , así como para su interacción (el tamaño de los efectos de LH y TH era medio, pero grande para la interacción, Tabla 10), lo que sugiere que la virulencia de los linajes evolucionados depende, tanto del ecotipo de huésped local en el que un linaje ha evolucionado, como del ecotipo huésped en el que se mide la virulencia. Los linajes evolucionados en $Di-2$ y $St-0$ son los menos virulentos en sus ecotipos locales, mientras que los linajes evolucionados en $Ei-2$ son los más virulentos (Figura 9C). Los linajes evolucionados en $St-0$, $Ler-0$, $Wt-1$, y $Di-2$ muestran síntomas más fuertes en huéspedes alternativos que en su huésped local, mientras que los linajes evolucionados en el ecotipo $Ei-2$ tienden a inducir más tolerancia a la infección en los linajes que no han sido seleccionados en la evolución experimental (Figura 9C). Por lo tanto, de la Figura 9C se puede concluir que los linajes más virulentos en su huésped local son menos virulentos en sus huéspedes alternativos.

3.2.4 Análisis de redes de infección

La matriz de infección mostrada en la Figura 10A tiene una temperatura $T = 13,3$, un valor que es significativamente inferior al esperado según el modelo nulo ($E(T) = 29,343 \pm 8,487$, $P = 0,019$), es decir, la matriz de infección se anida significativamente. La modularidad de la red bipartita mostrada en la Figura 10B era baja (0,192) y no significativamente diferente de lo esperado según el modelo nulo ($E(Q) = 0,222 \pm 0,049$, $P = 0,202$).

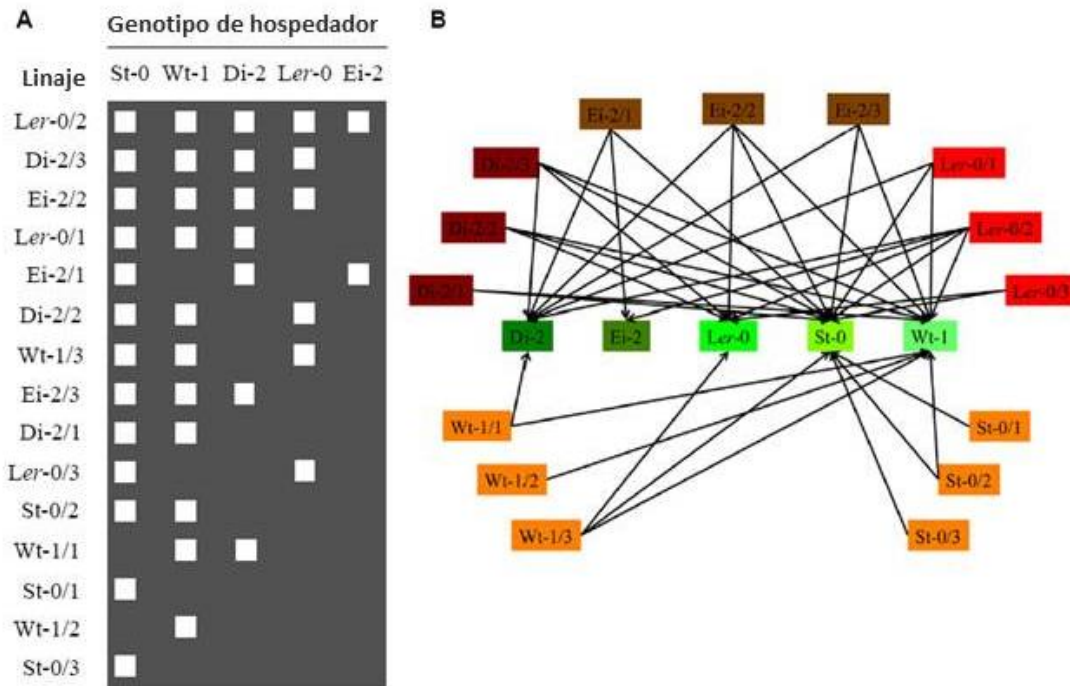


Figura 10. (A) Representación de la matriz de interacciones de los ecotipos del huésped con los linajes de virus evolucionados. Las filas representan los 15 linajes virales y las columnas representan los cinco ecotipos de huésped. Las celdas blancas representan combinaciones en las cuales la eficacia del virus evolucionado es significativamente mayor que la del virus ancestral TEV-At17b. **(B)** Representación bipartita de la matriz de infección.

La primera fila de la matriz corresponde al linaje Ler-0/2 (Figura 10A). Éste es el único linaje cuya eficacia aumentó a través de todos los ecotipos de huéspedes en comparación con el virus ancestral, aunque este linaje no tuvo cambios adaptativos fijos. El resto de linajes del virus evolucionado presentaron eficacias más bajas que el aislado ancestral en al menos un ecotipo del huésped. En un extremo, Di-2/3 y Ei-2/2 son peores que la cepa ancestral sólo en Ei-2. En el otro extremo, St-0/1, St-0/3 y Wt-1/2 son peores que el ancestral en todos los ecotipos menos uno. La primera columna en la matriz de infección (Figura 5A) corresponde al ecotipo St-0, que era el más susceptible, ya que era infectado con éxito por trece linajes virales evolucionados. Por otro lado, el ecotipo más resistente era Ei-2, que fue infectado con éxito sólo por dos de los linajes virales. Por lo tanto, estos resultados nos llevan a la conclusión de que durante el experimento de evolución la matriz de interacción virus-huésped no evolucionó hacia la modularidad, sino hacia un anidamiento significativo, tal y como predice el modelo gen a gen.

3.2.5 Evolución genómica

Todas las secuencias consenso de los linajes evolucionados, excepto el linaje Ler-0/2, mostraron al menos una mutación. En total, se observaron 79 mutaciones independientes que ocurrían en 62 sitios diferentes, con un rango de dos a ocho mutaciones por linaje (Figura 11).

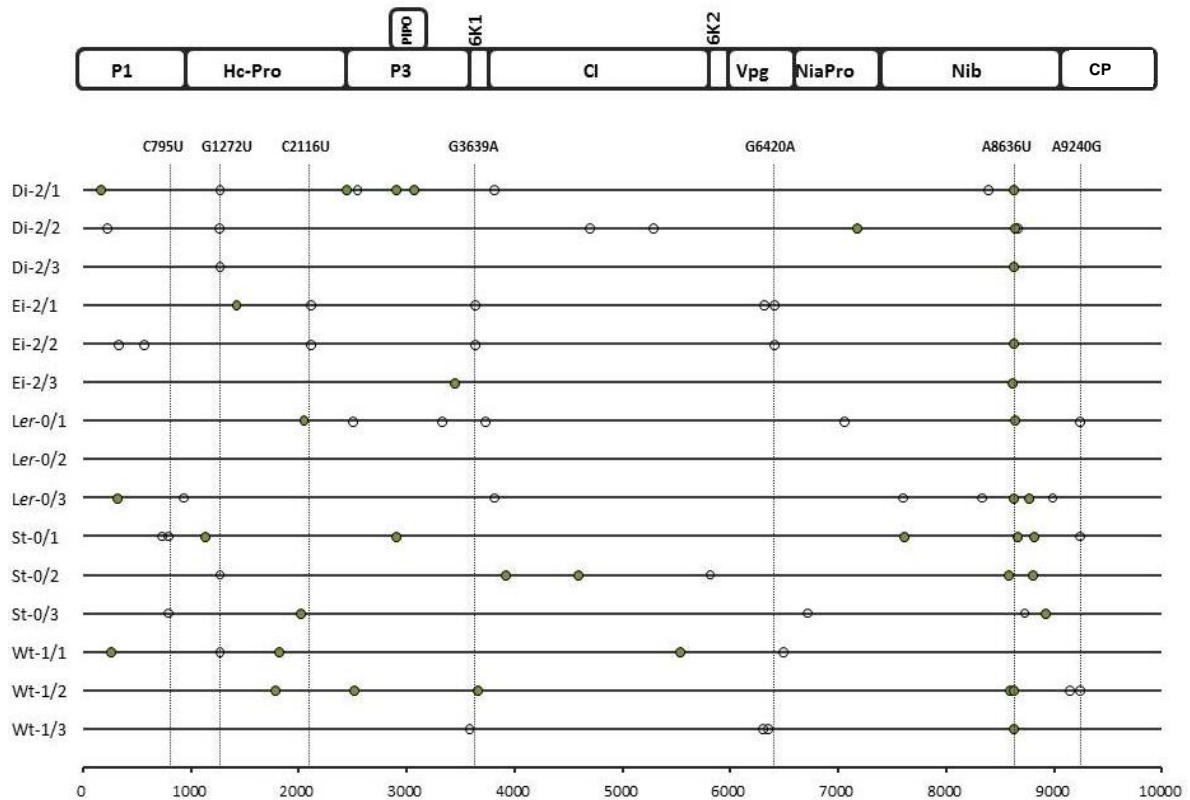


Figura 11. Esquema las mutaciones detectadas durante la evolución experimental en 14 linajes (el linaje Ler-0/2 no presentó ninguna mutación). En la parte superior se representa el genoma completo de TEV con la posición de sus 11 genes. Las líneas inferiores representan los genomas individuales de los linajes obtenidos tras el experimento de evolución. Los círculos vacíos representan mutaciones sinónimas y los círculos rellenos mutaciones no sinónimas. Las mutaciones convergentes están definidas por líneas verticales e indicadas en la zona superior de los linajes.

Del total de mutaciones, 43 eran mutaciones sinónimas y 36 no sinónimas. Veinticuatro mutaciones no eran únicas y de los 62 sitios polimórficos identificados, siete se presentaron en múltiples linajes independientes. Tres de estas mutaciones no únicas eran exclusivas de linajes Ei-2 (C2116U, G3639A y G6420A). La mutación C795U era compartida por dos linajes St-0. La mutación G1272U era compartida por los linajes Di-2 y St-0. La mutación A9240G era común a los linajes Ler-0/1, St-0/1 y Wt-1/2. Estas seis mutaciones convergentes eran sinónimas. Las mutaciones no sinónimas C2912A (A923D en el cistron P3 y L923I en el cistron solapante P3N-PIPO) eran compartidas por los linajes Di-2/1 y St-0/1. La mutación no sinónima C8636U (S2831L en el cistron CP) era compartida por todos los linajes evolucionados en Di-2, Ler-0 y por Wt-1/3. Además, los linajes Ei-2/2 (C8624U) y Ei-2/3 (U8623C) presentaban sustituciones en nucleótidos contiguos que afectaban al mismo codón en el cistron CP, pero que daban lugar a diferentes sustituciones de aminoácidos (S2827L y S2827P).

Las mutaciones convergentes se reflejan en el patrón de agrupamiento que se muestra en un árbol filogenético de máxima verosimilitud, donde grupos más diversos se alternan con grupos definidos por un huésped común (Figura 12).

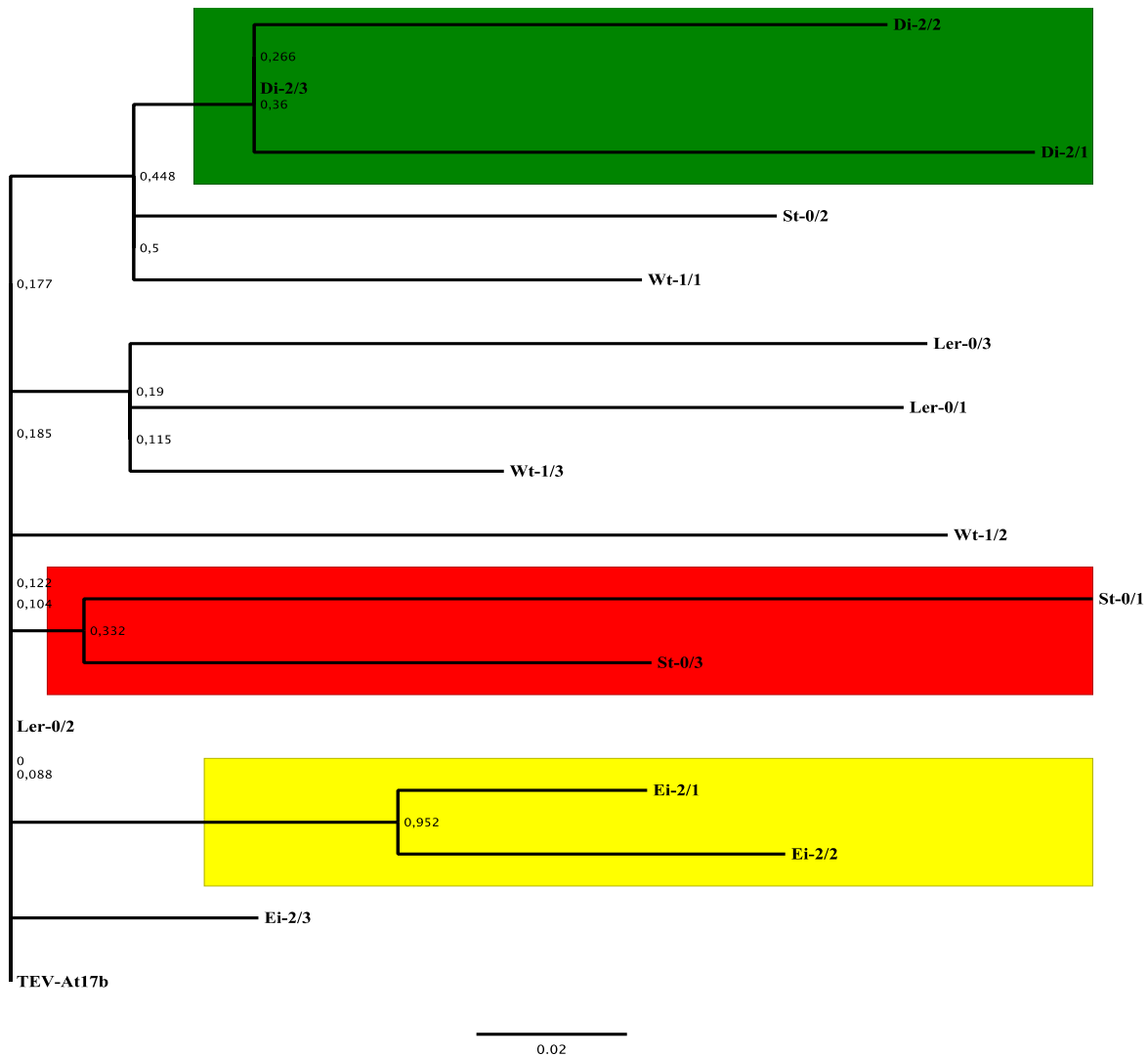


Figura 12. Árbol filogenético de máxima verosimilitud obtenido para las secuencias genómicas de los linajes evolucionados. Se empleó un modelo de sustitución de Kimura 2-parámetros. La secuencia ancestral TEV-At17b se incluyó para fines de enraizamiento. Los números en los nodos corresponden a los valores de apoyo *Bootstrap*. Los clústeres de aislados evolucionados en un huésped común se resaltan con colores diferentes.

Las mutaciones sinónimas y no sinónimas se distribuyen de manera uniforme entre las mutaciones comunes y únicas (prueba exacta de Fisher, $P = 0,260$). Tratando cada linaje como una observación y cada ecotipo de huésped como una subpoblación, la diversidad nucleotídica promedio dentro de huésped es $\hat{\pi}_S = 0,167 \pm 0,008$ (± 1 SD, 1000 pseudomuestras generadas mediante bootstrap). Por otra parte, la diversidad nucleotídica para toda la muestra, sin tener en cuenta la estructuración por ecotipos, es $\hat{\pi}_T = 0,187 \pm 0,014$. Por lo tanto, la estimación de la diversidad nucleotídica entre huéspedes es $\hat{\delta}_{ST} = 0,019 \pm 0,010$. Consecuentemente, la estimación de la proporción de

la diversidad nucleotídica entre huésped, conocida como el coeficiente de diferenciación nucleotídica¹⁷⁷, es $\hat{\delta}_{ST} = \hat{\delta}_{ST}/\hat{\pi}_T = 0,103 \pm 0,043$, un valor significativamente mayor que cero ($z = 2,395$, 1 cola $P = 0,004$). Por lo tanto, esto sugiere que esta pequeña pero significativa diferenciación genética ha sido generada entre los virus que se replican en diferentes ecotipos del huésped.

Para evaluar si la selección jugaba un papel en la diferenciación genética entre los ecotipos del huésped, se realizó una prueba D ²³⁴. El propósito de esta prueba es distinguir entre una secuencia de aminoácidos que evoluciona al azar y una cuya evolución no es aleatoria. Las mutaciones al azar son aquellas que no tienen ningún efecto en la supervivencia de un organismo. Si una población presenta un tamaño constante con una tasa de mutación constante, la población alcanzará un equilibrio de las frecuencias génicas. Este equilibrio tiene propiedades importantes, incluyendo el número de sitios segregantes o polimórficos y el número de nucleótidos diferentes entre pares de secuencias. El propósito de la prueba de Tajima es identificar secuencias que no encajan en el modelo de la Teoría neutral en el equilibrio entre mutación y deriva genética. En este caso, la prueba D mostró diferencias significativas ($D = -2,172$, $P = 0,015$), por lo que se rechazó la hipótesis de neutralidad de mutaciones.

3.2.6 Asociación entre diversidad molecular y variación genética para la eficacia.

Para testar la evolución adaptativa, observamos la correlación entre la diversidad genética de los virus evolucionados dentro de cada huésped local y la varianza genética de la eficacia $\sigma_G(W)$ entre los linajes paralelos. Si la diversidad genética fuera neutral, no se esperaría una correlación entre estos dos rasgos. Por el contrario, si las diferencias genéticas entre los linajes se tradujeran en diferencias en W , entonces esperaríamos una correlación positiva significativa. Para realizar esta prueba, primero se calculó la diversidad genética promedio entre los linajes dentro de cada huésped local (Figura 13A).

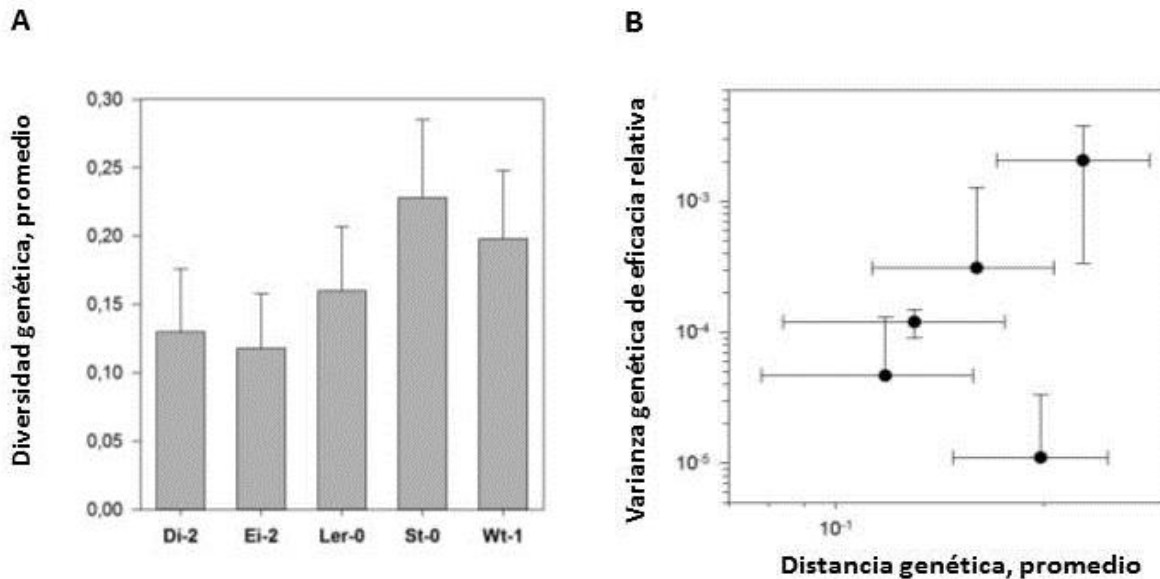


Figura 13. (A) Diversidad genética promedio del virus dentro de cada ecotipo de huésped. **(B)** Asociación entre la diversidad genética promedio y el componente genético de la varianza para la eficacia relativa (datos de la tabla 8). Los ejes en ambos paneles están en escala logarítmica. Las barras de error representan ± 1 SEM.

Se encontró que la diversidad genética más baja la tenían los linajes de ecotipo Ei-2 y la más alta los linajes de ecotipo St-0. La Figura 13B muestra la asociación entre la diversidad genética y la varianza genética de la eficacia ($\sigma_G(W)$). No se observó una correlación positiva significativa para todo el conjunto de datos ($r_S = 0,400$, 3 g.l., 1 cola $P = 0,253$). Sin embargo, al eliminar del análisis los datos correspondientes al ecotipo Wt-1, la correlación se vuelve significativa ($r_S = 1,000$, 2 g.l., 1 cola $P < 0,001$). Este resultado sugería que, con la excepción de los linajes evolucionados en Wt-1, las diferencias genómicas entre los linajes evolucionados en un huésped local dado explican buena parte de las diferencias genéticas en la eficacia relativa entre linajes. Por lo tanto, esto nos lleva a la conclusión de que algunas de las mutaciones observadas son beneficiosas y explican las diferencias en la eficacia entre los linajes en su ecotipo local.

3.2.7 Selección para la eficiencia traduccional

Una posible explicación para la convergencia en sitios sinónimos es que la selección para la eficiencia de la traducción daría lugar a la sustitución de codones poco utilizados por los sinónimos de los cuales la célula huésped tiene una gran reserva de ARNt.

La Tabla 11 incluye los codones ancestrales y mutados, así como la frecuencia de uso (f) para *A. thaliana*. Para cada una de las 62 mutaciones, se calculó el cambio relativo en el uso entre los codones evolucionados y ancestrales $C = f_{evolucionado} / f_{ancestral} - 1$.

Tabla 11. Lista completa de las mutaciones en los 15 linajes evolucionados independientemente.

Virus	Cambio	Cambio de codón	Cambio de	Proteína
	nucleotídico	(índice de uso de codones ¹)	aminoácido	
Di-2/1	A175G	<u>A</u> AC(20.9)/ <u>G</u> AC(17.2)	N11D	P1
	G1272U	UC <u>G</u> (9.3)/UC <u>U</u> (25.2)	sinónimo	HC-Pro
	G2451A	AU <u>G</u> (24.5)/AU <u>A</u> (12.6)	M769I	P3
	C2545U	<u>C</u> UG(9.8)/ <u>U</u> UG(20.9)	sinónimo	P3
	C2912A	G <u>C</u> U(28.3)/G <u>A</u> U(36.6)	A923D (P3), L923I (PIPO)	P3N-PIPO
	U3077C	UU <u>G</u> (20.9)/U <u>C</u> G(9.3)	L978S	P3
	G8388A	G <u>C</u> G(9.0)/G <u>C</u> A(17.5)	sinónimo	Nib
	C8636U	U <u>C</u> A(18.3)/U <u>U</u> A(12.7)	S2831L	CP
	Di-2/2	U231C	CA <u>U</u> (13.8)/CA <u>C</u> (8.7)	sinónimo
G1272U		UC <u>G</u> (9.3)/UC <u>U</u> (25.2)	sinónimo	HC-Pro
C4707U		CA <u>C</u> (8.7)/CA <u>U</u> (13.8)	sinónimo	CI
U5289C		CA <u>U</u> (13.8)/CA <u>C</u> (8.7)	sinónimo	CI
G7178A		A <u>G</u> A(19.0)/A <u>A</u> A(30.8)	R4345K	Nib
A8667G		A <u>A</u> A(30.8)/A <u>A</u> G(32.7)	sinónimo	CP
C8636U		U <u>C</u> A(18.3)/U <u>U</u> A(12.7)	S2831L	CP
Di-2/3	G1272U	UC <u>G</u> (9.3)/UC <u>U</u> (25.2)	sinónimo	HC-Pro
	C8636U	U <u>C</u> A(18.3)/U <u>U</u> A(12.7)	S2831L	CP
Ei-2/1	C1425U	CU <u>C</u> (16.1)/CU <u>U</u> (24.1)	sinónimo	HC-Pro
	C2116U	<u>C</u> UA(9.9)/ <u>U</u> UA(12.7)	sinónimo	HC-Pro
	G3639A	UU <u>G</u> (20.9)/UU <u>A</u> (12.7)	sinónimo	CI
	U6324C	AA <u>U</u> (22.3)/AA <u>C</u> (20.9)	sinónimo	NlaPro
	G6420A	CU <u>G</u> (9.8)/CU <u>A</u> (9.9)	sinónimo	NlaPro
Ei-2/2	U342C	CC <u>U</u> (18.7)/CC <u>C</u> (5.3)	sinónimo	P1
	U579C	G <u>C</u> U(28.3)/G <u>C</u> C(10.3)	sinónimo	P1
	C2116U	<u>C</u> UA(9.9)/ <u>U</u> UA(12.7)	sinónimo	HC-Pro
	G3639A	UU <u>G</u> (20.9)/UU <u>A</u> (12.7)	sinónimo	CI
	G6420A	CU <u>G</u> (9.8)/CU <u>A</u> (9.9)	sinónimo	NlaPro
	C8624U	U <u>C</u> A(18.3)/U <u>U</u> A(12.7)	S2827L	CP
Ei-2/3	G3456U	A <u>A</u> G(32.7)/AA <u>U</u> (22.3)	K1104N	P3
	U8623C	<u>U</u> CA(18.3)/ <u>C</u> CA(16.1)	S2827P	CP
Ler-0/1	C2050U	<u>C</u> UC(16.1)/ <u>U</u> UC(20.7)	L636F	HC-Pro
	G2511A	A <u>A</u> G(32.7)/A <u>A</u> A(30.8)	sinónimo	P3
	U3339C	AA <u>U</u> (22.3)/AA <u>C</u> (20.9)	sinónimo	P3
	G3738A	A <u>A</u> G(32.7)/A <u>A</u> A(30.8)	sinónimo	CI
	C7059U	G <u>U</u> C(12.8)/G <u>U</u> U(27.2)	sinónimo	Nib
	C8636U	U <u>C</u> A(18.3)/U <u>U</u> A(12.7)	S2831L	CP

	A9240G	GAA(34.3)/GAG(32.2)	sinónimo	CP
Ler-0/2		sin mutaciones		
Ler-0/3	A322U	AUU(21.5)/UUU(21.8)	I60F	P1
	G936A	UCG(9.3)/UCA(18.3)	sinónimo	P1
	C7602U	UUC(20.7)/UUU(21.8)	sinónimo	NIb
	C8331U	UUC(20.7)/UUU(21.8)	sinónimo	NIb
	C8636U	UCA(18.3)/UUA(12.7)	S2831L	CP
	G8779A	GCC(10.3)/ACC(10.3)	A2879T	CP
	U8985A	AUU(21.5)/AUA(12.6)	sinónimo	CP
St-0/1	G738A	AGG(11.0)/AGA(19.0)	sinónimo	P1
	C795U	GGC(9.2)/GGU(22.2)	sinónimo	P1
	A1139C	GAG(32.2)/GCG(9.0)	E332A	HC-Pro
	C2912A	GCU(9.0)/GAU(36.6)	A923D (P3), L923I (PIPO)	P3N-PIPO
	C7621U	CUC(16.1)/UUC(20.7)	L2493F	NIb
	U8669C	CUU(24.1)/CCU(18.7)	L2842P	CP
	G8821A	GAA(34.3)/AAA(30.8)	E2893K	CP
	A9240G	GAA(34.3)/GAG(32.2)	sinónimo	CP
St-0/2	G1272U	UCG(9.3)/UCU(25.2)	sinónimo	HC-Pro
	C3922U	CAU(13.8)/UAU(14.6)	H1260Y	CI
	C4601U	ACU(17.5)/AUU(21.5)	T1486I	CI
	C5811U	AAC(20.9)/AAU(22.3)	sinónimo	VPg
	C8587G	CAG(15.2)/GAG(32.2)	Q2815E	CP
	G8815A	GUG(17.4)/AUG(24.5)	V2891M	CP
St-0/3	C795U	GGC(9.2)/GGU(22.2)	sinónimo	P1
	A2024G	GAU(36.6)/GGU(22.2)	D627G	HC-Pro
	U6718C	UUA(12.7)/CUA(9.9)	sinónimo	NlaPro
	C8727U	UAC(13.7)/UAU(14.6)	sinónimo	CP
	A8924C	GAG(32.2)/GCG(9.0)	E2927A	CP
Wt-1/1	A269G	GAA(34.3)/GGA(24.2)	E42G	P1
	G1272U	UCG(9.3)/UCU(25.2)	sinónimo	HC-Pro
	A1822C	AUG(24.5)/CUG(9.8)	M560L	HC-Pro
	A5539G	AGC(11.3)/GGC(9.2)	S1799G	6K2
	G6492A	GGG(10.2)/GGA(24.2)	sinónimo	NlaPro
Wt-1/2	G1789A	GGU(22.2)/AGU(14.0)	G549S	HC-Pro
	A2522C	GAG(32.2)/GCG(9.0)	E793A	P3
	A3670U	AUG(24.5)/UUG(20.9)	M1176L	CI
	C8594U	UCA(18.3)/UUA(12.7)	S2817L	CP
	C8630U	ACA(15.7)/AUA(12.6)	T2829I	CP

	A9144G	AGA(19.0)/AGG(11.0)	sinónimo	CP
	A9240G	GAA(34.3)/GAG(32.2)	sinónimo	CP
Wt-1/3	G3585A	AAG(32.7)/AAA(30.8)	sinónimo	6K1
	C6306U	ACC(10.3)/ACU(17.5)	sinónimo	NlaPro
	G6348A	UCG(9.3)/UCA(18.3)	sinónimo	NlaPro
	C8636U	UCA(18.3)/UUA(12.7)	S2831L	CP

¹ tomado de base de datos <http://www.kazusa.or.jp/codon>; expresado como frecuencia por mil.

Los valores de $C > 0$ significan que la mutación transforma un codón en uno más utilizado, mientras que valores $C < 0$ implican que las mutaciones van a un codón más raro. Si la hipótesis de selección a nivel traduccional fuera cierta, entonces cabría esperar que el valor de \bar{C} para mutaciones convergentes sinónimas fuera positivo y significativamente mayor que el valor observado para todos los demás tipos de mutaciones.

En este caso, para mutaciones convergentes sinónimas $\bar{C} = 0,494 \pm 0,351$ ($\pm 1\text{SEM}$), mientras para el resto de las mutaciones $\bar{C} = -0,009 \pm 0,066$, cuyas diferencias eran significativas (prueba t para dos muestras, $t_{60} = 2,185$, 1 cola $P = 0,016$). Por lo tanto, las mutaciones convergentes sinónimas que se fijaron durante la evolución habrían dado lugar a codones que eran ~50% más utilizados por la maquinaria de traducción de *A. thaliana* que los codones ancestrales.

3.2.8 Discusión

3.2.8.1 Expansión de gama de huéspedes: especialistas y generalistas

La evolución de la gama de huéspedes de los virus de ARN ha recibido considerable atención debido a su implicación en las enfermedades infecciosas emergentes. Los virus de plantas tienen gamas de huéspedes muy variables: algunos son especialistas que infectan sólo una o pocas especies afines, mientras que otros son generalistas que infectan una amplia gama de huéspedes de diferentes grupos taxonómicos. La mayoría de los estudios previos sobre la evolución de la gama de huéspedes de virus de plantas exploraron el efecto sobre rasgos de eficacia de los virus que evolucionan en una única especie huésped^{5,6,19,42,201,248,267} o entre dos especies de huéspedes alternativos¹⁹. De este modo ha podido confirmarse con frecuencia la acción de la pleiotropía antagonista, es decir, la situación en la que los virus evolucionados se desenvuelven peor en el huésped original que sus antepasados. En consecuencia, la pleiotropía antagonista limita el rango de adaptación y promueve la evolución de la especialización ecológica¹⁹⁸. Otra observación, extensible a virus animales evolucionados en cultivos celulares, es que la alternancia frecuente entre distintas especies de huésped resulta en la evolución de virus

generalistas con aumento de eficacia en todos los huéspedes alternativos y sin aparente coste de eficacia¹⁹. Sin embargo, el efecto de las diferencias genéticas dentro de cada especie en la evolución de los virus de ARN de plantas ha recibido poca atención. Esto es particularmente sorprendente, dada la importancia de la diversidad genética del huésped en la aparición, propagación y prevalencia de las enfermedades infecciosas, algo que únicamente ha sido tratado en un trabajo que proviene del campo de la agronomía¹⁷⁰. Con este trabajo, se ha dado un paso más al analizar el efecto de las diferencias genéticas entre los ecotipos de *A. thaliana* en la evolución de un virus emergente. Las diferencias genéticas entre los ecotipos crean nuevos desafíos para el virus. St-0 y Wt-1 son los ecotipos en los que linajes evolucionados a nivel local muestran un mayor aumento de la eficacia, por lo que no es sorprendente que la contribución de la divergencia genética frente al paralelismo sea particularmente pronunciada en estos dos ecotipos. Una observación inesperada es que la magnitud de las mejoras de eficacia en los huéspedes locales no era dependiente de la constitución genética de los loci RTM. Dado que sólo los ecotipos que llevan alelos *rtm* eran susceptibles a la infección con la cepa ancestral de TEV, se esperaría más margen de mejora en la eficacia para los ecotipos Di-2 y Wt-1, algo que no ha sucedido. Una posible explicación estaría en que, durante la adaptación a *Ler-0*, TEV-At17b adquiriera la capacidad de infectar sistémicamente *Ler-0*, superando la resistencia mediada por RTM de tal manera que las diferencias en estos tres loci ya no representarían un desafío.

Como se describe en el primer capítulo de esta Tesis, los ecotipos de *A. thaliana* se clasificaron en dos grupos en función de su respuesta transcriptómica a la infección con la cepa ancestral TEV-At17b. La incorporación de estos dos grupos como un factor en los modelos lineales no resultó en mejoras en la capacidad explicativa (por ejemplo, para los datos de *W*, BIC = -7137,105 para el modelo aquí presentado y BIC = -4101,527 para el modelo que incorpora un factor adicional). Por lo tanto, el resultado de los experimentos de evolución no parece verse afectado por las diferencias en los síntomas o en los perfiles de transcripción de los ecotipos. De hecho, las diferencias de eficacia observadas entre distintos linajes pueden representar adaptación a una, o más probablemente, a varias de las pequeñas diferencias entre los componentes genéticos o bioquímicos de los ecotipos.

El análisis de la matriz de infección permite concluir que la adaptación a Di-2 y Ei-2 se produce sin ningún coste de eficacia respecto al ecotipo original *Ler-0*, mientras que los pases seriados en St-0 y Wt-1 suponen un coste significativo. Esto sugiere que los mecanismos de adaptación en Di-2 y Ei-2 son más similares a los de *Ler-0* y pueden ser más diferentes que los de St-0 y Wt-1. Cuando se estimó la eficacia de los linajes evolucionados en los cinco ecotipos, se encontró que los linajes evolucionados en Di-2 mostraron mayor eficacia para todos los huéspedes que el resto de los linajes. En este

sentido, los linajes evolucionados en Di-2 pueden ser considerados como los más generalistas, mientras que los linajes evolucionados en St-0 son los más especializados.

Por lo general, se asume que un generalista no puede evolucionar en presencia de compensaciones de eficacia a través de los huéspedes y por lo tanto un especialista en un huésped dado siempre podría superar al generalista en la competencia por el mismo huésped. Como consecuencia, no cabría esperar un linaje único que tuviera la eficacia más alta en todos los ecotipos. Sin embargo, Remold¹⁹⁸ presentó un modelo fundamentado en la pleiotropía epistática para explicar la evolución de los especialistas y generalistas sin coste. La pleiotropía epistática se produce cuando los fondos genéticos virales difieren en cómo el efecto de un alelo depende del huésped. Bajo pleiotropía epistática, las poblaciones virales pueden lograr la especialización o la generalización sin coste, dependiendo del huésped en el que evolucionan, a pesar de la existencia de compensaciones. De acuerdo con estas expectativas, se observó que ningún linaje viral fue superior en los cinco ecotipos y que algunos aislados pagaron un coste de eficacia, mientras que otros no lo hicieron.

Los datos anteriores han demostrado que la huella de la epistasia recíproca está omnipresente en el genoma de TEV¹³⁹ y que la epistasia depende de la especie del huésped en el que se evalúa¹⁴⁰. Esto significa que los linajes especializados no serían capaces de evolucionar hacia generalistas sin un coste, es decir, sin tener que cruzar un valle de eficacia.

Ningún linaje viral evolucionado era superior al resto de linajes en todos los ecotipos del huésped. Del mismo modo, no se observó un único ecotipo de huésped superior a todos los demás en resistencia al conjunto de linajes virales (Figura 10A). Bajo tales condiciones, podría ocurrir que los ecotipos de huéspedes y las frecuencias génicas de los virus fueran críticas para la evolución de las poblaciones naturales. Dicho de otro modo, la selección favorecería aquellos genotipos virales capaces de infectar a los ecotipos de huéspedes más susceptibles, como es el caso de St-0, donde se permitiría mayor replicación y acumulación viral. No obstante, estos ecotipos serían posteriormente desfavorecidos. Por lo tanto, en este sistema puede surgir la selección dependiente de frecuencia.

Debemos esperar modularidad en redes de infección cuando los huéspedes y patógenos preferentemente se infectan de forma cruzada dentro de un grupo²⁵⁴. Por el contrario, debemos esperar anidamiento si existen jerarquías en resistencia entre los huéspedes y en la capacidad de infección entre los virus²⁵⁴. Nuestro trabajo demostró que la red de infección (Figura 10B) no era modular, sino significativamente anidada. La falta de modularidad se puede considerar un indicio de que los ecotipos no son similares en su respuesta a la infección de ciertos linajes virales, por lo que cada ecotipo responde de una manera particular. El modelo de interacción *gen a gen* entre TEV y *A. thaliana*

predice la aparición de anidamiento. Este mecanismo implica que existan mutaciones que incrementen la eficacia en el nuevo huésped local, pero que no paguen un coste de eficacia en *Ler-0*. Por lo tanto, el conjunto de huéspedes que un linaje puede infectar serían subconjuntos de uno en otro.

Bajo la visión de un mecanismo de alelos coincidentes, la adquisición de la capacidad de infectar un ecotipo dado podría hacer que el virus perdiera por completo la capacidad de infectar *Ler-0*. Como ya se comentó anteriormente, sin embargo, existe un coste para los linajes que han experimentado el mayor aumento de la eficacia en su huésped local (*St-0* y *Wt-1*), pero no para los linajes que muestran aumentos menores de la eficacia en sus huéspedes locales (*Di-2* y *Ei-2*). Por lo tanto, nuestros datos no coinciden por completo con ninguno de estos dos modelos extremos, sino que podrían ser explicados mediante algún mecanismo intermedio por el cual los linajes de TEV desarrollaran la capacidad de infectar a nuevos huéspedes y perdieran su infectividad en el ecotipo original sólo bajo ciertas condiciones.

3.2.8.2 Equilibrio entre eficacia, virulencia e infectividad

Los linajes evolucionados en *Di-2* alcanzaron la eficacia más alta en todos los ecotipos de huéspedes, pero ese aumento de eficacia no estaba acompañado por aumento en infectividad ni en virulencia. Por el contrario, los linajes evolucionados en *Ei-2* tenían mayor eficacia en su huésped local comparando con todos los demás ecotipos, siendo los menos infecciosos pero los más virulentos. Los linajes evolucionados en *St-0* fueron los más infecciosos, pero al mismo tiempo menos virulentos, en su huésped local. De hecho, existe una correlación negativa significativa entre virulencia e infectividad promedio ($r_s = -0,900$, 3 g.l., $P = 0,037$). Desde la perspectiva de la permisividad a la infección para los linajes evolucionados en otros ecotipos, en promedio, la infectividad más alta se observó para *Ei-2*. Este genotipo también desarrolló los síntomas más débiles.

Dado que la virulencia no representa ninguna ventaja clara para el parásito, la explicación por la que la mayoría de los parásitos inducen síntomas en sus huéspedes es una cuestión pertinente. Una suposición común es que la virulencia es una consecuencia inevitable de la replicación del parásito¹⁴² y por lo tanto debe existir una asociación positiva entre virulencia y acumulación. Tal asociación se ha observado anteriormente en pimientos infectados por TEV⁶ y en nabos infectados por el virus del mosaico de la coliflor (CaMV)⁷⁰. En el presente trabajo, no se vio esta asociación. Del mismo modo, tampoco se había encontrado esta asociación para genotipos de TEV evaluados en su huésped natural y que diferían en mutaciones puntuales³³. Esta aparente contradicción sugiere que la asociación positiva puede ser dependiente del patosistema. Dos razones podrían explicar la falta de correlación positiva. En primer

lugar, las estimaciones de virulencia son demasiado ruidosas para obtener inferencias estadísticas fiables. En este sentido, la mayor parte de la variación fenotípica observada (71,9%) no se explica por las diferencias genéticas entre los aislados, sino que se atribuye al ruido. En segundo lugar, no existe una correlación y muchos otros factores que influyeran en la progresión de la infección viral explicarían la virulencia. En particular, la virulencia no dependería de la replicación en el huésped si la extensión del daño no fuera proporcional a la cantidad de partículas virales, como es el caso de una respuesta hipersensible¹⁶⁹, si la expresión de la vía de resistencia sistémica adquirida fuera costosa¹¹³, o si la asignación de recursos a la defensa perjudicara el crecimiento vegetativo o el esfuerzo reproductivo¹⁸⁶.

Una explicación adaptativa de la evolución de la virulencia es la hipótesis de equilibrio, que propone que la virulencia debe correlacionar positivamente con la transmisión¹⁰. Dependiendo de la forma de la función de virulencia-transmisión, la hipótesis de equilibrio predice que la virulencia puede evolucionar a niveles máximos o bien llegar a un óptimo intermedio. Lo último se produce cuando los costes de la virulencia no están compensados por los beneficios de un aumento adicional en la transmisión⁹⁰. En un ejemplo de virus de plantas, se encontró para CaMV una correlación positiva entre la transmisión y la virulencia⁷⁰. En contraste, en nuestro experimento hubo una correlación negativa significativa entre virulencia e infectividad. Durante nuestro experimento de transmisión artificial, los linajes fueron transmitidos mecánicamente independientemente de su virulencia, lo que podría haber atenuado el compromiso y dar lugar a la evolución independiente de virulencia e infectividad.

3.2.8.3 Bases moleculares de adaptación

La caracterización de las secuencias consenso del genoma completo de los linajes evolucionados reveló algunas características interesantes. En primer lugar, los aislados de TEV evolucionados contenían un número variable de sustituciones nucleotídicas, incluyendo cambios sinónimos y no sinónimos. En segundo lugar, se generó una diferenciación genética significativa entre los linajes durante el experimento de evolución. Se evaluó si esta divergencia era debida a la acción de la selección natural y se encontró un valor significativo y negativo de la D de Tajima. Valores negativos para D son compatibles con tres explicaciones: (i) la segregación de mutaciones levemente deletéreas, (ii) la selección purificadora, que es la eliminación selectiva de alelos perjudiciales y (iii) una rápida expansión de la población. En una población que se expande, las nuevas mutaciones pueden ser segregadas y se pueden observar como individuales. En nuestro caso, hemos observado 55 mutaciones individuales que han aumentado el número de sitios segregantes y han provocado valores de $D < 0$.

Tenemos otras evidencias que apoyan la evolución adaptativa a nivel genómico. Primero, se observaron cambios específicos y significativos de eficacia relativa en ecotipos de huéspedes. En segundo lugar, se observó una correlación positiva significativa entre la diversidad genética y la varianza de la eficacia relativa $\sigma_G(W)$. Esta asociación positiva entre la diversidad genética y las diferencias de eficacia entre los linajes se explica por el valor adaptativo de algunas de las mutaciones. En tercer lugar, observamos una serie de mutaciones convergentes, muchas de las cuales parecen ser dependientes del ecotipo del huésped. En cuarto lugar, se observó una diferenciación genómica significativa de las poblaciones virales que evolucionaron en los diferentes ecotipos del huésped.

Mientras que la evolución convergente en sitios no sinónimos se explica como consecuencia de presiones selectivas idénticas y por la existencia de vías adaptativas limitadas, la convergencia molecular en sitios sinónimos es más problemática de explicar. La evolución convergente en sitios sinónimos ha sido observada a menudo en trabajos de evolución experimental con virus de ARN^{1,26,28,55,179,199,261}, así como en el desarrollo de resistencias a medicamentos antivirales^{159,178}, pero estas observaciones aún son escasas para virus de plantas (e.g., Lafforgue *et al.*¹³⁵). Además, estudios que analizan los paisajes de mutación de virus de ARN han demostrado efectos significativos de eficacia asociados a mutaciones sinónimas^{1,33,54,214}, lo que apoya la idea de que asumir neutralidad en sustituciones sinónimas no es necesariamente válido, al menos para los genomas de ARN.

Tres explicaciones mutuamente no excluyentes se presentan para explicar la evolución convergente en sitios sinónimos: (i) la necesidad de preservar las estructuras secundarias de ARN reguladoras, (ii) la prevención de la formación de estructuras largas de ARN de doble cadena, lo que podría ser un objeto de silenciamiento de ARN y (iii) la selección de la eficiencia de traducción, que consistiría en promover la sustitución de codones raros por sinónimos para los que la célula huésped dispusiera de una mayor cantidad de ARNt. Sin descartar las hipótesis (i) y (ii), nuestros resultados sugieren que la selección para la eficiencia de la traducción podría explicar en gran parte las mutaciones sinónimas convergentes.

La mayoría de los experimentos con virus de plantas que analizan la evolución de virus generalistas y especialistas se basan en diferentes especies de huésped. En nuestro estudio, hemos analizado el efecto que puede tener la variabilidad genética intraespecífica respecto a la susceptibilidad en la adaptación viral. Algunos ecotipos de plantas seleccionaron virus más generalistas, mientras que otros ecotipos seleccionaron virus más especializados. Los huéspedes permisivos seleccionaron virus más especialistas, mientras que los huéspedes más restrictivos seleccionaron virus más generalistas. Ningún linaje viral evolucionado era superior a todos los demás en todos

los ecotipos y ningún ecotipo era resistente frente a todos los linajes evolucionados. Estas condiciones generan un anidamiento de la matriz de infección, donde se crean las condiciones para que la selección dependiente de frecuencia pueda operar a largo plazo.

A pesar de la similitud entre los huéspedes y el tiempo de evolución relativamente corto, los linajes independientes han acumulado variaciones genéticas y han divergido unos de otros. La caracterización genómica de los linajes evolucionados ha mostrado casos de mutaciones convergentes dependientes del huésped, incluyendo algunos cambios sinónimos que pueden contribuir a la mejora de la eficiencia en la traducción.

Por último, la eficacia relativa de las cepas evolucionadas es independiente de la virulencia y la infectividad. Por el contrario, virulencia e infectividad están vinculadas, de tal manera que los virus más virulentos se transmiten peor que los moderados. Este compromiso y la falta de correlación entre virulencia y eficacia relativa, tienen importantes implicaciones para la evolución de la virulencia en este patosistema.

3.3 Evaluación de efectos de las mutaciones convergentes fijadas durante la adaptación del virus a diferentes ecotipos de huésped

La variación genética en las poblaciones de huéspedes en relación a la susceptibilidad a la infección es un factor esencial para la propagación de un virus emergente. Los modelos predicen que la tasa de difusión se ralentiza al aumentar la frecuencia y la diversidad de alelos de resistencia en la población huésped. Este trabajo tiene por objeto proporcionar pruebas empíricas sobre este campo.

En el capítulo anterior, se ha descrito que las poblaciones de TEV evolucionados en cinco ecotipos de *A. thaliana* que difieren en su susceptibilidad a la infección, aumentaron su eficacia, virulencia e infectividad dentro del huésped de una manera compatible con un modelo gen a gen para las interacciones parásito-huésped: los ecotipos difíciles de infectar fueron infectados por virus generalistas, mientras que los ecotipos fáciles de infectar lo fueron por cualquier genotipo del virus. También se caracterizaron los genomas completos de los virus evolucionados y encontramos siete casos de mutaciones convergentes en varios linajes (Tabla 11). En este capítulo, se exploran las bases moleculares de la adaptación y diversificación de TEV en ecotipos de *A. thaliana*. Con este propósito, se clonaron varios genotipos de TEV que llevan las mutaciones putativamente adaptativas encontradas al final del experimento de evolución. En total se clonaron 14 genotipos: siete mutaciones individuales, seis mutantes dobles, y un triple mutante. Algunas de estas combinaciones aparecieron en los linajes al final del experimento de evolución (G1272U/C8636U, G1271U/A9240G, C8638U/A9240G, y C2116U/G3639A/G6420A, Tabla 11), mientras que otras no (C2116U/G3639A, C2116U/G6420A y G3639A/G6420A, Tabla 11), aunque podrían tener una existencia transitoria previa al punto temporal que ha sido caracterizado. Estas combinaciones no observadas han sido clonadas para explorar la topografía local del paisaje adaptativo de eficacia.

Para cada genotipo viral, hemos evaluado la eficacia dentro de huésped, tanto en el huésped local (simpátrico) como en los cuatro huéspedes alternativos (alopátricos).

Tabla 12. Lista de genotipos clonados de TEV. Todas las mutaciones se introdujeron partiendo del fondo genético del TEV-At17b.

Mutación	Cistron afectado	Linajes en los cuales ha sido encontrada la mutación
G1272U	<i>HC-Pro</i>	Di-2/1, Di-2/3, St-0/2, Wt-1/1
C8636U	<i>CP</i>	Di-2/1, Di-2/2, Di-2/3, Ler-0/1, Ler-0/3, Wt-1/3
A9240G	<i>CP</i>	Ler-0/1, St-0/2, Wt-1/2
G1272U/C8636U	<i>HC-Pro/CP</i>	Di-2/1, Di-2/2, Di-2/3
G1272U/A9240G	<i>HC-Pro/CP</i>	St-0/2
C8636U/A9240G	<i>CP/CP</i>	Ler-0/1
C2116U	<i>HC-Pro</i>	-
G3639A	<i>CI</i>	-
G6420A	<i>NlaPro</i>	-
C2116U/G3639A	<i>HC-Pro/CI</i>	-
C2116U/G6420A	<i>HC-Pro/NlaPro</i>	-
G3639A/G6420A	<i>CI/NlaPro</i>	-
C2116U/G3639A/G6420A	<i>HC-Pro/CI/NlaPro</i>	Ei-2/1, Ei-2/2, Ei-2/3
C795U	<i>P1</i>	St-0/1, St-0/3

3.3.1 Exploración de datos de eficacia dentro del huésped

La Tabla 13 muestra los valores de eficacia dentro de huésped estimados para cada genotipo en cada ecotipo de huésped alternativo. Para nueve genotipos mutantes, se han observado diferencias significativas en la eficacia dentro del huésped (prueba de Kruskal-Wallis, $P \leq 0,046$). Sin embargo, la prueba de probabilidad combinada de Fisher muestra que en general existen diferencias entre los ecotipos de huéspedes para todo el conjunto de datos ($\chi^2 = 107,167$, 28 g.l., $P < 0,001$).

En primer lugar, se exploraron con más detalle los cinco genotipos cuyos valores de eficacia son igualmente altos en todos los ecotipos de huéspedes. La mutación G1272U fue originalmente observada en varios linajes evolucionados de Di-2, St-0 y Wt-1, por lo tanto es probable que tuviera un efecto de eficacia independiente del huésped. Sin embargo, en todos los casos el efecto era pequeño y no significativo en comparación con el rendimiento del genotipo ancestral TEV-At17b en cada huésped (prueba de Mann-Whitney, $P \geq 0,05$ en todos los casos). Esta prueba es consistente con la observación anterior de que St-0 y Wt-1 habían sido huéspedes más permisivos que durante la

evolución experimental seleccionaron virus más especializados. El doble mutante G1272U/A9240G, detectado en el linaje St-0/2, tuvo un efecto deletéreo significativo en todos los ecotipos excepto en el huésped alopatrico Wt-1. Este genotipo alcanzó una alta frecuencia en el linaje St-0/2, bien porque estaba vinculado a otras mutaciones beneficiosas o porque se fijó a través de un proceso de cuello de botella. Los otros tres genotipos con efectos homogéneos en los huéspedes fueron C2116U, C2116U/G6420A y G3639A/G6420A. Ninguno de estos genotipos se observó en linajes evolucionados, pero fueron creados con el propósito de analizar la epistasia entre estas tres mutaciones en el genotipo triple mutante C2116U/G3639A/G6420A observado en todos los linajes de Ei-2.

Tabla 13. Eficacia relativa dentro de huésped de cada genotipo viral medido en cada ecotipo de *A. thaliana*. Los valores se expresan como media \pm 1 SEM (número de réplicas). Los asteriscos indican casos significativamente diferentes de TEV-At17b (prueba de Mann-Whitney; corrección FDR de pruebas múltiples). Las celdas sombreadas indican el huésped local. La última columna muestra el *P* valor de las pruebas de Kruskal-Wallis para las diferencias en la eficacia relativa dentro de huésped entre los ecotipos para cada genotipo viral.

Genotipo	Di-2	Ei-2	Ler-0	St-0	Wt-1	<i>P</i>
TEV-At17b	1,003 $\pm 0,008$ (6)	0,824 $\pm 0,109$ (5)	0,996 $\pm 0,007$ (9)	1.000 $\pm 0,2730$ (2)	0,999 $\pm 0,011$ (4)	0,858
G1272U	0,831 $\pm 0,237$ (2)	0,969 $\pm 0,015$ (7)	0,936 $\pm 0,060$ (6)	0,793 $\pm 0,076$ (6)	1,155 $\pm 0,000$ (1)	0,342
C8636U	1,062 $\pm 0,049$ (2)	1,028 $\pm 0,033$ (5)	0,979 $\pm 0,006^*$ (5)	0,722 $\pm 0,019^*$ (2)	0.582 $\pm 0,013^*$ (3)	0,020
A9240G	0,995 $\pm 0,007$ (4)	0,984 $\pm 0,017$ (10)	1,017 $\pm 0,018$ (7)	1,198 $\pm 0,068^*$ (8)	0,574 $\pm 0,007^*$ (2)	0,008
G1272U/ C8636U	0,891 $\pm 0,161$ (3)	0,842 $\pm 0,073$ (7)	1,020 $\pm 0,029$ (5)	0,697 $\pm 0,000$ (1)	0,552 $\pm 0,000^*$ (2)	0,046
G1272U/ A9240G	0,604 $\pm 0,013^*$ (4)	0,575 $\pm 0,011^*$ (4)	0,582 $\pm 0,011^*$ (2)	0,750 $\pm 0,012^*$ (3)	0,762 $\pm 0,214$ (2)	0,159
C8636U/ A9240G	0,829 $\pm 0,197$ (2)	0,803 $\pm 0,077$ (5)	0,760 $\pm 0,103$ (5)	1,215 $\pm 0,014^*$ (5)	not determined	0,017
C2116U	0,717 $\pm 0,109$ (4)	0,887 $\pm 0,054$ (7)	1,000 $\pm 0,013$ (5)	1,080 $\pm 0,095$ (5)	0,797 $\pm 0,241$ (2)	0,066
G3639A	0,964 $\pm 0,079$ (5)	0,941 $\pm 0,017^*$ (10)	0,985 $\pm 0,014$ (10)	1,251 $\pm 0,012^*$ (4)	0,678 $\pm 0,064^*$ (9)	<0,001
G6420A	0,992 $\pm 0,020$ (3)	0,763 $\pm 0,088$ (6)	1,007 $\pm 0,011$ (8)	0,918 $\pm 0,094$ (5)	0,635 $\pm 0,072^*$ (5)	0,007
C2116U/ G3639A	0,706 $\pm 0,103$ (2)	0,973 $\pm 0,011$ (7)	0,914 $\pm 0,088$ (4)	1,197 $\pm 0,013^*$ (5)	0,917 $\pm 0,114$ (4)	0,009
C2116U/ G6420A	0,742 $\pm 0,139$ (3)	0,790 $\pm 0,085$ (5)	0,984 $\pm 0,009$ (5)	0,985 $\pm 0,288$ (2)	1,036 $\pm 0,021$ (2)	0,138
G3639A/ G6420A	1,001 $\pm 0,014$ (4)	0,879 $\pm 0,086$ (4)	0,955 $\pm 0,078$ (6)	0,702 $\pm 0,004^*$ (3)	1,032 $\pm 0,000$ (1)	0,079
C2116U/ G3639A/ G6420A	0,897 $\pm 0,124$ (3)	0,643 $\pm 0,031^*$ (5)	0,886 $\pm 0,072$ (6)	not determined	0,555 $\pm 0,007^*$ (3)	0,015
C795U	0,927 $\pm 0,087$ (5)	0,948 $\pm 0,003^*$ (6)	0,949 $\pm 0,047$ (7)	1,260 $\pm 0,006^*$ (4)	0,553 $\pm 0,005^*$ (3)	0,001

Tres mutaciones proporcionaron ventajas de eficacia específica de huésped dentro de huésped. La mutación C8636U, observada en cuatro ecotipos de huéspedes (Tabla 13), sólo proporcionó un efecto beneficioso significativo en el ecotipo no permisivo Ler-0. La mutación A9240G, identificada en los ecotipos de huéspedes más permisivos, proporcionó un efecto beneficioso significativo sólo en St-0 y fue significativamente deletérea en Wt-1 (Tabla 13). La mutación C795U, observada en dos de los linajes de St-0, confirió un efecto beneficioso significativo en St-0 y fue significativamente deletérea en Ei-2 y Wt-1 (Tabla 13).

3.3.2 El efecto de la eficacia dentro del huésped es mayor en huéspedes simpátricos que en alopátricos

Los resultados anteriores muestran una compleja interacción entre la eficacia dentro del huésped de cada genotipo TEV y el ecotipo en el que se evalúa. Si los genotipos de TEV estuvieran igualmente bien adaptados a sus huéspedes simpátricos que a los huéspedes alopátricos, esperaríamos una correlación positiva entre estas dos medidas de eficacia dentro del huésped. En un hipotético caso de un perfecto ajuste en la eficacia de los huéspedes simpátricos y alopátricos, la regresión lineal de la eficacia promedio en el huésped alopátrico frente al huésped simpátrico sería igual a uno.

Las desviaciones de la pendiente hipotética pueden tomarse como indicativas de efectos en la eficacia en función del ecotipo en el que se han evaluado. Una pendiente por debajo de uno indicaría que cuanto más fuerte son los efectos en el huésped simpátrico, desproporcionadamente más débiles pueden ser los efectos en los huéspedes alopátricos, lo que sugeriría que las mutaciones contribuyen más bien a la especialización que al generalismo. Por el contrario, una pendiente superior a uno indicaría que cuanto más fuerte es el efecto en el huésped simpátrico, desproporcionadamente más fuerte lo es en el huésped alopátrico. Para probar esta hipótesis, primero calculamos la eficacia media dentro de los huéspedes en todos los huéspedes alternativos. En aquellos casos en los que existía más de un huésped local, también calculamos la eficacia media. La Figura 14 muestra la relación entre la eficacia dentro de huésped en huéspedes locales y alternativos.

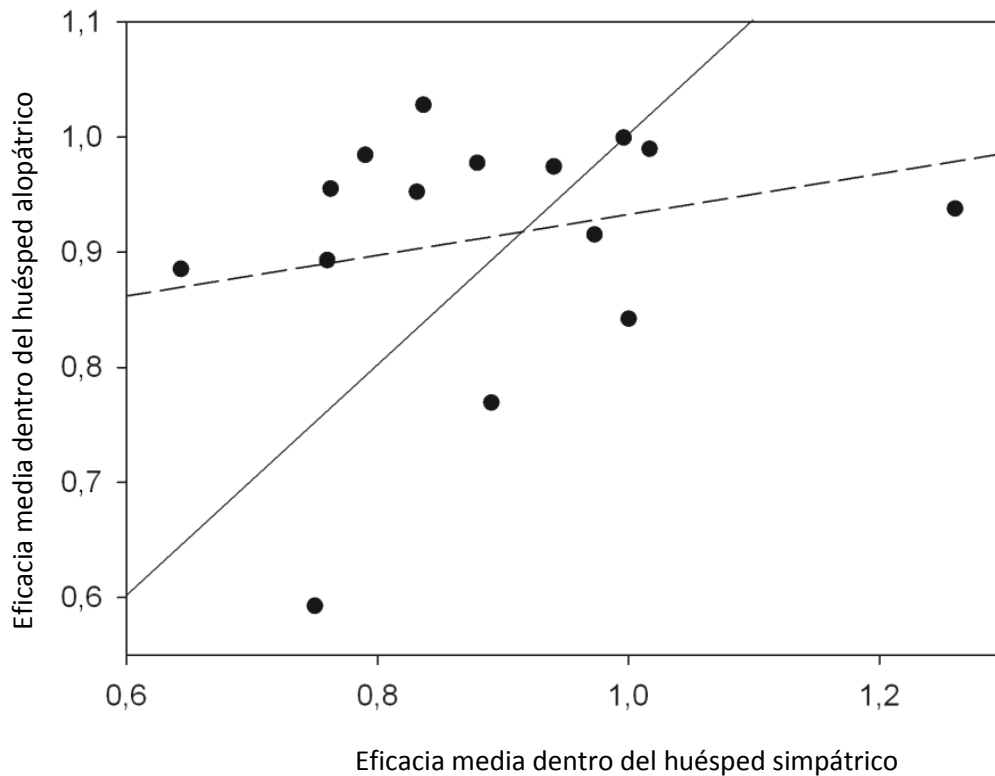


Figura 14. Relación entre la eficacia dentro del huésped en el huésped simpátrico y la eficacia media dentro del huésped en todos los huéspedes alopatricos. La línea continua representa la hipótesis nula de igualdad de efectos en todos los huéspedes. La línea discontinua representa el modelo lineal que mejor se ajusta a los datos.

La línea continua en la Figura 14 representa la hipótesis nula de que los efectos de la eficacia son idénticos a través de los huéspedes. La línea discontinua corresponde al mejor modelo de regresión lineal real. La pendiente de esta ecuación de regresión es $0,177 \pm 0,199$ (± 1 SEM), un valor que es significativamente menor que la expectativa nula de pendiente uno ($t_{13} = -4,116$, $P = 0,001$).

La Figura 14 también muestra que los genotipos se pueden clasificar en dos categorías respecto a la expectativa nula (línea continua). Para los ocho genotipos que se sitúan encima de la línea, los efectos de la eficacia promedio en los huéspedes alternativos son mayores de lo esperado dado su efecto en el ecotipo huésped local. Para los seis genotipos por debajo de la línea, los efectos de eficacia promedio en los huéspedes alternativos son más pequeños de lo esperado comparando con su valor en el ecotipo de huésped local.

Al analizar las propiedades de los genotipos en cada lado de la diagonal, se descubrió que aquellos genotipos que se sitúan por debajo de la diagonal procedían principalmente de la evolución en los ecotipos de huéspedes más permisivos (como se definió en el apartado 3.2.2: *Ler-0*, *St-0* y *Wt- 1*, Tabla 12: seis de los ocho casos). Lo contrario ocurría para aquellos genotipos por encima de la diagonal (*Di-2* y *Ei-2*, Tabla 12: cuatro

de 16 casos). De hecho, la prueba exacta de Fisher muestra que esta diferencia es significativa ($P = 0,032$). Por lo tanto, estos resultados confirmaron nuevamente la conclusión del Experimento 3.2 de que los huéspedes más permisivos seleccionaban genotipos virales especializados en infectar su huésped local, pero poco eficientes a la hora de infectar huéspedes alternativos. Por el contrario, los huéspedes locales restrictivos han seleccionado genotipos virales generalistas que infectan con éxito todo tipo de ecotipos de huéspedes alternativos. Este patrón anidado de interacciones corresponde al modelo gen-a-gen de interacciones virus-planta⁸⁶. Se espera que las redes de infección sean anidadas cuando existe una jerarquía de resistencia entre los huéspedes y la capacidad de infección entre los virus⁸⁶. Un mecanismo gen a gen implica que existen mutaciones que aumentan la eficacia en el nuevo huésped local pero no pagan un coste de eficacia en *Ler-0*, por lo tanto el conjunto de huéspedes que un aislado puede infectar son subconjuntos el uno del otro.

Muy pocos experimentos han probado si los efectos de las mutaciones adaptativas siguen siendo beneficiosos a través de un conjunto de diferentes entornos o son específicos del entorno. En un estudio pionero, Ostrowski *et al.*¹⁸⁴ encontró que las mutaciones fijadas en los linajes de *E. coli* adaptados a glucosa como única fuente de carbono también eran beneficiosas para otras cinco fuentes de carbono, concluyendo que el pleiotropismo positivo era una norma para la bacteria y que la adaptación local no limitaba la futura evolución. Al contrario, experimentos de evolución con virus de ARN han demostrado que las mutaciones que permiten ampliar la gama de huéspedes generalmente implican un coste en el huésped original^{6,19,72,74,199,242}, concluyendo que la norma para estos microparásitos sería el pleiotropismo negativo. Los resultados aquí presentados se expanden a las observaciones anteriores tanto de esta Tesis (ver apartado 3.2) como de otros autores^{6,19} y dan más apoyo a esta conclusión.

3.3.3 Patrones de epistasia y robustez de los paisajes adaptativos

Debido a sus genomas compactos y pautas de lectura solapantes, las epistásias son frecuentes en los virus de ARN⁸¹. En el capítulo anterior, se hicieron dos observaciones que apuntaban hacia un posible papel de la epistasia y su dependencia del ecotipo de huésped. En primer lugar, se encontró que las mutaciones G1272U, C8636U y A9240G eran comunes en linajes evolucionados en todos los ecotipos, excepto en Ei-2 (Tabla 13), pero sus combinaciones dobles fueron siempre exclusivas de un ecotipo de huésped dado. Como se observa en la Tabla 11, G1272U/C8636 se encuentra en todos los linajes evolucionados de Di-2, G1272U/A9240G en el linaje St-0/2 y C8636U/A9240G en el linaje *Ler-0/1*.

La segunda observación interesante implica a las mutaciones C2116U, G3639A y G6420A, sus tres combinaciones dobles y la triple mutante. Mientras que la triple

mutante se observó en todos los linajes evolucionados en Ei-2, no se observaron los genotipos intermedios ni en Ei-2 ni en otros ecotipos. Para evaluar el papel de la epistasia, y lo más importante, su dependencia del ecotipo de huésped, se construyeron dos pequeños paisajes de epistasia mostrados en la Figura 15.

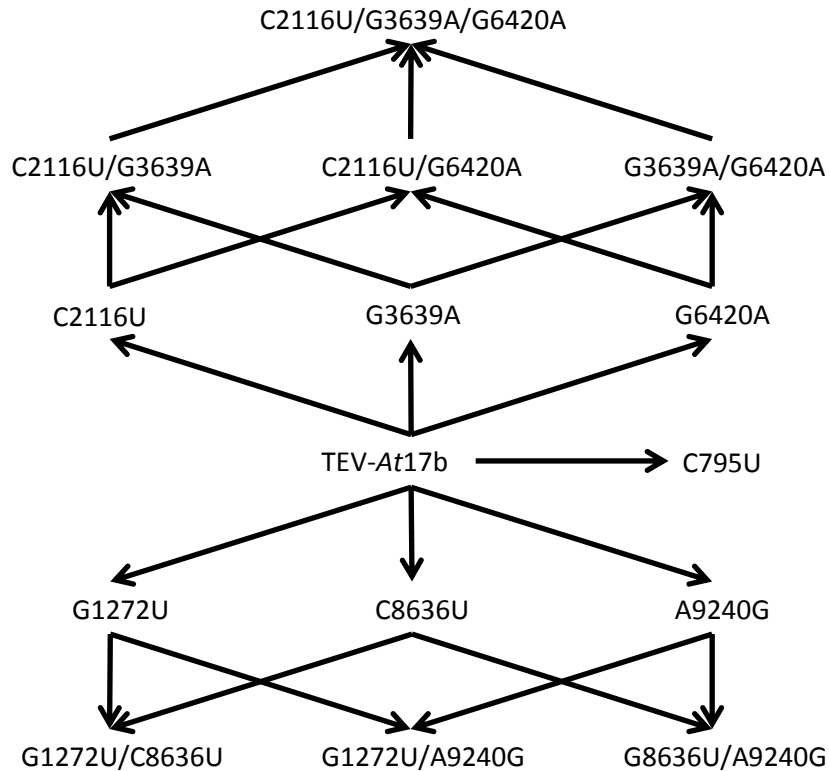


Figura 15. Paisajes de eficacia empíricos que conectan diferentes genotipos mutantes para los cuales ha sido evaluada la epistasia en los cinco ecotipos de huéspedes. Las flechas conectan genotipos que se distancian un paso mutacional.

La Tabla 14 muestra todos los valores de epistasia estimados. En cuanto a los genotipos individuales (filas de la Tabla 14), hay una gran variación en la magnitud y en el signo de la epistasia: ni una sola combinación de mutaciones tenía epistasia positiva o negativa en todos los ecotipos de huéspedes. En promedio, sólo la combinación C2116U/G6420A tenía una epistasia negativa significativa entre ecotipos de huéspedes. Desde el punto de vista de los ecotipos de plantas (columnas en la Tabla 14), la situación es similar: en todos los ecotipos se encuentran casos de genotipos de TEV con epistasia positiva y negativa.

Tabla 14. Coeficientes de epistasia (± 1 SD) evaluados para todos los mutantes dobles y triples que se muestran en la figura 9. Los casos significativos se indican con asteriscos (z-test; corrección FDR de múltiples pruebas). Las celdas sombreadas indican el huésped simpátrico en el que se detectó la mutación. La epistasia de signo se refiere a los casos en los que el signo del efecto de la eficacia depende del fondo genético. La epistasia recíproca de signo significa que el signo del efecto de la eficacia de una mutación está condicionado al estado de otro *locus* y vice versa. La última fila muestra la prueba de significación para la epistasia promedio (± 1 SEM) entre genotipos y la última columna muestra la prueba de significación para la epistasia promedio entre huéspedes (*t*-test).

Genotipo	Di-2	Ei-2	Ler-0	St-0	Wt-1	Epistasia promedio
G1272U/ C8636U	0,010 $\pm 0,640$	-0,303 $\pm 0,342$	0,100 $\pm 0,214$	0,124 $\pm 0,405$	-0,121 $\pm 0,037^*$ (signo)	-0,097 $\pm 0,150$
G1272U/ A9240G	-0,222 $\pm 0,363$	-0,480 $\pm 0,189^*$ (recip.)	-0,372 $\pm 0,177^*$	-0,200 $\pm 0,561$	0,099 $\pm 0,314$	-0,038 $\pm 0,079$
C8636U/ A9240G	-0,225 $\pm 0,350$	-0,276 $\pm 0,334$	-0,239 $\pm 0,279$	0,350 $\pm 0,614$	Sin determinar	0,026 $\pm 0,092$
C2116U/ G3639A	0,016 $\pm 0,393$	-0,033 $\pm 0,380$	-0,075 $\pm 0,227$	-0,154 $\pm 0,731$	0,375 $\pm 0,441$	0,052 $\pm 0,092$
C2116U/ G6420A	0,033 $\pm 0,509$	-0,026 $\pm 0,823$	-0,027 $\pm 0,294$	-0,007 $\pm 1,135$	0,529 $\pm 0,467$	-0,235 $\pm 0,098^*$
G3639A/ G6420A	0,047 $\pm 0,214$	0,008 $\pm 0,462$	-0,041 $\pm 0,244$	-0,447 $\pm 0,536$	0,599 $\pm 0,186^*$ (signo)	0,033 $\pm 0,167$
C2116U/ G3639A/ G6420A	0,213 $\pm 0,591$	-0,106 $\pm 0,507$	-0,110 $\pm 0,260$	Sin determinar	0,211 $\pm 0,432$	0,101 $\pm 0,108$
Epistasia promedio	-0,018 $\pm 0,059$	-0,174 $\pm 0,069^*$	-0,109 $\pm 0,058$	-0,056 $\pm 0,113$	0,282 $\pm 0,111^*$	

En Ei-2 y Ler-0, todos menos uno de los casos eran de epistasia negativa; por el contrario, en Wt-1 la epistasia fue mayoritariamente positiva. Por último, considerando combinaciones individuales de mutaciones, sólo cuatro casos han presentado epistasia significativa. La combinación de mutaciones G1272U/C8636U tuvo un coeficiente de epistasia significativamente negativo en el huésped alopátrico Wt-1. La combinación de mutaciones G1272U/A9240G estuvo implicada en interacciones epistáticas significativamente negativas en los huéspedes alopátricos Ei-2 y Ler-0. Finalmente, la combinación de mutaciones C3639A/G6420A estaba involucrada en epistasia positiva significativa en el huésped alopátrico Wt-1. En todos los demás casos, no se detectó ninguna desviación significativa de la hipótesis nula de efectos de eficacia multiplicativos.

Las interacciones negativas significativas entre los genotipos G1272U/C8636U y C3639A/G6420A eran del tipo epistasia de signo en Wt-1, lo que sugiere la existencia de perturbaciones locales en las proximidades de estas mutaciones en un paisaje aplanado de eficacia, aunque estas perturbaciones no son lo suficientemente fuertes como para crear un valle profundo. Es interesante que los dos únicos casos significativos de epistasia de signo se hayan detectado en el ecotipo Wt-1, probablemente debido a que este huésped puede imponer algunas restricciones que generan un cierto grado de robustez en el paisaje adaptativo. La forma de la interacción entre las mutaciones G1272U/A9240G depende del ecotipo de huésped en el que se evalúa. La epistasia es recíproca de signo en Ei-2, representando así una superficie local rugosa alrededor de estas mutaciones, con la existencia de dos picos locales; esto puede explicar por qué no se ha observado este doble mutante en linajes evolucionados en Ei-2. Por el contrario, este par de mutaciones muestra epistasia de magnitud cuando se evalúa en Ler-0, indicando un cambio descendente de la curvatura local de la superficie de eficacia en la vecindad de estas mutaciones, pero sin crear picos locales.

En conjunto, estos análisis sugieren que la topografía de los paisajes de eficacia representada en la figura 15 es fundamentalmente aplanada, si bien se presentan algunas crestas y valles en Wt-1 y Ei-2, respectivamente.

Esta dependencia de la topología del paisaje del virus de las condiciones ambientales se ha mostrado anteriormente para TEV¹⁴⁰ y el bacteriófago ID11³⁸. Por lo tanto, en lugar de pensar en una superficie de epistasia fija, deberíamos considerar mejor una superficie adaptativa altamente dinámica y flexible: el paisaje ondulante dinámico o *seascape*¹⁷³.

3.3.4 Discusión

El objetivo de este estudio fue evaluar los efectos en la eficacia de las mutaciones putativamente beneficiosas o de combinaciones de ellas, descritas en el capítulo anterior. Sorprendentemente, se encontró que la mayoría de estas mutaciones eran perjudiciales o neutrales en sus huéspedes simpátricos y sólo un número muy reducido tenía un efecto beneficioso específico de huésped. Por lo tanto, llegamos a la conclusión de que la mayoría de las mutaciones fijadas durante el experimento de evolución lo hicieron por deriva o por barridos selectivos, junto con la mutación beneficiosa que condicionó la selección. Con la construcción de dos paisajes de eficacia empíricos, se encontró que el signo y la magnitud de la epistasia eran variables entre las mutaciones y entre los huéspedes, de acuerdo con observaciones anteriores en TEV^{139,140}.

Una diferencia importante entre los estudios previos y los resultados aquí presentados debe ser especificada: en estudios previos con mutaciones al azar, la epistasia promedio era positiva y fundamentalmente de tipo recíproco de signo. En el presente estudio, los efectos de las mutaciones eran en su mayoría multiplicativos, con muy pocos casos de

epistasia significativa. De hecho, dos ejemplos (G1272U/A9240G medido en *Ler-0* y G1272U/C8636U evaluado en *Wt-1*) involucraron una mutación beneficiosa y una deletérea y dieron lugar a un genotipo perjudicial. En estos dos casos, la epistasia excluye la evolución espontánea de estos genotipos dentro de los ecotipos *Ler-0* y *Wt-1*, ya que después de la primera fijación de la mutación beneficiosa, la segunda mutación deletérea necesariamente generaría individuos de baja eficacia que se perderían durante la competencia dentro de huésped.

Según los criterios dados por Remold¹⁹⁸, las mutaciones G1272U y C2116U no tuvieron efectos pleiotrópicos, ya que no produjeron diferencias en la eficacia dentro de huésped en los ecotipos. Las cinco mutaciones restantes tenían un efecto pleiotrópico de signo, es decir, el signo del efecto de una mutación en la eficacia dentro de huésped depende del ecotipo actual. Asimismo, la mayoría de los múltiples mutantes analizados se clasifican en la categoría pleiotropía de signo no epistática, que siempre se traduce en un coste obligatorio de generalismo¹⁹⁸. El doble mutante C2116U/G6420A pertenece a la categoría sin epistasia y sin pleiotropía, que siempre resulta en generalismo sin coste¹⁹⁸. Finalmente, las mutaciones G1272U/A9240G y G1272U/C8636U son de pleiotropía de signo y de epistasia de signo. Bajo este escenario, las poblaciones pueden lograr la especialización o el generalismo sin coste, dependiendo del ecotipo en el que han evolucionado¹⁹⁸. Un escenario en el que, dependiendo del ecotipo en el que las poblaciones virales evolucionan, la selección puede favorecer la combinación genética produciendo generalistas sin coste, generalistas con coste o incluso especialistas, explica la evolución de un modelo gen-a-gen en la interacción virus-planta.

Globalmente, este estudio pone de manifiesto la necesidad de realizar análisis de genética inversa y análisis funcionales de las mutaciones observadas en la evolución de las poblaciones virales antes de concluir que son efectivamente beneficiosas y responsables de los cambios fenotípicos observados.

3.4 Transcriptoma de interacción gen a gen del virus con sus huéspedes nuevos

En capítulos anteriores, se han explorado las diferencias fenotípicas de los linajes de TEV-At17b evolucionados en diferentes ecotipos de *A. thaliana* y evaluado los efectos de las mutaciones convergentes fijadas durante la adaptación del virus a diferentes ecotipos de huésped. El presente capítulo continúa con la evaluación a nivel transcriptómico del grado en el que la diversidad genética del huésped condiciona el destino evolutivo del virus. En otras palabras, en este último capítulo se ha tratado de identificar similitudes y diferencias en el impacto que la infección de distintos ecotipos con virus experimentalmente evolucionados, tanto en simpatria como en alopatria. Las preguntas específicas que se plantearon son: (i) ¿favorece la variabilidad genética implicada en la respuesta del huésped a la infección viral la aparición de diversas estrategias de adaptación del virus? (ii) ¿Cuáles son las dianas moleculares de adaptación del virus a cada ecotipo del huésped? (iii) ¿Cómo la fase de evolución experimental y consecuente adaptación al nuevo ecotipo huésped determinan la interacción del virus evolucionado con el ecotipo huésped ancestral Ler-0? (iv) ¿Cómo de distintas son las respuestas transcriptómicas para los casos más extremos de virus especialista y generalista? Para abordar estas cuestiones, se realizaron una serie de pruebas utilizando la tecnología de micromatrices. Los ensayos de micromatrices de ARNm permiten cuantificar la expresión de casi todos los genes de *A. thaliana* durante la infección de la planta. Con este método, se analizaron tres plantas de *A. thaliana* para cada categoría de muestra. Se utilizaron plantas de todos los ecotipos infectadas con su virus localmente evolucionado, todos los linajes virales en su ecotipo original Ler-0 y también cada uno de los ecotipos de huésped infectados con los dos casos extremos de linajes virales evolucionados como especialista (St-0/1) y generalista (Ler-0/3). Además, a efectos de comparación, se utilizaron los datos obtenidos en el primer capítulo de la Tesis, los transcriptomas de los cinco ecotipos infectados con el linaje ancestral TEV-At17b.

Los resultados de las micromatrices de ARN se sometieron a cuatro evaluaciones estadísticas. En primer lugar, los perfiles transcriptómicos de los huéspedes infectados con cepas virales localmente adaptadas se contrastaron con las plantas no infectadas para ver la respuesta completa y específica a la infección en diferentes ecotipos. Asimismo, se compararon los linajes evolucionados con el linaje viral ancestral en su efecto sobre el mismo tipo de huésped. En tercer lugar, se evaluó el coste de la eficacia en el ecotipo de huésped original después de la adaptación del virus al nuevo ecotipo huésped. Finalmente, en el cuarto análisis, se abordaron los perfiles transcriptómicos de diferentes ecotipos infectados con los dos casos extremos de virus evolucionados como especialista (St-0/3) y generalista (Ler-0/1).

3.4.1 Análisis de la adaptación local: diferencias en la expresión de genes entre linajes virales evolucionados en diferentes ecotipos

La primera cuestión abordada aludía a en qué medida los tres linajes virales evolucionados en el mismo ecotipo de huésped convergían en la forma de interactuar con él. Para tal fin, se determinaron las similitudes en los perfiles de expresión génica entre los linajes virales evolucionados en un huésped común mediante un coeficiente de correlación de Pearson entre pares de linajes. La Tabla 15 muestra el promedio de estos coeficientes para los tres linajes evolucionados en el huésped común (Tabla 15).

Tabla 15. Promedio de coeficientes de correlación de Pearson para la expresión de genes de los linajes localmente adaptados.

	Di-2	Ei-2	Ler-0	St-0	Wt-1
Correlación promedio	0,615	0,857	0,647	0,784	0,842
Error estándar promedio	0,00428	0,00283	0,00411	0,00337	0,00295
P-valor	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

En todos los casos, las correlaciones han sido significativamente positivas, lo que sugiere un alto grado de convergencia evolutiva entre los linajes virales evolucionados en el mismo ecotipo de huésped. Sin embargo, el grado de convergencia fue variable entre los ecotipos: mientras que para Di-2 y Ler-0 el promedio de similitud entre los linajes virales fue ligeramente superior al 60%, lo era alrededor del 80% para los linajes evolucionados en Ei-2, St-0 y Wt-1. Esto sugiere que estos últimos pueden imponer restricciones más fuertes para la adaptación del virus y, por lo tanto, soluciones más restringidas y cambios evolutivos de las poblaciones virales necesariamente similares. A continuación, se evaluó si entre los linajes evolucionados en diferentes huéspedes había algún tipo de correlación respecto a la interacción con su huésped local. En otras palabras, se analizó si los patrones de expresión de genes entre diferentes ecotipos infectados con su huésped local tenían el mismo grado de similitud que los linajes evolucionados en el mismo ecotipo. Para ello, se estimaron los coeficientes de correlación de Pearson entre los promedios de la expresión génica de cada linaje en cada ecotipo (Tabla 16).

Tabla 16. Promedio de coeficientes de correlación de Pearson entre los ecotipos. Con asterisco están marcados los casos significativos $P < 0.001$.

	Di-2	Ei-2	Ler-0	St-0	Wt-1
Di-2	1,000	-0,074***	0,407***	-0,210***	0,021***
Ei-2		1,000	0,036***	0,468***	0,584***
Ler-0			1,000	-0,185***	0,260***
St-0				1,000	0,359***
Wt-1					1,000

Curiosamente, la magnitud de los coeficientes de correlación entre los ecotipos ha sido más baja que las correlaciones en linajes paralelos adaptados a un ecotipo común y además muy variable entre los ecotipos. Según lo mostrado en la Tabla 16, parejas tales como (Di-2, Ei-2) y (Di-2, St-0) muestran asociaciones muy débiles pero significativas. Además, los perfiles de expresión en St-0 se correlacionan negativamente a los observados en Di-2 y Ler-0, mientras que los perfiles de expresión entre los pares (Di-2, Ler-0), (Ei-2, Wt-1), (Ler-0, Wt-1) y (St-0, Wt-1) son todos significativamente positivos, con similitudes que oscilan entre el 26% y el 58%.

Es interesante que en dos de tres casos, los linajes que evolucionaron en St-0 hayan interactuado con su huésped local de una manera radicalmente diferente a la de otros linajes, lo que se muestra mediante las correlaciones negativas significativas entre los ecotipos. La Figura 16 muestra un dendrograma basado en los coeficientes de correlación que permite visualizar la similitud en la expresión génica inducida por la infección de los diferentes linajes virales evolucionados.

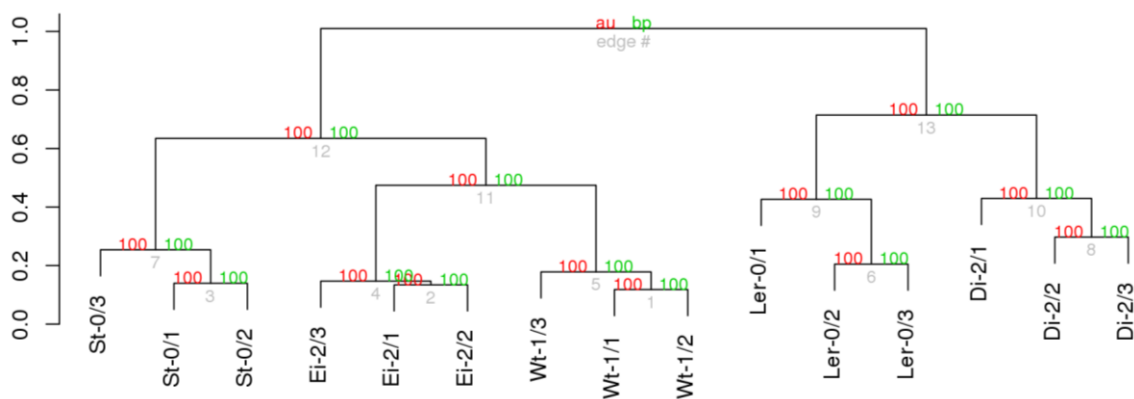


Figura 16. Dendrograma de distancias entre los ecotipos según sus correlaciones en respuesta transcriptómica a la infección con el virus localmente evolucionado. *Edge #* (color gris) representa si un clúster es más general (valores altos) o específico (valores bajos), *AU* (*Approximately unbiased*, color rojo) y *BP* (*Bootstrap probability*, color verde) son indicadores de significatividad con una escala entre 0 y 100. Valores más altos indican mayor número de muestras obtenidas por *bootstrap* corroborando ese agrupamiento. Es más robusto el valor proporcionado por *AU* que por *BP*.

En general, existen dos grupos de interacciones virus-ecotipo. En el primer grupo, los linajes evolucionados en *Ler-0* y *Di-2* forman un clúster que se subdivide en dos ramas, cada una agrupando linajes de un ecotipo. El segundo clúster incorpora linajes evolucionados en *Ei-2*, *St-0* y *Wt-1*, cada uno segregados en subgrupos separados definidos por el linaje. Dentro de este segundo grupo, los linajes evolucionados en *Ei-2* y *Wt-1* son más similares en su interacción con los huéspedes que los linajes evolucionados en *St-0*.

A continuación, se identificaron las categorías funcionales de los genes alterados para cada linaje viral en cada uno de los cinco ecotipos de plantas. Para este análisis, se utilizó el análisis conjunto de genes^{168,216}. En este análisis, se utilizaron todos los genes clasificados por su nivel de expresión diferencial de cada muestra. Comparando la lista de distribución funcional de genes de planta control procedente de la base de datos y la de planta infectada, buscamos conjuntos de genes cuyo nivel de expresión era diferente, aunque no necesariamente de manera significativa, en una de las dos listas. En otras palabras, se buscaron bloques de genes enriquecidos en determinadas zonas de la lista que comparten funciones y se identificaron las categorías funcionales de estos bloques, denominados términos de ontología génica (GO) para cada ecotipo. Los análisis generaron dos niveles de resultados: (i) todos los términos GO significativos y (ii) un grupo constituido sólo por los términos GO no redundantes y más específicos. Para identificar características consistentes entre linajes virales y ecotipos, se evaluaron las intersecciones de funciones no redundantes, teniendo en cuenta que, a pesar de una preselección de términos GO no redundantes, algunos términos podrían tener diferentes nombres pero apuntar a la misma función. Además, se asumían las intersecciones entre las réplicas paralelas como una imagen global de las similitudes. Incluso manteniendo este inconveniente en mente, se observó que los linajes evolucionados en los mismos ecotipos compartieron más categorías funcionales que los linajes evolucionados en diferentes ecotipos. Para los genes regulados positivamente, el número de categorías funcionales enriquecidas comunes para linajes del mismo ecotipo osciló entre 2 (para linajes evolucionados en *Di-2*) y 42 (para linajes evolucionados en *Ei-2*) y sólo había dos categorías funcionales compartidas por todos los linajes evolucionados (Figura 17A). Estas dos categorías son respuesta a iones de cadmio (GO:0046686) y fotorespiración (GO:0009853). Para las categorías reprimidas, el número de términos GO enriquecidos comunes para los tres linajes evolucionados en un huésped dado osciló entre 2 (linajes evolucionados en *St-0*) y 14 (linajes evolucionados en *Ler-0*) y no hubo categorías funcionales comunes para todos los linajes, únicamente en algunos casos entre pares de ecotipos. La mayor similitud de términos GO reprimidos se observó en los ecotipos *Ler-0* y *Di-2* (24), mientras que *Ei-2*, *Ler-0* y *Wt-1*, así como el par de ecotipos *Ei-2* y *St-0*, compartieron el mayor número de funciones sobre-expresadas (20 cada uno).

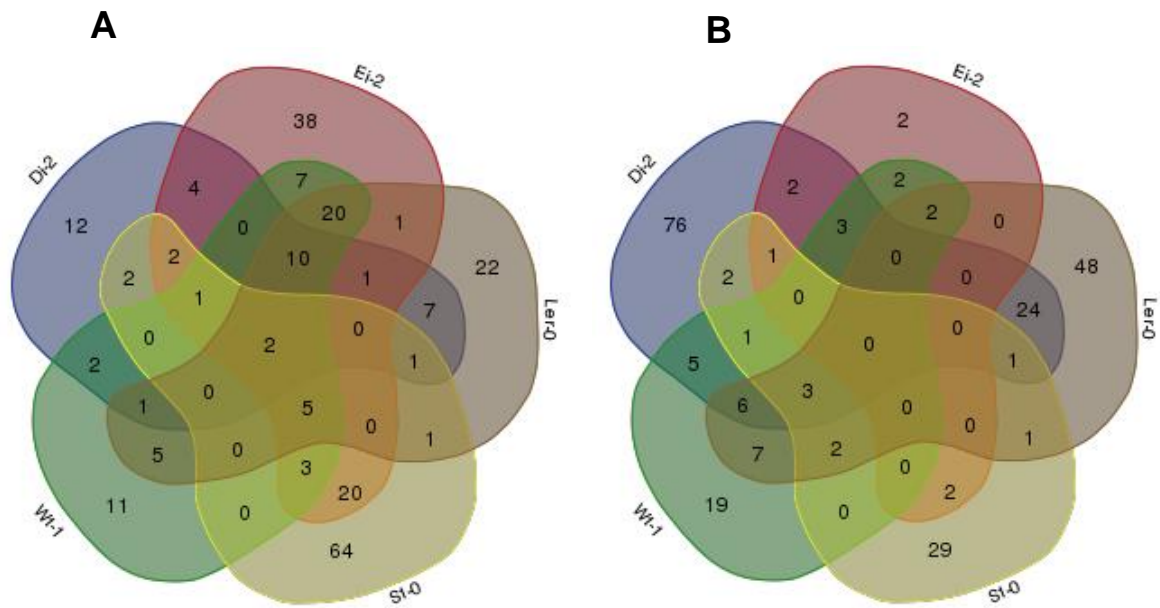


Figura 17. Diagramas de Venn de categorías funcionales enriquecidas sobre-expresadas **(A)** y reprimidas **(B)** en ecotipos infectados con linajes de virus localmente adaptados y contrastados con el ecotipo correspondiente inoculado solo con tampón.

<http://bioinformatics.psb.ugent.be/webtools/Venn>.

Para explorar en más detalle las similitudes y diferencias de las anotaciones funcionales, se procedió con el cálculo de los coeficientes de correlación de Pearson entre las categorías GO enriquecidas para cada linaje evolucionado. La Tabla 17 muestra los coeficientes de correlación promedio entre los linajes evolucionados en un ecotipo común (Tabla 17A) y entre ecotipos (Tabla 17B).

Tabla 17. Coeficientes de correlación promedio para categorías funcionales entre linajes (A) y entre pares de ecotipos (B). Con asterisco están marcados los casos significativos $P < 0.001$.

A

	Di-2	Ei-2	Ler-0	St-0	Wt-1
correlación media	0,709	0,919	0,833	0,721	0,793
error estándar medio	0,01917	0,01068	0,01514	0,01879	0,01553
P-valor	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

B

	Di-2	Ei-2	Ler-0	St-0	Wt-1
Di-2	1,000	0,279***	0,441***	-0,131***	0,250***
Ei-2		1,000	0,310***	0,136***	0,667***
Ler-0			1,000	-0,283***	0,581***
St-0				1,000	-0,001
Wt-1					1,000

Los linajes evolucionados en paralelo en un ecotipo común muestran correlaciones altamente significativas y positivas (Tabla 17A), lo que sugiere una fuerte convergencia en las funciones biológicas alteradas tras la infección. Al igual que en el caso de los perfiles de expresión génica, las combinaciones de diferentes ecotipos muestran más variabilidad que los linajes evolucionados en el mismo ecotipo, con pares de ecotipos que no muestran similitud en su respuesta a la infección (St-0 y Wt-1) o incluso asociaciones negativas (Di-2 y St-0; Ler-0 y St-0) y otros mostrando similitudes significativas que se extienden entre el 13,6% (Ei-2 y St-0) y el 66,7% (Ei-2 y Wt-1). Estos coeficientes de correlación se utilizaron para construir un dendrograma que muestra la similitud entre los linajes virales en las listas de términos GO generados en cada ecotipo (Figura 18).

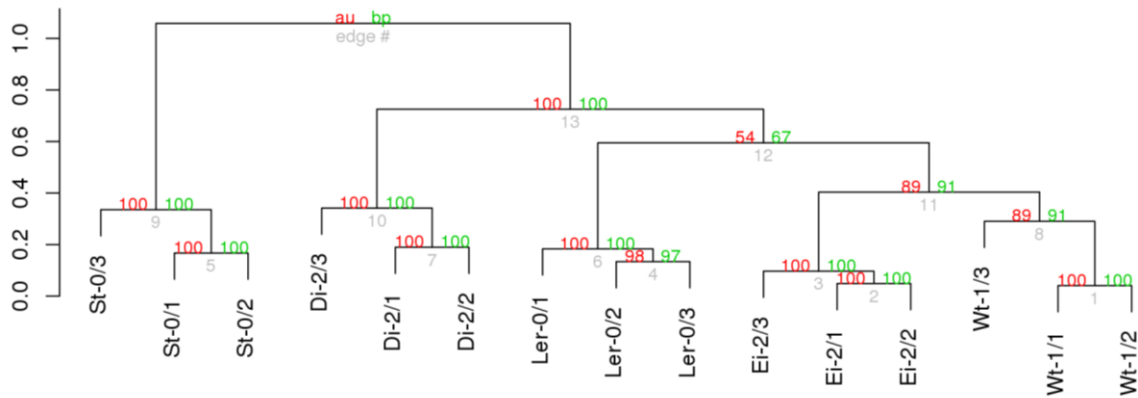


Figura 18. Correlación de distancias entre ecotipos en sus similitudes en respuesta funcional a la infección con el virus evolucionado localmente.

Edge # (color gris) representa si un clúster es más general (valores altos) o específico (valores bajos), *AU* (*Approximately Unbiased*, color rojo) y *BP* (*Bootstrap Probability*, color verde) son indicadores de significatividad con una escala entre 0 y 100 que proporciona el método de *bootstrap*. Valores más altos indican mayor número de muestras obtenidas por *bootstrap* corroborando ese agrupamiento. Es más robusto el valor proporcionado por AU que por BP.

Los linajes virales evolucionados en St-0 tenían un patrón de expresión funcionalmente diferente del resto de ecotipos, formando un grupo claramente independiente en el agrupamiento por correlaciones representado en la Figura 18. El resto de ecotipos forman un segundo grupo, no apreciándose subgrupos de ecotipos, sino que más bien el agrupamiento tenía una forma anidada. Cabe destacar que la agrupación funcional no fue la misma que la agrupación obtenida para los perfiles transcriptómicos ni para la respuesta funcional de la cepa ancestral TEV-At17b. Parece ser que el perfil funcional de los linajes evolucionados no es estrictamente dependiente del perfil de expresión génica, probablemente debido a la multifuncionalidad de los genes.

De este conjunto de análisis podemos concluir que los perfiles de expresión de ARNm de los huéspedes y sus correspondientes perfiles funcionales eran heterogéneos entre ecotipos, aunque se encontró un importante grado de paralelismo entre los linajes evolucionados en el mismo ecotipo.

En general, se pudo identificar dos tipos diferentes de respuesta transcriptómica. Con la excepción del ecotipo St-0, las correlaciones de expresión de genes observadas entre los linajes fueron similares a las de las respuestas tras la infección con el virus ancestral TEV-At17b (ver apartado 3.1.4). Esta observación implicaría que la evolución de las interacciones virus-huésped ha sido restringida por las características genéticas del huésped. Sin embargo, la respuesta funcional no apoya esta hipótesis, ya que la agrupación no tenía la misma ramificación que la respuesta transcriptómica al virus evolucionado. Estos resultados hacen pensar que los genes afectados en cada ecotipo codifican varias funciones agrupadas en distintas categorías en cada uno de los ecotipos.

3.4.2 Esclarecimiento de dianas específicas de ecotipo y generales de la adaptación viral

En esta segunda serie de análisis hemos tratado de identificar dianas de adaptación viral en diferentes huéspedes. En otras palabras, se comparó la respuesta transcriptómica de los ecotipos a la infección con los aislados localmente evolucionados con la respuesta a la infección con la cepa ancestral TEV-At17b.

En primer lugar, tanto la serie de muestras infectadas con el virus ancestral como la serie de muestras infectadas con el virus localmente evolucionado fueron simultáneamente normalizadas y los valores control (plantas inoculadas sólo con el tampón) se restaron de las muestras correspondientes. A continuación, se identificaron genes diferencialmente expresados entre plantas infectadas con los linajes virales evolucionados y plantas infectadas con el virus ancestral. La Figura 19 muestra el número de genes con expresión significativamente diferente y las intersecciones entre linajes evolucionados en el ecotipo común.

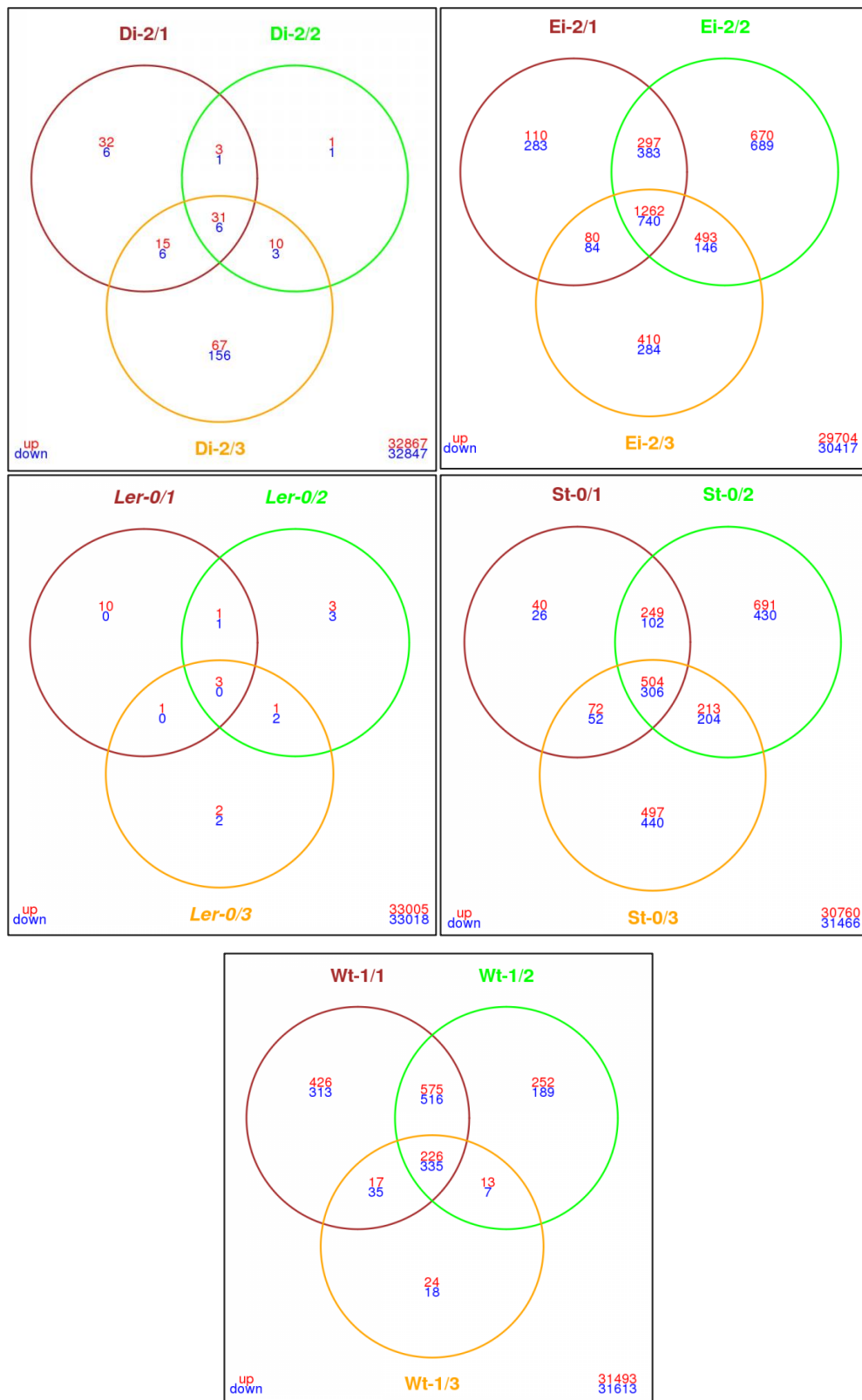


Figura 19. Diagramas de Venn de los genes diferencialmente expresados por la infección con el virus localmente evolucionado y el virus ancestral en los tres linajes de cada ecotipo. Cada círculo se corresponde con un linaje viral en un ecotipo dado. Las áreas de intersección indican el número de genes comunes diferencialmente expresados. Los números en rojo representan los genes sobre-expresados y en azul los reprimidos.

Los linajes virales evolucionados en los ecotipos Di-2 (37 genes compartidos por los tres linajes) y Ler-0 (3 genes compartidos por los tres linajes) indujeron el menor número de genes diferencialmente expresados como resultado de la evolución y adaptación del virus a estos ecotipos. Al contrario, los linajes evolucionados en Ei-2 indujeron la sobre-expresión/represión diferencial de una gran cantidad de genes en comparación con el aislado ancestral (2002 genes compartidos por los tres linajes). Una situación intermedia ha sido la causada por los linajes Wt-1 y St-0, donde el número total de genes afectados era parecido (810 y 561 genes redundantes entre las tres replicas, respectivamente). Cabe destacar que nuevamente la mayor cantidad de genes alterados siempre pertenece a la categoría de genes compartidos por los tres linajes independientes evolucionados en un mismo ecotipo (Figura 19), apuntando a una evolución similar determinada por el ecotipo de huésped.

A continuación, se realizó un análisis funcional utilizando el método del conjunto de genes. Se pudo identificar categorías funcionales enriquecidas para cada linaje. Para resumir los términos GO para los procesos biológicos, se determinaron súper-categorías funcionales utilizando la herramienta ReViGo²³² que se representan en las Tablas 18 y 19 para genes sobre-expresados y reprimidos, respectivamente.

Tabla 18. Grupos funcionales de términos GO sobre-expresados para los procesos biológicos. La columna de la izquierda representa súper-categorías de GO y el resto de columnas cada uno de los ecotipos. Las súper-categorías funcionales presentes en los ecotipos están marcadas con una cruz en la columna correspondiente al linaje.

Clústeres funcionales	Di-2	Ei-2	Ler-0	St-0	Wt-1
proceso catabólico de quitina	x				
organización del Golgi	x	x			
abscisión	x				
respuesta a la luz roja	x				
homeostasis de cationes	x		x		
transporte de iones de hierro	x				
proceso metabólico de maltosa	x		x		
asimilación de nitrato	x				
compuesto que contiene antocianina	x		x		
proceso de biosíntesis	x	x			x
regulación negativa de la actividad catalítica	x			x	
respuesta a nitrato	x				
proceso biosintético de myo-inositol hexakisfosfato	x				
respuesta celular a la inanición de iones de hierro	x				
ensamblaje del fotosistema II	x			x	
proceso metabólico de ácido para-aminobenzoico	x		x		
transporte transmembrana de drogas	x	x			
proliferación celular		x			
transporte mediado por vesículas		x			x
regulación positiva de la actividad catalítica		x			

organización o biogénesis de la pared celular	x			
respuesta a karrikin	x			
biogénesis de la pared celular	x	x		
plegamiento de proteínas	x		x	x
proceso metabólico de acetil-CoA	x			
crecimiento celular multidimensional	x			
desarrollo de las anteras	x			
proceso de biosíntesis de etileno	x			x
proceso metabólico de vitamina	x			
citocinesis por la formación de placa celular	x			
modificación de ARNm	x		x	
proceso metabólico celular	x			
proceso catabólico de toxina	x			
detección de estímulo biótico	x			x
vía de señalización del receptor de superficie celular	x			
iniciación de la replicación de ADN	x			
respuesta a ciclopentenona	x			x
homeostasis celular	x			
proceso de biosíntesis de ácido dicarboxílico	x			
proceso metabólico de ARNc	x			
proceso de biosíntesis de carotenoides	x			
regulación de la desfosforilación de proteínas	x			
transporte de ácidos monocarboxílicos	x			
respuesta de defensa por deposición de callosa	x			x
proceso biosintético de monofosfato ribonucleósido	x			
localización en orgánulos	x			
respuesta a ion de cadmio	x	x		
transporte de nucleósidos	x			
proceso catabólico de aminoácidos de la familia del				
aspartato	x			x
proteolisis		x		
transporte de agua		x		
organización de vacuolas		x		
respuesta a hongos simbioses		x		
ciclo del ácido tricarboxílico		x		
modificación de lípidos		x		
proceso catabólico de macromolécula		x		
respuesta a hexosa		x		
respuesta al estrés salino		x		
formación de exina de polen		x		
transporte de iones divalentes metálicos		x		
proceso de biosíntesis de glucosinolatos		x		
respuesta a la alta intensidad de luz			x	x
desarrollo del xilema			x	
fotosíntesis			x	
movimiento basado en filamentos de actina			x	
proceso de biosíntesis de peróxido de hidrógeno			x	
desumoylación de proteínas			x	

proceso metabólico secundario	x	
exportación del ARNm desde el núcleo	x	
proceso catabólico de polisacáridos	x	
proceso metabólico de ARNm	x	
proceso de biosíntesis del cofactor	x	
proceso de biosíntesis de metabolitos secundarios	x	x
proceso metabólico de alcohol	x	
destrucción de células de otro organismo	x	
respuesta a peróxido de hidrógeno	x	x
regulación de la organización de componente celulares	x	
determinación de simetría bilateral	x	
respuesta celular a escasez de fosfato	x	
regulación positiva de la transcripción de ADN	x	
transporte transmembrana	x	
modificación de la pared celular	x	x
proceso de biosíntesis de ácido jasmónico	x	x
proceso biosintético de ribonucleótido pirimidina	x	
germinación de las semillas	x	
regulación del desarrollo de crecimiento	x	
procesamiento del extremo 3' del ARN		x
cadena respiratoria de transporte de electrones		x
proceso catabólico de pectina		x
reconocimiento del polen		x
regulación negativa de la muerte celular programada		x
procesamiento de péptidos señal		x
miristoilación N-terminal de proteína		x
transporte de iones de fosfato		x
endocitosis		x
transporte de oligopéptidos		x
proceso biosintético de galactolípidos		x
proceso metabólico del glicerol		x

Tabla 19. Grupos funcionales de términos GO reprimidos para los procesos biológicos. La columna de la izquierda representa súper-categorías de GO y el resto de columnas cada uno de los ecotipos. Las súper-categorías funcionales presentes en los ecotipos están marcadas con una cruz en la columna correspondiente al linaje.

Clústeres funcionales	Di-2	Ei-2	Ler-0	St-0	Wt-0
proceso metabólico de acetil-CoA	x				
ritmo circadiano	x				
proliferación celular	x		x	x	
desarrollo de tejido	x				
respuesta a la radiación ionizante	x				
biogénesis de la subunidad pequeña del ribosoma	x				x
regulación positiva de la actividad hidrolasa	x		x		
maduración de proteína	x				x
fotorespiración	x			x	

transporte de factores de transcripción al núcleo	x			
señalización célula-célula	x			
recombinación mitótica	x	x		
movimiento basado en microtúbulos	x			
ciclo celular mitótico	x			
proceso de biosíntesis de ácido dicarboxílico	x			
respuesta a la molécula de origen bacteriano	x			
regulación de la proliferación celular	x			
respuesta al estrés del retículo endoplásmico	x			
homeostasis redox celular	x			
Proceso metabólico de ARNm	x			
plegamiento de proteínas	x	x		
organización del núcleo	x			
deneddylación de cullin	x			
síntesis de pseudouridina	x			
transcripción y elongación de ADN	x			
regulación de procesos metabólicos de nucleósidos	x			
elongación traduccional	x			x
proceso de catabolismo proteico proteasomal	x			
corte yempalme de ARN a través de la escisión endonucleolítica	x	x		x
regulación post-transcripcional de la expresión génica	x			x
proceso biosintético de glicerofosfolípidos	x			
modificación d peptidil-aminoácidos	x			
miristoilación N-terminal de proteínas	x		x	
traducción	x	x	x	x
respuesta de vernalización				
destrucción de células de otro organismo		x	x	
inclusión de proteínas en la matriz de peroxisomas		x		
abscisión		x		
proceso biosintético de myo-inositol hexakisfosfato		x		x
proceso catabólico de clorofila		x		
vía de señalización mediada por ácido giberélico		x		
proceso de latencia de semillas		x		
proceso de biosíntesis de etileno			x	x
transporte citoplasmático			x	
respuesta a la disminución de los niveles de oxígeno			x	
ensamblaje o desensamblaje de cromatina			x	x
desarrollo de anteras			x	
proceso basado en microtúbulos			x	x
movimiento basado en filamentos de actina			x	
citocinesis por formación de pared celular			x	
proceso catabólico de toxinas			x	
proceso de biosíntesis de ácido salicílico			x	
respuesta a ciclopentenona			x	
regulación negativa de la muerte celular programada			x	x
metilación de ARN			x	
vía de señalización retículo endoplásmico-núcleo			x	

especificación del destino celular	x	
respuesta a peróxido de hidrógeno	x	
proceso biosintético de ribonucleótido pirimidina	x	
aclimatación al calor	x	
detección de estímulo biótico		x
ensamblaje del complejo núcleo-proteosoma		x
regulación de la transición G2/M del ciclo celular mitótico		x
transporte mediado por vesículas del retículo endoplásmico al Golgi		x
respuesta a estímulo mecánico		x
proceso de biosíntesis de auxina		x
respuesta a la glucosa		x
explosión respiratoria implicada en la respuesta de defensa		x
regulación del proceso metabólico de peróxido de hidrógeno		x
iniciación de la replicación del ADN		x
cascada MAPK		x
regulación positiva del proceso de biosíntesis de flavonoides		x
proceso metabólico de celulosa de la pared celular		x
proceso de catabolismo proteico dependiente de ubiquitina		x
respuesta celular a la hipoxia		x
respuesta a la giberelina		x
cambio de fase vegetativa		x
direccionamiento de proteína al cloroplasto		x
recombinación de ADN		x
proceso metabólico de maltosa		x
transporte de electrones en el fotosistema I		x
proceso del ciclo celular		x
respuesta a la radiación gamma		x
desubiquitinación de proteínas		x
proceso catabólico nuclear de transcritos de RNAm		x
transporte polar de auxinas		x
especificación de polaridad del eje adaxial/abaxial		x
regulación del proceso de modificación de proteínas		x
morfogénesis de la hoja		x
regulación epigenética de la expresión génica		x

Tanto los clústeres de funciones sobre-expresadas como reprimidas tenían categorías comunes entre dos o más ecotipos (Tablas 18 y 19). Por ejemplo, se encontraron 24 categorías funcionales de genes positivamente reguladas. Las categorías sobre-expresadas que más se repitieron en varios ecotipos fueron la respuesta a heridas (GO:0009611; en Di-2, Ei-2 y Wt-1) y el plegamiento de proteínas (GO:0006457; en Ei-2, St-0 y Wt-1). Una de las categorías afectadas, tanto en el ecotipo Di-2 como Ler-0, fue el metabolismo de la maltosa (GO:0000023). El disacárido maltosa es un producto

fotosintético importante en *A. thaliana*¹⁰⁶ que se forma a partir de dos unidades de glucosa producidas cuando la encima amilasa descompone el almidón, un compuesto principal de almacenaje de carbohidratos en plantas. Además en estos ecotipos se modificaron los genes implicados en la homeostasis de cationes (GO:0055080), biosíntesis de antocianina (GO:0009718), formación de glucósidos de antocianinas intensamente coloreados, y finalmente el metabolismo del ácido p-aminobenzoico (GO:0046482), un intermediario en la síntesis de ácido fólico necesario para la amplificación de ADN y ARN.

Los ecotipos Ei-2 y Wt-1 han sobre-expresado los genes involucrados en los siguientes procesos: biosíntesis de etileno (GO:0009693), transporte mediado por vesículas (GO:0016192), detección de estímulos bióticos (GO:0009595), respuesta a ciclopentenona (GO:0010583), respuesta de defensa por deposición de callosa (GO:0052542) y catabolismo de aminoácidos de la familia del aspartato (GO:0009068). El etileno es una fitohormona involucrada en muchos puntos del ciclo de vida de la planta, incluyendo germinación de semillas, desarrollo de la raíz, senescencia de la flor, abscisión y maduración del fruto¹²³. La producción de etileno está fuertemente regulada por señales internas durante el desarrollo y como respuesta a estímulos ambientales de estreses bióticos (por ejemplo, patógenos) y abióticos (heridas, hipoxia, ozono o congelación)²⁴⁹. El transporte mediado por vesículas es un proceso de transporte celular por el cual las sustancias se mueven en vesículas delimitadas por membranas. Las sustancias transportadas pueden estar encerradas en el lumen de la vesícula o localizadas en su membrana. La detección de estímulos bióticos implica una serie de procesos en los que un estímulo causado o producido por un organismo vivo se recibe y convierte en una señal molecular. Las ciclopentanonas son derivados de ácidos grasos poliinsaturados, estructural y funcionalmente similares al ácido jasmónico, un regulador de la respuesta de las plantas al estrés abiótico y biótico, así como del desarrollo de la planta. En las primeras etapas de invasión de patógenos, el polisacárido vegetal callosa puede ser depositado en la pared celular y formar una barrera eficaz, aparte de servir como una matriz en la que se pueden depositar a su vez compuestos antimicrobianos¹⁵³. El catabolismo de aminoácidos de la familia del aspartato está implicado en la descomposición de estos residuos, que incluyen asparagina, aspartato, lisina, metionina y treonina.

Los ecotipos St-0 y Di-2 tenían dos clústeres comunes afectados positivamente: la regulación negativa de la actividad catalítica (GO:0043086) y el ensamblaje del fotosistema II (GO:0010207).

Los ecotipos St-0 y Wt-1 compartieron clústeres relacionados con el metabolismo, en concreto, la respuesta a luz de alta intensidad (GO:0009644) y la biosíntesis de metabolitos secundarios (GO:0044550). Asimismo, estos dos ecotipos también

mostraron una sobre-expresión significativa en categorías funcionales de respuesta inmune, como respuesta a peróxido de hidrógeno (GO:0042542) y biosíntesis de ácido jasmónico (GO:0009695). También en los ecotipos St-0 y Wt-1 se ha observado una represión en procesos que pueden causar alteraciones de celulosa y pectina en la pared celular (GO:0009827), que podrían provocar en último término la desintegración de la pared. Respecto a los ecotipos Di-2 y Ei-2, se ha observado una sobre-expresión de genes implicados en la organización de Golgi (GO:0007030), un proceso que se lleva a cabo a nivel celular y que da lugar al ensamblaje, organización de las partes constituyentes, o desensamblado del aparato de Golgi, así como también en genes implicados en el transporte de drogas a través de membranas (GO:0006855), en el que los compuestos se transportan a través de la membrana por medio de algún transportador o un poro. Ambas categorías están estrechamente relacionadas con la replicación viral y con el transporte de célula a célula de viriones o de complejos replicativos. La biogénesis de la pared celular de la planta (GO: 0009832) y la respuesta a ion cadmio (GO:0046686) eran las categorías comunes positivamente afectadas en los ecotipos Ei-2 y Ler-0. La respuesta de las plantas a cadmio evoca una serie de eventos paralelos y/o consecutivos a nivel molecular, fisiológico y morfológico²¹¹.

También se examinaron los efectos potenciales del cadmio en las interacciones huésped-patógeno y se demostró que en plantas de tabaco el cadmio induce la producción de factores antivirales putativos capaces de inhibir la propagación sistémica del virus del aclaramiento de las venas de nabo (TVCV)⁴⁵. En condiciones experimentales, sin estímulo obvio de cadmio, es posible que esta vía de señalización no responda exclusivamente a cadmio, sino que también pueda ser activada por la infección con el virus. La modificación del ARNm (GO:0016556) es la única categoría que los ecotipos Ei-2 y St-0 presentaron en común, pero este proceso puede tener una variedad de diferentes impactos en la interacción virus-huésped, que oscila entre la biosíntesis de proteínas y la resistencia de la planta a la infección.

La categoría reprimida más redundante entre ecotipos era la formada por genes implicados en la traducción (GO:0006412; Di-2, Ler-0, St-2 y Wt-1). Por otro lado, la categoría funcional relacionada - empalme del ARN a través de la escisión y unión endonucleolítica (GO:0000394), involucrada en el procesamiento de ARNt, ha sido suprimida en tres ecotipos (Ei-2, Di-2 y Wt-1). Los genes involucrados en proliferación celular (GO:0008283) fueron significativamente reprimidos en los ecotipos Ler-0, Di-2 y St-0.

La respuesta diferencialmente reprimida, provocada en Ei-2, incluye los procesos de destrucción de células de otro organismo (GO:0031640), en común con el ecotipo St-0, y una categoría funcional en común con Wt-1 - biosíntesis de mioinositol hexakisfosfato (GO:0010264), que está implicado en la regulación de la señalización intracelular y el

almacenamiento de fosfato y de compuestos minerales en semillas de plantas. Las categorías funcionales biogénesis de la subunidad ribosomal pequeña (GO:0042274), maduración de proteínas (GO:0051604), elongación de la traducción (GO:0006414) y regulación post-transcripcional de la expresión génica (GO:0010608), estaban reprimidos en los ecotipos Wt-1 y Di-2. Claramente, la respuesta frente a la infección viral en estos ecotipos ha sido la represión de la biogénesis de proteínas. La regulación negativa de la biosíntesis de proteínas también se contempló por la categoría de plegamiento de proteínas (GO:0006457) en los ecotipos Ler-0 y Di-2. Además, estos ecotipos reprimieron la expresión de genes que pertenecen a categorías de regulación positiva de la actividad hidrolasa (GO:0051345) y recombinación mitótica (GO:0006312). La foto respiración (GO:0009853) y miristoilación N-terminal de proteínas (GO:0006499) se vio afectada negativamente en los ecotipos St-0 y Di-2. La miristoilación puede dirigir reversiblemente las interacciones proteína-proteína y lípido-proteína y desempeña un papel esencial en la dirección a la membrana y en una variedad de vías de transducción de señal^{112,268}. Se ha propuesto que algunas proteínas virales, como las de geminivirus, pueden ser modificadas mediante anclaje de lípidos; estas proteínas estarían localizadas en plasmodesmos y ayudarían en el movimiento de célula a célula²⁰⁶. En los ecotipos Ler-0 y St-0, se observó la supresión de la regulación de la biosíntesis del etileno (GO:0009693) y la regulación negativa de la apoptosis (GO:0043069); ambas categorías probablemente permiten al virus evadir el sistema inmunológico. Finalmente, las categorías funcionales ensamblaje y desensamblaje de la cromatina (GO:0006333) y procesos basados en microtúbulos (GO:0007017), que se refieren a cualquier proceso celular que depende o altera los microtúbulos del citoesqueleto, han sido suprimidas tanto en Ler-0 como en Wt-1.

Para resumir, algunas dianas de adaptación viral eran comunes entre ecotipos. Estas dianas parecen activar las vías de defensa inmunitarias, como el proceso de biosíntesis de ácido jasmónico, la respuesta a peróxido de hidrógeno, la respuesta a ciclopentanona y a iones de cadmio y la detección de estímulos bióticos. Los procesos implicados directamente en el ciclo de vida del virus, como la organización del Golgi, el transporte mediado por vesículas, la modificación de ARNm, la modificación de la pared celular vegetal y la respuesta de defensa por la deposición de callosa, han sido enriquecidos en genes sobre-expresados. La biogénesis de los recursos básicos de la planta se incrementó en las plantas infectadas con el virus evolucionado, incluyendo categorías funcionales como la regulación negativa de la actividad catalítica, el proceso catabólico de aminoácidos de la familia del aspartato, el ensamblaje del fotosistema II, la respuesta a luz de alta intensidad y el proceso de biosíntesis de metabolitos secundarios.

Los ecotipos Ler-0, Ei-2 y Wt-1 han mostrado una respuesta negativa considerable en numerosos procesos de biosíntesis y modificación de proteínas: regulación positiva de

actividad hidrolasa, maduración de proteínas, biogénesis de la subunidad ribosomal pequeña, plegamiento de proteínas, elongación de la traducción, ensamblaje de ARN a través de escisión endonucleolítica y ligamiento y regulación post-transcripcional de la expresión génica y traducción. La proliferación celular se vio afectada negativamente por modificaciones de la cromatina y por supresión de recombinación mitótica. También se produjo una reducción en diversos ecotipos en los procesos de biosíntesis de etileno, de mioinositol hexakisfosfato y también de la fotorrespiración, necesarios para la supervivencia celular a largo plazo. Finalmente, también se observó represión durante la infección con el virus evolucionado en algunas categorías funcionales de la respuesta inmune, como la regulación negativa de la muerte celular programada o categorías que podrían ayudar en la replicación del virus, como la miristoilación N-terminal de proteínas y procesos basados en microtúbulos.

A pesar de que se detectaron varias categorías comunes entre pares de ecotipos, eran mucho más abundantes las categorías propias de cada ecotipo (Tablas 18 y 19). Estas categorías cubrieron un amplio espectro de funciones, lo que indica que las dianas de adaptación del virus han sido múltiples y diversas entre los ecotipos de huéspedes. El trastorno en la síntesis de los recursos celulares básicos, acompañado con el desarrollo de los síntomas por la infección viral, fue probablemente causado por la interferencia del virus en funciones vitales del huésped, pero que no son importantes para la supervivencia de la planta a corto plazo.

Aunque en todos los ecotipos se ha detectado al menos una categoría implicada en la respuesta inmune directa, este grupo funcional no presentó tanta diversidad como otros grupos funcionales, lo que indica que la resistencia de la planta no era el objetivo principal para la adaptación viral. Sin embargo, también es posible que se presentara una modificación no directa de las vías del sistema inmune de la planta durante la adaptación del virus, como por ejemplo es el caso de la modificación de lípidos, proceso implicado en los mecanismos de patogénesis y resistencia asociados a las interacciones planta-microorganismo²²⁰. En cuanto a los cambios en las interacciones virus-huésped causados por los virus adaptados a cada ecotipo de huésped, estos eran particularmente considerables para los linajes evolucionados en Ei-2, St-0 y Wt-1, como muestra el elevado número de genes diferencialmente afectados. Menos modificaciones transcriptómicas surgieron durante la evolución en Ler-0 y Di-2. Esta división en dos grupos refleja el esquema de similitudes en la expresión de genes de ecotipos infectados con el virus ancestral y remarca que el genotipo del huésped es crucial para el destino de la evolución viral.

En términos generales, el análisis de las categorías funcionales afectadas permitió el esclarecimiento de lo que sucede en el huésped a causa de la adaptación viral. Los cambios específicos en las interacciones, causados por diferentes linajes evolucionados,

tenían un perfil heterogéneo entre ecotipos, pero se detectaron también algunas intersecciones en categorías funcionales afectadas, lo que indicaba la existencia de objetivos más redundantes en la adaptación viral entre ecotipos. En cualquier caso, no hubo intersecciones entre todos los ecotipos, sino que más bien se formaron dos grupos de acuerdo con las similitudes de respuesta transcriptómica que presentan perfiles opuestos.

La composición de las categorías funcionales afectadas positiva y negativamente inducidas por el linaje *Ler-0/2* en su ecotipo local mostró cambios significativos en comparación con el virus ancestral (datos no mostrados), a pesar de la ausencia de mutaciones adaptativas fijadas en uno de sus linajes, por lo que los cambios transcriptómicos se atribuirían principalmente a subpoblaciones virales.

3.4.3 Efecto de la evolución en los ecotipos alternativos en el rendimiento en el ecotipo original *Ler-0*

En un tercer grupo de análisis, se evaluó si la fase experimental de evolución que llevó a la adaptación de los linajes virales a cada nuevo ecotipo podía estar asociada con un cambio en la forma en que los virus evolucionados interactuaban con el huésped original *Ler-0*. Para ello, se compararon los perfiles transcriptómicos entre la infección del huésped original *Ler-0* por el linaje ancestral *TEV-At17b* y por cada uno de los linajes evolucionados en los ecotipos alternativos.

La primera cuestión a abordar fue: ¿es la expresión de genes distinta a la causada por la infección con el virus original? Como respuesta a esta pregunta nos encontramos con pocos genes (o ninguno) diferencialmente expresados en los linajes virales (Figura 20).

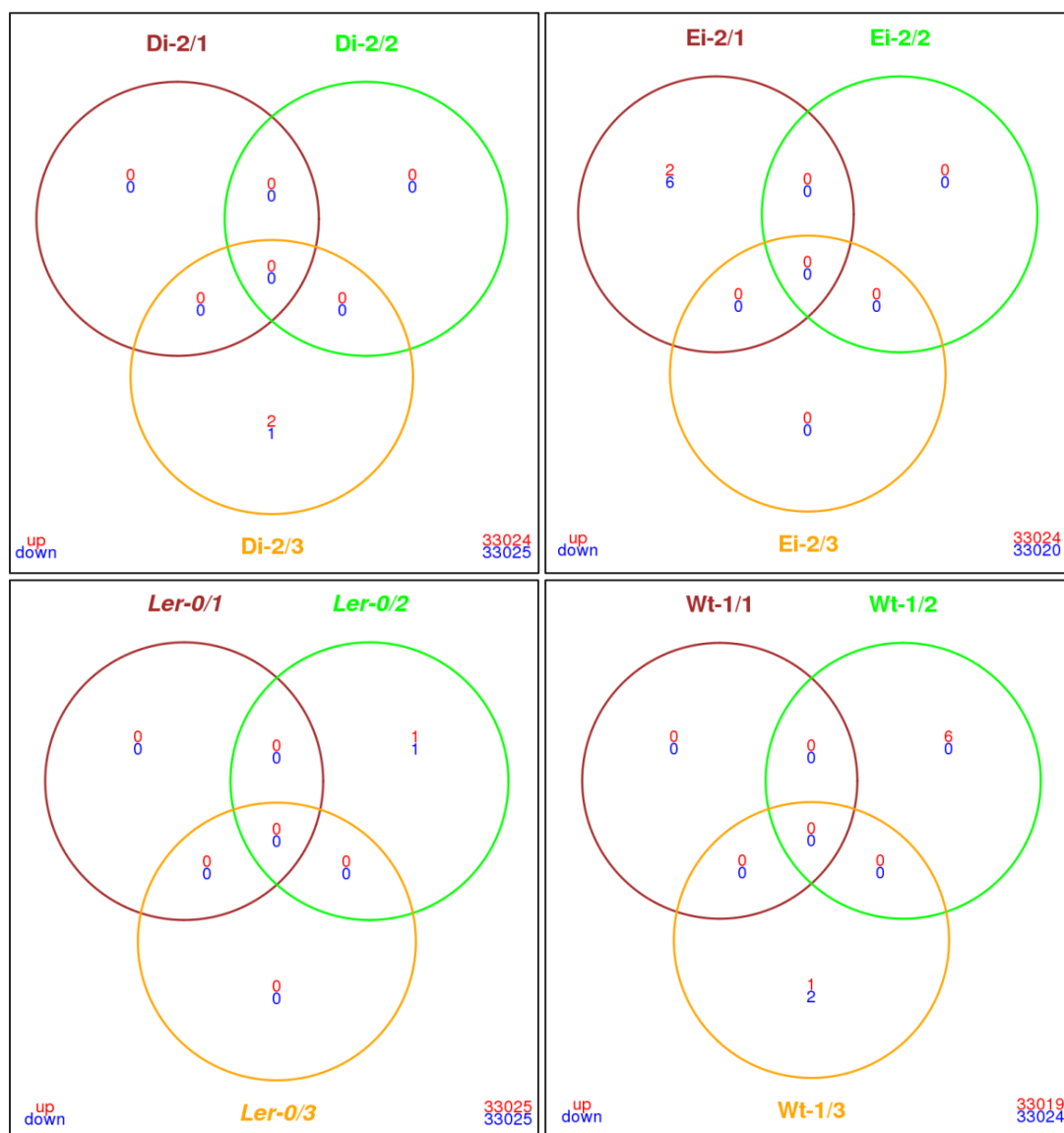


Figura 20. Diagramas de Venn de los genes diferencialmente expresados en cada linaje de virus evolucionado tras la infección del ecotipo original *Ler-0*. Cada círculo indica un linaje independiente de evolución viral. Las áreas combinadas de los círculos representan el número de genes comunes entre linajes independientes evolucionados en un ecotipo común. Los números en rojo representan los genes sobre-expresados y en azul los reprimidos. En los linajes evolucionados en *St-0*, no hubo genes diferencialmente expresados, por este motivo este ecotipo no está representado en el diagrama de intersecciones.

Uno de los dos genes afectados comunes en cuatro de los 15 linajes (*Di-2/3*, *Ei-2/1*, *Ler-0/2* y *Wt-1/2*) fue el factor de transcripción *PAP1*. Este gen sobre-expresado codifica un dominio putativo MYB de unión a ADN que contiene un factor de transcripción implicado en el metabolismo de antocianinas y eliminación de radicales, siendo esencial para la expresión mediada por sacarosa de dihidroflavonol reductasa. Además, el gen *PAP1* está implicado en el desarrollo de los síntomas de la infección por el virus del mosaico del tabaco (TMV)¹⁸⁵. También se observó un segundo gen de función desconocida (*At1g17147*) común en tres linajes (*Ei-2/1*, *Wt-1/3* y *Ler-0/2*) y que presentaba expresión disminuida.

El gen sobre-expresado *At1g17345* (Ei-2/1) pertenece a la familia de proteínas de respuesta a auxina parecidas a los *ARNs pequeños regulados positivamente por auxina* (SAUR); estos representan genes tempranos de respuesta a auxina presentes en muchas plantas, inclusive en *A. thaliana*, pero con funciones específicas en el desarrollo aún por determinar.

Wt-1/2 mostró seis genes sobre-expresados, por lo que fue uno de los linajes más modificados durante la adaptación al nuevo huésped. Uno de estos genes es el arriba descrito factor PAP1, mientras que otro codifica la proteína de dominio MYB PAP2 (*At1G66390*), casi idéntica a PAP1. Los genes restantes *At1g41870*, *At2g07530*, *At1g50691* y *At5g12100* tienen función desconocida, si bien los dos primeros tienen una similitud significativa de secuencia con retrotransposones tipo *gypsy* y el último contiene el motivo repetitivo de apentatricopéptido (PPRP). La función del motivo PPRP no está clara en *A. thaliana*, pero se ha demostrado que un gran número de este grupo de proteínas interactúan con mitocondrias y otros orgánulos¹⁵⁴ y que están posiblemente involucrados en la edición de ARN²³⁵. Ei-2/1 también mostró cinco genes reprimidos, aparte del gen común descrito arriba. Entre ellos, *At3g05741* y *At1g09360* codifican las proteínas de la superfamilia de inhibidores de la invertasa/pectina metilesterasa de plantas, que se localiza en el sistema de endomembranas. El gen *ATPUP19*, también reprimido en Ei2/1, codifica una permeasa de purinas 19 típica de *A. thaliana* y es miembro de una familia de proteínas relacionadas con el transportador de purinas PUP1, por lo que probablemente esté involucrado en el transporte de purinas y derivados de purina a través de la membrana plasmática. *At3g24065* y *At1g51230* codifican para proteínas de autoincompatibilidad de plantas ubicadas en el sistema de endomembranas que son miembros de la familia S1. Las proteínas de esta familia son un recurso importante de exogamia que se encuentra en las angiospermas. Se trata de un mecanismo que regula la aceptación o rechazo de polen y que está controlado por un solo gen multialélico S1 o por una compleja familia de genes, como es el caso de plantas de *Brassica*. La autofecundación se evita discriminando el polen que lleva el mismo alelo S que el estigma en el que aterriza⁸⁷.

Di-2/3 sobre-expresó el gen *At1g52000*, una proteína de la súperfamilia de lectina de unión a manosa e implicada en la respuesta al estrés salino. Wt-1/3 aumentó la expresión del gen *At1g22490*, que codifica el factor bHLH de la súperfamilia de proteínas de unión a ADN, un factor de transcripción localizado en el núcleo, que reconoce la secuencia específica y se une a ADN. La supresión de la expresión del gen *At3g15490*, que codifica la proteína reguladora de la actividad Vps4 en la vía de cuerpos multivesiculares, se observó durante la infección con el linaje Wt-1/3. El linaje Di-2/3 también reprimió al gen *A_84_P851955*, con función desconocida.

Para hacerse una idea de la imagen global de semejanzas de expresión génica entre los linajes en el ecotipo original Ler-0, se calcularon los coeficientes de correlación de Pearson entre las respuestas de Ler-0 a los linajes evolucionados en diferentes ecotipos y se formaron grupos de acuerdo con las semejanzas en la respuesta inducida por linajes virales en Ler-0.

La respuesta transcriptómica del ecotipo Ler-0 a los virus evolucionados en diferentes ecotipos mostró una clara correlación (73,6-89,9%) y reveló dos tipos de respuestas significativamente diferentes en las cuales no hubo más relación entre los linajes evolucionados en el ecotipo común que entre los evolucionados en diferentes ecotipos.

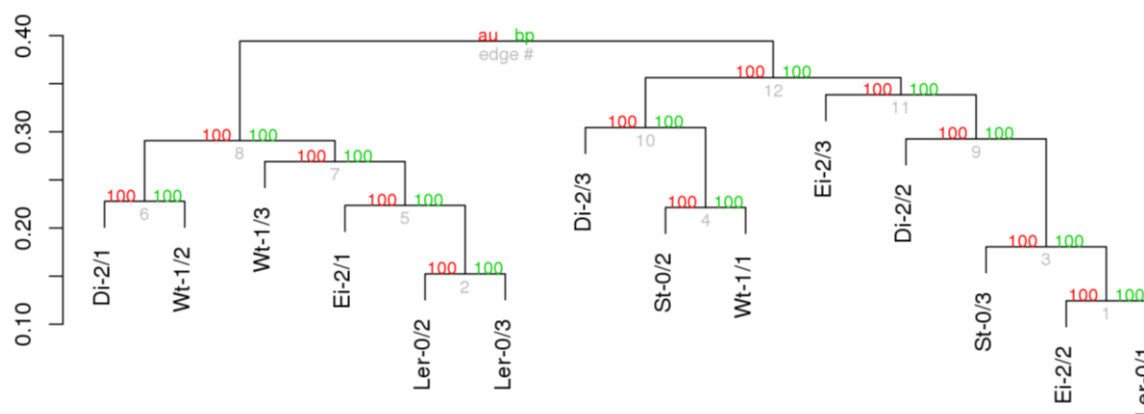


Figura 21. Dendrograma de los genes diferencialmente expresados en el ecotipo Ler-0 infectado con los linajes adaptados a nuevos huéspedes.

Edge # (color gris) representa si un clúster es más general (valores altos) o específico (valores bajos), *AU* (*Approximately unbiased*, color rojo) y *BP* (*Bootstrap probability*, color verde) son indicadores de significatividad con una escala entre 0 y 100 que proporciona el método de *bootstrap*. Valores más altos indican mayor número de muestras obtenidas por *bootstrap* corroborando ese agrupamiento. Es más robusto el valor proporcionado por AU que por BP.

A continuación se plantearon las siguientes preguntas: ¿existen categorías funcionales afectadas enriquecidas cuando el virus evolucionado en el nuevo huésped regresa al huésped original? Si es así, ¿pueden los ecotipos ser clasificados según el grado de similitud de enriquecimiento de categorías funcionales inducido por el virus evolucionado? En otras palabras, ¿hay ecotipos que requieren importantes cambios en el transcriptoma para que el virus se adapte a ellos y otros que requieren cambios menores? Se utilizó el método de análisis de conjunto de genes (GSA) para determinar los perfiles funcionales de genes afectados. La Tabla 20 resume el número de categorías funcionales enriquecidas en cada linaje tanto para los genes que incrementaron su expresión como para los genes que han sido reprimidos.

Tabla 20. Numero de categorías GO enriquecidas para los procesos biológicos.

	Reprimidos	No enriquecidos	Sobre- expresados
Di-2/1 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	28	1269	23
Di-2/2 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	109	1138	73
Di-2/3 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	97	1091	132
Ei-2/1 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	65	1236	19
Ei-2/2 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	81	1137	102
Ei-2/3 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	124	1192	4
<i>Ler-0</i> /1 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	135	1133	52
<i>Ler-0</i> /2 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	112	1195	13
<i>Ler-0</i> /3 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	129	1159	32
St-0/2 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	120	1128	72
St-0/3 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	131	1163	26
Wt-1/1 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	40	1264	16
Wt-1/2 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	316	1004	0
Wt-1/3 in <i>Ler-0</i> vs. ancestral en <i>Ler-0</i>	96	1198	26

Para visualizar la magnitud de los cambios en las categorías funcionales afectadas, cuando el ecotipo original *Ler-0* está infectado con los linajes virales evolucionados en los nuevos huéspedes, se calculó el porcentaje de genes afectados por cada ecotipo de nuevo huésped. Para ello, se determinó el número de genes afectados de todos los linajes evolucionados en el mismo ecotipo y se calculó el porcentaje de categorías afectadas en todas las categorías anotadas. La proporción de genes afectados y enriquecidos no era significativamente diferente para los linajes evolucionados en los diferentes ecotipos.

La respuesta transcriptómica era más homogénea que en el caso de la infección de diferentes ecotipos de huésped, lo que indica que el genotipo del huésped es crucial en la interacción virus-huésped y que el coste de la adaptación al huésped no seleccionado es similar para todos los linajes evolucionados. El análisis transcriptómico funcional fue en gran medida coherente con el análisis fenotípico de la eficacia viral en *Ler-0* (ver apartado 3.2.1), ya que los linajes evolucionados en los ecotipos Ei-2 y Di-2 habían afectado menos categorías funcionales de genes y, por lo tanto, no tenían ningún coste

de eficacia. Los linajes evolucionados en St-0 mostraron un descenso medio de la eficacia en Ler-0 de un 8,9%, siendo el linaje St-0/3 el más especializado, además de mostrar el mayor número categorías funcionalmente afectadas. La clasificación de perfiles funcionales según su grado de similitud no mostró una clara relación de los linajes evolucionados en el mismo ecotipo, sino que el dendograma muestra que la respuesta de Ler-0 al virus evolucionado es independiente del ecotipo en el que evolucionó el virus. Este resultado coincide con las evidencias descritas anteriormente en el apartado 3.2.2, donde se mostraba que los linajes evolucionaron fundamentalmente como generalistas, si bien se presentaban diferentes grados de especialización.

Los genes afectados tenían funciones diversas, algunas de ellas desconocidas, y sólo hubo un gen común para los cuatro linajes, el factor de transcripción PAP1, que ha sido descrito por Padmanabhan *et al.*¹⁸⁵ como directamente implicado en el desarrollo de síntomas de infección por el TMV.

3.4.4 Evaluación de las diferencias en la expresión génica de linajes generalistas y especialistas

Por último, se exploraron las respuestas transcriptómicas para los casos más extremos de especialización y generalización. Para ello, se infectaron los cinco ecotipos con estos linajes virales. Un virus especialista es aquel que replica bien en su huésped local, pero peor en otros huéspedes, mientras que un virus generalista se propaga con similar eficacia en diversos huéspedes. Por lo tanto, la comparación de la acumulación viral en plantas de diferentes ecotipos indica cómo de generalista o especialista es un virus. En la evolución experimental llevada a cabo en el presente trabajo, el linaje Ler-0/1 resultó ser el más generalista, mientras que el linaje St-0/3 fue el más especialista. Se analizaron las diferencias y similitudes en los genes alterados de ambos linajes y las categorías funcionales de estos genes. Se podría esperar que los virus especialistas y generalistas alteraran un conjunto diferente de genes del huésped y que los virus especializados mostrarán mayor grado de heterogeneidad.

El primer enfoque consistió en evaluar las diferencias de expresión génica en ambos casos. Para el linaje generalista Ler-0/1, los valores de expresión para los ecotipos infectados con este linaje se contrastaron con el ecotipo local. Como se muestra en la Figura 22, apenas hay genes diferencialmente expresados, excepto en el ecotipo Ei-2. Por tanto, las interacciones virus-huésped del virus generalista adaptado a Ler-0 parecen ser muy similares para todos los ecotipos de huéspedes, sin distinguir entre ecotipo local y alternativo, a excepción del ecotipo Ei-2.

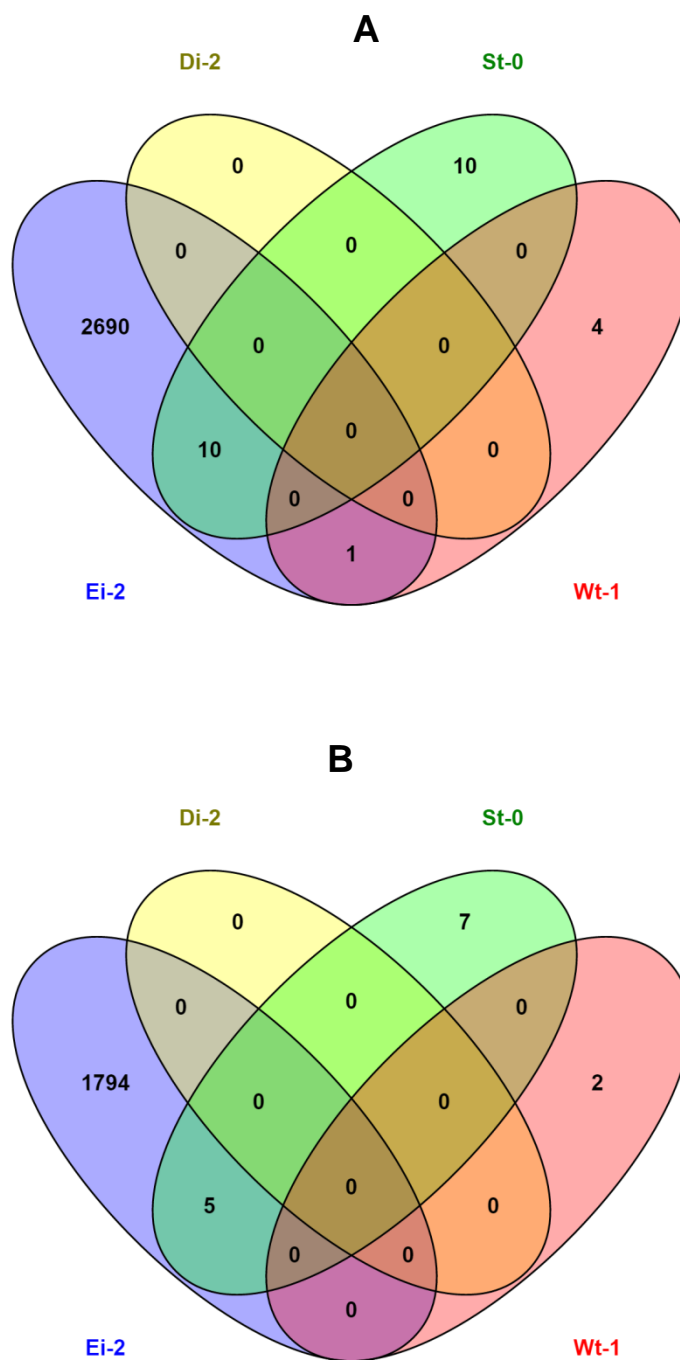


Figura 22. Diagrama de Venn del número de genes diferencialmente sobre-expresados **(A)** y reprimidos **(B)** como consecuencia de la infección con el linaje viral más generalista *Ler-0/1* en los ecotipos de huésped alternativos.

<http://bioinfofp.cnb.csic.es/tools/venny/index.html>¹⁸¹

Se realizó el mismo análisis para detectar genes diferencialmente expresados para el linaje viral más especialista *St-0/3*. En este caso, la expresión génica de los ecotipos infectados se contrastó con el transcriptoma del ecotipo *St-0* infectado con el linaje localmente adaptado *St-0/3*. Para todos los ecotipos, se obtuvo una serie de genes

diferencialmente expresados (Figura 23). El ecotipo más afectado en este caso fue también el ecotipo Ei-2, mientras que el ecotipo con cambios de menor importancia fue el ecotipo Di-2, si bien la magnitud de los cambios era similar en todos los ecotipos.

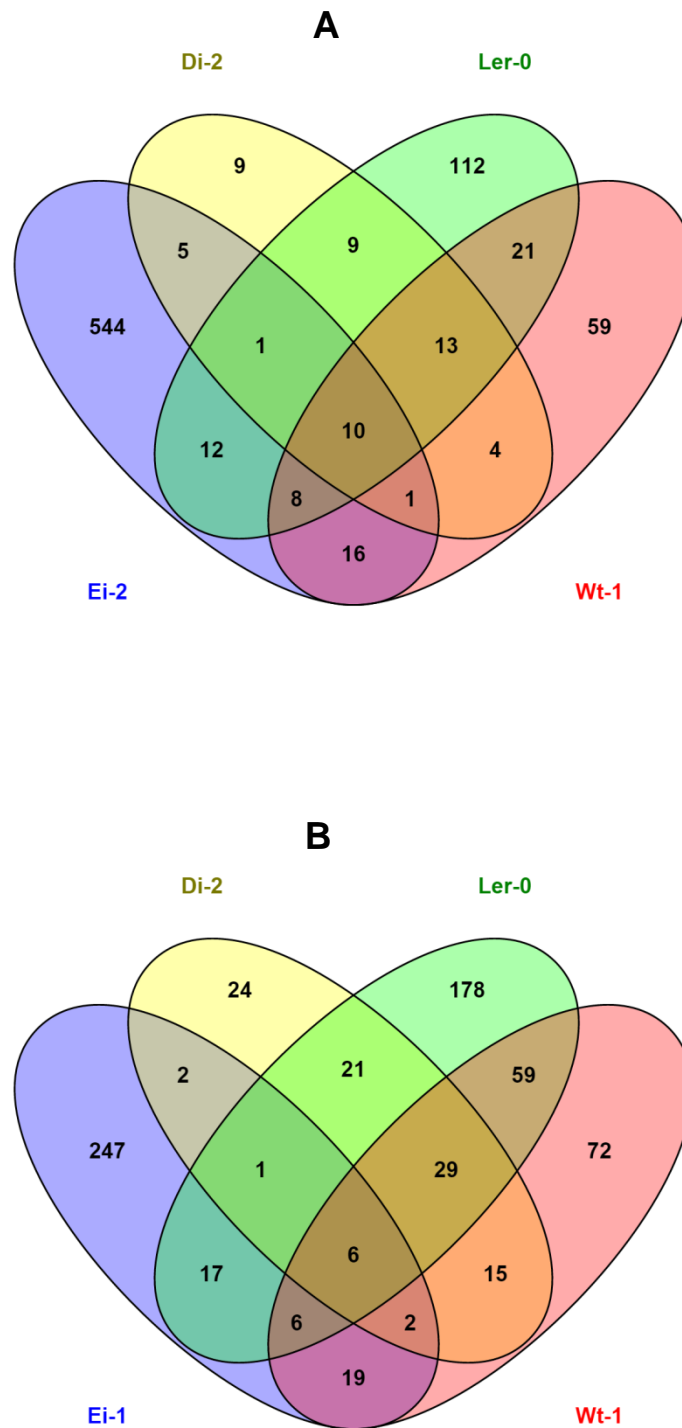


Figura 23. Diagrama de Venn del números de genes diferencialmente sobre-expresados **(A)** y reprimidos **(B)** como consecuencia de la infección con el linaje viral más especialista St-0/3 en los ecotipos de huésped alternativos.

<http://bioinfogp.cnb.csic.es/tools/venny/index.html>¹⁸¹

La respuesta transcriptómica de los ecotipos alternativos a la infección con el virus especialista es diferente de la del ecotipo local. Diez genes sobre-expresados y seis genes reprimidos son comunes para todos los ecotipos. *At5g43060* codifica una proteína

de la familia de proteasas de cisteína con repetición de granulina involucrada en el metabolismo de pectina de la pared celular y estrés salino. *At4g34180* también está implicado en el estrés salino y es una proteína de la familia ciclasa, que se localiza en la pared celular de la planta, pero cuya función molecular no se conoce todavía. La permeasa de purina 18 (ATPUP18) es un miembro de una familia de proteínas relacionadas con PUP1, transportador de purina, y está probablemente involucrado en el transporte de purina y derivados de purina, como citoquininas, a través de la membrana plasmática. WRKY26 es la proteína 26 de unión al ADN, que regula la actividad de factores de transcripción. XTH33 codifica una proteína localizada en la membrana que parece funcionar durante la modificación de la pared celular y cuya sobre-expresión resulta en una morfología celular anormal. *GLP5* codifica una proteína ubicada en plasmodesmos y está involucrada en la regulación del crecimiento de la raíz primaria mediante el control de la asignación mediada por floema de los recursos entre los meristemos primarios y laterales de la raíz. Esta proteína también está involucrada en procesos como respuesta al estrés, respuesta de resistencia adquirida, transporte del retículo al Golgi y transporte de aminoácidos. *At5g24490* es una proteína que constituye la unidad ribosomal 30S y está implicada en la traducción y en los procesos metabólicos primarios. El gen *SEN1* está asociado a la senescencia fuertemente inducida por carencia de fosfato, donde su transcripción suele estar diferencialmente regulada a nivel de estabilidad de ARNm a diferentes tiempos del día. El gen *SCPL51* codifica una proteína que tiene actividad carboxipeptidasa de tipo serina, que se encuentra en el sistema de endomembranas y está involucrada en la proteólisis. Para la sonda *TA31123_3702* no se encontró función anotada.

Varios genes reprimidos eran comunes para todos los ecotipos. *At2g44830* codifica una proteína de la súperfamilia de la proteína quinasa y está implicada en la fosforilación de aminoácidos. La proteína con dedos de zinc ZFP5 actúa en la regulación del desarrollo de tricomas y en la señalización de citoquinina. El gen *At1g80620* codifica una proteína de unión de ARN y constituyente estructural del ribosoma. *PDR1* codifica para una proteína multifuncional de membrana con actividad de unión de ATP y de ATPasa, que participa en procesos biosintéticos de cumarina, flavonoide, lignina y también está implicada en el transporte transmembrana y respuesta a heridas.

LHW codifica un activador transcripcional localizado en el núcleo que promueve la producción de células estela en los meristemos radiculares y es necesaria para establecer y mantener el número normal de células vasculares y el patrón en las raíces primarias y laterales. *At5g37970* es una proteína de la súperfamilia de metiltransferasas dependientes de S-adenosil-L-metionina pero cuya función biológica es desconocida.

Dada la circunstancia de que el número de genes diferencialmente expresados por la infección con el virus generalista *Ler-0/1* (con la excepción del ecotipo *Ei-2*) era muy

pequeño o nulo, fue imposible obtener intersecciones entre todos los ecotipos, si bien el ecotipo St-0, que tenía más de 20 genes afectados, compartía la mayor parte de estos genes con el ecotipo Ei-2. En consecuencia, con la excepción del ecotipo Ei-2, la respuesta transcriptómica entre ecotipos infectados con el linaje generalista fue homogénea, lo que indica que la respuesta del huésped a este linaje viral, incluyendo el huésped local, es independiente del ecotipo de acogida.

La cepa viral especialista St-0/3, por el contrario, tenía un mayor número de genes diferencialmente expresados, indicando que la respuesta a la infección en los ecotipos alternativos era diferente a la respuesta inducida en el ecotipo local. En cuanto a las diferencias de respuesta entre los ecotipos alternativos, los genes diferencialmente expresados eran en gran parte comunes entre ecotipos, indicando que a pesar de las diferencias en la respuesta a la infección viral entre los ecotipos alternativos y el ecotipo local, las diferencias entre los ecotipos alternativos eran mínimas.

En términos de funcionalidad de los genes afectados, en ambos linajes virales se encontraron diversas categorías enriquecidas. La magnitud de los cambios es similar en los linajes virales especialista y generalista. Sin embargo, el linaje especialista ha afectado en promedio casi al doble de categorías funcionales que el linaje viral generalista. La intersección de categorías funcionales se resume en la Figura 24 para el linaje generalista y en la Figura 25 para el especialista.

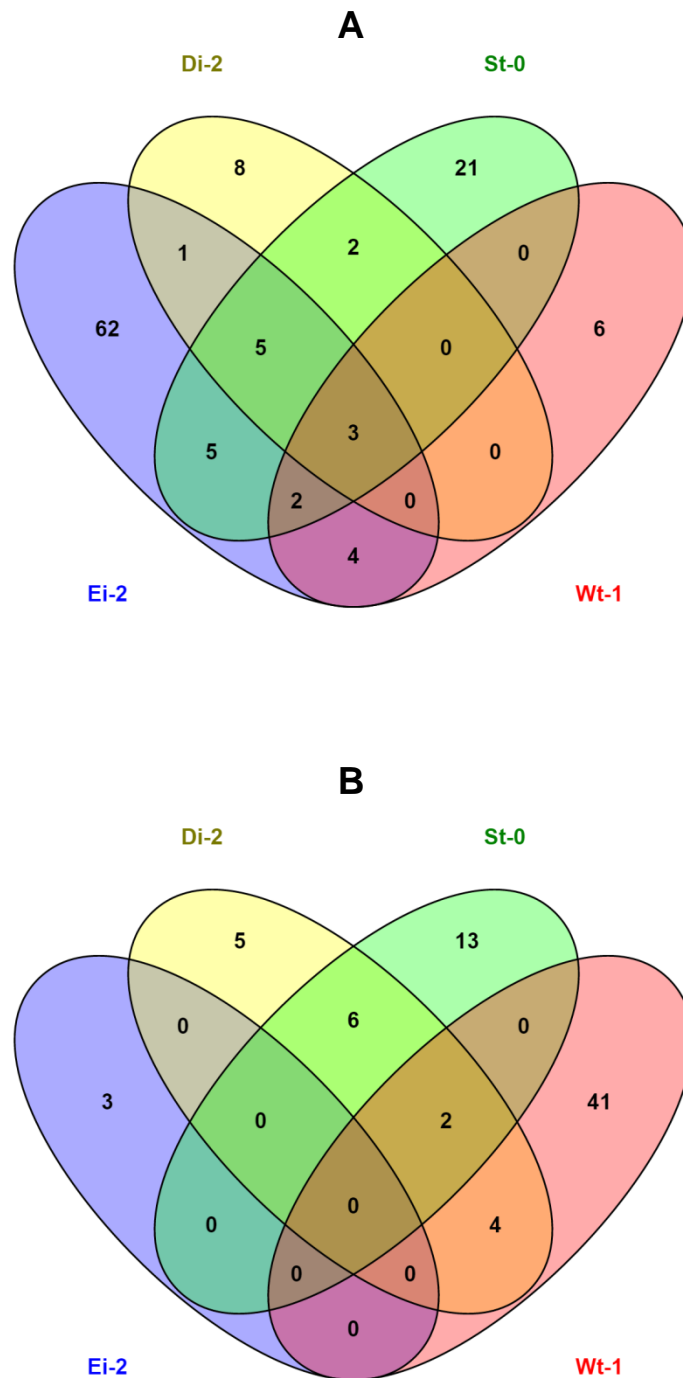


Figura 24. Diagramas de Venn de categorías funcionales enriquecidas para genes sobre-expresados **(A)** y reprimidos **(B)** como consecuencia de la infección con el linaje viral más generalista *Ler-0/1* en ecotipos de huéspedes alternativos.
<http://bioinfo.gp.cnb.csic.es/tools/venny/index.html>¹⁸¹

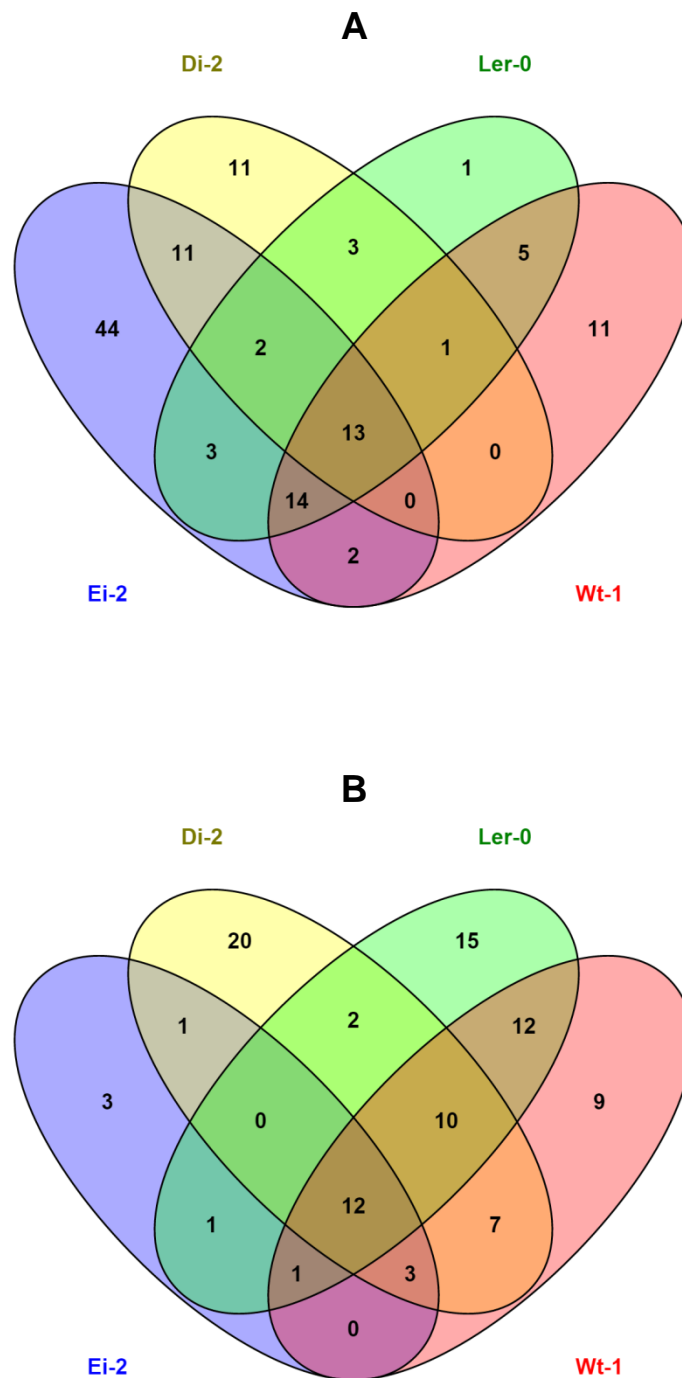


Figura 25. Diagramas de Venn de categorías funcionales enriquecidas para genes sobre-expresados **(A)** y reprimidos **(B)** como consecuencia de la infección con el linaje viral más especialista St-0/3 en ecotipos de huésped alternativos. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>¹⁸¹

Una parte de las categorías enriquecidas para el linaje más generalista Ler-0/1 es común entre todos los ecotipos o entre pares de ecotipos (Figura 24). Tres elementos eran comunes en todos los ecotipos para funciones aumentadas: regulación de la replicación del ADN (GO:0006275), proliferación celular (GO:0008283) e iniciación de la

replicación del ADN (GO:0006270). En categorías funcionales negativamente reguladas, no hubo intersecciones entre todos los ecotipos, pero sí entre pares de ecotipos. Las más redundantes incluían las categorías involucradas en importación de proteínas al núcleo (GO:0006606), respuesta al ion de cadmio (GO:0046686), traducción (GO:0006412), detección de estímulos bióticos (GO:0009595), biosíntesis del ácido salicílico (GO:0009697), respuesta a estrés salino (GO:0009651), biosíntesis de nucleótidos de purina (GO:0006164) y regulación negativa de respuesta de defensa (GO:0031348).

En el caso de la infección con el linaje específico, había un número mucho mayor de categorías comunes afectadas entre los ecotipos que en el caso del linaje generalista, probablemente debido a una distribución similar de los cambios inducidos por este linaje en los ecotipos. En total, 13 categorías funcionales estaban incrementadas por St-0/3 en todos los ecotipos, la mayoría de ellas en gran parte implicadas en la respuesta inmune. Una de ellas era estallido respiratorio implicado en la respuesta de defensa (GO:0002679), caracterizada por una fase de actividad metabólica elevada, durante la cual aumenta el consumo de oxígeno como parte de una respuesta de defensa y que, por un sistema dependiente de NADH, conduce a la producción de peróxido de hidrógeno, aniones superóxido y radicales hidroxilo. Las categorías funcionales directamente relacionadas con la infección incluyen proteína de direccionamiento a vacuola (GO:0006623), que es una característica importante para el mantenimiento de la organización y función celular¹⁷⁵, biosíntesis del etileno regulador del crecimiento (GO:0009693) y la categoría de ensamblaje de nucleosomas (GO:0006334). Otras categorías probablemente no relacionadas con la infección viral son respuesta a los estímulos mecánicos (GO:0009612), transporte del nitrato (GO:0015706) y respuesta al nitrato (GO:0010167), respuesta de defensa a los hongos (GO:0050832) y bacterias (GO:0042742). También a nivel de sobre-expresión se presentaban categorías funcionales para la regulación de proteínas como respuesta a la presencia de proteínas desnaturalizadas (GO:0051788), ensamblaje del complejo proteosomal en el núcleo (GO:0080129) y organización del Golgi (GO:0007030). En el caso de la infección con la cepa viral específica St-0/3, se encontraron doce categorías funcionales comunes reprimidas. Dos de ellas estaban relacionados con los nucleótidos, reguladores de muchos procesos celulares básicos: proceso de biosíntesis de ribonucleótido pirimidina (GO:0009220), reacciones químicas y vías resultantes en la formación del ribonucleótido pirimidina y metilación de ARN (GO:0001510). El grupo más grande, sin embargo, se asocia con la diferenciación de los tejidos y con el crecimiento de plantas como determinación de simetría bilateral (GO:0009855), iniciación de meristemas (GO:0010014), formación de patrones (GO:0010051), desarrollo del xilema (GO:0010089), biosíntesis de componentes del xilano de la pared celular (GO:0045492),

metabolismo del glucuronoxilano (GO:0010413), diferenciación celular del saco del óvulo embrionario (GO:0009560), especificación de polaridad de eje adaxial/abaxial (GO:0009944) y morfogénesis de la flor (GO:0048439). Otra categoría funcional reprimida fue la dirección de proteínas a mitocondrias (GO:0006626), un proceso integrado en una red funcional de diversos procesos fisiológicos.

El virus especialista St-0/3 activó principalmente los mecanismos de defensa y suprimió los procesos metabólicos y de crecimiento de planta en los huéspedes alternativos. En promedio, este linaje indujo una respuesta más fuerte en los nuevos huéspedes. Se encontraron muchos más genes diferencialmente expresados en el caso de la infección con el virus especialista que con el virus generalista. Esto indica que los perfiles de expresión de los huéspedes alternativos son diferentes al huésped local en el caso de la infección con el linaje especialista St-0/3, pero no respecto a la infección con Ler-0/1. Sin embargo, al observar las categorías funcionales afectadas en los nuevos huéspedes, se encuentran mucho más intersecciones en categorías funcionales afectadas por el virus especialista que por el virus generalista. Desde este punto de vista, la respuesta transcriptómica funcional a la infección con el virus especialista es más bien homogénea.

El ecotipo Ei-2 infectado con el linaje Ler-0/1 tuvo un perfil de respuesta transcriptómica excepcional, con una expresión diferencial mil veces más alta que en otros ecotipos. Sin embargo, el perfil funcional de los genes afectados era similar a otros ecotipos en cuanto a las funciones implicadas, incluso habiendo intersecciones entre los ecotipos en la mayoría de las categorías.

El análisis de categorías funcionales enriquecidas permitió identificar cambios en los huéspedes no seleccionados para ambos linajes. En el caso del linaje especializado St-0/3, había más genes comunes entre los huéspedes no seleccionados, mientras que la infección con el linaje generalista Ler-0/1 mostraba menos categorías en común entre los ecotipos. Sin embargo, la hipótesis de que el virus especializado conduce a una mayor heterogeneidad en la respuesta inmune entre los diferentes ecotipos de huéspedes no pudo ser rechazada. Esto es debido a que las categorías funcionales de los genes enriquecidos se componen de varios genes que pueden estar afectados de manera individual o pueden ser redundantes. Los resultados discrepantes del análisis de genes diferencialmente expresados apoyan la hipótesis inicial y deberían tener más peso en la evaluación de la hipótesis, porque a pesar de ser más detallados son menos complejos.

La respuesta común de los ecotipos a la infección viral era dependiente de la cepa viral infectante. Los mecanismos de defensa de las plantas se activaron en plantas infectadas con la cepa especializada y se reprimieron en plantas infectadas con la cepa viral más generalista. Los grupos de categorías funcionales de proliferación celular y replicación del ADN fueron sobre-expresados por la cepa generalista. Además, se

observó la regulación negativa del metabolismo de plantas y procesos de crecimiento para la cepa viral especialista, pero no para la generalista.

3.4.5 Discusión

Se realizaron varios análisis de datos de micromatrices para investigar a nivel transcriptómico diversas estrategias de adaptación del virus en huéspedes heterogéneos pero genéticamente muy relacionados.

Durante la evolución experimental los linajes virales cambiaron su interacción con el huésped cada uno de una manera diferente. Los perfiles transcriptómicos y funcionales de las plantas infectadas han demostrado que existe una heterogeneidad en la respuesta de ecotipos a los virus localmente evolucionados y que la correlación entre los ecotipos se mantiene en gran medida después de la evolución.

Las transcriptómicas de linajes virales evolucionados en el mismo ecotipo han sido más parecidas entre sí que su promedio comparando entre los ecotipos. Estos resultados sugieren la conclusión de que la genética del huésped juega un papel crucial en la adaptación del virus y que las pequeñas diferencias intra-específicas pueden determinar el destino de la evolución del virus.

El coste del regreso al huésped original *Ler-0* después de la adaptación del virus a nuevos huéspedes no parece ser importante, ya que apenas se detectaron genes diferencialmente expresados en ningún ecotipo y el análisis de categorías funcionales enriquecidas mostró una alta correlación entre ecotipos. En este punto, los análisis fenotípicos (ver apartado 3.2.2) están de acuerdo con los resultados de los análisis transcriptómicos; la mayoría de los linajes evolucionados no pagan ningún coste de eficacia ni al infectar al ecotipo original *Ler-0* ni a otros ecotipos no seleccionados. A pesar de las diferencias entre los linajes en sus características fenotípicas en el huésped local y alternativo, todos los linajes evolucionaron más bien como generalistas.

El linaje viral evolucionado como el más especialista *St-0/3* indujo una respuesta transcriptómica heterogénea en ecotipos no seleccionados. La respuesta funcional más redundante inducida por este linaje fue la activación de los mecanismos de defensa de los huéspedes y la supresión de categorías funcionales implicadas en metabolismo y en procesos de crecimiento vegetal.

La infección con el linaje viral más generalista *Ler-0/1* solo afectó la transcriptómica del ecotipo *Ei-2*, lo que indica que la evolución de la mayoría de virus generalistas se llevó a cabo en las dianas generales y comunes de la genética del huésped, que consistían principalmente en el incremento de la proliferación celular y la replicación del ADN, compensado por una supresión de las defensas del huésped y de procesos biosintéticos.

El análisis transcriptómico sugiere el papel crucial de la diversidad genética del huésped en el resultado de la adaptación del virus, observándose pocas diferencias en los perfiles transcriptómicos antes y después de la evolución experimental. En este sentido, el perfil funcional no siguió el mismo patrón de clasificación entre los ecotipos, lo que indica que durante la evolución no sólo han tenido lugar cambios directamente implicados en las interacciones virus-huésped, sino que también pueden haber cambiado auto-interacciones del huésped, siendo la mayoría de las vías celulares variables, ramificadas y complejas.

IV Discusión general

La comprensión de cómo sucede la emergencia de las enfermedades infecciosas y qué factores controlan su dinámica puede ayudar a predecir y controlar las epidemias. Los virus de ARN son la principal causa de enfermedades emergentes. Su capacidad para saltar de una especie a otra está favorecida por sus altas tasas de mutación, cortos tiempos de generación y grandes tamaños poblacionales, donde la diversidad genética generada de esta forma les permite adaptarse a nuevas condiciones ambientales, es decir, a nuevos huéspedes potenciales. Por otro lado, también hay que considerar la diversidad existente en los genes de resistencia del huésped como factor del que depende el éxito de la propagación viral. La interacción entre los genotipos de huésped y parásito ha sido estudiada en diferentes contextos, y por lo general, se considera que la diversidad genética para la resistencia a los patógenos en las poblaciones huéspedes reduce el riesgo de propagación de enfermedades infecciosas. En cualquier caso, muchas cuestiones acerca de esta compleja interacción aún están sin responder o carecen de pruebas empíricas. Una de ellas es, por ejemplo, cuánta diversidad es necesaria para la emergencia viral.

En trabajos previos a esta Tesis, se analizó la hipótesis de que la adaptación del virus a un nuevo huésped fuera concomitante con los cambios en el perfil de expresión génica del huésped. Agudelo-Romero *et al.*⁵ realizaron un experimento en el que TEV se adaptó al ecotipo susceptible *Ler-0* (*rtm1/rtm1 RTM2/RTM2 RTM3/RTM3*). La población ancestral de TEV era capaz de infectar sistémicamente plantas de *Ler-0*, aunque la infección progresaba de forma poco eficiente y asintomática. Sin embargo, tras 17 pases seriados en plantas *Ler-0*, el virus resultante (nombrado TEV-At17) había fijado seis mutaciones puntuales, mejorando su acumulación e induciendo síntomas visibles y más graves, que incluían retraso en el crecimiento, el grabado y la malformación de la hoja. El conjunto de genes alterados por la infección con TEV-At17 era casi tres veces más grande que el inducido por el TEV ancestral⁴. El análisis de los procesos biológicos enriquecidos con expresión alterada reveló que casi todos los procesos regulados por TEV también fueron regulados por la cepa TEV-At17, la cual adicionalmente suprimía los procesos metabólicos primarios como biosíntesis de clorofila, fijación de carbono y metabolismo de ADN, así como también los procesos de desarrollo y de ciclo celular. Esta observación llevó a Agudelo-Romero *et al.*⁵ a plantear la hipótesis de que el virus habría aumentado su eficacia mediante la adquisición de la capacidad de bloquear la activación de las defensas de las plantas. Si bien ambos virus afectaban diferencialmente la expresión de factores de transcripción principales y de proteínas altamente conectadas, en el caso del virus evolucionado la regulación se realizaba de una forma más centralizada, sobre elementos altamente conectados de la red de

regulación transcripcional²⁰⁴. Estos resultados apoyaban la hipótesis de que, mediante la adaptación a un nuevo huésped, los virus deben cambiar y profundizar la forma en que interactúan con los componentes de la red de regulación de la célula huésped.

Posteriormente, Lalić *et al.*¹³⁸ demostraron que TEV-At17 también era capaz de colonizar sistémicamente algunos ecotipos resistentes al virus TEV ancestral. Además, la infectividad, la acumulación y la gravedad de los síntomas también variaban entre los ecotipos. Estos tres rasgos se correlacionaban, de tal manera que los ecotipos donde la acumulación era mayor también eran más susceptibles y mostraban síntomas más fuertes¹³⁸.

El presente trabajo se conecta en este punto. El aislado viral utilizado como punto de partida está muy relacionado con el que analizaron Lalić *et al.*¹³⁸. El material de partida consistió en una planta paralela al último paso de la evolución experimental realizada por Agudelo-Romero *et al.*⁴. La planta entera, que había sido mantenida a -80 °C, fue pulverizada en nitrógeno líquido. Por un lado, se usó el polvo como material infeccioso para el inicio de la evolución experimental, y por otro lado, para la extracción de ARN y posterior análisis de la secuencia consenso de la población viral con la que fue infectada. La secuenciación del genoma consenso de la población viral permitió confirmar que tenía la misma secuencia previamente determinada para TEV-At17, salvo por la mutación adicional G6816A (M2224I), razón por la cual el nuevo genotipo se denominó TEV-At17b. Utilizando el material vegetal infeccioso resuspendido en tampón C (ver la sección de Materiales y Métodos) como muestra de partida, primero se comparó el fenotipo de la infección de TEV-At17b en siete ecotipos de *A. thaliana* (Col-0, Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1). Estos ecotipos se eligieron como representantes de varios grupos clasificados por Lalić *et al.*¹³⁸ por su grado de susceptibilidad a la infección con TEV-At17. Los resultados de este análisis preliminar confirmaron que la adaptación previa del virus ancestral al ecotipo Ler-0 no era específica de ecotipo y demostraron que el fenotipo de la infección era variable entre ecotipos. A continuación, se analizó el transcriptoma de seis ecotipos de *A. thaliana* (Di-2, Ei-2, Ler-0, Oy-0, St-0 y Wt-1) usando el método de micromatrices de ARN. El ecotipo Col-0 fue descartado de los análisis transcriptómicos debido a sus bajos niveles de acumulación viral. Se observaron diferencias en el perfil de expresión génica de los distintos ecotipos, al igual que ocurría con el perfil fenotípico; en otras palabras, la forma de percibir y responder a la infección variaba entre los ecotipos del huésped. Se clasificaron dos grupos claramente diferenciados en la expresión de categorías de genes funcionales como respuesta a la infección con TEV-At17b. El primer grupo se correspondía con los ecotipos que aumentaban fundamentalmente la expresión de genes de defensa (Di-2, Ler-0 y St-0). El segundo grupo incluía los ecotipos Ei-2 y Wt-1, caracterizados por la sobre-expresión de genes implicados en el estrés abiótico y en la construcción de nuevos tejidos. La mayoría

de las categorías funcionales comunes entre los ecotipos estaba involucrada en la respuesta al estrés, especialmente en la resistencia sistémica dependiente del etileno y una de ellas en el metabolismo de miARN (exonucleasa, *At3g50090*)¹⁹⁵. Cabía esperar que estas vías de defensa tuvieran relevancia en las infecciones virales en todos los ecotipos de huéspedes testados. En consecuencia, los resultados del análisis transcriptómico, junto con el análisis fenotípico, apoyaron la hipótesis de que el número de genes alterados en su expresión por la infección del virus dependía de la carga viral, ya que a mayor cantidad de virus, mayor era la perturbación generada sobre el metabolismo celular, y por lo tanto, más genes se veían afectados.

En resumen, en la primera parte del proyecto de Tesis se reveló que la interacción del virus TEV-*At17b* con los diferentes ecotipos de *A. thaliana* era heterogénea, dando lugar a diferencias en susceptibilidad, acumulación viral y gravedad de los síntomas. Esta heterogeneidad en el perfil fenotípico se correspondía a nivel transcriptómico con una regulación positiva o negativa en la expresión de diferentes conjuntos funcionales de genes.

En la segunda fase del proyecto de Tesis se utilizó el mismo patosistema (con excepción de los ecotipos Col-0 y Oy-0) para explorar si la evolución de TEV-*At17b* en ecotipos de *A. thaliana* daría lugar a la especialización o, si por el contrario, los nuevos virus evolucionados conservarían la capacidad de infectar todos los ecotipos de la misma forma que lo hacía TEV-*At17b*. Después del período de evolución experimental, que consistió en 15 pases seriados consistentes en tres semanas de infección, se evaluó infectividad, virulencia y eficacia relativa para cada población evolucionada en los distintos ecotipos de *A. thaliana* y se analizaron los perfiles transcriptómicos de los distintos ecotipos infectados. De esta forma, se podía analizar la potencial aparición de una interacción significativa entre el ecotipo del huésped y la población viral evolucionada. La magnitud de los cambios de eficacia de los virus, tras su evolución en huéspedes locales, no se vio condicionada ni por la constitución genética de los loci RTM, ni por las características iniciales de la infección con TEV-*At17b*. La resistencia determinada por los loci RTM ya habría sido posiblemente superada en la primera fase de evolución del virus TEV a *A. thaliana* Ler-0, por lo que ya no supondría un desafío para la adaptación de los virus a nuevos ecotipos. Si bien se observaron diferentes valores de eficacia entre distintos linajes evolucionados en el huésped común, a nivel transcriptómico los linajes se mostraron estrechamente relacionados. Las diferencias en eficacia pueden representar adaptación a pequeñas diferencias entre los componentes genéticos o bioquímicos de los ecotipos. Las estimas de eficacia también mostraban que todos los linajes evolucionaron como generalistas, es decir, se propagaron con éxito tanto en su ecotipo simpátrico como en los ecotipos alopátricos. Bajo el supuesto de adaptación local, se asume que un virus que evolucionara como generalista estaría en

desventaja por sus restricciones en la eficacia, y por lo tanto, un virus especialista siempre podría superar al generalista para un huésped dado. Como consecuencia, no sería esperable que un linaje adaptado a un único huésped se desarrollara como generalista. Sin embargo, Remold¹⁹⁸ propone que, bajo pleiotropía epistática, las poblaciones virales pueden lograr la especialización o generalización sin coste, porque las mismas modificaciones del genoma viral pueden tener diferentes efectos en diferentes huéspedes. De acuerdo con esta propuesta, se observó que ningún linaje viral era superior en los cinco ecotipos y que algunos aislados pagaban un coste de eficacia en el huésped no seleccionado, mientras que otros no lo hacían. A pesar de que, en términos generales, todos los linajes evolucionaban como generalistas, algunos linajes se especializaron más que otros. En este sentido, los huéspedes permisivos seleccionaron virus más especialistas, mientras que los huéspedes más restrictivos seleccionaron virus más generalistas. Asimismo, los ecotipos más permisivos inoculados con linajes localmente adaptados mostraron mayores cambios transcriptómicos que los más restrictivos, lo que no es sorprendente bajo la suposición de que estos linajes evolucionados en ecotipos permisivos lograban optimizar su interacción con el huésped modificándola y aumentando la eficacia relativa. Finalmente, el análisis de los casos más extremos de la evolución de generalista y especialista llevó a la conclusión de que el linaje más especialista inducía en promedio una respuesta transcriptómica más fuerte en nuevos huéspedes, mientras que el linaje generalista sólo afectaba diferencialmente al ecotipo Ei-2. Probablemente, la evolución del virus generalista se llevó a cabo a través de procesos generales y comunes en la genética del huésped, razón por la cual no se habrían visto diferencias entre ecotipos en la respuesta transcriptómica. En este contexto, los análisis de las matrices de infección permitían evaluar su anidamiento y modularidad. Para ello, la interacción entre los ecotipos de huéspedes y los genotipos virales fue modelada en el contexto de dos enfoques diferentes. En un extremo, existe el modelo gen a gen, donde un genotipo de parásito puede infectar a todos los genotipos de huésped y existe un genotipo de huésped universalmente susceptible⁸⁵. La resistencia se produce cuando un gen de "resistencia" del huésped se corresponde con al menos un gen de "avirulencia" del parásito. En esta situación, el polimorfismo en la infectividad y la resistencia se pueden mantener sólo si la virulencia paga un coste. En el lado opuesto, existe el modelo de los alelos coincidentes, que se basa en los sistemas de reconocimiento de lo propio y de lo extraño en los invertebrados. La infección no es posible a menos que el parásito posea todos los alelos que coinciden con los alelos del huésped⁹⁰. El enfoque clásico para el estudio de estos modelos ha consistido en inocular una serie de genotipos del huésped con una serie de genotipos del parásito y utilizar técnicas de ANOVA para evaluar si existe una interacción significativa entre huésped y parásito. El modelo gen a gen predice matrices de infectividad anidadas, ya que la gama

de huéspedes del virus especialista es un subconjunto de la gama de huéspedes del virus generalista. Por el contrario, el modelo de alelos coincidentes predice matrices de infección modulares. En este caso, la infección es probable para el virus y el huésped del mismo módulo, pero rara para los que pertenecen a distintos módulos. Cuando los huéspedes y los patógenos se infectan de forma cruzada dentro de un grupo, la matriz de interacciones es modular. Por el contrario, existen jerarquías en la resistencia dentro de los huéspedes y en la capacidad de infección entre los genotipos del virus. En el presente trabajo, la red de infección se mostró significativamente anidada. Sin embargo, los resultados de este trabajo no coinciden con ninguno de los modelos generales de infección, y por lo tanto, el mecanismo que la explica debe de ser intermedio entre ambos modelos.

A nivel molecular, todos los linajes obtenidos después de la fase de evolución, salvo uno evolucionado en *Ler-0*, mostraron sustituciones tanto sinónimas como no sinónimas. A pesar de que los perfiles transcriptómicos de los ecotipos eran muy variables entre sí, pero muy correlacionados entre los linajes evolucionados en el ecotipo común, se detectó una serie de mutaciones convergentes en varios linajes virales. La aparición de estos cambios, tanto sinónimos como no sinónimos, se podía atribuir a presiones idénticas de selección en los diversos huéspedes en conjunción con la limitación de vías adaptativas en el genoma viral. A partir de las mutaciones convergentes se llevaron a cabo estudios en mayor profundidad. En el caso particular de las mutaciones sinónimas, su aparición podría estar motivada por una mejora de la eficiencia traduccional, produciéndose cambios de codones raros por otros sinónimos presentes en mayor cantidad en la célula huésped. Otro posible factor de selección consistía en la necesidad de preservar las estructuras secundarias de ARN implicadas en regulación, así como en evitar la formación de estructuras largas de ARN de doble cadena, ya que podrían resultar un objetivo del silenciamiento de ARN. Para analizar el efecto epistático de las combinaciones de mutaciones, todas las mutaciones convergentes y sus combinaciones múltiples se clonaron y comprobaron en los huéspedes alopátricos (no locales) y simpátricos (locales). Se encontró que sólo un número reducido de las mutaciones tenía un efecto beneficioso en su huésped simpátrico. Este resultado lleva a la conclusión de que la mayoría de las mutaciones fijadas durante el experimento de evolución surgió por deriva o por barridos selectivos junto con la mutación beneficiosa que condicionó la selección. La construcción de dos paisajes de eficacia empíricos confirmó las observaciones anteriores en TEV^{139,140} al mostrar que el signo y la magnitud de la epistasia eran variables entre mutaciones y entre huéspedes. Sin embargo, mientras estudios previos mostraban epistasia promedio positiva y generalmente de signo recíproco, los resultados de esta Tesis muestran que los efectos mutacionales son fundamentalmente multiplicativos y se dan muy pocos

casos de epistasia significativa. La contradicción es solo aparente, ya que en los estudios previos las mutaciones analizadas eran elegidas al azar respecto a su historia evolutiva mientras que en el presente análisis eran mutaciones fijadas tras un proceso de evolución experimental, siendo algunas beneficiosas y las dianas directas de la selección y otras neutrales, o ligeramente deletéreas, en el fondo genético de las beneficiosas.

Llegando a este punto de la Tesis, aparte de comparar los resultados del análisis transcriptómico de las poblaciones evolucionadas con las características fenotípicas de las interacciones de linajes evolucionados en diferentes huéspedes, resultaba preciso identificar las dianas moleculares de adaptación del virus a los nuevos huéspedes. En este abordaje, se analizaron los cambios transcriptómicos y categorías funcionales de cada gen por causa de los linajes evolucionados en sus huéspedes nuevos y en comparación con plantas no infectadas e infectadas con el virus ancestral. También se analizaron los perfiles transcriptómicos de todos los linajes evolucionados en el ecotipo inicial *Ler-0* para determinar el coste de la adaptación al ecotipo nuevo en el ecotipo original. De esta forma, se obtuvo una aproximación de la magnitud de los cambios producidos en la interacción virus-huésped en diferentes contextos. El número de genes diferencialmente expresados era variable entre ecotipos, habiendo claramente más diferencias entre los ecotipos que entre los linajes evolucionados en el ecotipo común. Esto era una indicación de que las dianas de adaptación de los linajes virales en un ecotipo común son más parecidas y que las restricciones del huésped son cruciales para el destino de la evolución viral. Entre los ecotipos, al igual que ocurrió en el caso del virus ancestral, hubo dos tipos principales de respuesta funcional específica a la infección con el virus localmente evolucionado.

Las dianas comunes de la evolución viral en los distintos ecotipos resultaron ser procesos principales de la activación de las vías de defensa inmunitarias, procesos directamente implicados en el ciclo de vida del virus, biosíntesis de los recursos básicos de la planta como la regulación de la actividad catalítica, el proceso catabólico de aminoácidos y procesos de biosíntesis de metabolitos secundarios. Mientras los ecotipos de un grupo aumentaron ciertas funciones celulares con el fin de favorecer la propagación del virus, algunas de esas mismas funciones fueron reprimidas en el otro grupo de ecotipos con el mismo fin de mejorar la eficacia del virus. Es interesante que en todos los ecotipos se observara una represión considerable de la biosíntesis y de la modificación de proteínas, mientras que se aumentó la proliferación celular. Sin embargo, la mayor parte de la respuesta funcional de los ecotipos al virus evolucionado era específica de ecotipo.

En cuanto al coste en términos de eficacia que pueden pagar los linajes evolucionados al volver a infectar el huésped original *Ler-0*, hay que destacar que después de la adaptación del virus a nuevos huéspedes las modificaciones

transcriptómicas no resultaron ser grandes. Muestra de ello es que apenas se presentaban genes diferencialmente expresados en ningún ecotipo. Adicionalmente, el análisis de categorías funcionales de genes enriquecidos mostró una alta correlación entre los linajes, indicando una respuesta del ecotipo Ler-0 independiente del huésped en el cual evolucionó el virus. Los análisis de eficacia relativa en infecciones cruzadas, que apuntan a que todos los linajes evolucionaron como generalistas, coincidían en este punto con el análisis transcriptómico, ya que para la mayoría de los linajes evolucionados no hubo coste significativo de eficacia cuando infectaron el ecotipo original Ler-0, ni tampoco cuando infectaron otros ecotipos no seleccionados.

En este trabajo, se llevó a cabo la evolución de un virus emergente en diferentes ecotipos intra-específicos y se analizó el resultado de esta evolución en diferentes contextos con el objetivo de determinar el papel de la variabilidad genética del huésped en la susceptibilidad a la infección, así como estimar la variabilidad necesaria capaz de determinar el destino de la evolución viral. En este sentido, tanto el análisis fenotípico como el transcriptómico sugieren el papel crucial de la diversidad genética del huésped para el resultado de la adaptación del virus y demuestran que incluso los huéspedes que pertenecen a la misma especie dan lugar a una evolución viral divergente. Sin embargo, la evolución de los virus ha presentado una estrategia fundamentalmente generalista, probablemente debido a un carácter intrínsecamente permisivo de los huéspedes. Esto habría llevado a que no infringieran suficientes restricciones como para que el virus tuviera que recurrir a estrategias adaptativas más específicas de cada ecotipo. En cualquier caso, no es posible descartar que tiempos de evolución más largos causaran adaptaciones más específicas de huésped.

V Conclusiones finales

Las investigaciones de esta Tesis Doctoral permiten llegar a las siguientes conclusiones:

1. La heterogeneidad en el perfil fenotípico de interacción de diferentes ecotipos de *A. thaliana* con el virus TEV-At17b se corresponde a nivel transcriptómico con una alteración en la expresión de diferentes conjuntos funcionales de genes.
2. La evolución experimental de una población viral en un ecotipo determinado conlleva cambios en la interacción virus-huésped. La magnitud de estos cambios está claramente condicionada por la genética del huésped.
3. Los análisis fenotípicos y transcriptómicos de las interacciones virus-huésped sugieren el papel crucial de la diversidad genética del huésped en el resultado de la adaptación del virus. Incluso las diferencias intra-específicas, como las analizadas en el presente trabajo, pueden dar lugar a una evolución viral divergente.
4. Globalmente, la evolución de todos los linajes virales ha seguido una estrategia generalista, si bien se ha observado un cierto grado de especialización en algunos de ellos. Esta situación puede haber sido consecuencia de un carácter intrínsecamente permisivo de los huéspedes, aunque tampoco se puede descartar que tiempos de evolución más largos pudieran promover una mayor especialización.
5. El linaje viral con mayor grado de especialización inducía en promedio una respuesta transcriptómica más fuerte en los nuevos huéspedes que el linaje generalista. Esto sugiere que la evolución del virus generalista probablemente se llevó a cabo a través de procesos generales y comunes en la genética del huésped.
6. Las mutaciones convergentes observadas en varios linajes virales pueden ser atribuidas a presiones idénticas de selección en los distintos

huéspedes, además de evidenciar claras limitaciones en las estrategias adaptativas del genoma viral.

7. Sólo un número reducido de mutaciones tenía un efecto beneficioso en su huésped simpátrico, lo que sugiere que la mayoría de las mutaciones fijadas durante el experimento de evolución surgió por deriva o por barridos selectivos, junto con la potencial mutación beneficiosa que condicionó la selección.

En conjunto, el procedimiento experimental parece mostrar un efecto considerable de la deriva, posiblemente debida a los cuellos de botella presentes durante el proceso de inoculación. La comparación entre las condiciones naturales de infección y las mostradas aquí resulta relevante en un futuro para estimar con mayor precisión las dinámicas adaptativas virales.

8. El análisis de los efectos de los linajes virales evolucionados en distintos ecotipos sobre su ecotipo original no supuso ningún coste, salvo en algunos casos. De hecho, el análisis de las categorías funcionales de genes enriquecidos del ecotipo original Ler-0 mostraba una respuesta independiente del huésped en el cual había evolucionado el virus, lo que evidencia nuevamente el carácter generalista de la adaptación viral en el presente estudio.
9. La aparición de mutaciones sinónimas en los linajes virales evolucionados podría estar motivada por una mejora de la eficiencia traduccional en la célula huésped, como se ha podido evidenciar, o por la necesidad de preservar las estructuras secundarias de ARN.

Referencias

1. **Acevedo, A., L. Brodsky, and R. Andino.** 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**:686-690.
2. **Adams, M. J., J. F. Antoniw, and F. Beudoin.** 2005. Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Mol. Plant Pathol.* **6**:471-487.
3. **Agol, V. I.** 2006. Molecular mechanisms of poliovirus variation and evolution. *Curr. Top. Microbiol. Immunol.* **299**:211-259.
4. **Agudelo-Romero, P., P. Carbonell, I. de, I, J. Carrera, G. Rodrigo, A. Jaramillo, M. A. Perez-Amador, and S. F. Elena.** 2008. Changes in the gene expression profile of *Arabidopsis thaliana* after infection with Tobacco etch virus. *Viol. J.* **5**:92.
5. **Agudelo-Romero, P., P. Carbonell, M. A. Perez-Amador, and S. F. Elena.** 2008. Virus adaptation by manipulation of host's gene expression. *PLoS. One.* **3**:e2397.
6. **Agudelo-Romero, P., I. de, I, and S. F. Elena.** 2008. The pleiotropic cost of host-specialization in Tobacco etch potyvirus. *Infect. Genet. Evol.* **8**:806-814.
7. **Ahlquist, P., A. O. Noueir, W. M. Lee, D. B. Kushner, and B. T. Dye.** 2003. Host factors in positive-strand RNA virus genome replication. *J. Virol.* **77**:8181-8186.
8. **Ala-Poikela, M., E. Goytia, T. Haikonen, M. L. Rajamaki, and J. P. Valkonen.** 2011. Helper component proteinase of the genus Potyvirus is an interaction partner of translation initiation factors eIF(iso)4E and eIF4E and contains a 4E binding motif. *J. Virol.* **85**:6784-6794.
9. **Alan J.Cann.** 2005. Principles of Molecular Virology.
10. **Anderson, R. M. and R. M. May.** 1982. Coevolution of hosts and parasites. *Parasitology* **85 (Pt 2)**:411-426.
11. **Anindya, R., S. Chittori, and H. S. Savithri.** 2005. Tyrosine 66 of Pepper vein banding virus genome-linked protein is uridylylated by RNA-dependent RNA polymerase. *Virology* **336**:154-162.
12. **Babu, M., J. S. Griffiths, T. S. Huang, and A. Wang.** 2008. Altered gene expression changes in *Arabidopsis* leaf tissues and protoplasts in response to Plum pox virus infection. *BMC. Genomics* **9**:325.
13. **Bailer, S. M. and J. Haas.** 2009. Connecting viral with cellular interactomes. *Curr. Opin. Microbiol.* **12**:453-459.
14. **Barber, M. J.** 2007. Modularity and community detection in bipartite networks. *Phys. Rev. E. Stat. Nonlin. Soft. Matter Phys.* **76**:066102.
15. **Barrell, B. G., G. M. Air, and C. A. Hutchison, III.** 1976. Overlapping genes in bacteriophage phiX174. *Nature* **264**:34-41.

16. **Bascompte, J., P. Jordano, C. J. Melian, and J. M. Olesen.** 2003. The nested assembly of plant-animal mutualistic networks. *Proc. Natl. Acad. Sci. U. S. A* **100**:9383-9387.
17. **Beauchemin, C., N. Boutet, and J. F. Laliberte.** 2007. Visualization of the interaction between the precursors of VPg, the viral protein linked to the genome of turnip mosaic virus, and the translation eukaryotic initiation factor iso 4E in *Planta*. *J. Virol.* **81**:775-782.
18. **Beauchemin, C. and J. F. Laliberte.** 2007. The poly(A) binding protein is internalized in virus-induced vesicles or redistributed to the nucleolus during turnip mosaic virus infection. *J. Virol.* **81**:10905-10913.
19. **Bedhomme, S., G. Lafforgue, and S. F. Elena.** 2012. Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. *Mol. Biol. Evol.* **29**:1481-1492.
20. **Bedoya, L. C. and J. A. Daros.** 2010. Stability of Tobacco etch virus infectious clones in plasmid vectors. *Virus Res.* **149**:234-240.
21. **Belshaw, R., O. G. Pybus, and A. Rambaut.** 2007. The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res.* **17**:1496-1504.
22. **Benjamini, Y. and Hochberg Y.** 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*.
23. **Benjamini, Y. and Yekutieli D.** 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*.
24. **Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed.** 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* **19**:185-193.
25. **Boudart, G., E. Jamet, M. Rossignol, C. Lafitte, G. Borderies, A. Jauneau, M. T. Esquerre-Tugaye, and R. Pont-Lezica.** 2005. Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. *Proteomics.* **5**:212-221.
26. **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.
27. **Burgyan, J. and Z. Havelda.** 2011. Viral suppressors of RNA silencing. *Trends Plant Sci.* **16**:265-272.
28. **Cabanillas, L., M. Arribas, and E. Lazaro.** 2013. Evolution at increased error rate leads to the coexistence of multiple adaptive pathways in an RNA virus. *BMC. Evol. Biol.* **13**:11.
29. **Callaway, A., D. Giesman-Cookmeyer, E. T. Gillock, T. L. Sit, and S. A. Lommel.** 2001. The multifunctional capsid proteins of plant RNA viruses. *Annu. Rev. Phytopathol.* **39**:419-460.
30. **Carr, T. and Whitham S.A.** 2007. An emerging model system: *Arabidopsis* as a viral host plant. *Plant Cell Monogr.* **7**:159-183.

31. **Carrasco, J. L., G. Ancillo, E. Mayda, and P. Vera.** 2003. A novel transcription factor involved in plant defense endowed with protein phosphatase activity. *EMBO J.* **22**:3376-3384.
32. **Carrasco, P., J. A. Daros, P. Agudelo-Romero, and S. F. Elena.** 2007. A real-time RT-PCR assay for quantifying the fitness of tobacco etch virus in competition experiments. *J. Virol. Methods* **139**:181-188.
33. **Carrasco, P., I. de I, and S. F. Elena.** 2007. Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco Etch virus. *J. Virol.* **81**:12979-12984.
34. **Carrington, J. C. and D. D. Freed.** 1990. Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. *J. Virol.* **64**:1590-1597.
35. **Carrington, J. C. and K. L. Herndon.** 1992. Characterization of the potyviral HC-pro autoproteolytic cleavage site. *Virology* **187**:308-315.
36. **Carrington, J. C., P. E. Jensen, and M. C. Schaad.** 1998. Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. *Plant J.* **14**:393-400.
37. **Castello, M. J., J. L. Carrasco, M. Navarrete-Gomez, J. Daniel, D. Granot, and P. Vera.** 2011. A plant small polypeptide is a novel component of DNA-binding protein phosphatase 1-mediated resistance to plum pox virus in Arabidopsis. *Plant Physiol* **157**:2206-2215.
38. **Caudle, S. B., C. R. Miller, and D. R. Rokytka.** 2014. Environment determines epistatic patterns for a ssDNA virus. *Genetics* **196**:267-279.
39. **Chao, L., K. A. Hanley, C. L. Burch, C. Dahlberg, and P. E. Turner.** 2000. Kin selection and parasite evolution: higher and lower virulence with hard and soft selection. *Q. Rev. Biol.* **75**:261-275.
40. **Chapman, E. J., A. I. Prokhnevsky, K. Gopinath, V. V. Dolja, and J. C. Carrington.** 2004. Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes Dev.* **18**:1179-1186.
41. **Charron, C., M. Nicolai, J. L. Gallois, C. Robaglia, B. Moury, A. Palloix, and C. Caranta.** 2008. Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. *Plant J.* **54**:56-68.
42. **Chisholm, S. T., S. K. Mahajan, S. A. Whitham, M. L. Yamamoto, and J. C. Carrington.** 2000. Cloning of the Arabidopsis RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus. *Proc. Natl. Acad. Sci. U. S. A* **97**:489-494.
43. **Chisholm, S. T., M. A. Parra, R. J. Anderberg, and J. C. Carrington.** 2001. Arabidopsis RTM1 and RTM2 genes function in phloem to restrict long-distance movement of tobacco etch virus. *Plant Physiol* **127**:1667-1675.
44. **Chung, B. Y., W. A. Miller, J. F. Atkins, and A. E. Firth.** 2008. An overlapping essential gene in the Potyviridae. *Proc. Natl. Acad. Sci. U. S. A* **105**:5897-5902.
45. **Citovsky, V., S. Ghoshroy, F. Tsui, and D. Klessig.** 1998. Non-toxic concentrations of cadmium inhibit systemic movement of turnip vein clearing virus by a salicylic acid-independent mechanism. *Plant J.* **16**:13-20.

46. **Codoner, F. M., J. A. Daros, R. V. Sole, and S. F. Elena.** 2006. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS. Pathog.* **2**:e136.
47. **Coleman, J. R., D. Papamichail, S. Skiena, B. Futcher, E. Wimmer, and S. Mueller.** 2008. Virus attenuation by genome-scale changes in codon pair bias. *Science* **320**:1784-1787.
48. **Contreras-Paredes, C. A., L. Silva-Rosales, J. A. Daros, N. D. Alejandri-Ramirez, and T. D. Dinkova.** 2013. The absence of eukaryotic initiation factor eIF(iso)4E affects the systemic spread of a Tobacco etch virus isolate in *Arabidopsis thaliana*. *Mol. Plant Microbe Interact.* **26**:461-470.
49. **Cosgrove, M. S.** 2006. PHinDing a new histone "effector" domain. *Structure.* **14**:1096-1098.
50. **Cosson, P., V. Schurdi-Levraud, Q. H. Le, O. Sicard, M. Caballero, F. Roux, O. Le Gall, T. Candresse, and F. Revers.** 2012. The RTM resistance to potyviruses in *Arabidopsis thaliana*: natural variation of the RTM genes and evidence for the implication of additional genes. *PLoS. One.* **7**:e39169.
51. **Cosson, P., L. Sofer, Q. H. Le, V. Leger, V. Schurdi-Levraud, S. A. Whitham, M. L. Yamamoto, S. Gopalan, O. Le Gall, T. Candresse, J. C. Carrington, and F. Revers.** 2010. RTM3, which controls long-distance movement of potyviruses, is a member of a new plant gene family encoding a meprin and TRAF homology domain-containing protein. *Plant Physiol* **154**:222-232.
52. **Cosson, P., L. Sofer, V. Schurdi-Levraud, and F. Revers.** 2010. A member of a new plant gene family encoding a meprin and TRAF homology (MATH) domain-containing protein is involved in restriction of long distance movement of plant viruses. *Plant Signal. Behav.* **5**:1321-1323.
53. **Cotton, S., R. Grangeon, K. Thivierge, I. Mathieu, C. Ide, T. Wei, A. Wang, and J. F. Laliberte.** 2009. Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. *J. Virol.* **83**:10460-10471.
54. **Cuevas, J. M., P. Domingo-Calap, and R. Sanjuan.** 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. *Mol. Biol. Evol.* **29**:17-20.
55. **Cuevas, J. M., S. F. Elena, and A. Moya.** 2002. Molecular basis of adaptive convergence in experimental populations of RNA viruses. *Genetics* **162**:533-542.
56. **Cui, X., T. Wei, R. V. Chowda-Reddy, G. Sun, and A. Wang.** 2010. The Tobacco etch virus P3 protein forms mobile inclusions via the early secretory pathway and traffics along actin microfilaments. *Virology* **397**:56-63.
57. **Dardick, C.** 2007. Comparative expression profiling of *Nicotiana benthamiana* leaves systemically infected with three fruit tree viruses. *Mol. Plant Microbe Interact.* **20**:1004-1017.
58. **Decroocq, V., B. Salvador, O. Sicard, M. Glasa, P. Cosson, L. Svanella-Dumas, F. Revers, J. A. Garcia, and T. Candresse.** 2009. The determinant of potyvirus ability to overcome the RTM resistance of *Arabidopsis thaliana* maps to the N-terminal region of the coat protein. *Mol. Plant Microbe Interact.* **22**:1302-1311.

59. **DeFilippis V.R and Villareal L.P.** 2000. An introduction to evolutionary ecology of viruses, p. 126-208. *Viral Ecology*. Academic Press , New York.
60. **den Boon, J. A. and P. Ahlquist.** 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu. Rev. Microbiol.* **64**:241-256.
61. **Ding, S. W. and O. Voinnet.** 2007. Antiviral immunity directed by small RNAs. *Cell* **130**:413-426.
62. **Dolja, V. V., R. Haldeman, N. L. Robertson, W. G. Dougherty, and J. C. Carrington.** 1994. Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *EMBO J.* **13**:1482-1491.
63. **Dolja, V. V., R. Haldeman-Cahill, A. E. Montgomery, K. A. Vandenbosch, and J. C. Carrington.** 1995. Capsid protein determinants involved in cell-to-cell and long distance movement of tobacco etch potyvirus. *Virology* **206**:1007-1016.
64. **Domingo E.** 1999. *Encyclopedia of Virology*. Academic Press, London (UK).
65. **Domingo, E.** 1997. Rapid evolution of viral RNA genomes. *J. Nutr.* **127**:958S-961S.
66. **Domingo, E.** 2010. Mechanisms of viral emergence. *Vet. Res.* **41**:38.
67. **Domingo, E. and Holland J.J.** 1997. RNA virus mutations and fitness for survival. *Annual Review of Microbiology* **51**:151-178.
68. **Domingo, E., V. Martin, C. Perales, A. Grande-Perez, J. Garcia-Arriaza, and A. Arias.** 2006. Viruses as quasispecies: biological implications. *Curr. Top. Microbiol. Immunol.* **299**:51-82.
69. **Dougherty, W. G., J. C. Carrington, S. M. Cary, and T. D. Parks.** 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J.* **7**:1281-1287.
70. **Doumayrou, J., A. Avellan, R. Froissart, and Y. Michalakis.** 2013. An experimental test of the transmission-virulence trade-off hypothesis in a plant virus. *Evolution* **67**:477-486.
71. **Duffy, S., C. L. Burch, and P. E. Turner.** 2007. Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution* **61**:2614-2622.
72. **Duffy, S., P. E. Turner, and C. L. Burch.** 2006. Pleiotropic costs of niche expansion in the RNA bacteriophage phi 6. *Genetics* **172**:751-757.
73. **Dufresne, P. J., K. Thivierge, S. Cotton, C. Beauchemin, C. Ide, E. Ubalijoro, J. F. Laliberte, and M. G. Fortin.** 2008. Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles. *Virology* **374**:217-227.
74. **Dunoyer, P., C. H. Lecellier, E. A. Parizotto, C. Himber, and O. Voinnet.** 2004. Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. *Plant Cell* **16**:1235-1250.
75. **Edward C.Holmes.** 2009. *The Evolution and Emergence of RNA Viruses*.

76. **Eiamtanasate, S., M. Juricek, and Y. K. Yap.** 2007. C-terminal hydrophobic region leads PRSV P3 protein to endoplasmic reticulum. *Virus Genes* **35**:611-617.
77. **Elena, S. F., S. Bedhomme, P. Carrasco, J. M. Cuevas, I. de, I, G. Lafforgue, J. Lalić, A. Prosper, N. Tromas, and M. P. Zwart.** 2011. The evolutionary genetics of emerging plant RNA viruses. *Mol. Plant Microbe Interact.* **24**:287-293.
78. **Elena, S. F., P. Carrasco, J. A. Daros, and R. Sanjuan.** 2006. Mechanisms of genetic robustness in RNA viruses. *EMBO Rep.* **7**:168-173.
79. **Elena, S. F. and R. Sanjuan.** 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J. Virol.* **79**:11555-11558.
80. **Elena, S. F. and R. Sanjuan.** 2007. Virus Evolution: Insights from an Experimental Approach. *Annual Review of Ecology, Evolution, and Systematics* **38**:27-52.
81. **Elena, S. F., R. V. Sole, and J. Sardanyes.** 2010. Simple genomes, complex interactions: epistasis in RNA virus. *Chaos.* **20**:026106.
82. **Fellers, J., J. Wan, Y. Hong, G. B. Collins, and A. G. Hunt.** 1998. In vitro interactions between a potyvirus-encoded, genome-linked protein and RNA-dependent RNA polymerase. *J. Gen. Virol.* **79 (Pt 8)**:2043-2049.
83. **Flenniken, M. L.** 2014. Honey bee-infecting plant virus with implications on honey bee colony health. *MBio.* **5**:e00877-14.
84. **Flor H.H.** 1971. Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology* **9**:275-296.
85. **Flor, H. H.** 1956. The complementary genetic systems in flax and flax rush. *Adv. Genet* 29-54.
86. **Flores, C. O., J. R. Meyer, S. Valverde, L. Farr, and J. S. Weitz.** 2011. Statistical structure of host-phage interactions. *Proc. Natl. Acad. Sci. U. S. A* **108**:E288-E297.
87. **Foote, H. C., J. P. Ride, V. E. Franklin-Tong, E. A. Walker, M. J. Lawrence, and F. C. Franklin.** 1994. Cloning and expression of a distinctive class of self-incompatibility (S) gene from *Papaver rhoeas* L. *Proc. Natl. Acad. Sci. U. S. A* **91**:2265-2269.
88. **Fraenkel-Conrat, H. and B. Singer.** 1957. Virus reconstitution. II. Combination of protein and nucleic acid from different strains. *Biochim. Biophys. Acta* **24**:540-548.
89. **Fraenkel-Conrat, H., B. Singer, and R. C. Williams.** 1957. Infectivity of viral nucleic acid. *Biochim. Biophys. Acta* **25**:87-96.
90. **Frank, S. A.** 1996. Models of parasite virulence. *Q. Rev. Biol.* **71**:37-78.
91. **Frédéric Revers and Juan Antonio García.** 2015. Molecular Biology of Potyviruses, p. 102-199. *Advances in Virus Research.* Elsevier Inc.
92. **Fry J.D.** 1996. The evolution of host specialization: are tradeoffs overrated? *Am. Nat.* **148**:84-107.
93. **Furio, V., A. Moya, and R. Sanjuan.** 2005. The cost of replication fidelity in an RNA virus. *Proc. Natl. Acad. Sci. U. S. A* **102**:10233-10237.

94. **Futuyma D.J and Moreno G.** 1988. The evolution of ecological specialization. *Annu Rev Ecol Syst* **19**.
95. **Gabrenaite-Verkhovskaya, R., I. A. Andreev, N. O. Kalinina, L. Torrance, M. E. Taliansky, and K. Makinen.** 2008. Cylindrical inclusion protein of potato virus A is associated with a subpopulation of particles isolated from infected plants. *J. Gen. Virol.* **89**:829-838.
96. **Gadh, I. P. and V. Hari.** 1986. Association of tobacco etch virus related RNA with chloroplasts in extracts of infected plants. *Virology* **150**:304-307.
97. **Garcia-Arenal, F. and Fraile A.** 2013. Trade-offs in host range evolution of plant viruses. *Plant Pathology* **62**:2-9.
98. **George N.Agrios.** 2005. *Plant pathology.* Elsevier Academic Press.
99. **Gong, Z., H. Koiwa, M. A. Cushman, A. Ray, D. Bufford, S. Kore-eda, T. K. Matsumoto, J. Zhu, J. C. Cushman, R. A. Bressan, and P. M. Hasegawa.** 2001. Genes that are uniquely stress regulated in salt overly sensitive (sos) mutants. *Plant Physiol* **126**:363-375.
100. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17-37.
101. **Govier, D. A., B. Kassanis, and T. P. Pirone.** 1977. Partial purification and characterization of the potato virus Y helper component. *Virology* **78**:306-314.
102. **Grande-Perez, A., E. Lazaro, P. Lowenstein, E. Domingo, and S. C. Manrubia.** 2005. Suppression of viral infectivity through lethal defection. *Proc. Natl. Acad. Sci. U. S. A* **102**:4448-4452.
103. **Grangeon, R., M. Agbeci, J. Chen, G. Grondin, H. Zheng, and J. F. Laliberte.** 2012. Impact on the endoplasmic reticulum and Golgi apparatus of turnip mosaic virus infection. *J. Virol.* **86**:9255-9265.
104. **Grangeon, R., J. Jiang, and J. F. Laliberte.** 2012. Host endomembrane recruitment for plant RNA virus replication. *Curr. Opin. Virol.* **2**:683-690.
105. **Grangeon, R., J. Jiang, J. Wan, M. Agbeci, H. Zheng, and J. F. Laliberte.** 2013. 6K2-induced vesicles can move cell to cell during turnip mosaic virus infection. *Front Microbiol.* **4**:351.
106. **Grennan, A. K.** 2006. Regulation of starch metabolism in Arabidopsis leaves. *Plant Physiol* **142**:1343-1345.
107. **Grzela, R., L. Strokovska, J. P. Andrieu, B. Dublet, W. Zagorski, and J. Chroboczek.** 2006. Potyvirus terminal protein VPg, effector of host eukaryotic initiation factor eIF4E. *Biochimie* **88**:887-896.
108. **Gunasinghe U.B and Berger P.H.** 1991. Association of Potato Virus Y Gene Products with Chloroplasts in Tobacco. *Mol. Plant Microbe Interact.* **4**:452-457.
109. **Guo, D., M. L. Rajamaki, M. Saarma, and J. P. Valkonen.** 2001. Towards a protein interaction map of potyviruses: protein interaction matrixes of two potyviruses based on the yeast two-hybrid system. *J. Gen. Virol.* **82**:935-939.

110. **Gutierrez, S., M. Yvon, E. Pirolles, E. Garzo, A. Fereres, Y. Michalakis, and S. Blanc.** 2012. Circulating virus load determines the size of bottlenecks in viral populations progressing within a host. *PLoS. Pathog.* **8**:e1003009.
111. **Hafren, A., D. Hofius, G. Ronnholm, U. Sonnewald, and K. Makinen.** 2010. HSP70 and its cochaperone CPIP promote potyvirus infection in *Nicotiana benthamiana* by regulating viral coat protein functions. *Plant Cell* **22**:523-535.
112. **Hayashi, N., M. Matsubara, Y. Jinbo, K. Titani, Y. Izumi, and N. Matsushima.** 2002. Nef of HIV-1 interacts directly with calcium-bound calmodulin. *Protein Sci.* **11**:529-537.
113. **Heidel, A. J., J. D. Clarke, J. Antonovics, and X. Dong.** 2004. Fitness costs of mutations affecting the systemic acquired resistance pathway in *Arabidopsis thaliana*. *Genetics* **168**:2197-2206.
114. **Hofius, D., A. T. Maier, C. Dietrich, I. Jungkunz, F. Bornke, E. Maiss, and U. Sonnewald.** 2007. Capsid protein-mediated recruitment of host DnaJ-like proteins is required for Potato virus Y infection in tobacco plants. *J. Virol.* **81**:11870-11880.
115. **Holmes, E. C.** 2003. Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol.* **11**:543-546.
116. **Hong, Y. and A. G. Hunt.** 1996. RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. *Virology* **226**:146-151.
117. **Huang, T. S., T. Wei, J. F. Laliberte, and A. Wang.** 2010. A host RNA helicase-like protein, AtRH8, interacts with the potyviral genome-linked protein, VPg, associates with the virus accumulation complex, and is essential for infection. *Plant Physiol* **152**:255-266.
118. **Huibers, R. P., M. de Jong, R. W. Dekter, and A. G. Van den.** 2009. Disease-specific expression of host genes during downy mildew infection of *Arabidopsis*. *Mol. Plant Microbe Interact.* **22**:1104-1115.
119. **Ishihara, T., N. Sakurai, K. T. Sekine, S. Hase, M. Ikegami, D. Shibata, and H. Takahashi.** 2004. Comparative analysis of expressed sequence tags in resistant and susceptible ecotypes of *Arabidopsis thaliana* infected with cucumber mosaic virus. *Plant Cell Physiol* **45**:470-480.
120. **Jiang, J. and J. F. Laliberte.** 2011. The genome-linked protein VPg of plant viruses—a protein with many partners. *Curr. Opin. Virol.* **1**:347-354.
121. **Johansen, I. E., W. G. Dougherty, K. E. Keller, D. Wang, and R. O. Hampton.** 1996. Multiple viral determinants affect seed transmission of pea seedborne mosaic virus in *Pisum sativum*. *J. Gen. Virol.* **77 (Pt 12)**:3149-3154.
122. **John J. Gillespie.** 1991. *The Causes of Molecular Evolution.* Oxford University Press.
123. **Johnson, P. R. and J. R. Ecker.** 1998. The ethylene gas signal transduction pathway: a molecular perspective. *Annu. Rev. Genet.* **32**:227-254.
124. **Jungkunz, I., K. Link, F. Vogel, L. M. Voll, S. Sonnewald, and U. Sonnewald.** 2011. AtHsp70-15-deficient *Arabidopsis* plants are characterized by reduced growth, a constitutive cytosolic protein response and enhanced resistance to TuMV. *Plant J.* **66**:983-995.

125. **Kassanis, B.** 1941. Transmission of tobacco etch viruses by aphides. *Ann. Appl. Biol.* **28**:238-243.
126. **Kasschau, K. D. and J. C. Carrington.** 2001. Long-distance movement and replication maintenance functions correlate with silencing suppression activity of potyviral HC-Pro. *Virology* **285**:71-81.
127. **Kawecki T.J.** 1994. Accumulation of deleterious mutations and the evolutionary cost of being a generalist. *Am. Nat.* **144**:833-838.
128. **Kekarainen, T., H. Savilahti, and J. P. Valkonen.** 2002. Functional genomics on potato virus A: virus genome-wide map of sites essential for virus propagation. *Genome Res.* **12**:584-594.
129. **Kennedy J.S, Day M.F, and Eastop V.F.** 1962. A conspectus of aphids as vectors of plant viruses. *The Canadian Entomologist* 144-145.
130. **Khan, M. A., H. Yumak, D. R. Gallie, and D. J. Goss.** 2008. Effects of poly(A)-binding protein on the interactions of translation initiation factor eIF4F and eIF4F.4B with internal ribosome entry site (IRES) of tobacco etch virus RNA. *Biochim. Biophys. Acta* **1779**:622-627.
131. **Kimura, M.** 1983. Diffusion model of intergroup selection, with special reference to evolution of an altruistic character. *Proc. Natl. Acad. Sci. U. S. A* **80**:6317-6321.
132. **Klein, P. G., R. R. Klein, E. Rodriguez-Cerezo, A. G. Hunt, and J. G. Shaw.** 1994. Mutational analysis of the tobacco vein mottling virus genome. *Virology* **204**:759-769.
133. **Kouyos, R. D., O. K. Silander, and S. Bonhoeffer.** 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* **22**:308-315.
134. **Kwon, S. J., H. C. Jin, S. Lee, M. H. Nam, J. H. Chung, S. I. Kwon, C. M. Ryu, and O. K. Park.** 2009. GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in Arabidopsis. *Plant J.* **58**:235-245.
135. **Lafforgue, G., F. Martinez, J. Sardanyes, I. de I, Q. W. Niu, S. S. Lin, R. V. Sole, N. H. Chua, J. A. Daros, and S. F. Elena.** 2011. Tempo and mode of plant RNA virus escape from RNA interference-mediated resistance. *J. Virol.* **85**:9686-9695.
136. **Lain, S., J. L. Riechmann, and J. A. Garcia.** 1990. RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* **18**:7003-7006.
137. **Laliberte, J. F. and H. Sanfacon.** 2010. Cellular remodeling during plant virus infection. *Annu. Rev. Phytopathol.* **48**:69-91.
138. **Lalić, J., P. Agudelo-Romero, P. Carrasco, and S. F. Elena.** 2010. Adaptation of tobacco etch potyvirus to a susceptible ecotype of Arabidopsis thaliana capacitates it for systemic infection of resistant ecotypes. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **365**:1997-2007.
139. **Lalić, J. and S. F. Elena.** 2012. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity (Edinb.)* **109**:71-77.

140. **Lalić, J. and S. F. Elena.** 2013. Epistasis between mutations is host-dependent for an RNA virus. *Biol. Lett.* **9**:20120396.
141. **Lenski, R. E.** 1991. Quantifying fitness and gene stability in microorganisms. *Biotechnology* **15**:173-192.
142. **Lenski, R. E. and R. M. May.** 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. *J. Theor. Biol.* **169**:253-265.
143. **Leonard, S., D. Plante, S. Wittmann, N. Daigneault, M. G. Fortin, and J. F. Laliberte.** 2000. Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. *J. Virol.* **74**:7730-7737.
144. **Leonard, S., C. Viel, C. Beauchemin, N. Daigneault, M. G. Fortin, and J. F. Laliberte.** 2004. Interaction of VPg-Pro of turnip mosaic virus with the translation initiation factor 4E and the poly(A)-binding protein in planta. *J. Gen. Virol.* **85**:1055-1063.
145. **Lerich, A., M. Langhans, S. Sturm, and D. G. Robinson.** 2011. Is the 6 kDa tobacco etch viral protein a bona fide ERES marker? *J. Exp. Bot.* **62**:5013-5023.
146. **Li, J. L., R. S. Cornman, J. D. Evans, J. S. Pettis, Y. Zhao, C. Murphy, W. J. Peng, J. Wu, M. Hamilton, H. F. Boncristiani, Jr., L. Zhou, J. Hammond, and Y. P. Chen.** 2014. Systemic spread and propagation of a plant-pathogenic virus in European honeybees, *Apis mellifera*. *MBio.* **5**:e00898-13.
147. **Lin, L., Y. Shi, Z. Luo, Y. Lu, H. Zheng, F. Yan, J. Chen, J. Chen, M. J. Adams, and Y. Wu.** 2009. Protein-protein interactions in two potyviruses using the yeast two-hybrid system. *Virus Res.* **142**:36-40.
148. **Lively, C. M.** 2010. The effect of host genetic diversity on disease spread. *Am. Nat.* **175**:E149-E152.
149. **Llave, C., K. D. Kasschau, and J. C. Carrington.** 2000. Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. *Proc. Natl. Acad. Sci. U. S. A* **97**:13401-13406.
150. **Longdon, B., J. D. Hadfield, C. L. Webster, D. J. Obbard, and F. M. Jiggins.** 2011. Host phylogeny determines viral persistence and replication in novel hosts. *PLoS. Pathog.* **7**:e1002260.
151. **Lu, S. and B. R. Cullen.** 2004. Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. *J. Virol.* **78**:12868-12876.
152. **Luhua, S., S. Ciftci-Yilmaz, J. Harper, J. Cushman, and R. Mittler.** 2008. Enhanced tolerance to oxidative stress in transgenic Arabidopsis plants expressing proteins of unknown function. *Plant Physiol* **148**:280-292.
153. **Luna, E., V. Pastor, J. Robert, V. Flors, B. Mauch-Mani, and J. Ton.** 2011. Callose deposition: a multifaceted plant defense response. *Mol. Plant Microbe Interact.* **24**:183-193.
154. **Lurin, C., C. Andres, S. Aubourg, M. Bellaoui, F. Bitton, C. Bruyere, M. Caboche, C. Debast, J. Gualberto, B. Hoffmann, A. Lechary, M. Le Ret, M. L. Martin-Magniette, H. Mireau, N. Peeters, J. P. Renou, B. Szurek, L. Tacconnat, and I. Small.** 2004. Genome-wide analysis of Arabidopsis pentatricopeptide

- repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**:2089-2103.
155. **Mahajan, S. K., S. T. Chisholm, S. A. Whitham, and J. C. Carrington.** 1998. Identification and characterization of a locus (RTM1) that restricts long-distance movement of tobacco etch virus in *Arabidopsis thaliana*. *Plant J.* **14**:177-186.
 156. **Manrubia S.C and E. Lazaro.** 2005. *Viral evolution.* Elsevier B. V *Physics of life* **3**:65-92.
 157. **Martin, G. B., A. J. Bogdanove, and G. Sessa.** 2003. Understanding the functions of plant disease resistance proteins. *Annu. Rev. Plant Biol.* **54**:23-61.
 158. **Martinez, F. and J. A. Daros.** 2014. Tobacco etch virus protein P1 traffics to the nucleolus and associates with the host 60S ribosomal subunits during infection. *J. Virol.* **88**:10725-10737.
 159. **Martinez-Picado, J., M. P. DePasquale, N. Kartsonis, G. J. Hanna, J. Wong, D. Finzi, E. Rosenberg, H. F. Gunthard, L. Sutton, A. Savara, C. J. Petropoulos, N. Hellmann, B. D. Walker, D. D. Richman, R. Siliciano, and R. T. D'Aquila.** 2000. Antiretroviral resistance during successful therapy of HIV type 1 infection. *Proc. Natl. Acad. Sci. U. S. A* **97**:10948-10953.
 160. **Mayhew, D. E. and R. E. Ford.** 1974. Detection of ribonuclease-resistant RNA in chloroplasts of corn leaf tissue infected with maize dwarf mosaic virus. *Virology* **57**:503-509.
 161. **Merits, A., M. L. Rajamaki, P. Lindholm, P. Runeberg-Roos, T. Kekarainen, P. Puustinen, K. Makelainen, J. P. Valkonen, and M. Saarma.** 2002. Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells. *J. Gen. Virol.* **83**:1211-1221.
 162. **Michael W. Pfaffl.** 2004. Quantification strategies in real-time PCR, p. 87-112. *In* SA Bustin (ed.), *A-Z of quantitative PCR.* International University Line (IUL), La Jolla, CA, USA.
 163. **Michella, D. J.** 1985. Host life-history variation in response to parasitism. *Parasitology.*
 164. **Michon, T., Y. Estevez, J. Walter, S. German-Retana, and O. Le Gall.** 2006. The potyviral virus genome-linked protein VPg forms a ternary complex with the eukaryotic initiation factors eIF4E and eIF4G and reduces eIF4E affinity for a mRNA cap analogue. *FEBS J.* **273**:1312-1322.
 165. **Milne, R. G.** 1988. Species concept should not be universally applied to virus taxonomy--but what to do instead? *Intervirology* **29**:254-259.
 166. **Minskaia, E., T. Hertzog, 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. *Proc. Natl. Acad. Sci. U. S. A* **103**:5108-5113.
 167. **Miralles, R., A. Moya, and S. F. Elena.** 1997. Is group selection a factor modulating the virulence of RNA viruses? *Genet. Res.* **69**:165-172.
 168. **Montaner, D. and J. Dopazo.** 2010. Multidimensional gene set analysis of genomic data. *PLoS. One.* **5**:e10348.

169. **Morel, J. B. and J. L. Dangl.** 1997. The hypersensitive response and the induction of cell death in plants. *Cell Death. Differ.* **4**:671-683.
170. **Mundt, C. C.** 2002. Use of multiline cultivars and cultivar mixtures for disease management. *Annu. Rev. Phytopathol.* **40**:381-410.
171. **Murphy, J. F., P. G. Klein, A. G. Hunt, and J. G. Shaw.** 1996. Replacement of the tyrosine residue that links a potyviral VPg to the viral RNA is lethal. *Virology* **220**:535-538.
172. **Murray, K. E. and D. J. Barton.** 2003. Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J. Virol.* **77**:4739-4750.
173. **Mustonen, V. and M. Lassig.** 2009. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* **25**:111-119.
174. **Nagasaki, N., R. Tomioka, and M. Maeshima.** 2008. A hydrophilic cation-binding protein of *Arabidopsis thaliana*, AtPCaP1, is localized to plasma membrane via N-myristoylation and interacts with calmodulin and the phosphatidylinositol phosphates PtdIns(3,4,5)P(3) and PtdIns(3,5)P(2). *FEBS J.* **275**:2267-2282.
175. **Nakamura, K. and K. Matsuoka.** 1993. Protein targeting to the vacuole in plant cells. *Plant Physiol* **101**:1-5.
176. **Navarro, L., P. Dunoyer, F. Jay, B. Arnold, N. Dharmasiri, M. Estelle, O. Voinnet, and J. D. Jones.** 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* **312**:436-439.
177. **Nei, M.** 1982. Evolution of human races at the gene level. *Prog. Clin. Biol. Res.* **103 Pt A**:167-181.
178. **Nijhuis, M., R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schipper, S. Gulnik, and C. A. Boucher.** 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. *AIDS* **13**:2349-2359.
179. **Novella, I. S., S. Zarate, D. Metzgar, and Ebendick-Corp.** 2004. Positive selection of synonymous mutations in vesicular stomatitis virus. *J. Mol. Biol.* **342**:1415-1421.
180. **Ohta, T. and Y. Ina.** 1995. Variation in synonymous substitution rates among mammalian genes and the correlation between synonymous and nonsynonymous divergences. *J. Mol. Evol.* **41**:717-720.
181. Oliveros, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams. 2007.
- Ref Type: Computer Program
182. **Olivier Voinnet.** 2005. Induction and suppression of RNA silencing: insights from viral infections. *Nature Reviews Genetics* **6**:206-220.
183. **Oruetxebarria, I., D. Guo, A. Merits, K. Makinen, M. Saarma, and J. P. Valkonen.** 2001. Identification of the genome-linked protein in virions of Potato virus A, with comparison to other members in genus Potyvirus. *Virus Res.* **73**:103-112.

184. **Ostrowski, E. A., D. E. Rozen, and R. E. Lenski.** 2005. Pleiotropic effects of beneficial mutations in *Escherichia coli*. *Evolution* **59**:2343-2352.
185. **Padmanabhan, M. S., S. P. Goregaoker, S. Golem, H. Shiferaw, and J. N. Culver.** 2005. Interaction of the tobacco mosaic virus replicase protein with the Aux/IAA protein PAP1/IAA26 is associated with disease development. *J. Virol.* **79**:2549-2558.
186. **Pagan, I., C. Alonso-Blanco, and F. Garcia-Arenal.** 2008. Host responses in life-history traits and tolerance to virus infection in *Arabidopsis thaliana*. *PLoS. Pathog.* **4**:e1000124.
187. **Parrish, C. R., E. C. Holmes, D. M. Morens, E. C. Park, D. S. Burke, C. H. Calisher, C. A. Laughlin, L. J. Saif, and P. Daszak.** 2008. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol. Mol. Biol. Rev.* **72**:457-470.
188. **Perrin, N. and P. Christe.** 1996. On host life-history response to parasitism. *Oikos.*
189. **Pita, J. S., J. R. de Miranda, W. L. Schneider, and M. J. Roossinck.** 2007. Environment determines fidelity for an RNA virus replicase. *J. Virol.* **81**:9072-9077.
190. **Pita, J. S. and M. J. Roossinck.** 2013. Mapping viral functional domains for genetic diversity in plants. *J. Virol.* **87**:790-797.
191. **Plotkin, J. B. and G. Kudla.** 2011. Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* **12**:32-42.
192. **Poelwijk, F. J., S. Tanase-Nicola, D. J. Kiviet, and S. J. Tans.** 2011. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.* **272**:141-144.
193. **Puustinen, P. and K. Makinen.** 2004. Uridylylation of the potyvirus VPg by viral replicase NIb correlates with the nucleotide binding capacity of VPg. *J. Biol. Chem.* **279**:38103-38110.
194. **Puustinen, P., M. L. Rajamaki, K. I. Ivanov, J. P. Valkonen, and K. Makinen.** 2002. Detection of the potyviral genome-linked protein VPg in virions and its phosphorylation by host kinases. *J. Virol.* **76**:12703-12711.
195. **Ramachandran, V. and X. Chen.** 2008. Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science* **321**:1490-1492.
196. **Rancurel, C., M. Khosravi, A. K. Dunker, P. R. Romero, and D. Karlin.** 2009. Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. *J. Virol.* **83**:10719-10736.
197. **Redondo, E., R. Krause-Sakate, S. J. Yang, H. Lot, O. Le Gall, and T. Candresse.** 2001. Lettuce mosaic virus pathogenicity determinants in susceptible and tolerant lettuce cultivars map to different regions of the viral genome. *Mol. Plant Microbe Interact.* **14**:804-810.
198. **Remold, S.** 2012. Understanding specialism when the Jack of all trades can be the master of all. *Proc. Biol. Sci.* **279**:4861-4869.

199. **Remold, S. K., A. Rambaut, and P. E. Turner.** 2008. Evolutionary genomics of host adaptation in vesicular stomatitis virus. *Mol. Biol. Evol.* **25**:1138-1147.
200. **Restrepo-Hartwig, M. A. and J. C. Carrington.** 1994. The tobacco etch potyvirus 6-kilodalton protein is membrane associated and involved in viral replication. *J. Virol.* **68**:2388-2397.
201. **Rico, P., P. Ivars, S. F. Elena, and C. Hernandez.** 2006. Insights into the selective pressures restricting Pelargonium flower break virus genome variability: Evidence for host adaptation. *J. Virol.* **80**:8124-8132.
202. **Robaglia, C. and C. Caranta.** 2006. Translation initiation factors: a weak link in plant RNA virus infection. *Trends Plant Sci.* **11**:40-45.
203. **Rocher Hull.** 2009. Matthews' Plant Virology.
204. **Rodrigo, G., J. Carrera, V. Ruiz-Ferrer, F. J. del Toro, C. Llave, O. Voinnet, and S. F. Elena.** 2012. A meta-analysis reveals the commonalities and differences in *Arabidopsis thaliana* response to different viral pathogens. *PLoS. One.* **7**:e40526.
205. **Rodríguez-Gironés M.A and Santamaría L.** 2006. A new algorithm to calculate the nestedness temperature of presence-absence matrices. *Biogeography* **33**:924-935.
206. **Rojas, M. R., H. Jiang, R. Salati, B. Xoconostle-Cazares, M. R. Sudarshana, W. J. Lucas, and R. L. Gilbertson.** 2001. Functional analysis of proteins involved in movement of the monopartite begomovirus, Tomato yellow leaf curl virus. *Virology* **291**:110-125.
207. **Rojas, M. R., F. M. Zerbini, R. F. Allison, R. L. Gilbertson, and W. J. Lucas.** 1997. Capsid protein and helper component-proteinase function as potyvirus cell-to-cell movement proteins. *Virology* **237**:283-295.
208. **Roossinck, M. J.** 2008. *Plant Virus Evolution.* Springer Berlin, Germany.
209. **Rowland, O. and J. D. Jones.** 2001. Unraveling regulatory networks in plant defense using microarrays. *Genome Biol.* **2**:REVIEWS1001.
210. **Saenz, P., B. Salvador, C. Simon-Mateo, K. D. Kasschau, J. C. Carrington, and J. A. Garcia.** 2002. Host-specific involvement of the HC protein in the long-distance movement of potyviruses. *J. Virol.* **76**:1922-1931.
211. **Sanità di Toppi L., G. R.** 1999. Response to cadmium in higher plants. *Environmental and Experimental Botany* **Volume 41**:105-130.
212. **Sanjuan, R., P. Agudelo-Romero, and S. F. Elena.** 2009. Upper-limit mutation rate estimation for a plant RNA virus. *Biol. Lett.* **5**:394-396.
213. **Sanjuan, R., J. M. Cuevas, V. Furio, E. C. Holmes, and A. Moya.** 2007. Selection for robustness in mutagenized RNA viruses. *PLoS. Genet.* **3**:e93.
214. **Sanjuan, R., A. Moya, and S. F. Elena.** 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. U. S. A* **101**:8396-8401.
215. **Sanjuan, R., M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw.** 2010. Viral mutation rates. *J. Virol.* **84**:9733-9748.

216. **Sartor, M. A., G. D. Leikauf, and M. Medvedovic.** 2009. LRpath: a logistic regression approach for identifying enriched biological groups in gene expression data. *Bioinformatics*. **25**:211-217.
217. **Sasaki, A. and M. A. Nowak.** 2003. Mutation landscapes. *J. Theor. Biol.* **224**:241-247.
218. **Schaad, M. C., P. E. Jensen, and J. C. Carrington.** 1997. Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *EMBO J.* **16**:4049-4059.
219. **Schaad, M. C., A. D. Lellis, and J. C. Carrington.** 1997. VPg of tobacco etch potyvirus is a host genotype-specific determinant for long-distance movement. *J. Virol.* **71**:8624-8631.
220. **Shah, J.** 2005. Lipids, lipases, and lipid-modifying enzymes in plant disease resistance. *Annu. Rev. Phytopathol.* **43**:229-260.
221. **Shcherbakov, D. V. and M. B. Garber.** 2000. [Overlapping genes in bacterial and bacteriophage genomes]. *Mol. Biol. (Mosk)* **34**:572-583.
222. **Shukla, D. D., C.W.Ward, and A.A.Brunst.** 1994. *The Potyviridae*. Wallingford: CAB International.
223. **Shukla, D. D. and C. W. Ward.** 1989. Identification and classification of potyviruses on the basis of coat protein sequence data and serology. Brief review. *Arch. Virol.* **106**:171-200.
224. **Siaw, M. F., M. Shahabuddin, S. Ballard, J. G. Shaw, and R. E. Rhoads.** 1985. Identification of a protein covalently linked to the 5' terminus of tobacco vein mottling virus RNA. *Virology* **142**:134-143.
225. **Simmonds, P.** 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. *J. Gen. Virol.* **85**:3173-3188.
226. **Simmonds, P. and D. B. Smith.** 1999. Structural constraints on RNA virus evolution. *J. Virol.* **73**:5787-5794.
227. **Smyth, D. R.** 1999. Gene silencing: plants and viruses fight it out. *Curr. Biol.* **9**:R100-R102.
228. **Smyth, G. K.** 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**:Article3.
229. **Sonia Laín, JoséLuis Riechmann, Enrique Méndez, and Juan Antonio García.** 1988. Nucleotide sequence of the 3' terminal region of plum pox potyvirus RNA. *Virus Research* **10**:325-341.
230. **Springbett, A. J., K. MacKenzie, J. A. Woolliams, and S. C. Bishop.** 2003. The contribution of genetic diversity to the spread of infectious diseases in livestock populations. *Genetics* **165**:1465-1474.
231. **Stearns, S. C.** 1976. Life-history tactics: a review of the ideas. *Q. Rev. Biol.* **51**:3-47.
232. **Supek, F., M. Bosnjak, N. Skunca, and T. Smuc.** 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS. One.* **6**:e21800.

233. **Szittyta, G., D. Silhavy, A. Molnar, Z. Havelda, A. Lovas, L. Lakatos, Z. Banfalvi, and J. Burgyan.** 2003. Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. *EMBO J.* **22**:633-640.
234. **Tajima, F.** 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585-595.
235. **Takenaka, M., D. Verbitskiy, A. Zehrmann, and A. Brennicke.** 2010. Reverse genetic screening identifies five E-class PPR proteins involved in RNA editing in mitochondria of *Arabidopsis thaliana*. *J. Biol. Chem.* **285**:27122-27129.
236. **Tavert-Roudet, G., A. Abdul-Razzak, B. Doublet, J. Walter, T. Delaunay, S. German-Retana, T. Michon, O. Le Gall, and T. Candresse.** 2012. The C terminus of lettuce mosaic potyvirus cylindrical inclusion helicase interacts with the viral VPg and with lettuce translation eukaryotic initiation factor 4E. *J. Gen. Virol.* **93**:184-193.
237. **Tenllado, F. and J. R. Diaz-Ruiz.** 2001. Double-stranded RNA-mediated interference with plant virus infection. *J. Virol.* **75**:12288-12297.
238. **The Arabidopsis Genome Initiative.** 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**:796-815.
239. **Thivierge, K., S. Cotton, P. J. Dufresne, I. Mathieu, C. Beauchemin, C. Ide, M. G. Fortin, and J. F. Laliberte.** 2008. Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. *Virology* **377**:216-225.
240. **Thomas, C. L., D. Schmidt, E. M. Bayer, R. Dreos, and A. J. Maule.** 2009. *Arabidopsis* plant homeodomain finger proteins operate downstream of auxin accumulation in specifying the vasculature and primary root meristem. *Plant J.* **59**:426-436.
241. **Tromas, N. and S. F. Elena.** 2010. The rate and spectrum of spontaneous mutations in a plant RNA virus. *Genetics* **185**:983-989.
242. **Turner, P. E. and S. F. Elena.** 2000. Cost of host radiation in an RNA virus. *Genetics* **156**:1465-1470.
243. **Van der Biezen, E. A. and J. D. Jones.** 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* **23**:454-456.
244. **Vasi, F., M. Travisano, and R. E. Lenski.** 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am. Nat.* **144**:432-456.
245. **Verchot, J. and J. C. Carrington.** 1995. Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. *J. Virol.* **69**:1582-1590.
246. **Vijayapalani, P., M. Maeshima, N. Nagasaki-Takekuchi, and W. A. Miller.** 2012. Interaction of the trans-frame potyvirus protein P3N-PIPO with host protein PCaP1 facilitates potyvirus movement. *PLoS. Pathog.* **8**:e1002639.
247. **Walkey, D. G. A.** 1991. *Applied Plant Virology*. Chapman & Hall.

248. **Wallis, C. M., A. L. Stone, D. J. Sherman, V. D. Damsteegt, F. E. Gildow, and W. L. Schneider.** 2007. Adaptation of plum pox virus to a herbaceous host (*Pisum sativum*) following serial passages. *J. Gen. Virol.* **88**:2839-2845.
249. **Wang, K. L., H. Li, and J. R. Ecker.** 2002. Ethylene biosynthesis and signaling networks. *Plant Cell* **14 Suppl**:S131-S151.
250. **Ward, C. W. and D. D. Shukla.** 1991. Taxonomy of potyviruses: current problems and some solutions. *Intervirology* **32**:269-296.
251. **Wei, T., T. S. Huang, J. McNeil, J. F. Laliberte, J. Hong, R. S. Nelson, and A. Wang.** 2010. Sequential recruitment of the endoplasmic reticulum and chloroplasts for plant potyvirus replication. *J. Virol.* **84**:799-809.
252. **Wei, T. and A. Wang.** 2008. Biogenesis of cytoplasmic membranous vesicles for plant potyvirus replication occurs at endoplasmic reticulum exit sites in a. *J. Virol.* **82**:12252-12264.
253. **Wei, T., C. Zhang, X. Hou, H. Sanfacon, and A. Wang.** 2013. The SNARE protein Syp71 is essential for turnip mosaic virus infection by mediating fusion of virus-induced vesicles with chloroplasts. *PLoS. Pathog.* **9**:e1003378.
254. **Weitz, J. S., T. Poisot, J. R. Meyer, C. O. Flores, S. Valverde, M. B. Sullivan, and M. E. Hochberg.** 2013. Phage-bacteria infection networks. *Trends Microbiol.* **21**:82-91.
255. **Whitham, S. A., R. J. Anderberg, S. T. Chisholm, and J. C. Carrington.** 2000. Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. *Plant Cell* **12**:569-582.
256. **Whitham, S. A., S. Quan, H. S. Chang, B. Cooper, B. Estes, T. Zhu, X. Wang, and Y. M. Hou.** 2003. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana* plants. *Plant J.* **33**:271-283.
257. **Whitham, S. A. and Y. Wang.** 2004. Roles for host factors in plant viral pathogenicity. *Curr. Opin. Plant Biol.* **7**:365-371.
258. **Whitham, S. A., M. L. Yamamoto, and J. C. Carrington.** 1999. Selectable viruses and altered susceptibility mutants in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A* **96**:772-777.
259. **Whitham, T. G., J. K. Bailey, J. A. Schweitzer, S. M. Shuster, R. K. Bangert, C. J. LeRoy, E. V. Lonsdorf, G. J. Allan, S. P. DiFazio, B. M. Potts, D. G. Fischer, C. A. Gehring, R. L. Lindroth, J. C. Marks, S. C. Hart, G. M. Wimp, and S. C. Wooley.** 2006. A framework for community and ecosystem genetics: from genes to ecosystems. *Nat Rev Genet.* **7**:510-523.
260. **Whitlock, M. C.** 1996. The Red Queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. *Am. Nat* **148**:65-77.
261. **Wichman, H. A., M. R. Badgett, L. A. Scott, C. M. Boulianne, and J. J. Bull.** 1999. Different trajectories of parallel evolution during viral adaptation. *Science* **285**:422-424.
262. **Wise, R. P., M. J. Moscou, A. J. Bogdanove, and S. A. Whitham.** 2007. Transcript profiling in host-pathogen interactions. *Annu Rev Phytopathol.* **45**:329-369.

263. **Witwer, C., S. Rauscher, I. L. Hofacker, and P. F. Stadler.** 2001. Conserved RNA secondary structures in Picornaviridae genomes. *Nucleic Acids Res.* **29**:5079-5089.
264. **Woolhouse, M. E., D. T. Haydon, and R. Antia.** 2005. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol. Evol.* **20**:238-244.
265. **Woolhouse, M. E., L. H. Taylor, and D. T. Haydon.** 2001. Population biology of multihost pathogens. *Science* **292**:1109-1112.
266. **Yang, C., R. Guo, F. Jie, D. Nettleton, J. Peng, T. Carr, J. M. Yeakley, J. B. Fan, and S. A. Whitham.** 2007. Spatial analysis of *Arabidopsis thaliana* gene expression in response to Turnip mosaic virus infection. *Mol. Plant Microbe Interact.* **20**:358-370.
267. **Yarwood, C. E.** 1970. Reversible host adaptation in Cucumber mosaic virus. *Phytopathology* **60**:1117-1119.
268. **Zacharias, D. A., J. D. Violin, A. C. Newton, and R. Y. Tsien.** 2002. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**:913-916.
269. **Zilian, E. and E. Maiss.** 2011. Detection of plum pox potyviral protein-protein interactions in planta using an optimized mRFP-based bimolecular fluorescence complementation system. *J. Gen. Virol.* **92**:2711-2723.