



Análisis de relaciones estructura-función en el virus del arabesco del *Pelargonium*

Memoria presentada por
AURORA CASTAÑO SANSANO

para optar al grado de
DOCTORA INGENIERA AGRÓNOMA

Director:
Dr. CARMEN HERNÁNDEZ FORT

Tutor:
Dr. ISMAEL RODRIGO BRAVO

VALENCIA, 2009

Doña Carmen Hernández Fort, Doctora en Ciencias Biológicas, Científico Titular del Consejo Superior de Investigaciones Científicas en el Instituto de Biología Molecular y Celular de Plantas (Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas) de Valencia,

CERTIFICA:

que Doña Aurora Castaño Sansano ha realizado bajo su dirección el trabajo que con el título “Análisis de relaciones estructura-función en el virus del arabesco del *Pelargonium*”, presenta para optar al grado de Doctora Ingeniera Agrónoma.

Para que así conste a los efectos oportunos, firma el presente certificado en Valencia a 15 de Abril de 2009.

Firmado: Carmen Hernández Fort

A mi iaia

A mis padres

A mi hermana

A Ferran

Abreviaturas

AMV: *Alfalfa mosaic virus* (virus del mosaico de la alfalfa)

AnFBV: *Angelonia flower break virus* (virus de la rotura del color de la flor de la angelonia)

BMMV: *Bean mild mosaic virus* (virus del mosaico suave de la judía)

BMoV: *Blackgram mottle virus* (virus del moteado del *Blackgram*)

BMV: *Brome mosaic virus* (virus del mosaico del bromo)

BNYVV: *Beet necrotic yellow vein virus* (virus de las venas necróticas y amarillas de la remolacha)

BPYV: *Beet pseudo-yellows virus* (virus del pseudo-amarilleamiento de la remolacha)

BSMV: *Barley stripe mosaic virus* (virus del mosaico estriado de la cebada)

BTE: barley translational enhancer (estimulador de la traducción de la cebada)

BWYV: *Beet western yellows virus* (virus del amarilleamiento de la remolacha occidental)

BYDV: *Barley yellow dwarf virus* (virus del enanismo amarillo de la cebada)

CaMV: *Cauliflower mosaic virus* (virus del mosaico de la coliflor)

CarMV: *Carnation mottle virus* (virus del moteado del clavel)

CCFV: *Cardamine chlorotic fleck virus* (virus del salpicado clorótico del cardamomo)

cDNA: copy DNA (DNA copia)

CIRV: *Carnation italian ringspot virus* (virus de las manchas en anillo del clavel italiano)

CITE: cap independent translation enhancer (estimulador de la traducción independiente de cap)

CLCV: *Cotton leaf curl virus* (virus del rizado de la hoja del algodón)

CLSV: *Cucumber leaf spot virus* (virus de las manchas de las hojas del pepino)

CMV: *Cucumber mosaic virus* (virus del mosaico del pepino)

CNV: *Cucumber necrosis virus* (virus de la necrosis del pepino)

CPMoV: *Cowpea mottle virus* (virus del moteado del chícharo)

CPMV: *Cowpea mosaic virus* (virus del mosaico del chícharo)

CP: coat protein (proteína de cubierta o proteína de la cápsida)

crTMV: *crucifer-infecting tobamovirus* (tobamovirus que infecta crucíferas)

CTV: *Citrus tristeza virus* (virus de la tristeza de los cítricos)

CuSBV: *Cucumber soil borne virus* (virus del pepino transmitido por el suelo)

CYLCV: *Cestrum yellow leaf curl virus* (virus de la hoja amarilla y rizada del *Cestrum*)

CyRSV: *Cymbidium ringspot virus* (virus de las manchas anulares del *Cymbidium*)

CYSDV: *Cucurbit yellow stunning disorder virus* (virus del enanismo amarillento de las cucurbitáceas)

DdDp: DNA dependent-DNA polimerase (DNA polimerasa DNA dependiente)

DdRp: DNA dependent-RNA polimerase (RNA polimerasa DNA dependiente)

DGB: double gene block (bloque de dos genes)

DI RNA: defective interfering RNA (RNA defectivo interferente)

dsRNA: double stranded RNA (RNA de doble cadena)

dsDNA: double stranded DNA (DNA de doble cadena)

ELV: *Elderberry latent virus* (virus latente de la baya del saúco)

FMV: *Figwort mosaic virus* (virus del mosaico del Figwort)

FS: frameshift (desplazamiento de la pauta de lectura)

GaMV: *Galisonga mosaic virus* (virus del mosaico del *Galisonga*)

GFP: green fluorescent protein (proteína de fluorescencia verde)

gRNA: genomic RNA (RNA genómico)

HCRSV: *Hibiscus chlorotic ringspot virus* (virus de las manchas cloróticas en anillo del hibisco)

ICTV: International Committee on Taxonomy of Viruses (Comité Internacional de Taxonomía de Virus)

INSV: *Impatiens necrotic spot virus* (virus de las manchas necróticas del *Impatiens*)

IRES: internal ribosome entry site (sitio de entrada interna de ribosomas)

JINRV: *Japanese iris necrotic ringspot virus* (virus de los anillos necróticos del iris japonés)

MCMV: *Maize chlorotic mottle virus* (virus del moteado clorótico del maíz)

MNSV: *Melon necrotic spot virus* (virus de las manchas necróticas del melón)

MP: movement protein (proteína de movimiento)

mRNA: messenger RNA (RNA mensajero)

NLVCV: Nootka lupine vein-clearing virus (virus del aclaramiento de las venas del lupino Nootka)

ORF: open reading frame (pauta abierta de lectura)

PABP: polyA-binding protein (proteína de unión a poliA)

PARS: polypurine A-rich sequence (secuencia rica en polipurinas A)

PCRPV: *Pelargonium chlorotic ring pattern virus* (virus del anillo clorótico del *Pelargonium*)

PCSV: *Peanut chlorotic streak virus* (virus del rayado clorótico del cacahuete)

PCV: *Peanut clump virus* (virus del macizo del cacahuete)

PeIRSV: *Pelargonium ringspot virus* (virus de la mancha clorótica del *Pelargonium*)

PFBV: *Pelargonium flower break virus* (virus de la rotura del color de la flor del *Pelargonium*)

pgRNA: pregenomic RNA (RNA pregenómico)

PLCV: *Pelargonium leaf curl virus* (virus del rizado de la hoja del *Pelargonium*)

PLPV: *Pelargonium line pattern virus* (virus del arabesco del *Pelargonium*)

PLRV: *Potato leafroll virus* (virus de la hoja enrollada de la patata)

PMMV: *Pepper mild mottle virus* (virus del moteado atenuado del pepino)

PMTV: *Potato mop-top virus* (virus mop-top de la patata)

PMV: *Panicum mosaic virus* (virus del mosaico del *Panicum*)

PoLV: *Photos latent virus* (virus latente del *Photos*)

PPV: *Plum pox virus* (virus de la sharka)

PSNV: *Pea stem necrosis virus* (virus de la necrosis del tallo del guisante)

PTGS: post-transcriptional gene silencing (silenciamiento génico postranscripcional)

PVM: *Potato virus M* (virus M de la patata)

PVX: *Potato virus X* (virus X de la patata)

PVY: *Potato virus Y* (virus Y de la patata)

PZSV: *Pelargonium zonale spot virus* (virus del punteado del geranio zonal)

RCNMV: *Red clover necrotic mosaic virus* (virus del mosaico necrótico del trébol rojo)

RdDp: RNA dependent-DNA polimerase (DNA polimerasa RNA dependiente)

RdRp: RNA dependent-RNA polimerase (RNA polimerasa RNA dependiente)

RT: readthrough (lectura a través del codón de parada)

RTBV: *Rice tungro bacilliform virus* (virus baciliforme del tungro del arroz)

RTSV: *Rice tungro spherical virus* (virus esférico del tungro del arroz)

RYMV: *Rice yellow mottle virus* (virus del moteado amarillo del arroz)

satRNA: satellite RNA (RNA satélite)

SBMV: *Southern bean mosaic virus* (virus del mosaico del frijol del sur)

SCV: *Saguaro cactus virus* (virus del cactus Saguaro)

sgRNA: subgenomic RNA (RNA subgenómico)

sORF: small ORF (pequeña ORF)

SPCSV: *Sweet potato chlorotic stunt virus* (virus del enanismo clorótico del boniato)

ssRNA: single stranded RNA (RNA de simple cadena)

ssDNA: single stranded DNA (DNA de simple cadena)

SPMV: *satellite Panicum mosaic virus* (virus satélite del mosaico del *Panicum*)

TAV: translational transactivator protein (proteína de transactivación traduccional)

TBSV: *Tomato bushy stunt virus* (virus del enanismo arbustivo del tomate)

TCV: *Turnip crinkle virus* (virus del arrugamiento del nabo)

TE: translational enhancer (estimulador de la traducción)

TEV: *Tobacco etch virus* (virus del grabado del tabaco)

TGB: triple gene block (bloque de tres genes)

TLS: tRNA-like structure (estructura tipo tRNA)

TMGMV: *Tobacco mild green mosaic virus* (virus del mosaico verde atenuado del tabaco)

TMV: *Tobacco mosaic virus* (virus del mosaico del tabaco)

TNV: *Tobacco necrosis virus* (virus de la necrosis del tabaco)

ToRSV: *Tomato ringspot virus* (virus de las manchas anulares del tomate)

TRSV: *Tobacco ringspot virus* (virus de las manchas anulares del tabaco)

TSWV: *Tomato spotted wilt virus* (virus del bronceado del tomate)

TuMV: *Turnip mosaic virus* (virus del mosaico del nabo)

TYMV: *Turnip yellow mosaic virus* (virus del mosaico amarillo del nabo)

UPD: upstream pseudoknot domain (dominio pseudonudo aguas arriba)

UTR: untranslated region (región no traducible)

VPg: viral protein genome-linked (proteína viral unida al genoma)

Resumen

RESUMEN

ANÁLISIS DE LAS RELACIONES ESTRUCTURA-FUNCIÓN EN EL VIRUS DEL ARABESCO DEL PELARGONIUM

Una prospección realizada recientemente en España ha puesto de manifiesto que el virus del arabesco del Pelargonium (*Pelargonium line pattern virus*, PLPV) es el agente de tipo viral más frecuente en geranio (*Pelargonium* spp.) con porcentajes de incidencia que oscilan entre el 40 y el 90% dependiendo del área geográfica examinada. Una situación similar es previsible en países de nuestro entorno y, probablemente, en otros más alejados. Se trata de un virus con partículas isométricas y un genoma monopartido de RNA de simple cadena y polaridad positiva cuyas infecciones naturales parecen restringidas a especies del género *Pelargonium*, aunque experimentalmente puede infectar especies muy diversas. Dada la escasez de datos sobre el PLPV y la importancia que el virus está adquiriendo como patógeno, en esta tesis hemos pretendido profundizar en sus características biológicas y moleculares.

El primer objetivo abordado en este trabajo ha sido determinar la secuencia nucleotídica completa de su RNA genómico (gRNA). Esta molécula está constituida por 3883 nt y, mediante análisis *in silico*, inicialmente identificamos seis marcos abiertos de lectura (ORFs) que potencialmente codificaban proteínas de 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), y 37 kDa (p37). Estas ORFs están flanqueadas por una región no traducible (UTR) en 5' inusualmente corta, con sólo 6 nt, y por una 3' UTR de 246 nt. Tanto la organización de las ORFs en el RNA viral como la mayoría de sus productos potenciales se asemejaban mucho a los implicados en replicación (p27 y p87), movimiento (p7) y encapsidación (p37) de miembros del género *Carmovirus* (familia *Tombusviridae*). A pesar de las semejanzas que compartía el PLPV con miembros de dicho género, los carmovirus producen dos RNAs subgenómicos (sgRNAs) para la síntesis de las proteínas de movimiento y de cubierta, mientras que un análisis Northern indicó que el PLPV genera un solo sgRNA de aproximadamente 1.7 kb presumiblemente implicado en la traducción de las proteínas p7, p6 y p37. Además, la región central del genoma de los carmovirus contiene dos pequeñas ORFs normalmente en distintas fases de lectura y con codones de iniciación canónicos que codifican dos proteínas de movimiento. Sin embargo, solo la p7 potencialmente codificada por el PLPV mostraba homología significativa con

proteínas de movimiento mientras que no se encontraron homólogos claros para la p6 (cabe reseñar que tampoco para la p13). Además esta p6 debía ser expresada a partir de una ORF con un codón de iniciación no canónico y/o, alternativamente, mediante un mecanismo de “-1 frameshift” (FS) o de corrimiento de pauta de lectura que se produciría inmediatamente aguas arriba del codón de parada de la ORF (p7), generándose una proteína de 12 kDa (p7-FS ó p12). Esta posibilidad estaba apoyada por la presencia, inmediatamente aguas arriba de dicho codón de parada, de un heptanucleótido cuya secuencia se ajusta al motivo característico de este tipo de mecanismo, y por la posible formación de una estructura de tipo horquilla aguas abajo del mencionado codón. Es destacable que tanto la organización de ORFs como las estrategias de expresión génica predichas para el sgRNA del PLPV eran muy similares a las descritas para el único sgRNA del virus del mosaico del Panicum (PMV), que constituye el género *Panicovirus* dentro de la familia *Tombusviridae*.

Además del PLPV, se habían descrito un conjunto de pequeños virus isométricos cuyas infecciones naturales parecían también restringidas a especies del género *Pelargonium*, entre los que se encontraban el virus de las manchas anulares de *Pelargonium* (*Pelargonium ringspot virus*, PelRSV) y el virus del anillo clorótico de *Pelargonium* (*Pelargonium chlorotic ring pattern virus*, PCRPV). Curiosamente el PCRPV codifica potencialmente cinco proteínas muy similares a las implicadas en replicación (p27 y p87), movimiento (p7 y p9) y encapsidación (p37) del género *Carmovirus*, pero produce un único sgRNA presumiblemente implicado en la traducción de las ORFs alejadas del extremo 5' del gRNA y podría utilizar un mecanismo -1 FS para la traducción de una ORF interna de pequeño tamaño, al igual que había sido predicho para el PLPV. Algunos datos indican que estas características podrían ser compartidas por el PelRSV y por el virus latente de la baya del saúco (*Elderberry latent virus*, ELV) y, sobre la base de estas características comunes y de similitudes de secuencia, se propuso incluir a los cuatro agentes virales en un nuevo género dentro de la familia *Tombusviridae*, designado provisionalmente como *Pelarspovirus*.

Las peculiaridades genómicas del PLPV (y su posible adscripción a un nuevo género) merecían sin duda ser estudiadas en mayor profundidad, por ello el siguiente objetivo fue la generación de un clon infeccioso del virus para llevar a cabo un análisis detallado de las relaciones estructura-función en el patógeno. Con esta finalidad, fusionamos un cDNA viral de longitud completa al promotor de la RNA polimerasa del fago T7 y lo ligamos a un vector de clonación de alto número de copias. Los

transcritos generados *in vitro* a partir de la construcción resultante se inocularon mecánicamente en distintos huéspedes experimentales. Los resultados mostraron que los RNAs del PLPV sintetizados *in vitro* daban lugar a infecciones indistinguibles de aquellas establecidas por el aislado viral de partida en plantas de *C. quinoa*, *N. clevelandii* y *N. benthamiana*. Disponer de esta herramienta nos permitió desarrollar estrategias para obtener información acerca del papel de las distintas ORFs en el ciclo biológico del virus y de los mecanismos que median su expresión.

Mediante experimentos de traducción *in vitro* y bioensayos de distintos mutantes identificamos las ORFs biológicamente activas del PLPV y obtuvimos datos acerca de la(s) etapa(s) del ciclo infeccioso en la(s) que están involucradas. Las funciones predichas inicialmente para las ORFs (p27), (p87), presuntamente implicadas en replicación, y para las ORFs (p7) y (p37), en movimiento y encapsidación, respectivamente, eran consistentes con los resultados de este trabajo, mientras que las ORFs (p13) y (p6), inicialmente identificadas *in silico*, no eran traducidas *in vitro* ni requeridas *in vivo*, de modo que probablemente carecen de significado biológico. En su lugar se identificó una nueva ORF, localizada en la región central del genoma, que se traducía a partir de un codón débil de iniciación (GUG) y que daba lugar a una proteína de tamaño 9.7 kDa implicada en el movimiento del virus. Esta proteína, además, mostraba una notable identidad de secuencia (44,4%) con otra de tamaño 9.4 kDa, potencialmente codificada por una ORF del PCRPV también ubicada en la región central del genoma y cuyo codón de inicio era, asimismo, un triplete no-AUG, lo que ampliaba las semejanzas entre estos dos virus, ambos miembros provisionales del género propuesto *Pelarspovirus*.

Una vez determinada la función de las proteínas codificadas por el PLPV se inició un estudio sobre las estrategias que utiliza este patógeno para la expresión de todos sus genes. Ensayos de traducción *in vitro* confirmaron que: i) el gRNA dirige la síntesis de las proteínas p27 y p87 de manera eficiente (la última mediante un mecanismo de lectura a través del codón de parada débil del gen p27), pese a carecer de una estructura cap en su extremo 5', como ocurre en el resto de miembros de la familia *Tombusviridae*, y presentar una secuencia líder extremadamente corta (6 nt) y, ii) el sgRNA actúa como mRNA para la producción de las tres proteínas restantes, p7, p9.7 y p37, mediante escapes al proceso de rastreo o *leaky-scanning*. Estos procesos parecen estar favorecidos por el contexto subóptimo del codón de inicio del gen p7, por el codón débil de inicio (GUG) del gen

p9.7, y por la ausencia de codones AUG adicionales en cualquier pauta de lectura en la región que precede al codón de inicio del gen p37 en el sgRNA.

Por último se llevó a cabo un análisis de la variabilidad genética del PLPV para obtener información acerca de las presiones selectivas que operan sobre el genoma del virus. Para ello, se partió de una colección de aislados provenientes de plantas de *P. zonale* o *P. peltatum* con distintos orígenes geográficos y/o con diferentes fechas de recolección. La región genómica examinada, de 1817 nt, comprendía una región parcial del gen de la RdRp, los genes de las proteínas de movimiento (p7 y p9.7) y de la CP, y la 3' UTR. El análisis de las secuencias obtenidas ha permitido determinar algunos de los factores que actúan limitando la heterogeneidad viral como son la conservación de las secuencias aminoacídicas de determinados dominios de las proteínas, y la preservación del plegamiento de regiones reguladoras del genoma viral. El estudio de covariaciones dentro y entre proteínas también ha permitido identificar una red potencial de interacciones específicas probablemente requerida para mantener la conformación correcta de las proteínas virales y para conducir al RNA viral desde el proceso de replicación hasta su disseminación en la planta. Adicionalmente, mediante pases seriados de un aislado del PLPV en el huésped experimental *C. quinoa*, se ha demostrado que este patógeno es capaz de evolucionar de forma muy rápida y reproducible cuando es transferido a una nueva especie vegetal. Esta observación pone de manifiesto el papel crítico que desempeñan las presiones selectivas ejercidas por el huésped en la estructura genética de las poblaciones del virus y sugiere que la tasa de mutación del PLPV es inusualmente elevada.

RESUM

ANÀLISI DE LES RELACIONS ESTRUCTURA-FUNCIÓ EN EL VIRUS DE L'ARABESC DEL PELARGONIUM

Una prospecció realitzada recentment a Espanya ha posat de manifest que el virus de l'arabesc del Pelargonium (*Pelargonium line pattern virus*, PLPV) és l'agent de tipus viral més freqüent en gerani (*Pelargonium* spp.) amb percentatges d'incidència que oscil·len entre el 40 i el 90% depenent de l'àrea geogràfica examinada. Una situació similar és previsible en països del nostre entorn i, probablement, en altres més allunyats. Es tracta d'un virus amb partícules isomètriques i un genoma monopartit de RNA de simple cadena i polaritat positiva les infeccions naturals del qual semblen restringides a espècies del gènere *Pelargonium*, encara que experimentalment pot infectar espècies molt diverses. Donada l'escassetesa de dades sobre el PLPV i la importància que el virus està adquirint com a patògen, en aquesta tesi hem pretès aprofundir en les seues característiques biològiques i moleculars.

El primer objectiu abordat en aquest treball ha estat determinar la seqüència nucleotídica completa del seu RNA genòmic (gRNA). Aquesta molècula està constituïda per 3883 nt i, mitjançant anàlisi *in silico*, inicialment identificarem sis marcs oberts de lectura (ORFs) que potencialment codificaven proteïnes de 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), i 37 kDa (p37). Aquestes ORFs estan flanquejades per una regió no traduïble (UTR) en 5' inusualment curta, amb només 6 nt, i per una 3' UTR de 246 nt. Tant l'organització de les ORFs en l' RNA viral com la majoria dels seus productes potencials s'assemblaven molt als implicats en replicació (p27 i p87), moviment (p7) i encapsidació (p37) de membres del gènere *Carmovirus* (família *Tombusviridae*). A pesar de les semblances que compartia el PLPV amb membres d'aquest gènere, els carmovirus produeixen dos RNAs subgenòmics (sgRNAs) per a la síntesi de les proteïnes de moviment i de coberta, mentre que un anàlisi Northern va indicar que el PLPV genera un sol sgRNA d'aproximadament 1.7 kb presumiblement implicat en la traducció de les proteïnes p7, p6 i p37. A més, la regió central del genoma dels carmovirus conté dos xicotets ORFs normalment en diferents fases de lectura i amb codons d'iniciació canònics que codifiquen dues proteïnes de moviment. No obstant això, solament la p7 potencialment codificada pel PLPV mostrava homologia significativa amb proteïnes de moviment

mentre que no es van trobar homòlegs clars per a la p6 (cal ressenyar que tampoc per a la p13). A més aquesta p6 del PLPV havia d'ésser expressada a partir d'una ORF amb un codó d'iniciació no canònic i/o, alternativament, mitjançant un mecanisme de "-1 frameshift" (FS) o de corriment de pauta de lectura que es produiria immediatament aigües dalt del codó de parada de la ORF (p7), generant-se una proteïna de 12 kDa (p7-FS o p12). Aquesta possibilitat es veia recolzada per la presència, immediatament aigües a dalt de l'esmentat codó, d'un heptanucleòtid la seqüència del qual s'ajusta al motiu característic d'aquest tipus de mecanisme, i per la possible formació d'una estructura de tipus forqueta aigües abaix de l'esmentat codó. És remarcable que tant l'organització de ORFs com les estratègies d'expressió gènica predites per al sgRNA del PLPV eren molt similars a les descrites per a l'únic sgRNA del virus del mosaic del Panicum (PMV), que constitueix el gènere *Panicovirus* dins de la família *Tombusviridae*.

A més del PLPV, s'havien descrit un conjunt de xicotets virus isomètrics les infeccions naturals dels quals semblaven també restringides a espècies del gènere *Pelargonium*, entre els quals es trobaven el virus de les taques anulars de *Pelargonium* (*Pelargonium ringspot virus*, PeIRSV) i el virus de l'anell cloròtic del *Pelargonium* (*Pelargonium chlorotic ring pattern virus*, PCRPV). Curiosament el PCRPV codifica potencialment cinc proteïnes molt similars a les implicades en replicació (p27 i p87), moviment (p7 i p9) i encapsidació (p37) del gènere *Carmovirus*, però produeix un únic sgRNA presumiblement implicat en la traducció de les ORFs allunyades de l'extrem 5' del gRNA i podria utilitzar un mecanisme -1 FS per a la traducció d'una ORF interna de poca grandària, igual que havia estat predit per al PLPV. Algunes dades indiquen que aquestes característiques podrien ser compartides pel PeIRSV i pel virus latent de la baia del saüc (*Elderberry latent virus*, ELV) i, sobre la base d'aquestes característiques comunes i de similituds de seqüència, es va proposar incloure als quatre agents virals en un nou gènere dins de la família *Tombusviridae*, amb el nom provisional de *Pelarspovirus*.

Les peculiaritats genòmiques del PLPV (i la seua possible adscripció a un nou gènere) mereixien sens dubte ser estudiades en major profunditat, per això el següent objectiu va ser la generació d'un clon infeccios del virus per a portar a terme un anàlisi detallat de les relacions estructura-funció en el patogen. Amb esta finalitat, fusionarem un cDNA viral de longitud completa al promotor de l' RNA polimerasa del fago T7 i ho lligarem a un vector de clonació d'alt nombre de còpies. Els transcrits generats *in vitro* a partir de la construcció resultant es van inocular mecànicament en diferents hostes

experimentals. Els resultats van mostrar que els RNAs del PLPV sintetitzats *in vitro* donaven lloc a infeccions indistingibles d'aquelles establertes per l'aïllat viral de partida en plantes de *C. quinoa*, *N. clevelandii* i *N. benthamiana*. Disposar d'aquesta eina ens va permetre desenvolupar estratègies per a obtenir informació sobre el paper de les distintes ORFs en el cicle biològic del virus i dels mecanismes que intervien en la seua expressió.

Mitjançant experiments de traducció *in vitro* i bioassajos de distints mutants identificàrem les ORFs biològicament actives del PLPV i vam obtenir dades sobre la/les etapa(es) del cicle infecció en la/les que estan involucrades. Les funcions predites inicialment per a les ORFs (p27), (p87), presumptivament implicades en replicació, i per les ORFs (p7) i (p37), en moviment i encapsidació, respectivament, eren consistents amb els resultats d'aquest treball, mentre que les ORFs (p13) i (p6), inicialment identificades *in silico*, no eren traduïdes *in vitro* ni requerides *in vivo*, de manera que probablement careixen de significat biològic. En el seu lloc es va identificar una nova ORF, localitzada en la regió central del genoma, que es traduïa a partir d'un codó feble d'iniciació (GUG) i que donava lloc a una proteïna de grandària 9.7 kDa implicada en el moviment del virus. Aquesta proteïna a més, mostrava una notable identitat de seqüència (44,4%) amb una altra de grandària 9.4 kDa, potencialment codificada per una ORF del PCRPV també situada en la regió central del genoma i que el seu codó d'inici era, així mateix, un triplet no-AUG, el que ampliava les semblances entre aquests dos virus, ambdós membres provisionals del gènere proposat *Pelarspovirus*.

Una vegada determinada la funció de les proteïnes codificades pel PLPV es va iniciar un estudi sobre les estratègies que utilitza aquest patògen per a l'expressió de tots els seus gens. Assajos de traducció *in vitro* van confirmar que: I) el gRNA dirigeix la síntesi de les proteïnes p27 i p87 de manera eficient (l'última mitjançant un mecanisme de lectura a través del codó de parada feble del gen p27), malgrat mancar d'una estructura cap en el seu extrem 5', com ocorre en la resta de membres de la família *Tombusviridae*, i presentar una seqüència capdavantera extremadament curta (6 nt) i, II) el sgRNA actua com mRNA per a la producció de les tres proteïnes restants, p7, p9.7 i p37, mitjançant escapes al procés de rastrege ribosomal o *leaky-scanning*. Aquests processos semblen estar afavorits pel context subòptim del codó d'inici del gen p7, pel codó feble d'inici (GUG) del gen p9.7, i per l'absència de codons AUG addicionals en qualsevol pauta de lectura en la regió que precedeix al codó d'inici del gen p37 en el sgRNA.

Finalment es va portar a terme un anàlisi de la variabilitat genètica del PLPV per tal d'obtenir informació sobre les pressions selectives que operen sobre el genoma del virus. Per a tal efecte, es va partir d'una col·lecció d'aïllats provinents de plantes de *P. zonale* o *P. peltatum* amb diferents orígens geogràfics i/o amb diferents dates de recollida. La regió genòmica examinada, de 1817 nt, comprenia una regió parcial del gen de la RdRp, els gens de les proteïnes de moviment (p7 i p9.7) i de la CP, i la 3' UTR. L'anàlisi de les seqüències obtingudes ha permès determinar alguns dels factors que actuen limitant l'heterogeneïtat viral com són la conservació de les seqüències aminoacídiques de determinats dominis de les proteïnes, i la preservació del plegament de regions reguladores del genoma viral. L'estudi de covariacions dins i entre proteïnes també ha permès identificar una suposada xarxa d'interaccions específiques probablement requerida per a mantenir la conformació correcta de les proteïnes virals i per a conduir l'RNA viral des del procés de replicació fins a la seua disseminació en la planta. Addicionalment, mitjançant passades seriadades d'un aïllat del PLPV en l'hoste experimental *C. quinoa*, s'ha demostrat que aquest patògen és capaç d'evolucionar de forma molt ràpida i reproduïble quan és transferit a una nova espècie vegetal. Aquesta observació posa de manifest el paper crític que ocupen les pressions selectives exercides per l'hoste en l'estructura genètica de les poblacions del virus i suggereix que la taxa de mutació del PLPV és inusualment elevada.

SUMMARY

ANALYSIS OF STRUCTURE-FUNCTION RELATIONSHIPS IN PELARGONIUM LINE PATTERN VIRUS

A recent survey made in Spain showed that the Pelargonium line pattern virus (PLPV) is the most frequent viral agent in geranium (*Pelargonium* spp.), with percentages of incidence ranging from 40 to 90 % depending on the examined geographical area. A similar situation is predictable in neighbouring countries and probably worldwide. PLPV produces isometric particles and has a single stranded monopartite RNA genome of positive polarity. Natural infections of the virus seem to be restricted to species of the genus *Pelargonium*, though diverse plant species can become experimentally infected. Due to the scarce data on PLPV and to the relevance that the virus is acquiring as pathogen, in this thesis we have intended to deep our knowledge on its biological and molecular characteristics.

The first objective of this work has been to determine the complete nucleotide sequence of PLPV genomic RNA (gRNA). This molecule comprises 3883 nt and, by *in silico* analysis, we initially identified six open reading frames (ORFs) potentially encoding proteins of 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), and 37 kDa (p37). These ORFs are flanked by an unusually short untranslatable region (UTR) at the 5' side with just 6 nt, and by a 3' UTR of 246 nt. The arrangement of the ORFs in the gRNA as well as the majority of its potential products were very similar to those involved in replication (p27 and p87), movement (p7) and encapsidation (p37) of members of the genus *Carmovirus* (family *Tombusviridae*). However, some unique features distinguished PLPV from carmoviruses. The latter ones produce two subgenomic RNAs (sgRNAs) for the synthesis of the movement and coat proteins, while a Northern blot analysis revealed that PLPV generates a single sgRNA, of approximately 1.7 kb, presumably involved in the translation of the proteins p7, p6 and p37. In addition, the central region of the carmovirus genome contains two small ORFs usually in different reading frames and with canonical initiation codons which code for two movement proteins. However, only the putative p7 of PLPV showed significant homology with movement proteins, while clear homologous could not be found for p6 (and neither for p13). Moreover, the ORF potentially encoding p6 had a non canonical initiation codon though, alternatively, this protein could be expressed through a -1 frameshift (FS) mechanism that would take place immediately upstream of

the stop codon of the ORF (p7), giving rise to a protein of 12 kDa (p7-FS or p12). This possibility was supported by the presence, immediately upstream of the mentioned stop codon, of a heptanucleotide whose sequence fits the canonical frameshifting motif. In addition, the formation of a stem-loop downstream of the putative shifty site was predicted, a type of structure which seems to favour frameshifting. Outstandingly, the arrangement of ORFs as well as the anticipated strategies of gene expression from PLPV sgRNA were very similar to those described for the single sgRNA of the *Panicum mosaic virus* (PMV), the only member of the genus *Panicovirus* in the family *Tombusviridae*.

Besides PLPV, a group of small isometric viruses whose natural infections also seem restricted to species from the genus *Pelargonium* had been described. This group included Pelargonium ringspot virus (PeIRSV) and Pelargonium chlorotic ring pattern virus (PCRPV). Remarkably, PCRPV potentially encodes five proteins very similar to those involved in replication (p27 and p87), movement (p7 and p9) and encapsidation (p37) in the genus *Carmovirus*. However, as predicted for PLPV, PCRPV generates a single sgRNA that is likely involved in the translation of 5'-distal ORFs of the gRNA and could use a -1 FS mechanism for the expression of an internal small ORF. Some data indicate that these characteristics could be shared by PeIRSV and Elderberry latent virus, (ELV) and, on the bases of these common characteristics and of sequence similarities, it was proposed to group the four viral agents in a prospective new genus in the family *Tombusviridae*, named *Pelarspovirus*.

The genomic peculiarities of PLPV (and its possible adscription to a new genus) deserved to be studied in depth, thus the next objective was to obtain an infectious clone of the virus in order to carry out a detailed analysis of the structure-function relationships in the pathogen. To this aim, a full length viral cDNA was fused to the promoter of bacteriophage T7 RNA polymerase and subsequently ligated to a high-copy-number vector. The transcripts generated *in vitro* from the resulting construct were mechanically inoculated in different experimental hosts. These *in vitro* synthesized PLPV RNAs gave rise to infections indistinguishable to those established by the parental virus in *C. quinoa*, *N. clevelandii* and *N. benthamiana* plants. The availability of this tool has allowed us to develop strategies aimed at obtaining information on the role of the different ORFs in the biological cycle of the virus and on the mechanisms that mediate their expression.

Through *in vitro* translation experiments and bioassays of different mutants, we identified the biologically active ORFs of PLPV and obtained data about the step(s) of the infectious cycle in which they are implicated. The results of this study were consistent with the anticipated functions of ORFs

(p27) and (p87), involved in replication, and ORFs (p7) and (p37), involved in movement and encapsidation, respectively, while ORFs (p13) and (p6), initially identified *in silico*, were not translated *in vitro* nor required *in vivo*, so they likely lack biological significance. On the other hand, a new ORF located in the central region of the genome was identified, which bears a weak initiation codon (GUG) and encodes a protein of 9.7 kDa involved in the movement of the virus. Remarkably, this PLPV protein shares notable sequence identity (44.4%) with a 9.4 kDa protein potentially encoded by a centrally located ORF of PCRPV. The initiation codon of such ORF would also correspond to a non-AUG triplet, further extending the resemblances among PLPV and PCRPV, both tentative members of the prospective genus *Pelarspovirus*.

Once the function of the proteins encoded by PLPV was determined, we investigated the strategies that this pathogen employs for translation of all its genes. The layout of *in vitro* translation experiments confirmed that: i) the gRNA directs efficient translation of proteins p27 and p87 (the latter by read-through of the weak stop codon of gene p27), despite lacking a cap structure at the 5' end, as most (if not all) members of the family *Tombusviridae*, and having an extremely short leader sequence (6nt), and, ii) the sgRNA acts as mRNA for the production of the three remaining proteins p7, p9.7 and p37, through *leaky-scanning* processes. These processes seem to be favoured by the suboptimal context of the starting codon of the gene p7, by the weak codon of initiation (GUG) of the gene p9.7, and by the absence of additional AUG codons in any reading frame in the region preceding the initiation codon of the gen p37 in the sgRNA.

Finally, in order to get insights into the selective pressures that operate on the PLPV genome, the molecular diversity of the virus was assessed using a collection of isolates from different geographical origins, *Pelargonium* spp., and collecting times. The genomic region examined was 1817 nt long and comprised the complete movement (p7 and p9.7) and coat protein genes as well as flanking segments including the 3'-proximal region of the RdRp gene and the entire 3' UTR. The results of sequence analyses have allowed to identify some constraints limiting virus heterogeneity such as the conservation of the aminoacidic sequences of specific domains of the proteins and the preservation of certain foldings of RNA regulatory regions. Covarying amino acids within and between proteins have allowed to infer a potential network of specific interactions probably required to maintain the correct conformation of the viral proteins and to drive the viral RNA from the replication process to its dissemination in the plant. Additionally, through serial passages of a PLPV variant in the experimental

host *C. quinoa*, we have proved that this pathogen is able to evolve in a very fast and reproducible fashion when it is transferred to a new plant species. This result highlights the strong selective pressures imposed by the host on the genetic structure of PLPV populations and suggests that the virus is endowed with an usually elevated mutation rate.

Índice

ÍNDICE

INTRODUCCIÓN GENERAL	1
1. ANTECEDENTES DE LA VIROLOGÍA	3
2. FAMILIA <i>TOMBUSVIRIDAE</i>	7
2.1. Género <i>Carmovirus</i>	13
2.1.1. Organización genómica	15
2.1.2. Mecanismos de expresión génica	17
2.2. Género <i>Panicovirus</i>	18
2.2.1. Organización genómica.....	19
2.2.2. Mecanismos de expresión génica.....	20
2.3. Características de las proteínas implicadas en replicación	21
2.4. Características de las proteínas implicadas en movimiento	22
2.5. Características de la proteína de cubierta.....	24
3. TRADUCCIÓN EN EUCARIOTAS	26
3.1. Iniciación de la traducción en eucariotas	26
3.1.1. Modelo del bucle cerrado en mRNAs eucariotas	27
3.1.2. Rastreo o barrido (<i>scanning</i>)	28
3.1.3. Codón de inicio	28
3.2. Elongación	29
3.3. Terminación	30
4. TRADUCCIÓN DE VIRUS DE PLANTAS	30
4.1. Modelo del bucle cerrado en mRNAs virales	30
4.2. Mecanismos no convencionales de traducción en RNAs virales	34
4.2.1. Corrimiento en la pauta de lectura del ribosoma (<i>ribosomal frameshifting</i>)....	35
4.2.2. Lectura a través de codones de terminación débiles (<i>readthrough</i>).....	37
4.2.3. Reiniciación.....	38
4.2.4. Rastreo ribosomal discontinuo (<i>shunting</i>).....	39
4.2.5. Escape al proceso de rastreo ribosomal (<i>leaky scanning</i>).....	40
4.2.6. Iniciación interna	43
5. VARIABILIDAD Y EVOLUCIÓN DE LOS VIRUS	48
5.1. Heterogeneidad genética en las poblaciones de virus de RNA	48
5.2. Fuentes de variación genética	49
5.3. Factores que determinan la estructura genética de las poblaciones virales.....	52
5.3.1. Selección natural	52
5.3.2. Deriva genética	53
5.4. Coevolución	54
6. VIROSIS DEL GERANIO	55
6.1. El virus del arabesco del Pelargonium (PLPV).....	56
6.1.1. Propiedades biológicas y modo de transmisión	56
6.1.2. Propiedades moleculares	57
OBJETIVOS	59

CAPÍTULO I	63
Complete nucleotide sequence and genome organization of Pelargonium line pattern virus and its relationship with the family <i>Tombusviridae</i>	65
CAPÍTULO II	83
Biological activity of transcripts from cDNA of Pelargonium line pattern virus.....	85
CAPÍTULO III	89
Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus.....	91
CAPÍTULO IV	101
Genetic structure and evolution of Pelargonium line pattern virus populations.....	103
DISCUSIÓN GENERAL	131
CONCLUSIONES	151
BIBLIOGRAFÍA	157

Introducción General

1. ANTECEDENTES DE LA VIROLOGÍA

Los orígenes de la disciplina científica conocida hoy día como Virología apenas se remontan a finales del siglo XIX, cuando se empezaron a llevar a cabo las primeras investigaciones científicas centradas en el estudio de enfermedades de las plantas cuyo agente causal era un virus. En esta época, los resultados obtenidos en microbiología médica habían establecido sólidamente la teoría de que las enfermedades infecciosas son causadas por gérmenes microscópicos que podían ser cultivados en medios nutritivos especiales y aislados por medio de filtros muy finos capaces de retener a dichos microorganismos. Sin embargo, en 1892 el ruso Dimitri Ivanovsky demostró que el agente causal de la enfermedad vegetal conocida como mosaico del tabaco, descrita en 1886 por Adolf Mayer, era capaz de pasar a través de filtros a prueba de bacterias y no podía ser observado bajo el microscopio ni cultivado en medios artificiales. Estos resultados llevaron a Ivanovsky a especular que el mosaico del tabaco era causado por un germen productor de toxinas, que serían las que atravesaban el filtro bacteriológico. En 1898, el holandés Martinus Beijerinck realizó experimentos más rigurosos que demostraron que el agente del mosaico del tabaco podía multiplicarse en los tejidos de las plantas infectadas, una observación que descartaba la posibilidad de que se tratara de una toxina. Beijerinck lo denominó "*contagium vivum fluidum*" y concluyó que debía tratarse de una sustancia disuelta o soluble, capaz de replicarse pero solamente cuando se encontraba incorporada al protoplasma de la célula viva infectada.

En 1935 el químico Wendell Stanley pudo aislar el virus causante del mosaico del tabaco (*Tobacco mosaic virus*, TMV, *Tobamovirus*) y purificarlo en forma cristalina. Este hecho representó la primera cristalización de un material biológico supuestamente vivo. En 1937 Bawden y Pirie pusieron de manifiesto que el TMV estaba compuesto por proteína y ácido ribonucleico (RNA), y no fue hasta 1952 cuando se demostró que la información genética reside en los ácidos nucleicos. Hershey y Chase infectaron *Escherichia coli* con un virus bacteriano y el resultado indicó que solamente el DNA viral penetraba en las bacterias y éste era suficiente para causar una infección efectiva. La proteína de la cubierta viral (*coat protein*, CP) no era necesaria para producir la infección y su función parecía ser la de proteger al ácido nucleico y permitir la adsorción del virus a las células. Los resultados de este experimento sirvieron para demostrar que el DNA constituye el material genético hereditario. En 1956

4 Introducción General

Fraenkel-Conrat y colaboradores demostraron la infectividad del RNA del TMV y el papel protector de la CP, poniendo de manifiesto que el RNA, al igual que el DNA, también era portador de la información genética. Estos descubrimientos marcaron el inicio de la Virología moderna.

Un virus puede definirse como un conjunto de una o más moléculas de ácidos nucleicos, normalmente encapsidada/s en una/s cubierta/s de protección de proteínas o lipoproteínas, capaz de multiplicarse en el interior de la célula hospedadora. Se trata de parásitos intracelulares que utilizan la maquinaria metabólica de la célula huésped para poder propagarse. Es un proceso que incluye varias síntesis separadas y el ensamblaje posterior de todos los componentes para dar origen a nuevas partículas infecciosas.

Los virus muestran una gran diversidad en tamaño y forma. Resulta difícil establecer un concepto clasificador, de hecho hay muchos criterios para ello, entre los que destacan: i) la forma de la partícula vírica o virión, ii) la presencia o no de cubierta proteica o lipoproteica, iii) el tipo de ácido nucleico que poseen. Atendiendo a este último criterio, los virus pueden presentar genomas de RNA de simple cadena y polaridad positiva (ssRNA+), simple cadena y polaridad negativa (ssRNA-), RNA de doble cadena (dsRNA), o genomas de DNA que a su vez pueden ser de simple (ssDNA) o de doble cadena (dsDNA) (Fauquet *et al.*, 2005). Dependiendo del tipo de ácido nucleico, los virus seguirán distintos pasos para su replicación. En los virus de ssRNA+ el genoma actúa directamente como RNA mensajero (mRNA) para la traducción de las proteínas virales que se encargarán de replicar el RNA viral. A su vez, servirá de molde para la síntesis de la hebra complementaria, un proceso catalizado por la RNA polimerasa-RNA dependiente (*RNA dependent-RNA polymerase*, RdRp) codificada por el propio patógeno. A partir de las hebras complementarias se generarán nuevas hebras positivas, que serán el material genético de la progenie viral. En los virus de ssRNA- se sintetiza inicialmente un RNA de polaridad positiva que sirve, por un lado, como molde para la producción de nuevas moléculas de RNA con polaridad negativa que constituyen el genoma de la progenie y, por otro, como mRNA para la traducción de proteínas virales. Para sintetizar el RNA de polaridad positiva se utiliza la RdRp viral que está contenida en el propio virión en este tipo de virus, al igual que en virus de dsRNA, donde también será la encargada de sintetizar los mRNAs virales y los RNAs de doble hebra genómicos de la progenie. En los virus de ssDNA, el DNA viral

monocatenario pasa a ser bicatenario por medio de una DNA polimerasa-DNA dependiente (*DNA dependent-DNA polymerase*, DdDp) de la célula huésped. A partir de los dsDNAs se generan los mRNAs virales, un proceso mediado por una RNA polimerasa-DNA dependiente (*DNA dependent-RNA polymerase*, DdRp) del huésped y, a su vez, los dsDNAs sirven como molde para la síntesis de nuevas moléculas de ssDNA que constituyen el genoma de la progenie. Por último, en el caso de los virus de dsDNA se han de sintetizar los mRNAs, normalmente mediante una DdRp de la célula huésped, que dirigen la traducción de proteínas virales. La multiplicación del dsDNA para generar la progenie es catalizada habitualmente por una DpDd del huésped aunque existen casos en que la maquinaria replicacional es codificada enteramente por el virus. En el caso de virus de plantas, algunos de los mRNAs virales actúan como intermediarios replicativos, sirviendo para la síntesis de nuevas moléculas de dsDNA, un proceso de retrotranscripción asistido por una DNA polimerasa-RNA dependiente (*RNA dependent-DNA polymerase*, RdDp) codificada por el propio virus. Los retrovirus de animales siguen un proceso similar para su replicación, con la diferencia importante de que la forma infecciosa es el ssRNA⁺ y de que el intermediario de dsDNA se integra en el genoma del huésped (Fauquet *et al.*, 2005). Los virus de ssRNA⁺ se pueden clasificar en tres grupos dependiendo de las similitudes que muestran sus correspondientes RdRps: supergrupo I (virus tipo picorna), supergrupo II (virus tipo carmo y tipo flavi) y supergrupo III (virus tipo alfa) (Koonin y Dolja, 1991; Koonin y Dolja, 1993).

Los genomas virales raramente superan las 10 kilobases (kb) debido a restricciones derivadas de la necesidad de adaptarse fácilmente al medio y de conseguir una replicación y encapsidación rápida y eficiente. El número de proteínas que codifican es variable, normalmente entre cuatro y diez, que a su vez suelen ser multifuncionales. Los virus utilizan diferentes estrategias para conseguir que todos sus genes sean accesibles a la maquinaria de síntesis proteica de la célula huésped. Dichas estrategias incluyen la partición de su genoma en varios componentes, la síntesis de RNAs subgenómicos (sgRNAs) a partir de los RNAs genómicos (gRNAs), el procesamiento proteolítico de una única poliproteína, el solapamiento de pautas abiertas de lectura (*open reading frames*, ORFs), etc. Además de las regiones codificantes, el genoma viral contiene regiones nucleotídicas no traducibles (*unstranlated*

6 Introducción General

regions, UTRs) en los extremos 5' y 3' que desempeñan un papel importante en la replicación y traducción viral.

En la actualidad hay descritas más de 800 especies de virus de plantas (Fig. 1) que se clasifican en 20 familias y 88 géneros, lo que refleja la gran variabilidad existente entre los mismos (Fauquet *et al.*, 2005). La mayoría de ellos (más del 80%) presentan genomas de ssRNA+.



Fig. 1. Síntomas inducidos por diferentes virus en distintas especies vegetales. (A) Virus de las manchas necróticas en anillo del ciruelo. (B) Virus de la rotura del color del tulipán. (C) Virus del moteado del pimiento dulce. (D) Virus de los anillos cloróticos del pimiento Morrón (E) Virus Y de la patata. (F) Virus de la tristeza de los cítricos. (G) Virus de la psoriasis de los cítricos.

El virus objeto de estudio de esta tesis doctoral es el virus del arabesco del *Pelargonium* (Pelargonium line pattern virus, PLPV). Actualmente no está adscrito formalmente a ningún grupo viral aunque, como se pondrá de manifiesto en esta memoria, se puede considerar un miembro de la familia *Tombusviridae* que comparte propiedades con especies de los géneros *Carmovirus* y *Panicovirus*. A continuación se describen las características generales de la familia *Tombusviridae* y, más detalladamente, las de los géneros mencionados.

2. FAMILIA *TOMBUSVIRIDAE*

La familia *Tombusviridae* está constituida por ocho géneros: *Tombusvirus* (15 especies definitivas y 1 especie provisional), *Avenavirus* (1 especie), *Carmovirus* (14 especies definitivas y 8 provisionales), *Necrovirus* (6 especies definitivas y 2 provisionales), *Dianthovirus* (3 especies definitivas y 3 provisionales), *Machlomovirus* (1 especie), *Aureusvirus* (2 especies) y *Panicovirus* (1 especie definitiva y 1 especie provisional) (Lommel *et al.*, 2005).

Los componentes de esta familia presentan una distribución mundial, con una mayor incidencia en las regiones templadas. Su gama de huéspedes naturales incluye a plantas monocotiledóneas o dicotiledóneas, pero no ambas a la vez. La gama de huéspedes experimentales suele ser muy amplia para cada especie viral. La infección a menudo está limitada al sistema radicular, pero en otros casos, dependiendo del huésped, el virus puede invadir la planta sistémicamente. Se transmiten por inoculación mecánica y por propagación vegetativa; además se encuentran en ambientes naturales (aguas superficiales y tierras de cultivo) por lo que, en general, no necesitan vectores para su transmisión, aunque en varios casos se ha descrito la transmisión de virus por medio de vectores biológicos. La sintomatología causada por la infección de estos virus presenta con frecuencia rasgos comunes e incluye moteado, arrugamiento, necrosis y deformación de las hojas. Sin embargo, las infecciones pueden ser también asintomáticas y la inducción o no de efectos fenotípicos parece estar condicionada por el aislado del virus, la especie vegetal y/o las condiciones ambientales.

El genoma de los miembros de esta familia, con excepción del género *Dianthovirus* (que contienen dos gRNAs), está constituido por una molécula de ssRNA⁺ con un tamaño que oscila entre 3.7 y 4.8 kb (Fig. 2). El extremo 3' de dicho RNA no está poliadenilado y el extremo 5' no parece estar protegido. Sólo en el caso del virus del moteado del clavel (*Carnation mottle virus*, CarMV, *Carmovirus*), del virus del mosaico necrótico del trébol rojo (*Red clover necrotic mosaic virus*, RCNMV, *Dianthovirus*), y del virus del moteado clorótico del maíz (*Maize chlorotic mottle virus*, MCMV, *Machlomovirus*) se ha apuntado la existencia de una estructura cap, pero revisiones recientes de esta cuestión en el RCNMV han mostrado que el extremo 5' carece de esta estructura (Mizumoto *et al.*, 2006; Iwakawa *et al.*, 2007), y es posible que la situación sea similar en el CarMV y MCMV.

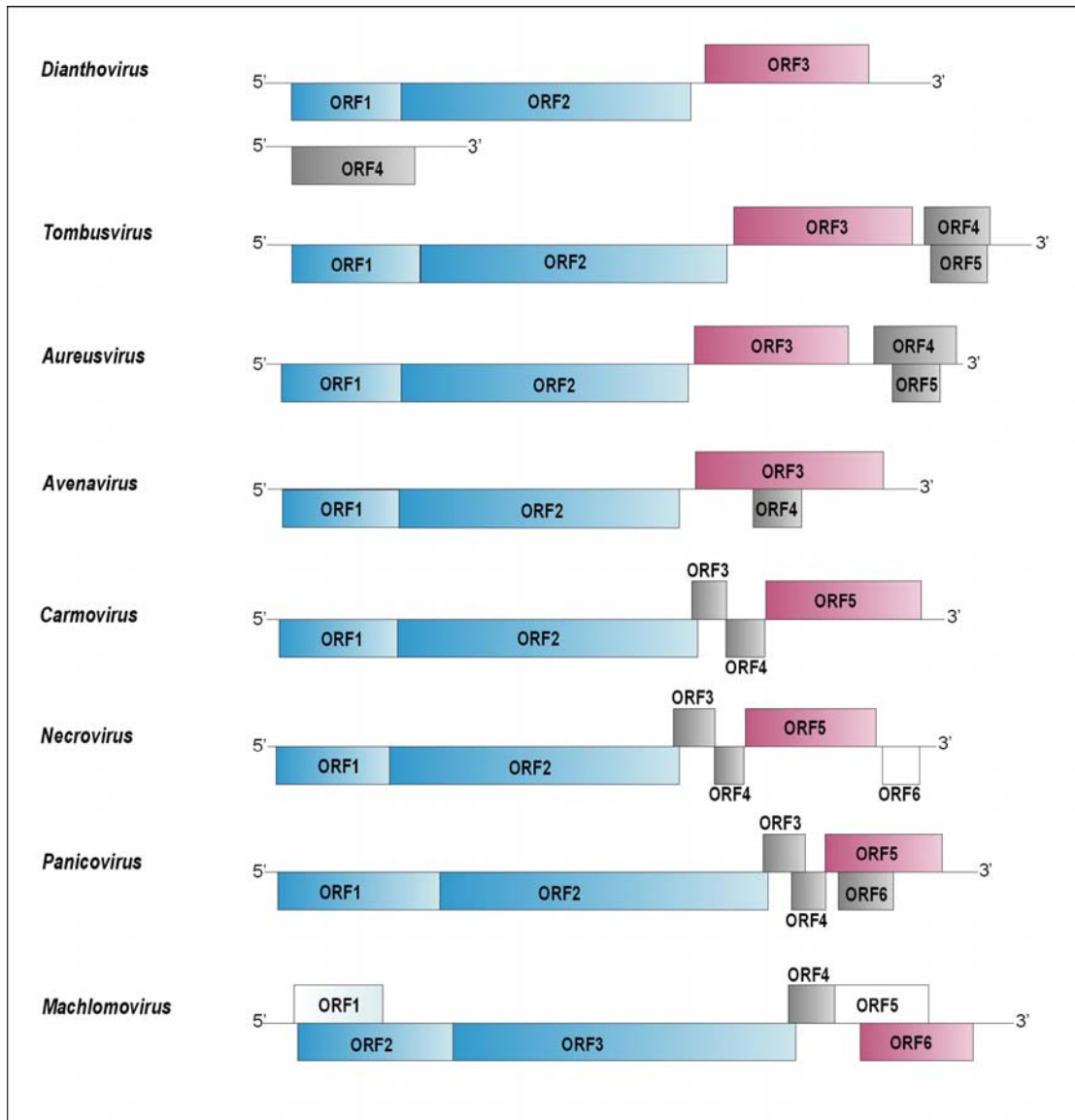


Fig. 2. Representación de la organización genómica de las especies tipo de los géneros pertenecientes a la familia *Tombusviridae*. Las diferentes ORFs se representan como cajas de distintos colores que indican la función de las proteínas que codifican: i) azul: proteínas implicadas en replicación, ii) rosa: proteína de cubierta, iii) gris: proteínas implicadas en movimiento, iv) blanco: ORF de función desconocida.

La organización genómica difiere entre los géneros de esta familia (Fig. 2), siendo el rasgo común de todos ellos la producción de una RdRp muy conservada que es traducida a partir de una ORF que, con excepción del género *Dianthovirus*, está interrumpida por un codón de parada débil que no es reconocido en determinadas rondas de traducción produciéndose un proceso de lectura a través (*readthrough*). Los miembros del género *Dianthovirus* expresan la RdRp mediante un corrimiento de -1 nt en la pauta de lectura (*frameshift -1*) de los ribosomas que han empezado la traducción en el codón de inicio de la ORF1. Los genomas de los virus

pertenecientes a los géneros *Dianthovirus* y *Avenavirus* poseen cuatro ORFs, mientras que el resto de los géneros de la familia contienen al menos cinco. Los géneros *Machlomovirus* y *Panicovirus* pueden extender uno de los genes de las proteínas de movimiento (MPs) por la presencia de un codón de terminación débil que permite la lectura a través (*Machlomovirus*), o de un heptanucleótido que facilita un proceso de corrimiento de lectura (*Panicovirus*), dando lugar a la generación de nuevas ORFs. Las ORFs 1 y 2 son expresadas a partir del gRNA, mientras que las situadas en posición 3'-proximal se expresan a partir de sgrNAs. Con la excepción del MCMV (Fig. 2), el gen de la RdRp es la ORF más cercana al extremo 5'. Para todos los géneros, la localización del gen que codifica la CP es interna o 3'-proximal (Lommel *et al.*, 2005).

Morfológicamente todos los miembros de la familia *Tombusviridae* se caracterizan por presentar partículas formadas por 180 subunidades de la misma proteína (CP) con simetría icosaédrica (T=3). Los viriones de los distintos géneros presentan dos orígenes filogenéticos diferentes. Por un lado, los géneros *Aureusvirus*, *Avenavirus*, *Carmovirus*, *Dianthovirus* y *Tombusvirus*, producen partículas con una forma redondeada, superficie granular y un diámetro de 32-35 nm. Cada subunidad proteica se pliega en tres dominios estructurales distintos: i) R, el dominio interno N-terminal que contiene un gran número de residuos proteicos cargados positivamente y que interacciona con el RNA, ii) S, el dominio armazón que constituye el esqueleto de la cápsida y que se conecta mediante un brazo amino terminal al dominio R, y iii) P, el dominio C-terminal que se proyecta hacia fuera y proporciona al virus su apariencia granular (Fig. 3A). Los dominios P se agrupan en parejas para formar 90 proyecciones sobre la superficie de la partícula, y los contactos entre dímeros son importantes para el ensamblaje y la estabilización de la estructura del virión. El dominio S forma una estructura cilíndrica compuesta por hebras β . Dos sitios de unión a Ca^{++} estabilizan los contactos entre dominios S. Los viriones de los géneros *Machlomovirus*, *Necrovirus* y *Panicovirus*, se componen de cubiertas proteicas que carecen del dominio P, lo que les proporciona una superficie de aspecto no granulado (Fig. 3B). Su diámetro oscila entre los 30-32 nm (Lommel *et al.*, 2005).

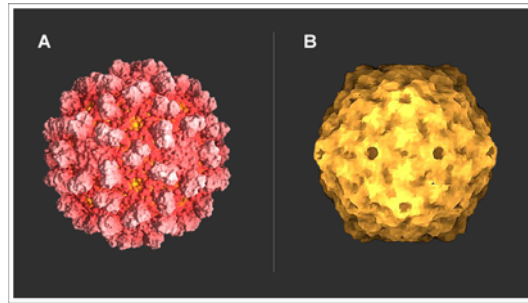


Fig. 3. (A) Representación de una partícula vírica de un miembro del género *Tombusvirus*, con proyecciones de los dominios P que proporcionan al virión el característico aspecto granuloso. (B) Representación de una partícula vírica de un miembro del género *Necrovirus*, que se compone de cápsidas proteicas que carecen del dominio P.

Las MPs de los componentes de esta familia también tienen diferentes orígenes filogenéticos: i) *Avenavirus*, *Carmovirus*, *Machlomovirus*, *Necrovirus* y *Panicovirus*, codifican una proteína de 7-9 kDa que está acompañada en todos los casos, excepto en el género *Avenavirus*, por una proteína de 6-12 kDa, ii) *Tombusvirus* y *Aureusvirus* codifican una proteína de tamaño 22-27 kDa, junto con una adicional de 14-19 kDa que también está implicada en movimiento y en el aumento de la severidad de los síntomas y, iii) *Dianthovirus* producen una proteína de 35 kDa. Además de las ORFs implicadas en el movimiento viral, existen ORFs muy próximas en el genoma y de función desconocida en los géneros *Necrovirus* y *Machlomovirus* (Lommel *et al.*, 2005).

La replicación tiene lugar en el citoplasma celular, posiblemente en vesículas membranosas que pueden estar asociadas al retículo endoplasmático o a orgánulos como peroxisomas, mitocondrias, o más raramente, cloroplastos. Los viriones se acumulan en el citoplasma y vacuolas, algunas veces en forma cristalina (Lommel *et al.*, 2005).

Los virus de la familia *Tombusviridae* poseen en la 3' UTR secuencias y/o estructuras que están implicadas en la replicación y la traducción viral. Algunos de estos motivos estructurales funcionan como promotores que son reconocidos por la RdRp para el comienzo de la síntesis de las cadenas de RNA de polaridad negativa a partir del extremo 3' de las de polaridad positiva. Dichos promotores pueden corresponder tanto a elementos de estructura primaria como secundaria e incluso terciaria (Turner y Buck, 1999; Wang y Wong, 2004; White y Nagy, 2004). En los géneros *Carmovirus* y *Tombusvirus*, el promotor mínimo incluye los 19 últimos nucleótidos del extremo 3' del gRNA de polaridad positiva que se pliegan adoptando una

estructura de tipo horquilla flanqueada en 3' por una secuencia de simple cadena, CUGCCC, en el caso de *Carmovirus* (Simon y Howell, 1986; Song y Simon, 1995; Stupina y Simon, 1997; Carpenter y Simon, 1998) y CCC en el caso de *Tombusvirus* (White y Nagy, 2004). Estas bases desapareadas especifican el sitio de iniciación de la síntesis de las cadenas de RNA de polaridad negativa y, aunque la función de la horquilla no se ha determinado de forma precisa, se ha observado que su desestabilización da lugar a una disminución notable en los niveles de replicación en plantas y protoplastos (Havelda *et al.*, 1995; Fabian *et al.*, 2003). Una vez sintetizada la cadena negativa, la RdRp debe reconocer secuencias/estructuras específicas en su 3' UTR (que corresponde a la 5' UTR del ssRNA+) para la síntesis de la cadena de polaridad positiva (Panavas *et al.*, 2002). Además de los promotores se han descrito elementos estimuladores de la replicación (*replication enhancers*) que actúan potenciando la síntesis de cadenas positivas, y elementos supresores de la replicación que disminuyen la síntesis de cadenas negativas; ambos tipos de elementos probablemente son en gran parte responsables de la asimetría entre cadenas positivas y negativas que caracteriza a la replicación viral (Ray y White, 1999; Nagy *et al.*, 1999b, 2001; Pogany *et al.*, 2003; Zhang *et al.*, 2004a, 2004b; Panavas y Nagy, 2005; Na y White, 2006)

Los virus de esta familia codifican una o más proteínas que son traducidas a partir de sgRNAs. La producción de sgRNAs es una estrategia desarrollada por los virus para subsanar la escasa o nula eficiencia de la traducción de ORFs internas o 3'-terminales, originada por la dependencia de los ribosomas eucarióticos del extremo 5' del mRNA. Se han propuesto diferentes mecanismos para la síntesis de los sgRNAs: i) iniciación interna, en el que la replicasa inicia la transcripción del sgRNA internamente a partir del gRNA de polaridad negativa, ii) síntesis discontinua, descrita en el caso concreto de la familia *Nidoviridae*, y iii) terminación prematura de la síntesis de cadenas negativas que a su vez servirán de molde para la síntesis de sgRNAs de polaridad positiva (revisado por Miller y Koev, 2000; White, 2002). El mecanismo de iniciación interna es, aparentemente, el más extendido para la síntesis de los sgRNAs. Como norma general, las secuencias promotoras para la generación de sgRNAs mediante este mecanismo suelen localizarse justo antes del sitio de inicio de la transcripción (Johnston y Rochon, 1995; van der Vossen *et al.*, 1995; Wang y Simon, 1997; Koev *et al.*, 1999), aunque en algunos casos se ha descrito la participación de secuencias

12 Introducción General

alejadas físicamente (Balmori *et al.*, 1993; Koev y Miller, 2000). En el género *Carmovirus*, concretamente en el virus del arrugamiento del nabo (*Turnip crinkle virus*, TCV) y en el virus de las manchas cloróticas en anillo del hibisco (*Hibiscus chlorotic ringspot virus*, HCRSV), se ha demostrado la actividad promotora de las regiones que preceden a los sitios de inicio de los sgRNAs (Wang y Simon, 1997; Koev *et al.*, 1999; Li y Wong, 2006). Estas regiones, de aproximadamente 100 nt, pueden plegarse adoptando estructuras secundarias del tipo horquilla que son muy estables debido a que contienen varios pares de bases G:C consecutivos en la base del tallo. La adopción de este tipo de conformación parece ser crítica para el mecanismo de transcripción de los sgRNAs (Wang *et al.*, 1999; Li y Wong, 2006). Los gRNAs y sgRNAs del TCV y del HCRSV presentan el mismo motivo de secuencia en su extremo 5' ($G_{2-3}(A/U)_{3-7}$), que además está conservado en otros carmovirus. Este motivo puede ser requerido para la replicación y transcripción de los RNAs virales, facilitando el reconocimiento por parte de la polimerasa viral o permitiendo una interacción adecuada con la cadena negativa para la síntesis de la positiva. Numerosos virus de plantas de polaridad positiva también presentan similitud de secuencia entre los extremos 5' de los gRNAs y sus correspondientes sgRNAs (Marsh *et al.*, 1989; Kelly *et al.*, 1994; Zavriev *et al.*, 1996; Vives *et al.*, 2002). Aunque el modelo de iniciación interna parece operar en el caso de carmovirus, la síntesis de sgRNAs de miembros de los géneros *Tombusvirus* y *Aureusvirus* parece seguir un modelo de terminación prematura (Lin y White, 2004; Lin *et al.*, 2007; Wang *et al.*, 2008; Xu y White, 2008), de modo que virus relacionados pueden diferir en los mecanismos de generación de mRNAs para la traducción de ORFs internas o 3' proximales.

Por otro lado, en algunos géneros se han descrito RNAs defectivos interferentes (*defective interfering RNAs*, DI RNAs), virus satélites y/o RNAs satélites (*satellite RNAs*, satRNAs) asociados a determinadas especies (revisado por Pelczyk *et al.*, 2006). Los DI RNAs se originan a partir del genoma del virus correspondiente mediante un mecanismo de síntesis discontinua de la RdRp, de modo que normalmente presentan una o varias deleciones internas y carecen de capacidad codificante, por lo que precisan del virus parental para poder replicarse y moverse. De forma similar, los virus satélites y satRNAs necesitan estar asociados a un virus auxiliar, aunque no presentan homología de secuencia con el mismo o bien esta es muy limitada. La diferencia entre los virus satélites y los satRNAs es que los primeros codifican su

propia CP, mientras que los segundos utilizan la CP del virus auxiliar. Todos estos tipos de moléculas pueden alterar los síntomas inducidos por el virus con el que están asociados, ya sea atenuándolos o intensificándolos.

2.1. Género Carmovirus

Tal como se ha indicado anteriormente, el género *Carmovirus* está constituido por 14 especies definitivas y 8 provisionales (Tabla 1).

Tabla 1. Especies pertenecientes al género *Carmovirus*

Especies definitivas	
<i>Virus del ajo transmitido por el agua</i>	AWBV
<i>Virus del mosaico suave de la judía</i>	BMMV
<i>Virus del salpicado clorótico del cardamomo</i>	CCFV
<i>Virus del moteado del clavel</i>	CarMV
<i>Virus del moteado del chícharo</i>	CPMoV
<i>Virus del pepino transmitido por el suelo</i>	CuSBV
<i>Virus del mosaico de la Galisonga</i>	GaMV
<i>Virus de las manchas cloróticas en anillo del Hibiscus</i>	HCRSV
<i>Virus de las manchas necróticas del melón</i>	MNSV
<i>Virus de la rotura del color de la flor de Pelargonium</i>	PFBV
<i>Virus del cactus Saguaro</i>	SCV
<i>Virus del arrugamiento del nabo</i>	TCV
<i>Virus de las malas hierbas transmitido por el agua</i>	WWBV
<i>Virus de los anillos necróticos del iris japonés</i>	JINRV
Especies provisionales	
<i>Virus de la necrosis del vástago del guisante</i>	PNSV
<i>Virus del moteado del Blackgram</i>	BMoV
<i>Virus latente de la baya del saúco</i>	EILDV
<i>Virus del moteado de la glicina</i>	GMoV
<i>Virus de la necrosis apical del narciso</i>	NTNV
<i>Virus del plantain 6</i>	PIV-6
<i>Virus asintomático de la tefrosia</i>	TeSV
<i>Virus de la necrosis de la calabaza</i>	SqNV

Este género, cuya especie tipo es el CarMV, pasó a ser reconocido como grupo taxonómico en 1991 por el Comité Internacional para la Taxonomía de los Virus (*International Committee*

on *Taxonomy of Viruses*, ICTV) (Morris, 1991). La gama de huéspedes naturales de cada especie es reducida y restringida a dicotiledóneas, si bien la de huéspedes experimentales es amplia. Suelen causar importantes pérdidas económicas en las cosechas, especialmente en las regiones templadas (Russo *et al.*, 1994). Todos ellos se transmiten por inoculación mecánica y por propagación vegetativa, y en algunos casos también se ha descrito la transmisión por vectores. Así, se ha demostrado que el hongo *Olpidium bornovanum* actúa como vector del virus del pepino transmitido por el suelo (*Cucumber soil borne virus*, CuSBV) y del virus de las manchas necróticas del melón (*Melon necrotic spot virus*, MNSV) (Martelli, 1994), mientras que otros miembros del género pueden ser transmitidos por escarabajos, como el virus del moteado del chícharo (*Cowpea mottle virus*, CPMoV), el virus del mosaico suave de la judía (*Bean mild mosaic virus*, BMMV), el virus del moteado del *Blackgram* (*Blackgram mottle virus*, BMoV), y el TCV (van Regenmortel y Mahy, 2004), o por trips, como se ha descrito para el virus de la rotura del color de la flor de *Pelargonium* (*Pelargonium flower break virus*, PFBV) (Krczal *et al.*, 1995).

Actualmente se conoce la secuencia completa de una especie provisional del género, el virus de la necrosis del tallo del guisante (*Pea stem necrosis virus*, PSNV) (Suzuki *et al.*, 2002), y de diez especies definitivas: CarMV (Guilley *et al.*, 1985), TCV (Carrington *et al.*, 1989), MNSV (Riviere y Rochon, 1990), virus del salpicado clorótico del cardamomo (*Cardamine chlorotic fleck virus*, CCFV) (Skotnicki *et al.*, 1993), CPMoV (You *et al.*, 1995), virus del cactus Saguaro (*Saguaro cactus virus*, SCV) (Weng y Xiong, 1997), virus del mosaico del *Galisonga* (*Galisonga mosaic virus*, GaMV) (Ciuffreda *et al.*, 1998), HCRSV (Huang *et al.*, 2000), virus de los anillos necróticos del iris japonés (*Japanese iris necrotic ring virus*, JINRV) (Takemoto *et al.*, 2000) y PFBV (Rico y Hernández, 2004). Además, también se conoce la secuencia completa del virus del aclaramiento de las venas del lupino Nootka (Nootka lupine vein-clearing virus, NLVCV), y del virus de la rotura del color de la flor de la angelonia (*Angelonia flower break virus*, AnFBV) que han sido propuestos como nuevas especies de este género (Adkins *et al.*, 2006; Robertson *et al.*, 2007).

El genoma de estos virus se encapsida en partículas isométricas de 32-35 nm de diámetro que poseen una superficie regular vista al microscopio electrónico, con apariencia granular. El estudio de las características estructurales de los viriones de carmovirus ha sido abordada de

forma detallada con la cristalización, análisis con rayos-X y reconstrucción tridimensional del HCRSV (Doan *et al.*, 2003; Lee *et al.*, 2003). Recientemente, también se ha llevado a cabo un análisis con rayos-X del MNSV (Wada *et al.*, 2008).

2.1.1. Organización genómica

El gRNA de las especies del género *Carmovirus*, cuyo tamaño oscila entre 3.88 y 4.45 kb, contiene al menos cinco ORFs que codifican proteínas implicadas en replicación, movimiento y encapsidación viral, además de UTRs en los extremos 5' y 3' (Fig. 4). La ORF más próxima al extremo 5', ORF1, codifica una proteína de 27-29 kDa y termina con un codón ámbar cuya lectura a través permite continuar la traducción hasta el codón de terminación de la ORF2 dando lugar a una proteína de 85-88 kDa que contiene los motivos conservados de las RdRps virales. En la parte central del genoma se sitúan dos ORFs pequeñas cuya traducción da lugar a dos proteínas de tamaños 7-8 y 8-12 kDa, respectivamente, implicadas en el movimiento del virus. La ORF5, cercana al extremo 3', codifica la CP, de tamaño 36-41 kDa (Lommel *et al.*, 2005).

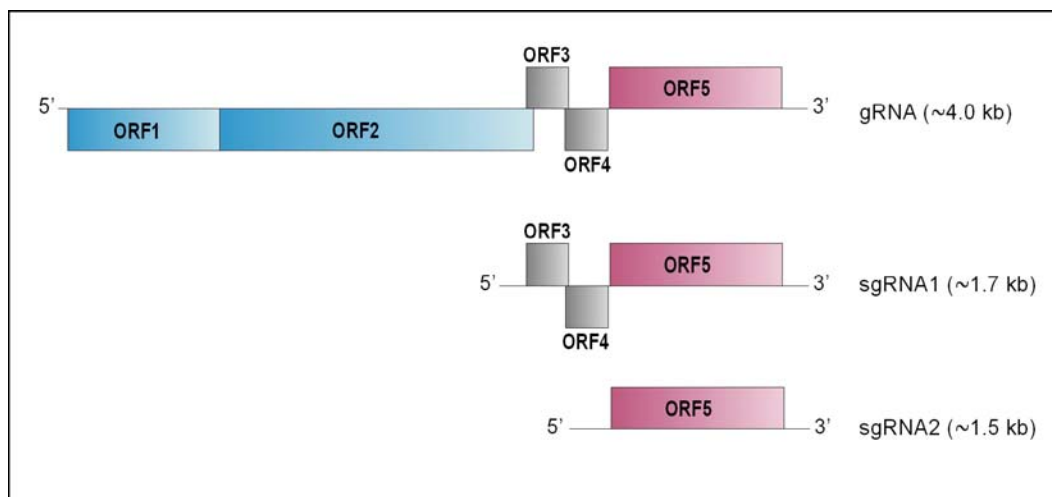


Fig. 4. Representación de la organización genómica de la mayoría de los miembros del género *Carmovirus*. Las distintas ORFs se representan con recuadros de diferente color dependiendo de la función de las proteínas que codifican. Las ORFs 1 y 2 codifican proteínas implicadas en replicación y se expresan a partir del gRNA. Las ORFs 3 y 4 codifican proteínas implicadas en movimiento y son expresadas a partir del sgRNA1. La ORF 5 codifica la CP viral y es traducida a partir del sgRNA2.

A pesar de la simplicidad estructural de estos patógenos, parece existir cierta variabilidad en el número de ORFs y en los mecanismos de expresión génica entre los diferentes miembros del grupo. En el caso del HCRSV, se ha confirmado que, además de las cinco ORFs presentes en el resto de carmovirus, contiene una ORF adicional cerca del extremo 5' del gRNA que codifica una proteína de 23 (p23) kDa, y otra, en el extremo 3', que codifica una proteína de 27 (p27) kDa y dos isoformas de ésta que comparten su extremo C-terminal y que poseen tamaños de 25 (p25) y 22.5 (p22.5) kDa (Huang *et al.*, 2000; Liang *et al.*, 2002b; Koh *et al.*, 2006; Zhou *et al.*, 2006). Se ha demostrado que la proteína p23, cuyo gen se encuentra imbricado dentro de la ORF1, está presente en tejido infectado y parece ser indispensable para la replicación del HCRSV en su huésped natural (*Hibiscus cannabinus*), pero no en el huésped experimental *Chenopodium quinoa* (Liang *et al.*, 2002b). Experimentos adicionales han puesto de manifiesto que esta proteína no une ácidos nucleicos *in vitro* ni actúa como un supresor del silenciamiento génico postranscripcional (*post-transcriptional gene silencing*, PTGS), y se ha sugerido que esta proteína podría estar interaccionando con factores específicos del huésped, o bien podría actuar como un activador en *cis* o *trans* que regularía de forma directa o indirecta la expresión de genes implicados en replicación viral, e incluso podría interaccionar directamente con el complejo replicativo del virus (Liang *et al.*, 2002b). Por otro lado, mediante experimentos de genética reversa se ha llegado a la conclusión de que la p27 está implicada en la exacerbación de los síntomas inducidos por el virus en su huésped natural, mientras que la p25 probablemente esté asociada al movimiento sistémico del patógeno (Zhou *et al.*, 2006).

Además de estas ORFs adicionales en el HCRSV, se ha postulado la existencia de una ORF de función desconocida en el CPMoV (You *et al.*, 1995). Cabe señalar también que la ORF2 del CarMV posee un codón de terminación débil que podría permitir la lectura hasta un codón de terminación fuerte situado aguas abajo en el gRNA. El proceso de doble lectura a través daría lugar a una proteína de 98 kDa cuya producción se ha observado en condiciones especiales de traducción *in vitro*, y, aparentemente, en protoplastos aunque su relevancia *in vivo* no ha sido demostrada (Harbison *et al.*, 1985). Además, en el caso del MNSV, las ORFs 3 y 4 no son solapantes, a diferencia de lo que ocurre en el resto de carmovirus, y están separadas por un codón de terminación débil cuya lectura a través daría lugar a una proteína de 14 kDa, aunque esta proteína no ha sido detectada *in vivo* (Russo *et al.*, 1994).

2.1.2. Mecanismos de expresión génica

La regulación de la expresión génica de los carmovirus se produce tanto a nivel transcripcional, mediante la generación de sgRNAs, como a nivel post-transcripcional, mediante la utilización de distintos mecanismos de traducción tanto canónicos como no-canónicos.

Los carmovirus producen dos sgRNAs, 3' co-terminales con respecto al gRNA. El mayor de ellos (sgRNA1), presenta un tamaño aproximado de 1.7 kb, es bicistrónico y está implicado en la traducción de las dos proteínas de movimiento, mientras que el otro (sgRNA2), de tamaño aproximado 1.5 kb, es responsable de la traducción de la CP (revisado en Lommel *et al.*, 2005). Estos dos sgRNAs han sido detectados en plantas infectadas por el CarMV (Carrington y Morris, 1985), el TCV (Carrington *et al.*, 1989), el MNSV (Riviere y Rochon, 1990), el CPMoV (Kim y Bozarth, 1992), el SCV (Weng y Xiong, 1997), el HCRSV (Huang *et al.*, 2000) y el PFBV (Rico y Hernández, 2006). En general, sólo se encapsida el gRNA, aunque en algunos casos también los sgRNAs pueden hacerlo con una relativa eficiencia, como es el caso del CarMV (Guilley *et al.*, 1985; Díez *et al.*, 1999), del SCV (Weng y Xiong, 1997) y del HCRSV (Huang *et al.*, 2000).

Ensayos de traducción *in vitro* han permitido confirmar la producción de todas las proteínas codificadas tanto por los gRNAs como por los sgRNAs del CarMV, del TCV y del HCRSV (Carrington y Morris, 1985; Harbison *et al.*, 1985; Carrington *et al.*, 1987; Hacker *et al.*, 1992; Huang *et al.*, 2000).

Los mecanismos de traducción empleados por los virus de este género para la expresión de sus distintas ORFs son diversos. Como ya se ha indicado antes, las dos primeras ORFs de carmovirus se traducen a partir del gRNA, la segunda de ellas mediante un mecanismo de lectura a través de un codón de terminación débil. Normalmente para que este mecanismo sea eficiente, es necesaria la secuencia consenso AA(A/G)-UAG-G(G/U)(G/A) flanqueando al codón de terminación de la ORF1 (UAG normalmente), una secuencia que se encuentra conservada en muchos virus de RNA de plantas que siguen esta estrategia de traducción (Dugeon *et al.*, 1999). La expresión de las proteínas implicadas en movimiento se lleva a cabo a partir del sgRNA1 probablemente mediante un proceso de escape al rastreo ribosomal (*leaky-scanning*), aunque este punto no ha sido confirmado (Russo *et al.*, 1994). Por último, la

CP se traduce a partir del sgRNA2, por un mecanismo convencional de rastreo ribosomal. No obstante cabe señalar que en el caso del HCRSV se ha identificado un sitio de entrada interna de ribosomas (*internal ribosome entry site*, IRES) para la traducción de la CP que parece ser funcional desde el gRNA y los sgRNAs (Koh *et al.*, 2003, 2006).

2.2. Género Panicovirus

El género *Panicovirus* está constituido por una especie definitiva, el virus del mosaico del *Panicum* (*Panicum mosaic virus*, PMV) y una especie provisional (Tabla 2).

Tabla 2. Especies pertenecientes al género *Panicovirus*

Especies definitivas	
<i>Virus del mosaico del Panicum</i>	PMV
Especies provisionales	
<i>Molinia streak virus</i>	MoSV

Los primeros datos del PMV se recogen en 1953. La gama de huéspedes naturales de este virus está restringida a especies gramíneas de la tribu *Paniceae*, dentro de la familia de las *Poaceae* (Lommel *et al.*, 2005), siendo conocidos los daños causados por el virus en *Panicum virgatum*, *Stenotaphrum secundatum* y *Eremochloa aphiuroides*.

La gama de huéspedes experimentales se amplía a un mayor número de plantas gramíneas. El virus se transmite por inoculación mecánica y por la replantación en suelos ya infectados por el virus. El PMV puede causar infecciones sólo, junto a un virus satélite (*satellite Panicum mosaic virus*, SPMV) y/o junto a satRNAs (Desvoyes y Scholthof, 2000; Omarov *et al.*, 2005; Makino *et al.*, 2006; Qi *et al.*, 2008; Qi y Scholthof, 2008). El virus típicamente induce un mosaico sistémico en la planta infectada, aunque los síntomas más severos, que incluyen un moteado clorótico, atrofiamiento de la planta y pérdida en rendimiento del cultivo, aparecen en las infecciones mixtas del virus y de su virus satélite (Fig. 5).

El genoma del PMV se encapsida en partículas isométricas de 30 nm de diámetro que poseen una superficie estructural regular vista al microscopio electrónico (Lommel *et al.*, 2005).

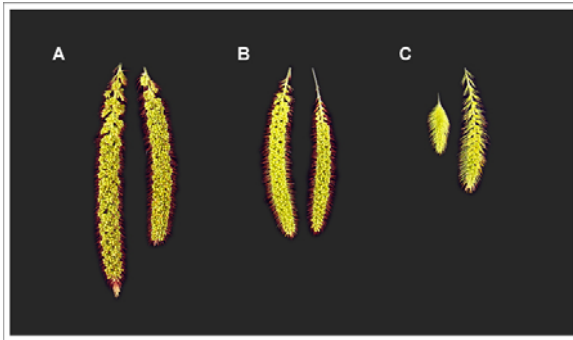


Fig. 5. (A) Panículas de mijo sanas. (B) Panículas de mijo infectadas con el PMV. (C) Panículas de mijo coinfectadas con el PMV y su virus satélite.

2.2.1. Organización genómica

El PMV posee un genoma de ssRNA+ que está constituido por 4326 nt y contiene 6 ORFs (Turina *et al.*, 1998, 2000) (Fig. 6). La ORF más próxima al extremo 5', ORF1, codifica una proteína de 48 kDa y termina en un codón ámbar cuya lectura a través permite continuar la traducción hasta el codón de terminación de la ORF2 dando lugar a una proteína de 112 kDa. Ambas proteínas son necesarias para la replicación del virus. Se ha detectado un único sgRNA de 1475 nt en plantas y protoplastos infectados con el PMV, a partir del cual se traducen el resto de ORFs (Turina *et al.*, 1998, 2000). La ORF3 codifica una proteína de 8 kDa y la ORF4 codifica una proteína de 6.6 kDa. Ambas proteínas están implicadas en el movimiento del virus. Las ORFs 5 y 6, cercanas al extremo 3' e imbricadas una dentro de la otra, codifican la CP, de 26 kDa, y una proteína de 15 kDa, respectivamente. Esta última no presenta homología con ninguna otra de la base de datos y parece estar implicada también en el movimiento del virus (Turina *et al.*, 2000).

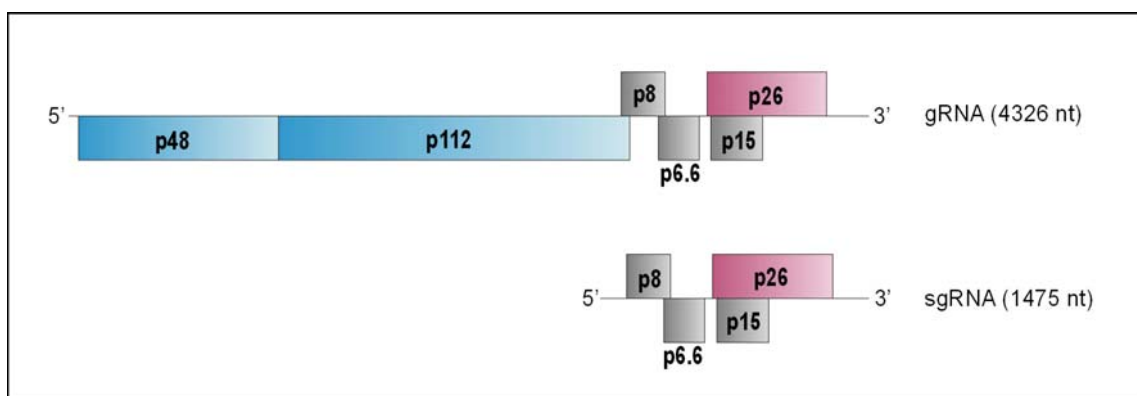


Fig. 6. Representación de la organización genómica del PMV. Las distintas ORFs se representan con recuadros de diferente color dependiendo de la función de las proteínas que codifican. Las ORFs 1 y 2 codifican proteínas implicadas en replicación y se expresan a partir del gRNA. Las ORFs 3, 4, 5 y 6 son expresadas a partir del sgRNA. Las ORFs 3, 4 y 6 codifican proteínas implicadas en movimiento y la ORF5 codifica la CP viral.

2.2.2. Mecanismos de expresión génica

Al igual que ocurre con los carmovirus, la regulación de la expresión génica del PMV se produce tanto a nivel transcripcional, mediante la producción de un sgRNA, como a nivel traduccional.

Las proteínas p48 y p112, implicadas en replicación, son traducidas a partir del gRNA por un mecanismo de lectura a través similar al que emplean otros miembros de la familia *Tombusviridae* (Turina *et al.*, 1998, 2000). El resto de ORFs son traducidas a partir del sgRNA, 3' co-terminal con respecto al gRNA. La expresión de al menos cuatro proteínas a partir de un único sgRNA es algo bastante inusual en virus de plantas (y, en general, en eucariotas). Las estrategias de traducción empleadas por este sgRNA para la expresión de todos sus genes deben incluir una gran variedad de mecanismos que no han sido totalmente determinados (Turina *et al.*, 1998, 2000). Estudios de traducción *in vitro* han permitido confirmar la producción de las proteínas p8, p6.6, CP y p15 a partir del sgRNA. Los niveles de acumulación de las proteínas p6.6 y p15 son bajos, en el primer caso posiblemente debido a la baja eficiencia de su codón de inicio que no es el canónico AUG, sino GUG. Ambas proteínas, p6.6 y p15 podrían traducirse por un mecanismo de escape al proceso de rastreo de las ORFs (p8) y (CP), respectivamente (Turina *et al.*, 2000). Además de éstas, una ORF adicional podría traducirse, por un mecanismo de corrimiento de - 1 nt en la pauta de lectura durante la traducción de la proteína p8, dando lugar a un polipéptido de 14.6 kDa. La presencia de un heptanucleótido (5'-CAAUUUC-3') característico de este tipo de mecanismo (XXXYYYZ, siendo X cualquier base, Y normalmente A o U y Z, cualquier base a excepción de G), inmediatamente aguas arriba del codón de parada de la ORF (p8) facilitaría el proceso de corrimiento y, por tanto, apoyaría la idea de que esta proteína pudiera sintetizarse, aunque experimentos de traducción *in vitro* y bioensayos con mutantes no han permitido corroborar su producción (Turina *et al.*, 1998, 2000). Por último, la traducción de la CP y de la p15 podría estar mediada por la presencia de un IRES, aunque este punto aún no ha sido abordado experimentalmente (Turina *et al.*, 1998, 2000; Batten *et al.*, 2006a).

2.3. Características de las proteínas implicadas en replicación

Como se ha indicado anteriormente, un rasgo común de los miembros de la familia *Tombusviridae* es la implicación de dos proteínas, productos de la traducción de las ORFs 1 y 2, en la replicación del genoma viral y en la transcripción de sgRNAs (excepto en el género *Machlomovirus* donde las proteínas presumiblemente implicadas en estos procesos son el producto de traducción de las ORFs 2 y 3). Mientras que la proteína codificada por la ORF2 corresponde a una RdRp muy conservada, la función de la proteína codificada por la ORF1 no está del todo clara. Estudios realizados con las proteínas correspondientes de los géneros *Tombusvirus* y *Dianthovirus* han mostrado por un lado, sus propiedades de unión a RNA, como es el caso de las proteínas p33 y p92 del virus del enanismo arbustivo del tomate (*Tomato bushy stunt virus*, TBSV, *Tombusvirus*) (Rajendran y Nagy, 2003) y, por otro, su direccionamiento a membranas específicas de las células huésped, pudiendo inducir una proliferación de las mismas. Este direccionamiento a membranas se ha descrito para la proteína p33 del virus de las manchas anulares del *Cymbidium* (*Cymbidium ringspot virus*, CyRSV, *Tombusvirus*) y del virus de la necrosis del pepino (*Cucumber necrosis virus*, CNV, *Tombusvirus*) que están asociadas a peroxisomas (Navarro *et al.*, 2004; Panavás *et al.*, 2005), para las proteínas p36 y p95 del virus de las manchas en anillo del clavel italiano (*Carnation italian ringspot virus*, CIRV, *Tombusvirus*), que se localizan en mitocondrias (Weber-Lofti *et al.*, 2002), y para las proteínas p27 y p88 del RCNMV, que aparecen asociadas al retículo endoplasmático (Turner *et al.*, 2004). Además, la función de este tipo de proteínas parece estar regulada por fosforilación, al menos en el caso de la proteína p33 del CNV (Shapka *et al.*, 2005; Stork *et al.*, 2005). La interacción entre la RdRp y el producto de la ORF1 ha sido demostrada *in vitro* e *in vivo* para el TBSV, una interacción que parece ser esencial para la replicación del virus (Rajendran y Nagy, 2004, 2006; Panavas *et al.*, 2005).

En el caso de los carmovirus, las proteínas p27-29 y p85-88 son necesarias y suficientes para la replicación viral (Hacker *et al.*, 1992; White *et al.*, 1995), y se acumulan *in vivo* en una proporción relativa de aproximadamente 20:1 (White *et al.*, 1995). La p85-p88 contiene los ocho motivos característicos de las RdRps virales de los componentes del supergrupo II, incluyendo el triplete GDD altamente conservado. Estas proteínas sólo contienen dos de los siete motivos de helicasa de ácidos nucleicos encontrados en las RdRps de los supergrupos I y

III (Koonin y Dolja, 1991; Koonin y Dolja, 1993), lo que sugiere que esta función, requerida para la replicación viral, es desempeñada por factores del huésped. El estudio de la actividad enzimática de la RdRp ha sido complicado debido a que normalmente se encuentra asociada a estructuras membranosas en las células y, por tanto, su aislamiento resulta a menudo problemático. A pesar de esto, se han desarrollado métodos que permiten la purificación del complejo replicativo a partir de tejido infectado con el TCV (Song y Simon, 1994) y el HCRSV (Wang y Wong, 2004).

En el género *Panicovirus*, las proteínas codificadas a partir de las ORFs 1 y 2 del PMV son de tamaño algo superior a las codificadas por otros miembros de la familia *Tombusviridae*, concretamente 48 y 112 kDa, respectivamente. Ambas proteínas son requeridas para la replicación del virus y de su virus satélite (Batten *et al.*, 2006b). Experimentos de purificación parcial de la RdRp del PMV indican que tanto la p48 como p112 están asociadas a estructuras membranosas en las células (Batten *et al.*, 2006b). La proteína p112, además de contener el motivo característico GDD, presenta un dominio (CD) común con la p48 (localizado entre los aminoácidos 306 y 405) que está conservado en miembros de la familia *Tombusviridae*, particularmente en los géneros *Carmovirus*, *Necrovirus* y *Tombusvirus*. Aminoácidos hidrofóbicos de este dominio parecen ser importantes en la replicación, posiblemente por dirigir el complejo de la replicasa a membranas celulares. Además, otros residuos del dominio parecen contribuir a la invasión sistémica de la planta huésped por el virus (Batten *et al.*, 2006b).

2.4. Características de las proteínas implicadas en movimiento

La dispersión de un virus en plantas desde el sitio inicial de la infección está mediada por las MPs. Estas proteínas desempeñan un papel fundamental en el movimiento intra e intercelular. Este último, también llamado movimiento célula a célula, se realiza a través de los plasmodesmos, canales membranosos que interconectan las células vegetales. Las MPs se pueden clasificar en: i) miembros de la superfamilia de las 30K que engloba a 18 géneros de virus de plantas entre los que se encuentran *Bromovirus*, *Alfamovirus* o *Ilarvirus*, y que tiene como representante a la MP del TMV, ii) las pequeñas MPs de *Geminivirus* o *Carmovirus*, menores de 12 kDa (codificadas por el denominado bloque de dos genes -*doble gene block*,

DGB- en carmovirus), iii) el grupo codificado por el bloque de tres genes (*triple gene block*, TGB) de *Potexvirus*, *Carlavirus*, *Allexvirus*, *Foveavirus*, *Hordevirus*, *Benyvirus*, *Pomovirus* y *Pecluvirus* y, iv) las grandes MPs de *Tymovirus*, de 69-85 kDa. A pesar de la variabilidad entre las distintas familias en la secuencia, número y tamaño de este tipo de proteínas, todas ellas presentan características funcionales comunes: i) son capaces de unir ácidos nucleicos y de conducir al virus desde el citoplasma de la célula infectada al de la célula vecina hasta alcanzar el sistema vascular (Citovsky *et al.*, 1990; Donald *et al.*, 1997; Vaquero *et al.*, 1997; Fujita *et al.*, 1998; Herranz y Pallás, 2004; Herranz *et al.*, 2005; Sánchez-Navarro *et al.*, 2006), ii) interaccionan con los plasmodesmos celulares y actúan aumentando el tamaño de exclusión molecular en algunos casos (Wolf *et al.*, 1989; Nouery *et al.*, 1994; Ding *et al.*, 1995; Crawford y Zambryski, 2001; Roberts *et al.*, 2001) y, iii) pueden interaccionar con elementos del citoesqueleto y con microtúbulos (Karasev *et al.*, 1992; Heinlein *et al.*, 1995; McLean *et al.*, 1995; Boyko *et al.*, 2000).

Como se ha indicado anteriormente, las proteínas implicadas en movimiento en el género *Carmovirus* corresponden a dos pequeños polipéptidos de 7-8 kDa y 8-12 kDa. Bioensayos con mutantes del TCV demostraron que las dos proteínas, p8 y p9, eran requeridas para el movimiento viral (Hacker *et al.*, 1992; Li *et al.*, 1998). La p8 posee propiedades de unión al RNA, y estudios de mutagénesis dirigida han mostrado que un solo cambio de aminoácido afecta a su capacidad de unión al ácido nucleico y a la virulencia inducida por el patógeno en el huésped *Arabidopsis thaliana* (Wobbe *et al.*, 1998). Las proteínas homólogas del CarMV (p7 y p9) han sido también extensamente estudiadas. Se ha descrito que la p7 posee propiedades de unión a RNA *in vitro*, y se compone de tres dominios: i) central, con una estructura de α -hélice y responsable de la unión al RNA, ii) C-terminal, que se pliega en forma de hoja β muy estable, y, iii) N-terminal sin una estructura concreta (Marcos *et al.*, 1999; Vilar *et al.*, 2001). La proteína p9 contiene dos dominios transmembrana y se ha observado *in vitro* su inserción a membranas del retículo endoplásmico adoptando una topología que dispone a las porciones N- y C-terminales en la región citosólica celular (Vilar *et al.*, 2002; Saurí *et al.*, 2005). Las proteínas de movimiento del MNSV también parecen presentar características similares ya que por un lado se ha descrito que la proteína p7A posee propiedades de unión a RNA (Navarro *et*

al., 2006) y, por otro, se ha observado *in vitro* la inserción en membranas del retículo endoplásmico de la proteína p7B (Martínez-Gil *et al.*, 2007).

Con la información obtenida acerca de las MPs del CarMV y del MNSV se ha propuesto un modelo topológico y funcional para las mismas. En este modelo, la p7 (o p7A en el MNSV) se une al RNA viral y esta unión induce un cambio conformacional en esta proteína soluble que hace que su dominio C-terminal de hoja β pueda a su vez interactuar con el dominio C-terminal expuesto en el citoplasma de la proteína p9 (o p7B en el MNSV), formando así un complejo ternario en la membrana (Vilar *et al.*, 2002). Este modelo reproduce la topología previamente descrita para la MP de 30 kDa del TMV (Brill *et al.*, 2000). Cabe señalar que la proteína p7B del MNSV, además de participar directamente en movimiento, podría actuar también como un supresor débil del PTGS (Genovés *et al.*, 2006).

Con respecto a la localización subcelular de estas proteínas, estudios realizados mediante la fusión de la proteína verde fluorescente (GFP) a las proteínas p8 y p9 del TCV, muestran una localización nuclear para la p8 y citoplasmática para la p9 en protoplastos y plantas de *A. thaliana* (Cohen *et al.*, 2000). El significado biológico de la localización nuclear de la p8, que está mediado por dos motivos que no están presentes en proteínas homólogas, está por determinar. Para el CarMV, ha sido descrito que la proteína p7 se localiza principalmente en el citosol, aunque en estadios tardíos de la infección viral también puede ser detectada en las fracciones correspondientes a las paredes celulares. Este resultado es consistente con el modelo anteriormente propuesto (García-Castillo *et al.*, 2003).

En el caso del PMV, se ha estudiado la localización subcelular de la proteína p8, una de las tres proteínas que parecen mediar el movimiento del virus, observándose que está asociada preferentemente con la pared celular (Turina *et al.*, 1998, 2000), lo que resulta coherente con su función de transporte.

2.5. Características de la proteína de cubierta

Además de su papel estructural, la CP de los virus de plantas puede estar implicada en procesos tales como: i) la replicación viral (Mahajan *et al.*, 1996; Qiu y Scholthof, 2001; Aparicio *et al.*, 2003; Asurmendi *et al.*, 2004; Bol, 2005; Malik *et al.*, 2005), ii) determinación del hospedador, modulación de síntomas o modo de transmisión (Rao y Grantham, 1996; Wang y

Simon, 1999; Soto *et al.*, 2005), iii) movimiento célula a célula y sistémico (Taliensky y García-Arenal, 1995; Ryabov *et al.*, 1999; Sareila *et al.*, 2004; Omarov *et al.*, 2005; Rao y Cooper, 2006) o, iv) supresión del PTGS (Liu *et al.*, 2004; Lu *et al.*, 2004; Cañizares *et al.*, 2008).

En el caso del género *Carmovirus*, los viriones del TCV son requeridos para el movimiento a larga distancia pero el movimiento célula a célula puede tener lugar en ausencia de CP en algunos huéspedes, aunque no en otros (Heaton *et al.*, 1991; Hacker *et al.*, 1992). Este tipo de proteínas poseen sitios de unión a Ca^{++} , y se ha observado que mutaciones en aminoácidos de la CP del TCV implicados en dicha unión dan lugar a una disminución en la capacidad del virus para moverse célula a célula (Laakso y Heaton, 1993). En concordancia con estos resultados, un estudio posterior mostró que otras mutaciones puntuales en el dominio de unión a Ca^{++} de la CP del TCV, además de modificar la capacidad del virus para unir iones, da lugar a alteraciones en la estructura secundaria de la proteína o en la conformación de la partícula, lo que afecta al movimiento (célula a célula o sistémico) y a la aparición de los síntomas (Lin y Heaton, 1999).

Estudios realizados con el CarMV muestran que la CP se localiza principalmente en el citosol, aunque en estadios tempranos de la infección también está presente en fracciones membranosas y en la pared celular (García-Castillo *et al.*, 2003).

Otra función de la CP de carmovirus parece ser la supresión del PTGS, como se ha puesto de manifiesto para la proteína del TCV, del HCRSV, del MNSV y del PFBV (Qu *et al.*, 2003; Thomas *et al.*, 2003; Genovés *et al.*, 2006; Meng *et al.*, 2006; Martínez-Turiño y Hernández, 2009).

En lo que respecta al género *Panicovirus*, la CP del PMV podría participar de la replicación viral, ya que estudios en protoplastos indican que la ausencia de esta proteína repercute negativamente en los niveles de acumulación del gRNA. Sin embargo, no es descartable que este efecto sea consecuencia del papel protector de la CP, pues la encapsidación evitará la degradación del gRNA. La CP parece estar también implicada en el movimiento del virus, aunque esta implicación podría ser directa o indirecta, ya que parece regular los niveles de acumulación de la p8, una proteína que, como ya se ha indicado anteriormente, asiste al virus en su transporte (Turina *et al.*, 2000). La CP del PMV se localiza preferentemente en la pared celular, fracciones membranosas y citosol (Turina *et al.*, 2000).

3. TRADUCCIÓN EN EUCARIOTAS

Aunque durante mucho tiempo se pensó que la regulación de la expresión génica en organismos eucariotas se producía fundamentalmente a nivel transcripcional, cada vez existen más datos que establecen firmemente la importancia de procesos post-transcripcionales en dicha regulación. Entre ellos ocupa un lugar destacado el proceso de traducción, cuyo estudio ha experimentado un resurgimiento en los últimos años con el descubrimiento de nuevos mecanismos de control de la síntesis proteica (Gale *et al.*, 2000; Merrick, 2004).

La inmensa mayoría de los mRNAs eucarióticos son monocistrónicos, es decir, dirigen la traducción de una sola proteína, y contienen, además de la secuencia codificante, UTRs en 5' y 3', con elementos reguladores cuya identificación y caracterización está permitiendo entender mejor el control de la traducción. Asimismo, los mRNAs eucariotas presentan un residuo 7-metilguanosa (m^7GpppX), denominado cap, en el extremo 5' y una cola poliA en el extremo 3', y ambas estructuras son determinantes importantes para la eficiencia traduccional. La maquinaria necesaria en el proceso de traducción incluye ribosomas, tRNAs y proteínas no ribosomales llamadas eIFs, eEFs y eRFs, que son factores de iniciación, elongación y terminación, respectivamente (Gallie, 1996). En plantas, el proceso de traducción es muy similar al de otros organismos eucariotas (Fütterer y Hohn, 1996), aunque se han descrito diferencias significativas en algunas partes de la maquinaria traduccional, especialmente en la diversidad de factores de iniciación de la traducción (Browning, 2004).

El proceso de traducción está sumamente regulado y coordinado con el resto de procesos celulares, sufriendo alteraciones en ciertas fases del ciclo celular o bajo determinadas condiciones de estrés (Duncan y Hershey, 1989; Brostrom y Brostrom, 1998; Holcik y Sonenberg, 2005).

3.1. Iniciación de la traducción en eucariotas

La traducción puede ser regulada en cada una de las tres etapas de que consta el proceso: iniciación, elongación y terminación. Sin embargo, el paso limitante en la traducción de la mayoría de los mRNAs es la etapa de iniciación.

La iniciación de la traducción de mRNAs eucariotas se ajusta habitualmente al modelo de rastreo o barrido (*scanning*). En esta etapa el complejo de preiniciación 43S (formado por los

factores de iniciación eIF3, eIF1A, eIF2-GTP, el iniciador Met-tRNA y la subunidad pequeña del ribosoma 40S) junto al complejo proteico eIF4F se une al extremo 5' del mensajero (que contiene la estructura cap) (Fig. 7). La unión del complejo eIF4F al mRNA depende fundamentalmente del componente eIF4E (que forma parte del complejo), ya que es el que reconoce y facilita la unión a la estructura cap en 5' (Gale *et al.*, 2000).

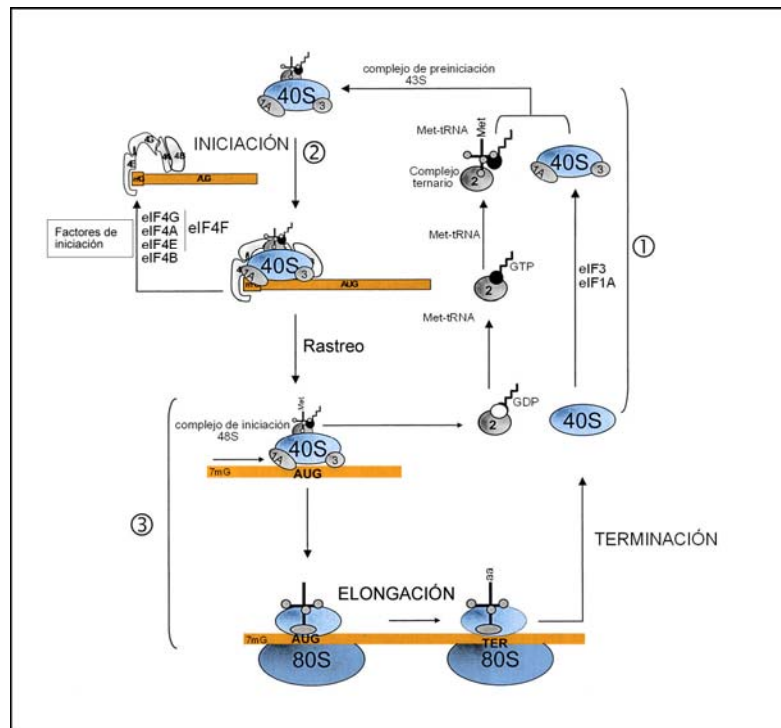


Fig. 7. Ilustración esquemática de la traducción de un mRNA eucariota. 40S y 80S son la subunidad pequeña del ribosoma y el ribosoma de elongación, respectivamente. **1.-** Formación del complejo ternario (formado por el iniciador Met-tRNA, GTP y eIF2) y unión a la subunidad 40S (asociada a factores de iniciación) para constituir el complejo de preiniciación 43S. **2.-** Ensamblaje de componentes en el mRNA. **3.-** Rastreo del ribosoma hasta alcanzar el primer codón de inicio, normalmente AUG, reciclaje del factor eIF2-GDP y unión de la subunidad ribosomal 60S. TER es el codón de terminación (basado en Gale *et al.*, 2000).

3.1.1. Modelo del bucle cerrado en mRNAs eucariotas

Como se ha indicado anteriormente, tanto la estructura cap en 5' como la cola poliA en 3' son necesarias para que se produzca una traducción eficiente de los mRNAs celulares. Concretamente, la interacción del factor eIF4E con la estructura cap y la interacción de la proteína de unión a poliA (PABP) con la cola poliA facilita el inicio de la traducción. Mediante la gran proteína adaptadora eIF4G (o eIFiso4G), que une eIF4E (o eIFiso4E) y PABP

simultáneamente, el mRNA adopta una conformación de bucle cerrado (Jacobson, 1996; Sachs *et al.*, 1997; Gallie, 1998). Esta conformación se cree que aumenta la afinidad entre los factores de iniciación, el ribosoma y el mRNA para iniciar el rastreo, y probablemente facilita el reciclaje de la maquinaria traduccional para iniciar nuevas rondas de traducción sobre el mismo mRNA (Pestova *et al.*, 2001) (Fig. 8).

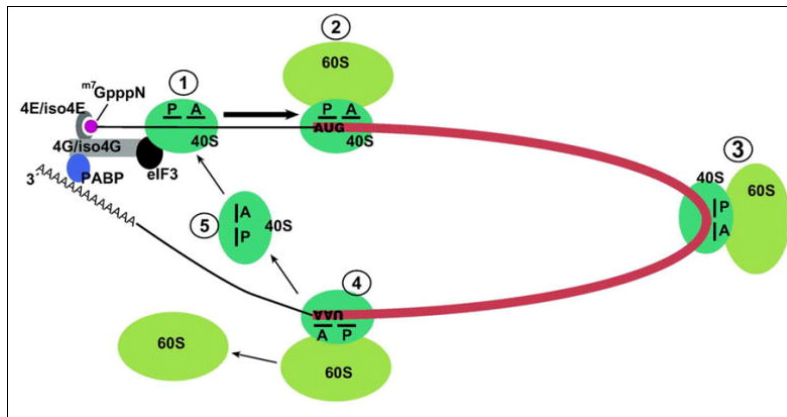


Fig. 8. Modelo del bucle cerrado del mRNA durante la traducción. 1.- Ensamblaje de los componentes de la maquinaria traduccional en el mRNA. 2.- Rastreo. 3.- Elongación. 4.- Terminación. 5.- Reciclaje de la subunidad pequeña del ribosoma 40S (basado en Dreher y Miller, 2006).

3.1.2. Rastreo o barrido (*scanning*)

El complejo de preiniciación comienza con el rastreo de la secuencia líder en dirección lineal 5'-3' hasta encontrar un codón apropiado, normalmente AUG, a partir del cual iniciará la traducción. Por medio del factor eIF5, se activa la actividad GTPasa del eIF2-GTP, liberando eIF2-GDP de la subunidad 40S que así puede unirse a la subunidad 60S. Como resultado se forma el ribosoma 80S que contiene el iniciador Met-tRNA (Fig. 7).

3.1.3. Codón de inicio

En eucariotas, en más del 90% de los casos, el codón AUG más próximo al extremo 5' del mRNA es el que sirve de iniciador. La eficiencia de reconocimiento de dicho codón está modulada por el contexto que lo flanquea (Kozak, 1986, 2002). En plantas, el contexto óptimo es (A/G)aaAUGGc en dicotiledóneas y (A/G)ccAUGGc en monocotiledóneas (Joshi *et al.*, 1997; Lukaszewicz *et al.*, 2000). Las posiciones cruciales en ambos casos son una purina en posición

-3 y una guanina en posición +4 (donde la A del codón de inicio representa la posición +1), al igual que ocurre en otros eucariotas (Kozak, 2002).

A pesar de que el triplete normalmente reconocido por los ribosomas para el inicio de la síntesis proteica es el AUG, en algunos casos el codón de inicio puede ser un triplete no-AUG. Se trata de codones débiles de iniciación que difieren en una posición respecto al canónico AUG y han sido descritos tanto en mensajeros virales como celulares (Kozak, 1989; Peabody, 1989; Gordon *et al.*, 1992). Concretamente en plantas, el uso de este tipo de codones se describió por primera vez en el virus del mosaico de la coliflor (*Cauliflower mosaic virus*, CaMV, *Caulimovirus*) (Schultze *et al.*, 1990) y, posteriormente, en el virus baciliforme del tungro del arroz (*Rice tungro bacilliform virus*, RTBV, *Tungrovirus*) (Rothnie *et al.*, 1994).

3.2. Elongación

Esta etapa incluye todas las reacciones comprendidas entre la formación del primer enlace peptídico y la adición del último aminoácido a la cadena polipeptídica (Fig. 7). Es la etapa más rápida de la biosíntesis de las proteínas y requiere la presencia del complejo de iniciación, el aminoacil-tRNA codificado por el siguiente codón del mRNA, los factores de elongación eEF1 y eEF2 y GTP (Riis *et al.*, 1990). Esta etapa se inicia cuando un segundo aminoacil-tRNA, cuyo anticodón es complementario al codón situado a continuación del iniciador, ocupa el sitio A del ribosoma, que se encuentra libre. El siguiente paso es la formación de un enlace peptídico entre el aminoácido que ocupa el sitio P (generalmente metionina) y el nuevo aminoácido que ocupa el sitio A. La reacción de formación del enlace peptídico está catalizada por el enzima peptidil-transferasa. Como consecuencia de la formación de este enlace peptídico, el segundo tRNA queda unido por un extremo al dipéptido formado y por el otro a su codón complementario. A continuación se produce la translocación del ribosoma, que implica el desplazamiento del ribosoma a lo largo del mRNA en sentido 5' - 3'. Como este desplazamiento es exactamente de tres bases, el primer tRNA abandona el ribosoma, y el peptidil-tRNA, que todavía se mantiene unido a su codón, pasa a ocupar el sitio P, quedando el sitio A libre. En estas condiciones, otro aminoacil-tRNA se puede incorporar al sitio A, de manera que el proceso de elongación de la cadena proteica puede continuar, repitiéndose el ciclo.

3.3. Terminación

Esta etapa tiene lugar cuando el ribosoma llega a un lugar del mRNA donde se encuentra un codón de terminación (UAA, UGA o UAG), que no es reconocido por ningún tRNA, pero sí por factores de terminación de naturaleza proteica que se sitúan en el sitio A y hacen que la peptidil-transferasa separe, por hidrólisis, la cadena polipeptídica del tRNA (Zhouraleva *et al.*, 1995) (Fig. 7).

4. TRADUCCIÓN DE VIRUS DE PLANTAS

Dado el pequeño tamaño de su genoma, la información genética de los virus está extremadamente compactada. Por esta razón, han desarrollado un conjunto de estrategias, a menudo combinadas, para expresar sus múltiples genes. Entre dichas estrategias destaca la utilización de una serie de mecanismos de traducción no-canónicos que les permiten “descomprimir” su condensada información genética y traducir más de un gen a partir de un mismo mRNA, como se describe más adelante (apartado 4.2). Además de posibilitar que todos los genes sean accesibles a la maquinaria de síntesis proteica del hospedador, estos mecanismos les permiten a los RNAs virales una regulación tanto cualitativa (es decir, seleccionar qué ORFs han de traducirse) como cuantitativa (es decir, controlar las cantidades relativas de proteínas sintetizadas a partir de cada ORF) de la expresión génica, un punto crucial para la progresión de una infección viral.

Por otra parte, a diferencia de los mRNAs eucariotas, con una estructura cap en el extremo 5' y una cola poliA en el extremo 3', muchos mRNAs virales carecen de uno o de ambos tipos de estructuras en sus extremos (Fauquet *et al.*, 2005). Este es el caso de los mRNAs de más del 80% de virus de plantas y, pese a ello, parece que también pueden adoptar una conformación de bucle cerrado y acceder eficientemente a la maquinaria de traducción del hospedador.

4.1. Modelo del bucle cerrado en mRNAs virales

Como se ha indicado anteriormente, la mayoría de los virus de plantas tienen un genoma ssRNA+, y aproximadamente sólo un 20% de estos virus contienen una estructura cap en 5' y una cola poliA en 3', mientras que el tipo de estructuras presentes en el resto de virus son

variables. El extremo 5' puede contener un polipéptido denominado VPg (*viral protein genome-linked*) unido covalentemente o un grupo fosfato, y el extremo 3' puede formar una estructura de tipo tRNA (*tRNA-like structure*, TLS) o adoptar una conformación formada por varias estructuras de tipo horquilla. De qué modo estos RNAs no canónicos interactúan con factores de iniciación de la traducción no está claro en muchas ocasiones, pero distintos datos indican que el modelo de bucle cerrado es también aplicable en estos casos, con una variedad de interacciones novedosas que sustituyen los puentes moleculares que se establecen en mRNAs celulares convencionales. A continuación se describen las diferentes posibilidades:

i) mRNAs virales con una estructura cap en el extremo 5', pero que carecen de una cola poliA en el extremo 3'. Más de un tercio de los géneros de virus de plantas se ajustan a este modelo. La cola poliA funciona como estimulador de la traducción al actuar sinérgicamente con la estructura cap en 5' del mRNA. Sin embargo, en los virus que carecen de cola poliA se ha demostrado que la 3' UTR es capaz de desempeñar funciones similares. Algunos ejemplos los encontramos en el virus del mosaico de la alfalfa (*Alfalfa mosaic virus*, AMV, *Alfamovirus*), el virus del mosaico amarillo del nabo (*Turnip yellow mosaic virus*, TYMV, *Tymovirus*) o el TMV (Fig. 9A, 9B y 9C). En el caso del AMV, su 3' UTR contiene más de nueve tripletes del tipo AUG que pueden reclutar algún factor proteico y/o los ribosomas, favoreciendo la traducción (Hann *et al.*, 1997). También se ha propuesto un escenario diferente donde la traducción de los RNAs del AMV podría estar regulada por medio de su CP, al unirse ésta a la 3' UTR y al factor proteico eIF4G (Bol, 2005; Krab *et al.*, 2005). Esta interacción CP-RNA reemplazaría así a la interacción PABP-poliA en una traducción normal (Neeleman *et al.*, 2001, 2004; Krab *et al.*, 2005) (Fig. 9A).

Por su parte, las regiones 3' terminales de los RNAs del TYMV y del TMV adoptan una TLS. La observación de sinergismo entre los extremos de estos RNAs apoya la idea de que son traducidos en un formato de bucle cerrado, aunque las uniones moleculares que lo permiten no están claras (Gallie y Walbot, 1990; Matsuda y Dreher, 2004). En el caso del TYMV, probablemente el TLS aminoacilado junto al factor de iniciación eEF1A podría interactuar con el extremo 5'-cap del mRNA viral, aunque este punto no está inequívocamente demostrado (Dreher *et al.*, 1999; Matsuda y Dreher, 2004) (Fig. 9B). En el caso del TMV, se ha definido una región compuesta por cinco pseudonudos en serie. Dos de ellos, que forman parte del TLS,

están implicados en replicación (Takamatsu *et al.*, 1990), mientras que los tres restantes, que constituyen el dominio denominado UPD (*upstream pseudoknot domain*), podrían ser responsables de incrementar la eficiencia traduccional de un modo similar a la cola poliA (Gallie y Walbot, 1990; Leathers *et al.*, 1993), ya que una proteína de 102 kDa que se une a la región 5' líder del TMV se une también al pseudonudo 3' distal del UPD (Tanguay y Gallie, 1996) (Fig. 9C). Una situación similar se ha descrito para el virus del mosaico del bromo (*Brome mosaic virus*, BMV, *Bromovirus*) (Gallie y Kovayashi, 1994).

ii) mRNAs virales con una cola poliA en 3' y sin estructura cap en 5'. Los mRNAs de una cuarta parte de los géneros de virus de plantas presentan una cola poliA en 3', pero en el extremo 5' están covalentemente unidos a la proteína VPg codificada por el propio virus. En estos casos, la traducción se lleva a cabo por un mecanismo independiente de cap. El virus del grabado del tabaco (*Tobacco etch virus*, TEV, *Potyvirus*) (Fig. 9D) y el virus del mosaico del nabo (*Turnip mosaic virus*, TuMV, *Potyvirus*), ambos de la familia *Potyviridae*, son ejemplos de este tipo, y se ha demostrado que poseen secuencias en la 5' UTR capaces de dirigir el inicio de la traducción aún en ausencia de cap (Basso *et al.*, 1994; Niepel y Gallie, 1999). Estos elementos, que actúan como IRES, reclutan directamente factores de iniciación y/o los ribosomas y se explicarán con más detalle en el apartado 4.2.6. En el caso del TEV se ha definido un elemento IRES que, junto al factor de iniciación eIF4G, podría interactuar con la cola poliA vía PABP, adoptando una conformación de bucle cerrado (Gallie, 2001; Khan *et al.*, 2008) (Fig. 9D). Respecto a la proteína VPg, se ha especulado acerca de su papel en el reclutamiento de factores de traducción, como se ha observado en el caso del TuMV (Khan *et al.*, 2006; Beauchemin *et al.*, 2007), reemplazando así a la estructura cap en una traducción convencional (Thivierge *et al.*, 2005). Sin embargo, su participación en este proceso no está muy clara, ya que esta proteína no es necesaria en la etapa de traducción (Basso *et al.*, 1994; Niepel y Gallie, 1999), y no es descartable que la interacción de ésta con diferentes factores esté involucrada en otras etapas del ciclo vital del virus (Gao *et al.*, 2004; Sato *et al.*, 2005).

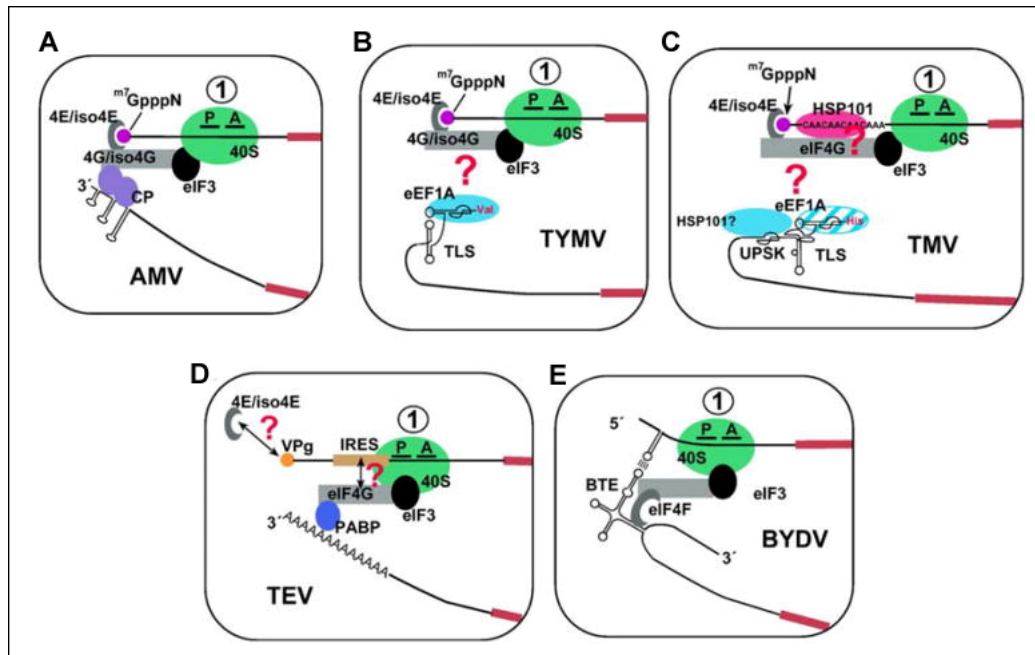


Fig. 9. Modelo del bucle cerrado en mRNAs virales con diferentes estructuras en sus extremos (basado en Dreher y Miller, 2006).

iii) mRNAs virales con los extremos no modificados. Los RNAs virales de las familias *Tombusviridae* y *Luteoviridae* carecen de una estructura cap en el extremo 5' y de una cola poliA en el extremo 3' y, sin embargo, se traducen eficientemente. En estos virus, existen secuencias en la region 3' de su genoma que actúan como estimuladores de la traducción (*translational enhancer*, TE), reclutando factores de iniciación y/o subunidades ribosomales y facilitando la expresión de ORFs situadas en posición 5' proximal. En el caso del virus del enanismo amarillo de la cebada (*Barley yellow dwarf virus*, BYDV, *Luteovirus*), se ha definido una secuencia de 17 nt en la 3' UTR, denominada BTE (*barley translational enhancer*) (Guo *et al.*, 2001), que además está conservada en miembros de los género *Luteovirus*, *Dianthovirus* (Mizumoto *et al.*, 2003; Iwakawa *et al.*, 2007) y *Necrovirus* (Meuleweater *et al.*, 2004; Shen y Miller, 2004). Este TE permite la expresión de genes, por un mecanismo independiente de cap, desde el gRNA y los sgRNAs tanto *in vivo* como *in vitro* (Wang *et al.*, 1997). Probablemente la unión del factor de iniciación eIF4E al BTE y su posterior transferencia a la 5' UTR por medio de interacciones RNA-RNA, permita el inicio de la traducción (Guo *et al.*, 2001; Treder *et al.*, 2008) (Fig. 9E).

Dentro de la familia *Tombusviridae*, se ha identificado una secuencia TE en la 3' UTR del PMV (género *Panicovirus*) que es requerida para una traducción eficiente de sus proteínas al menos *in vitro* (Batten *et al.*, 2006a). En el género *Carmovirus* se han descrito diferentes TEs, en el TCV (Qu y Morris, 2000; Fabian y White, 2004), el HCRSV (Koh *et al.*, 2002) y el MNSV (Truniger *et al.*, 2008). En el TCV se ha definido recientemente un modelo por el cual la subunidad ribosomal 60S se une al TE de la 3' UTR del virus, y el conjunto a su vez interacciona con la 5' UTR a la que previamente se ha unido el complejo de preiniciación 43S (Stupina *et al.*, 2008). En el TBSV (género *Tombusvirus*) se ha identificado una región de aproximadamente 400 nt, compuesta por tres dominios contiguos (RIII, R3.5 y RIV), denominada CITE (*cap independent translation enhancer*). Estudios *in vivo* han demostrado que esta región permite la traducción del virus por un mecanismo independiente de cap (Wu y White, 1999). Los dominios RIII y RIV son esenciales en replicación (Oster *et al.*, 1998), mientras que el dominio R3.5 es esencial en la traducción viral (Oster *et al.*, 1998; Wu y White, 1999). Este dominio adopta una estructura tipo Y, que está muy conservada en el género *Tombusvirus* (Fabian y White, 2004). Se ha demostrado que una de las horquillas terminales de esta estructura interacciona mediante pares Watson-Crick con otra estructura tipo horquilla de la 5' UTR del gRNA, adoptando una conformación de bucle cerrado. Análisis *in silico* permiten predecir interacciones similares en los sgRNAs del TBSV, y en los RNAs de todos los miembros del género *Tombusvirus* y de distintos representantes de otros géneros de la familia *Tombusviridae* (Fabian y White, 2004).

4.2. Mecanismos no convencionales de traducción en RNAs virales

Los mecanismos de traducción no-canónicos que han desarrollado los virus para expresar todos sus genes están representados en la Fig. 10 y se explican con más detalle a continuación. Trabajos recientes indican que ciertos mRNAs celulares también utilizan alguno de estos mecanismos, al menos bajo determinadas condiciones (Calkhoven *et al.*, 2000; Harding *et al.*, 2000; Morris y Geballe, 2000; Rogers *et al.*, 2004; Chappell *et al.*, 2006; Mauro *et al.*, 2007; Lewis y Holcik, 2008; Spriggs *et al.*, 2008). Sin embargo, estas estrategias fueron descritas inicialmente en virus, subrayando el papel fundamental que han desempeñado (y

siguen desempeñando) los sistemas virales en la disección de los aspectos que contribuyen a la traducción no convencional de mRNAs.

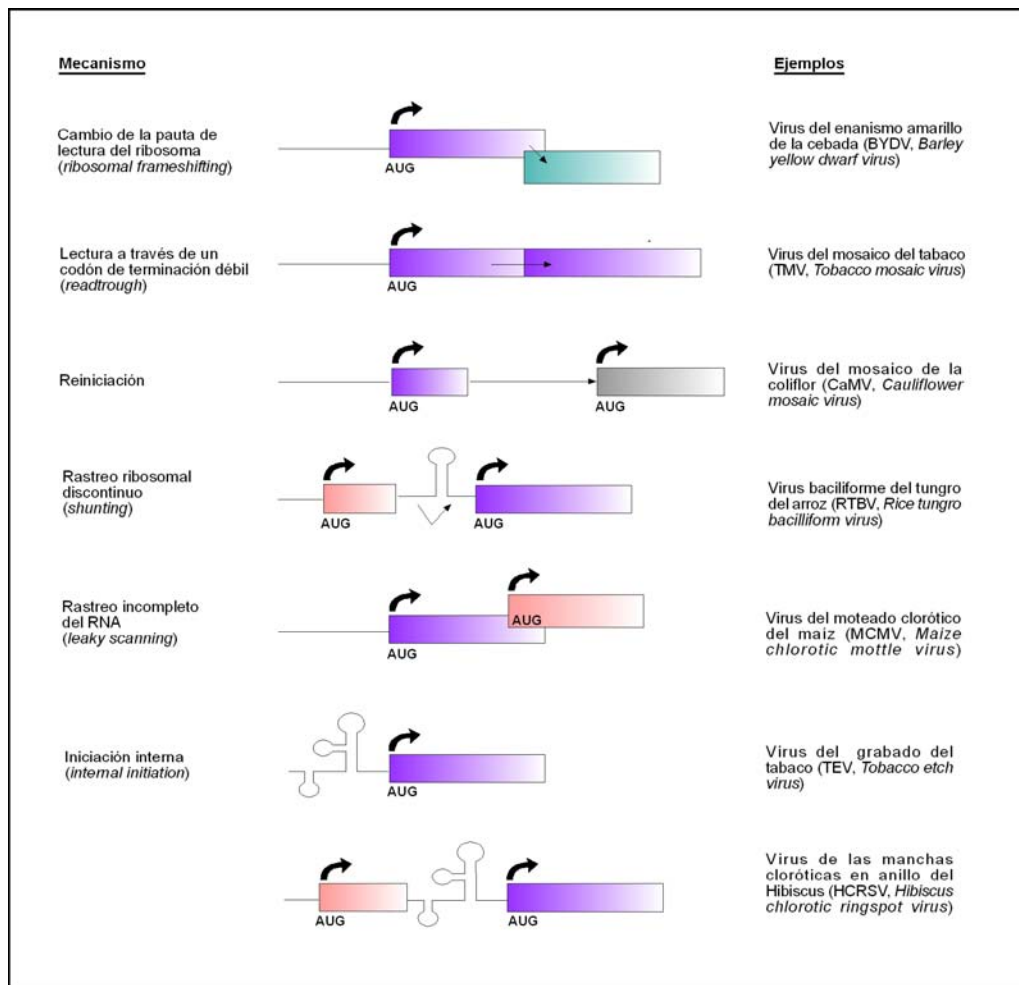


Fig. 10. Mecanismos no convencionales de traducción de RNAs virales. El codón de iniciación (AUG) de la primera ORF se indica en todos los casos y está precedido por una 5' UTR. Las estrategias de rastreo ribosomal discontinuo e iniciación interna dependen de la presencia de estructuras secundarias (simbolizadas por estructuras del tipo horquilla); en el último caso además se incluyen dos representaciones con un IRES situado en la 5' UTR o en una posición interna, respectivamente. Las flechas de trazo grueso indican sitios de inicio de traducción independientes.

4.2.1. Corrimiento en la pauta de lectura del ribosoma (*ribosomal frameshifting*)

En este mecanismo, señales que se encuentran en el mRNA provocan que el ribosoma, que ha iniciado la traducción en una pauta de lectura, cambie a otra durante la fase de elongación y continúe el proceso (Fig. 10) (Giedroc *et al.*, 2000; Baranov *et al.*, 2002). Los requerimientos para que ocurra este cambio son esencialmente dos: una secuencia de siete nucleótidos que

se ajusta al consenso X-XXY-YYZ (siendo X cualquier base, Y normalmente A o U, y Z, cualquier base a excepción de G) y que facilita el deslizamiento del ribosoma, y un elemento próximo altamente estructurado situado entre 5 y 8 nt aguas abajo de donde se produce el deslizamiento. Este elemento puede ser una horquilla simple o un pseudonudo más complejo (Chamorro *et al.*, 1992; Shen y Tinoco, 1995; Ten Dam *et al.*, 1990, 1995; Giedroc *et al.*, 2000; Giedroc y Cornish, 2009). Se postula que este tipo de estructuras podría servir para que el ribosoma realice una pausa facilitando así el desplazamiento, o que constituya un sitio de reconocimiento para la unión de alguna proteína celular involucrada en el proceso. Además, un codón de parada se encuentra con frecuencia inmediatamente aguas abajo de las señales de desplazamiento y parece tener un efecto positivo en el proceso, probablemente porque contribuye a la parada del ribosoma (Brault y Miller, 1992; Di *et al.*, 1993). En algún caso se ha descrito también la contribución de secuencias situadas a larga distancia del sitio de desplazamiento (Mohan *et al.*, 1995; Wang y Miller, 1995). La eficiencia del proceso de corrimiento puede oscilar entre 1-30%, dependiendo de las señales en el mensajero y de los sistemas de ensayo empleados (Dreher y Miller, 2006).

Este tipo de mecanismo ha sido descrito para virus de la familia *Luteoviridae* (Di *et al.*, 1993) y para miembros de los géneros *Sobemovirus* (Makinen *et al.*, 1995) y *Dianthovirus* (Kim y Lommel, 1994; Ryabov *et al.*, 1994). Todos ellos utilizan una estrategia de corrimiento de -1 nt en la pauta de lectura para regular la expresión de la RdRp. También un mecanismo de este tipo, pero menos caracterizado, podría ser el responsable de la traducción de una proteína de 12 kDa que presenta propiedades de unión a RNA en el virus M de la patata (*Potato virus M*, PVM, *Carlavirus*). Esta proteína se traduciría por un corrimiento de -1 nt en la pauta de lectura del gen de la CP, al alcanzarse el codón de parada de dicho gen. Este mecanismo requeriría sólo una secuencia de 4 nt localizada inmediatamente aguas arriba del mencionado codón, que presumiblemente induce la parada (Gramstat *et al.*, 1994). Otros dos casos bastante inusuales donde se ha descrito este mecanismo conciernen al virus de la tristeza de los cítricos (*Citrus tristeza virus*, CTV, *Closterovirus*), donde los ribosomas emplean un mecanismo de corrimiento en la pauta de lectura de +1 nt para expresar su RdRp (Karasev *et al.*, 1995; Cevik *et al.*, 2008), y al TuMV (*Potyvirus*) donde los ribosomas parecen experimentar un corrimiento de +2 nt en la pauta de lectura para expresar una pequeña proteína (PIPO) cuya secuencia

codificante está imbricada dentro del gen P3 (implicado en replicación y/o determinación de síntomas) y está conservada en todos los potyvirus de la base de datos (Chung *et al.*, 2008).

4.2.2. Lectura a través de codones de terminación débiles (*readthrough*)

Este mecanismo ocurre durante la traducción de algunas ORFs de ciertos mRNAs virales, generándose una segunda proteína con una extensión en el extremo C-terminal (*readthrough domain*) debido a la supresión de un codón de parada (Fig. 10). Esta supresión ocurre cuando el anticodón de un aminoacil-tRNA hibrida con el codón de parada de la primera ORF en el sitio ribosomal A, y se produce transferencia peptídica en lugar de inserción de un factor de disociación en dicho sitio. La eficiencia de este mecanismo suele oscilar entre 1-10% y depende del contexto del codón de parada y de la presencia de tRNAs supresores que pueden decodificar el codón de parada correspondiente.

Esta estrategia es empleada por distintos virus para producir la RdRp o extensiones de la CP que están normalmente involucradas en la transmisión por vectores (Tamada y Kusume, 1991; Miller *et al.*, 1995; Zaccomer *et al.*, 1995). El caso más estudiado corresponde a la RdRp del TMV. El gRNA del TMV produce dos proteínas implicadas en replicación de tamaños 126 y 183 kDa, respectivamente. La segunda proteína, la RdRp viral, es expresada por un mecanismo de supresión de la terminación del gen de la p126. La secuencia UAGCARYYA (siendo el UAG el codón de parada del gen que codifica la p126, R una purina e Y una pirimidina) parece clave para que, en algunas rondas de traducción, los ribosomas continúen el proceso al alcanzar el codón de parada del gen de la p126 (Skuzeski *et al.*, 1991; Namy *et al.*, 2001). También, los miembros de la mayoría de géneros de la familia *Tombusviridae* utilizan este mecanismo para regular la expresión de la RdRp y, además, en el género *Machlomovirus* este mecanismo puede ser empleado para extender el gen de la MP (comentado en el apartado 2 de esta tesis). Componentes de la familia *Luteoviridae*, o de los géneros *Furovirus* (no asignado a familia), *Benyvirus* (no asignado a familia) y *Pomovirus* (no asignado a familia) presentan un codón de parada débil en el gen de la CP, de modo que puede generarse una segunda proteína correspondiente a la CP con una extensión en el extremo C-terminal. La CP de miembros de la familia *Luteoviridae* por sí sola es capaz de formar viriones infecciosos, pero la proteína con la extensión C-terminal es necesaria para la transmisión del virus por áfidos

(Brault *et al.*, 1995, 2005), de forma similar a lo que ocurre con el virus mop-top de la patata (*Potato mop-top virus*, PMTV, *Pomovirus*), que también necesita la extensión C-terminal de la CP para ser transmitido por hongos (Reavy *et al.*, 1998). Sin embargo, en el virus de las venas necróticas y amarillas de la remolacha (*Beet necrotic yellow vein virus*, BNYVV, *Benyvirus*) el dominio *readthrough* es requerido para la formación de viriones (Schmitt *et al.*, 1992), además de facilitar la transmisión por hongos del virus (Tamada *et al.*, 1996; Reavy *et al.*, 1998). Las señales que dirigen este mecanismo en los virus de la familia *Luteoviridae* son diferentes a las del TMV, concretamente una secuencia de entre 8 y 16 nt, rica en citosinas, próxima al codón de parada del gen de la CP y otra secuencia localizada 700 nt aguas abajo de dicho codón (Brown *et al.*, 1996).

4.2.3. Reiniciación

Este mecanismo requiere la presencia de una (o varias) ORF(s), normalmente pequeña(s) (*short ORF*, sORF), que es traducida por los ribosomas y termina antes del comienzo de la ORF principal (Fig. 10). Cuando los ribosomas llegan al codón de parada de la sORF, las subunidades ribosomales 40S reinician la síntesis proteica en el siguiente codón de iniciación (Luukkonen *et al.*, 1995). Para que este mecanismo pueda ocurrir, en general es necesario que el tamaño de la sORF no exceda de los 30 codones para evitar en lo posible la disociación de factores esenciales, y que la distancia que separa a la sORF de la ORF principal sea superior a 70 nt para permitir que, después del primer proceso de traducción, la subunidad 40S pueda volver a unirse al complejo eIF2-GTP-Met-tRNA, necesario para el reconocimiento del siguiente codón de inicio y la traducción de la ORF correspondiente (Meijer y Thomas, 2002).

En miembros del género *Begomovirus* (familia *Geminiviridae*) se ha propuesto la utilización de los mecanismos de escape al proceso de rastreo ribosomal (que se explica más adelante) y de reiniciación combinados para la síntesis de algunas de sus proteínas (Shivaprasad *et al.*, 2005).

Una variante particular de este tipo de mecanismo ocurre en el CaMV (*Caulimovirus*) y otros pararetrovirus relacionados, como son el virus del mosaico del Figwort (*Figwort mosaic virus*, FMV, *Caulimovirus*) o el virus del rayado clorótico del cacahuete (*Peanut chlorotic streak virus*, PCSV, *Caulimovirus*), que codifican proteínas que promueven la reiniciación después de la

traducción de una o más ORFs largas. En el caso concreto del CaMV, la proteína TAV (*translational transactivator protein*) codificada por la ORFVI del virus, parece activar la traducción de las ORFs largas consecutivas presentes en el mRNA policistrónico 35S del virus (Bonneville *et al.*, 1989; Fütterer y Hohn, 1991; Park *et al.*, 2001).

Además de en RNAs virales, este mecanismo se ha propuesto también para la traducción de algunos mRNAs celulares que presentan secuencias líder con sORFs que desempeñan un papel en la regulación de la traducción de la(s) ORF(s) localizada(s) aguas abajo (Grant *et al.*, 1995; Morris y Geballe, 2000; Sarrazin *et al.*, 2000; Gaba *et al.*, 2001; Lu *et al.*, 2004; Iacono *et al.*, 2005).

4.2.4. Rastreo ribosomal discontinuo (*shunting*)

Este mecanismo, fue descrito por primera vez en el CaMV (Fütterer *et al.*, 1990, 1993) y, más tarde, en el RTBV (*Tungrovirus*) (Fütterer *et al.*, 1996). En estos virus, el proceso opera sobre la secuencia líder de su RNA pregenómico (pgRNA), que contiene múltiples sORFs en una región muy estructurada. El plegamiento que adopta esta región junto a la presencia de sORFs constituye una barrera para la migración lineal de los ribosomas (Ryabova *et al.*, 2000; Pooggin *et al.*, 2000). La sORF 5'-proximal, que termina inmediatamente aguas arriba de la región estructurada, y esta región en sí misma parecen ser imprescindibles para que se produzca el mecanismo de rastreo ribosomal discontinuo (Fig. 11). Algunos ribosomas al alcanzar el elemento de estructura secundaria "saltan" y reanudan el proceso aguas abajo, comenzando la síntesis proteica en la siguiente ORF larga (ORFVII en el CaMV y ORFI en el RTBV). Se ha determinado, además, una serie de parámetros que son críticos en este proceso como son el tamaño óptimo de la sORF, que oscila entre 2 y 15 codones (Hemmings-Mieszczak *et al.*, 2000; Pooggin *et al.*, 2000), y la distancia entre el codón de parada de la sORF y la base del tallo de la estructura secundaria, que ha de ser entre 5 y 10 nt (Dominguez *et al.*, 1998; Pooggin *et al.*, 2000).

Junto al CaMV y al RTBV, la mayoría de pararetrovirus de plantas contienen en la región líder sORFs seguidas por zonas muy estructuradas (Pooggin *et al.*, 1999). Posiblemente este mecanismo esté conservado entre virus de la familia *Caulimoviridae* (Pooggin *et al.*, 1999), con excepción del virus de la hoja amarilla y rizada del *Cestrum* (*Cestrum yellow leaf curl virus*,

CYLVCV, *Caulimovirus*) que presenta una región líder diferente (Stavolone *et al.*, 2003). Este mecanismo podría tener también lugar en al menos un virus RNA de plantas, el virus esférico del tungro del arroz (*Rice tungro spherical virus*, RTSV, *Waikavirus*, *Sequiviridae*) (indicado como resultado no publicado en Ryabova *et al.*, 2006). Curiosamente este virus y el RTBV coexisten en la naturaleza y causan, mediante interacción sinérgica, la enfermedad del tungro en el arroz. Una hipótesis atractiva es que quizá estos dos virus han coevolucionado desarrollando una estrategia de traducción común por medio del intercambio de su información genética.

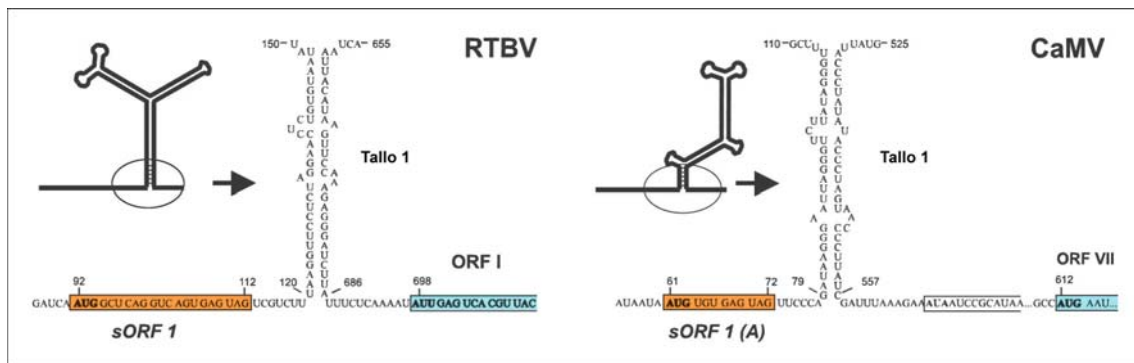


Fig. 11. Representación esquemática de la región líder de los pgRNAs del RTBV y del CaMV, donde se produce el mecanismo de rastreo ribosomal discontinuo. La sORF aparece sobre un fondo naranja, a continuación se representa la región de estructura secundaria y, por último, el inicio de la ORF larga que traducen los ribosomas después del “salto” y que se muestra sobre fondo azul. En el caso del CaMV, el rectángulo previo a la ORFVII, delimita un sitio alternativo de inicio de la traducción caracterizado por un codón no canónico (AUA) en lugar del AUG convencional (basado en Pooggin *et al.*, 2006).

Aunque el mecanismo de rastreo ribosomal discontinuo no parece estar muy extendido en la naturaleza, además de en los mencionados virus de plantas, también se ha descrito en ciertos virus de animales (Latorre *et al.*, 1998; Yueh y Schneider, 2000; Sen *et al.*, 2004) y en algún mensajero celular (Rogers *et al.*, 2004; Chappell *et al.*, 2006; Mauro *et al.*, 2007).

4.2.5. Escape al proceso de rastreo ribosomal (*leaky scanning*)

Este mecanismo es utilizado frecuentemente por virus de plantas y permite la traducción de dos o, excepcionalmente, tres proteínas a partir de un único mRNA. Se caracteriza porque una fracción de los ribosomas que comienza el rastreo lineal desde el extremo 5' evita el primer

codón de inicio, ya sea porque esta flanqueado por un contexto subóptimo o porque es un codón de iniciación débil, continuando el rastreo hasta alcanzar otro codón más apropiado donde se inicia la síntesis proteica (Fig. 10). En este mecanismo puede estar involucrada una misma ORF con sitios de inicio diferentes o, alternativamente, ORFs solapantes u ORFs consecutivas (Fig. 12).

La eficiencia de reconocimiento del primer codón va a depender de la secuencia que lo flanquea, de las condiciones del sistema de traducción y de otras señales del mRNA (Dinesh-Kumar y Miller, 1993). Así, tanto el virus de la hoja enrollada de la patata (*Potato leafroll virus*, PLRV, *Luteovirus*) como el BYDV (cepa PAV) presentan dos ORFs solapantes que son traducidas por este mecanismo (Fig. 12) y, sin embargo, las proteínas correspondientes se sintetizan en proporciones 100:1 y 1:7, respectivamente (Dinesh-Kumar *et al.*, 1992; Dinesh-Kumar y Miller, 1993).

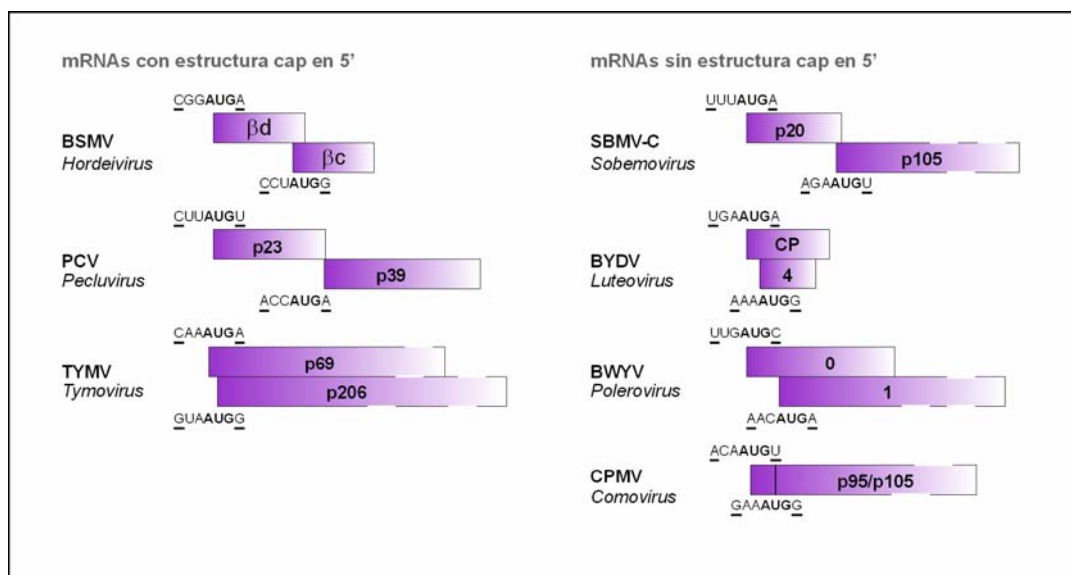


Fig. 12. Virus que utilizan el mecanismo de escape al proceso de rastreo ribosomal para expresar algunos de sus genes. El género al que pertenecen cada uno de los virus aparece en cursiva. Sólo se muestran las ORFs implicadas en el mecanismo, que han sido representadas por medio de rectángulos (con línea discontinua si el tamaño de la ORF no está a escala). El codón de inicio de cada ORF y el contexto de secuencia que lo flanquea aparecen especificados (basado en Dreher y Miller, 2006).

Entre los virus que utilizan este mecanismo para traducir dos proteínas a partir de sitios de inicio diferentes en una misma ORF se encuentran el virus de la sharka (PPV, *Plum pox virus*,

Potyvirus) (Riechmann *et al.*, 1991; Simon-Buela *et al.*, 1997) o el virus del mosaico del chícharo (CPMV, *Cowpea mosaic virus*, *Comovirus*) (Holness *et al.*, 1989) (Fig. 12).

La expresión de dos ORFs solapantes mediante el mecanismo de escape al proceso de rastreo se ha descrito en el caso del BYDV (*Luteovirus*) (Dinesh Kumar-y Miller, 1993), del TYMV (*Tymovirus*) (Weiland y Dreher, 1989), del CNV (*Tombusvirus*) (Rochon y Johnston, 1991), del virus del mosaico del frijol del sur (*Southern bean mosaic virus*, SBMV, *Sobemovirus*) (Sivakumaran y Hacker, 1998), del virus del amarilleamiento de la remolacha occidental (*Beet western yellows virus*, BWYV, *Polerovirus*) (Veidt *et al.*, 1998), del virus del mosaico estriado de la cebada (*Barley stripe mosaic virus*, BSMV, *Hordeivirus*) (Zhou and Jackson, 1996), y del virus X de la patata (*Potato virus X*, PVX, *Potexvirus*) (Verchot *et al.*, 1998) (ejemplos en Fig. 12).

Por último, entre los virus que utilizan esta estrategia para la expresión de ORFs independientes se encuentra el virus del macizo del cacahuete (*Peanut clump virus*, PCV, *Pecluvirus*) (Herzog *et al.*, 1995) (Fig. 12) y el RTBV (Fig. 13).

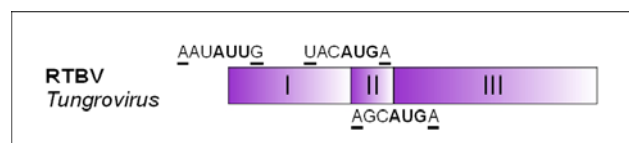


Fig. 13. Representación de las ORFs involucradas en un mecanismo de escape al proceso de rastreo ribosomal en el pgRNA del RTBV. Se especifica el codón de inicio de cada ORF y el contexto que lo flanquea.

El RTBV representa un caso muy destacable de escape al proceso de rastreo ya que mediante este mecanismo se producen tres proteínas independientes a partir de un mismo mRNA. Las ORFs que codifican dichas proteínas están solapadas por medio de codones de inicio y parada del tipo AUGA (Futterer *et al.*, 1997). La ORFI presenta un codón de iniciación débil AUU y se ha estimado que un 10% de los ribosomas que han iniciado el rastreo desde el extremo 5' del mRNA comienza la traducción de dicha ORF, mientras que el resto continúa el rastreo hasta alcanzar el codón de inicio de la ORFII (Futterer *et al.*, 1996, 1997). El contexto subóptimo que flanquea a este último codón permite, a su vez, que una fracción de los ribosomas continúe el rastreo hasta alcanzar el codón de inicio de la ORFIII, que se encuentra

en un contexto más favorable (Fig. 13). La región que han de recorrer los ribosomas hasta alcanzar la ORFIII es de aproximadamente 1 kb y carece de codones AUG adicionales en cualquier pauta de lectura, un rasgo compartido por badnavirus relacionados. Esto sugiere que el mecanismo de escape al proceso de rastreo que da lugar a tres proteínas distintas está conservado en dos géneros diferentes de pararetrovirus baciliformes (Poogin *et al.*, 1999).

Al igual que ocurre con algunas de las estrategias descritas anteriormente, el mecanismo de escape al proceso de rastreo ribosomal también es empleado por algunos mRNAs celulares (Touriol *et al.*, 2003; Wang y Rothnagel, 2004; Iacono *et al.*, 2005; Zhou y Song, 2006).

4.2.6. Iniciación interna

En contraste con los mecanismos descritos anteriormente que se caracterizan por un rastreo ribosomal dependiente del extremo 5', el mecanismo de iniciación interna está mediado por la presencia de una región en el mRNA, denominada IRES como se ha comentado anteriormente, que recluta directamente a los ribosomas y factores de iniciación para comenzar la síntesis proteica.

Los IRES se descubrieron por primera vez en virus animales, concretamente en picornavirus (Jang *et al.*, 1988; Pelletier y Sonenberg, 1988). La 5' UTR de estos virus presenta una gran complejidad estructural y se postuló que su traducción podía tener lugar mediante un mecanismo inusual, distinto del utilizado en la mayoría de mRNAs eucariotas donde los ribosomas comienzan el barrido desde el extremo 5' del mensajero. Esta hipótesis se comprobó mediante el diseño de mRNAs bicistrónicos y el análisis de la eficiencia de su traducción: sólo se observaba expresión del segundo gen cuando un segmento de la 5' UTR de los picornavirus estaba presente entre los dos cistrones, indicando que el IRES era activo incluso fuera de su contexto genético (Jang *et al.*, 1988; Pelletier y Sonenberg, 1988; Martínez-Salas, 1999; Hellen y Sarnow, 2001). Posteriormente este tipo de elementos han sido descritos en otros mRNAs virales, tanto de animales (Jang, 2006; Fraser y Doudna, 2007; Martínez-Salas *et al.*, 2008) como de plantas (Wong *et al.*, 2008), y en algunos mRNAs celulares (Dinkova *et al.*, 2005; Graber y Holcik, 2007; Lewis y Holcik, 2008; Mardanova *et al.*, 2008; Spriggs *et al.*, 2008).

Entre los virus de plantas donde se han descrito elementos IRES se encuentra el TEV (*Potyvirus*, *Potyviridae*), el PLRV (*Luteovirus*, *Luteoviridae*), el TMV (*Tobamovirus*, sin asignar a familia) y el HCRSV (*Carmovirus*, *Tombusviridae*), todos ellos virus de ssRNA+.

Los elementos IRES pueden localizarse en la 5' UTR o internamente en el genoma. Los miembros de la familia *Potyviridae* presentan elementos IRES en su 5' UTR. Los RNAs de esta gran familia se asemejan a los RNAs de picornavirus, ya que también codifican una proteína VPg que se une covalentemente al extremo 5' y una cola poliA en el extremo 3'. Como ya se ha comentado anteriormente, la proteína VPg podría interactuar con el factor proteico eIF4E y dicha unión podría tener un papel en la traducción viral, aunque este punto no está aún claro. Los potyvirus difieren de los picornavirus en que sus VPgs son bastante más grandes y su 5' UTR es bastante más corta (alrededor de 150 nt en potyvirus frente a 600-1200 nt en picornavirus) y menos estructurada. El diseño de mensajeros bicistrónicos con la 5' UTR del TEV entre los dos cistrones, reveló traducción *in vitro* del segundo gen, sugiriendo que dicha región actuaba como un IRES (Gallie, 2001). El IRES del TEV se ha cartografiado en un segmento de 143 nt donde se ha definido un pseudonudo (PK1), entre los nt 38 y 75, importante para la traducción (Zeenko y Gallie, 2005; Khan *et al.*, 2008) (Fig. 15). Un bucle (L3) de PK1 es complementario a los nt 1117-1123 nt del 18S rRNA. Las mutaciones en L3 que perturban esta complementariedad reducen sustancialmente la eficiencia de la traducción, mientras que las mutaciones que la conservan no alteran dicha eficiencia. Esta interacción quizá favorece la unión de la subunidad ribosomal 40S al RNA viral, facilitando el proceso de traducción (Zeenko y Gallie, 2005). Además de este pseudonudo, una secuencia aguas arriba del mismo, entre los nt 28 y 37, también parece necesaria (Fig. 15). Experimentos adicionales han corroborado que la 5' UTR confiere al virus capacidad de traducción *in vivo* por un mecanismo independiente de cap y que además involucra específicamente al factor de iniciación eIF4G (Gallie, 2001; Ray *et al.*, 2006; Khan *et al.*, 2008) (Fig. 15)

Las 5' UTRs de otros potyvirus, como el virus Y de la patata (*Potato virus Y*, PVY) (Fig. 15), el PPV o el TuMV, también estimulan la traducción por un mecanismo independiente de cap en diversos sistemas de ensayo. En construcciones bicistrónicas, la 5' UTR del PVY promueve *in vitro* la traducción del segundo cistrón (Levis y Astier-Manifacier, 1993) y, además, se ha

demostrado que los últimos 55 nt de dicha región son los más importantes para que tenga lugar el proceso de iniciación interna en protoplastos de plantas de tabaco (Yang *et al.*, 1997).

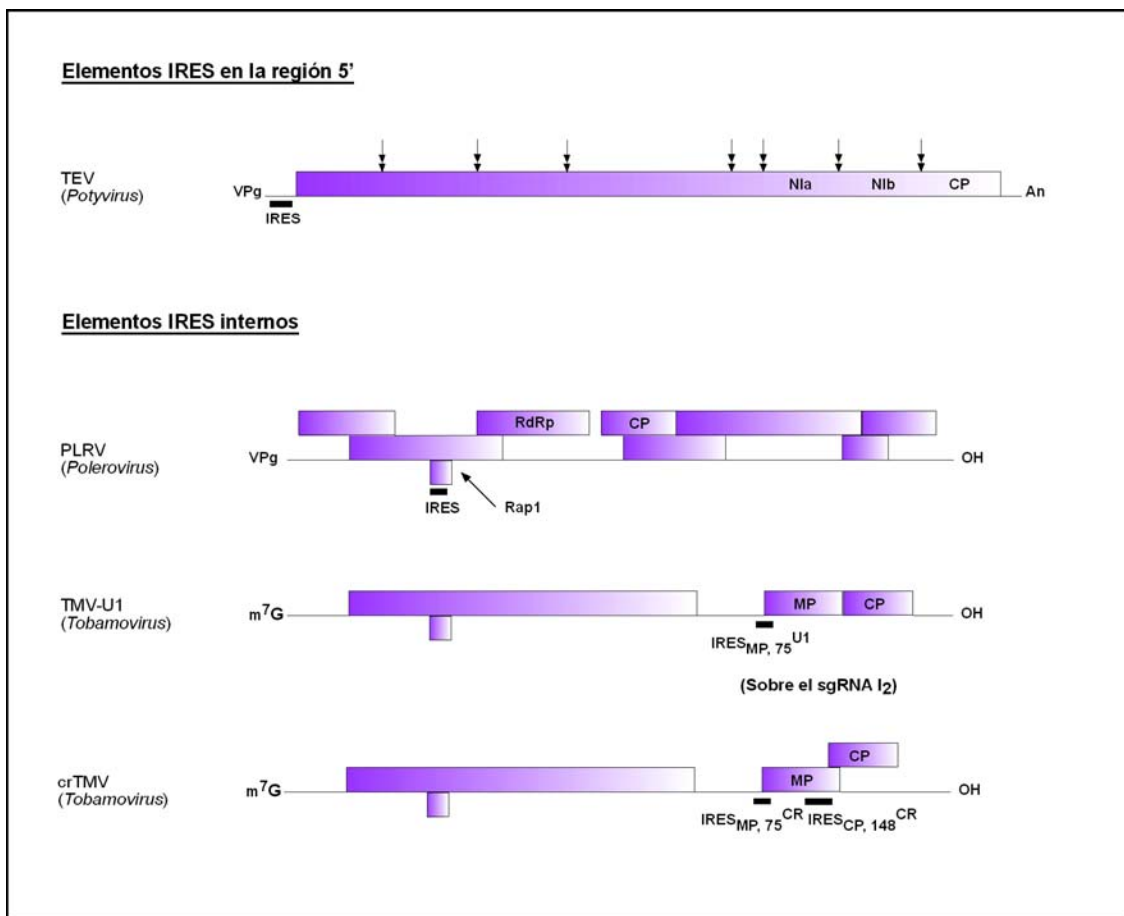


Fig. 14. Organización genómica de virus de plantas que utilizan el mecanismo de iniciación interna para la expresión de alguno de sus genes. Los elementos IRES en el genoma viral están representados por medio de pequeños rectángulos negros. Las flechas con doble cabeza del TEV indican las posiciones de corte de la poliproteína. En el TMV-U1 y en el crTMV la MP es traducida a partir del sgRNA I₂, mientras que en el crTMV la traducción de la CP puede ocurrir a partir del sgRNA I₂ y del gRNA (basado en Kneller *et al.*, 2006).

Los RNAs virales de los géneros *Polevirus* y *Enamovirus*, dentro la familia *Luteoviridae*, tienen también una proteína VPg unida covalentemente al extremo 5', por lo que cabría esperar que estos virus contuvieran elementos IRES en la 5' UTR. En el caso del polevirus PLRV (Fig. 14), los tres genes más próximos al extremo 5' se traducen a partir del gRNA, mientras que el resto de ORFs localizadas aguas abajo, lo hacen a partir de sgRNAs (Tacke *et al.*, 1990;

Ashoub *et al.*, 1998). Hasta el momento no se ha identificado ningún elemento en la 5' UTR capaz de mediar la traducción por un mecanismo independiente de cap; sin embargo se ha descrito una secuencia que permite la traducción de una proteína de 5 kDa, denominada Rap1, codificada por un gen imbricado dentro de la ORF2 (Jaag *et al.*, 2003) (Fig. 14). La función específica de Rap1 se desconoce, pero es indispensable para la replicación viral. La secuencia, que actúa como IRES, se localiza alrededor de 1500 nt aguas abajo del extremo 5' del gRNA (Jaag *et al.*, 2003), y se caracteriza por un motivo rico en polipurinas, GGAGAGAGAGG, situado 22 nt aguas abajo del codón de inicio del gen de la Rap1. Este motivo, junto con los primeros 22 nt del gen de la Rap1, es suficiente para mediar la traducción por un mecanismo independiente de cap *in vitro*, pero no *in vivo* (Jaag *et al.*, 2003).

Entre los virus que presentan elementos IRES localizados internamente en el genoma, se encuentran el TMV y el HCRSV. La presencia de estos elementos permite a estos virus la traducción de genes internos o 3'-proximales. El TMV (cepa U1, TMV-U1) presenta cuatro ORFs, las dos primeras codifican proteínas implicadas en replicación y se traducen a partir del gRNA, mientras que los genes internos y 3'-proximales codifican la MP y la CP, respectivamente, y se traducen a partir de sgRNAs. El sgRNA I₂ del TMV-U1 es dicistrónico, y no presenta estructura cap en 5' (Hunter *et al.*, 1983; Joshi *et al.*, 1983). Se ha identificado un elemento IRES, con un tamaño de 75 nt, localizado aguas arriba del gen de la MP (y que se corresponde con la 5' UTR del sgRNA I₂) denominado IRES_{MP, 75}^{U1} (Fig. 14) que promueve la traducción de la MP. Sin embargo, la misma secuencia es incapaz de mediar la expresión de dicha proteína a partir del gRNA, probablemente por problemas de inaccesibilidad de los ribosomas a dicha región (Skulachev *et al.*, 1999). Un elemento funcionalmente similar, IRES_{MP, 75}^{CR}, está presente en el tobamovirus que infecta crucíferas (crTMV) (Skulachev *et al.*, 1999) (Fig. 14). Este último además contiene otro elemento IRES que media la traducción de la CP a partir del sgRNA I₂ y del gRNA (Ivanov *et al.*, 1997; Dorokhov *et al.*, 2002). Este IRES, llamado IRES_{CP, 148}^{CR}, está constituido por 148 nt y está localizado en la región del genoma que codifica la MP aunque también solapa 25 nt con el gen de la CP (Ivanov *et al.*, 1997) (Fig. 14). El análisis de la secuencia y de la estructura de IRES_{CP, 148}^{CR} revela la presencia de diferentes tallos y bucles flanqueados por dos regiones de polipurinas ricas en A (*polypurine A-rich sequences*, PARS), necesarias para la actividad IRES (Fig. 15). Se ha demostrado además

que una versión simplificada de la PARS, consistente en tan sólo una secuencia $(GAAA)_{16}$, también promueve la iniciación interna de forma muy eficiente. Experimentos *in vivo* indican que el IRES_{CP, 148}^{CR} favorece la producción temprana de CP, lo que parece facilitar el movimiento sistémico del virus durante la infección de plantas crucíferas (Asurmendi *et al.*, 2004; Dorokhov *et al.*, 2006).

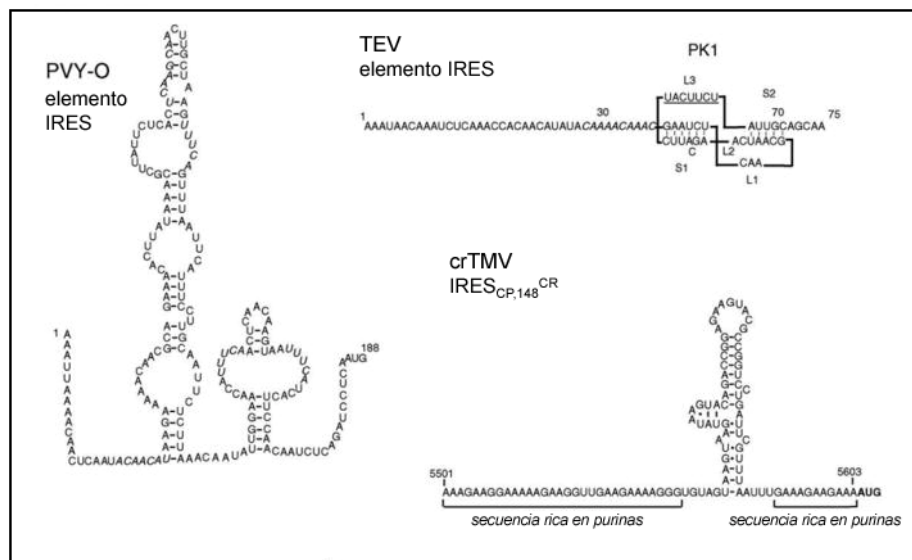


Fig. 15. Elementos IRES de algunos virus de plantas, localizados en la región 5' UTR (TEV, PVY) o en regiones intergénicas (crTMV). En el TEV se han representado los primeros 75 nt del elemento IRES, que incluye los 45 nt del pseudonudo PK1 con la secuencia complementaria al 18S rRNA en L3 subrayada (basado en Kneller *et al.*, 2006).

En el HCRSV se ha demostrado también la existencia de un IRES que dirige la traducción de CP desde el gRNA y los sgRNAs. Este elemento, que actúa sinérgicamente con la 3' UTR, consta de 100 nt y está localizado inmediatamente aguas arriba del gen de la CP (Koh *et al.*, 2003, 2006). Al diseccionar el elemento IRES, se observó que en la primera mitad se localiza una secuencia de seis nucleótidos necesaria para su actividad mientras que, en la segunda mitad, la secuencia primaria y un motivo GNRA no son esenciales para dicha actividad. Al igual que otros IRES, el del HCRSV promueve la expresión del segundo cistron en construcciones bicistrónicas homólogas y heterólogas, y dicha expresión se ve incrementada cuando se introduce la 3' UTR del virus, lo que verifica la cooperación entre el elemento IRES y esta región no codificante (Koh *et al.*, 2003, 2006).

5. VARIABILIDAD Y EVOLUCIÓN DE LOS VIRUS

El análisis de las estructuras genéticas y de la evolución de las poblaciones resulta crucial en el área de la Biología. Con el desarrollo de la disciplina conocida como Evolución Molecular, que surge de la interacción entre aproximaciones experimentales, descriptivas y teóricas y que actualmente ha experimentado un gran impulso gracias a la Bioinformática, se han podido determinar las relaciones evolutivas entre distintos organismos basándose en el análisis de genes y genomas, y de proteínas y proteomas, prescindiendo de detalles fenotípicos. En el caso de los virus, esto es especialmente interesante dada su alta tasa de variación genética y la facilidad con que se pueden obtener sus secuencias genómicas, permitiendo estudiar las bases de evolución de muchas familias y grupos de virus. El conocimiento de la dinámica evolutiva de los virus es esencial para el desarrollo de estrategias estables de control, ya que en muchos casos las metodologías utilizadas no son efectivas debido a que las poblaciones virales presentan un grado de diversidad genética mayor que el observado en organismos superiores (García-Arenal *et al.*, 2003; Timm y Roggendor, 2007; McBurney y Ross, 2008).

5.1. Heterogeneidad genética de las poblaciones de virus de RNA

Los virus tienen que enfrentarse con frecuentes cambios ambientales ya que cada nuevo hospedador supone una modificación de las condiciones en las que debe llevar a cabo su ciclo infeccioso. En el caso de vertebrados, el sistema inmunitario produce además moléculas para impedir la multiplicación del patógeno, una situación a la que los virus deben adaptarse para poder persistir. Como consecuencia, los virus han desarrollado una gran capacidad para generar diversidad genética como estrategia evolutiva, en detrimento del mantenimiento de una alta fidelidad de la copia.

Muchos virus (más del 80% en el caso de virus de plantas) poseen un genoma de RNA y son capaces de variar y, por tanto, de evolucionar más rápidamente que los que tienen genomas de DNA. Este fenómeno es en gran parte debido a que las RdRps presentan una mayor tasa de error que las DdDps, a consecuencia de la ausencia (o muy baja actividad) de las funciones correctoras y reparadoras (Domingo y Holland, 1997; Castro *et al.*, 2005; Agol, 2006). La elevada tasa de error de la replicación del RNA vírico junto a los cortos tiempos de

generación, puede dar lugar a una gran cantidad de mutantes que difieren en uno o más nucleótidos, lo que se traduce en una gran heterogeneidad genética (Domingo, 1994).

Como consecuencia de la variación genética, las poblaciones de virus de RNA están formadas por distribuciones dinámicas de genomas víricos no idénticos pero genéticamente relacionados denominados cuasiespecies (Domingo *et al.*, 1985). Este término fue acuñado por primera vez por Eigen, Schuster y colaboradores para referirse a las primeras estructuras autorreplicativas (Eigen y Schuster, 1981; Eigen *et al.*, 1981; Eigen y Biebricher, 1988). Las cuasiespecies se hallan sometidas a un continuo proceso de variación, competición y selección que permiten al virus adaptarse rápidamente en función del ambiente (Eigen, 1996). A medida que progresa la infección se imponen, por competición y selección, las variantes con mayor eficacia biológica. El reparto de variantes presente en una cuasiespecie no es, pues, arbitrario sino que se centra en torno a una secuencia maestra (la de mayor eficacia biológica) que generalmente coincide con la secuencia consenso o promedio de la población. Esta última se define estadísticamente y tendría en cada posición el nucleótido más frecuente del conjunto de moléculas de la cuasiespecie. En un ambiente constante puede darse un equilibrio poblacional que conduce a una estasis evolutiva, mientras que las perturbaciones ambientales suelen disparar una rápida evolución viral (Domingo, 1994; Domingo *et al.*, 1996).

5.2. Fuentes de variación genética

La variación genética se genera mediante errores producidos durante la replicación de los genomas. En los virus existen tres posibles fuentes de error: mutación, recombinación y reordenamiento génico.

i) Mutación. Es un mecanismo clave en la generación de variación genética. Se define como el proceso por el cual nucleótidos que no estaban presentes en la secuencia molde son incorporados en la cadena hija durante la replicación. En la mayoría de los casos se trata de sustituciones nucleotídicas aunque también pueden producirse deleciones o inserciones. Dada la importancia de este mecanismo, es de sumo interés conocer la frecuencia con la que ocurren. Se define como *tasa de mutación* a la cantidad de errores que se producen por nucleótido (o por genoma) y por ronda de síntesis o, dicho en otras palabras, es la frecuencia con que ocurre una mutación durante la replicación del genoma y es un excelente parámetro

para medir el grado de mutabilidad entre los diferentes virus (De Filippis y Villareal, 1999; Malpica *et al.*, 2002; Duffy *et al.*, 2008). Las estimaciones de la tasa de mutación para virus de RNA que infectan a mamíferos oscilan entre 10^{-3} - 10^{-5} incorporaciones erróneas por nt y por ronda de replicación, lo que resulta en aproximadamente un error por genoma por ciclo replicativo (Drake y Holland, 1999). Una estimación de la tasa de mutación de un virus de plantas, el TMV, indica que dicha tasa es similar a la de los virus líticos de RNA de mamíferos (Malpica *et al.*, 2002).

Como se ha indicado anteriormente las elevadas tasas de mutación que muestran los virus de RNA pueden reflejar una estrategia evolutiva (Drake, 1993), aunque esta propiedad también conlleva algunos costes como una elevada proporción de mutantes deletéreos y la pérdida de infectividad, fenómenos que pueden conducir a una rápida extinción de la población (Holland *et al.*, 1990). En cualquier caso, se debe diferenciar entre la tasa de mutación y la frecuencia de mutación observada, ya que las dos variables pueden diferir debido a que una fracción desconocida de los mutantes generados es deletérea y otra puede ser eliminada de la población mediante selección o por deriva génica (ver apartado 5.3). En este contexto, varios estudios han demostrado baja diversidad nucleotídica en al menos dos miembros del género *Tobamovirus*, en el virus del moteado atenuado del pepino (*Pepper mild mottle virus*, PMMV) y el virus del mosaico verde atenuado del tabaco (*Tobacco mild green mosaic virus*, TMGMV), y en un closterovirus, el virus del enanismo amarillento de las cucurbitáceas (*Cucurbit yellow stunting disorder virus*, CYSDV) (Rodríguez-Cerezo *et al.*, 1989; Fraile *et al.*, 1996; Rubio *et al.*, 2001; Marco y Aranda, 2005). Igualmente, se ha observado baja diversidad entre aislados separados geográficamente del virus del pseudo-amarilleamiento de la remolacha (*Beet pseudo-yellows virus*, BPYV, *Crinivirus*) (Rubio *et al.*, 1999, 2001), y entre aislados de una misma región del virus del enanismo clorótico del boniato (*Sweet potato chlorotic stunt virus*, SPCSV, *Crinivirus*) (Alicai *et al.*, 1999). Sin embargo, el virus del moteado amarillo del arroz (*Rice yellow mottle virus*, RYMV, *Sobemovirus*), el virus del rizado de la hoja del algodón (*Cotton leaf curl virus*, CLCV, *Begomovirus*) y un satRNA del virus del mosaico del pepino (*Cucumber mosaic virus*, CMV, *Cucumovirus*) presentan niveles mayores de variación. En el caso del satRNA del CMV y de algunos *Geminivirus*, dicha variabilidad está asociada a la adaptabilidad de los mismos a sus huéspedes (Roossink, 1997).

ii) Recombinación. Se puede definir como un proceso mediante el cual el material genético de diferentes variantes es intercambiado, dando lugar a la formación de un ácido nucleico derivado de dos ácidos nucleicos parentales o más. Este tipo de mecanismo genera variantes genéticas con mayores cambios que las resultantes de mutaciones puntuales. Aunque en un principio el fenómeno de la recombinación no fue muy explorado en virus, está siendo estudiado ampliamente en la actualidad tanto en virus de animales (Prlic *et al.*, 2004; Althaus y Bonhoeffer, 2005) como en virus de plantas (Nagy y Bujarski, 1993; Revers *et al.*, 1996; Olsthoorn *et al.*, 2002; Chare y Holmes, 2006). En estos últimos, el fenómeno de recombinación parece ser una fuente de variación muy relevante que contribuye de forma significativa a la evolución del virus de RNA (Lai, 1992; Simon y Bujarski, 1994; Roossinck, 1997; Nagy *et al.*, 1999a; Worobey y Holmes, 1999; Alejska *et al.*, 2001, 2005a, 2005b; García-Arenal *et al.*, 2001) y de DNA (Navas-Castillo *et al.*, 2000; García-Andrés *et al.*, 2007).

La recombinación puede ser homóloga, cuando el donador de secuencias reemplaza una región homóloga del aceptor dejando su estructura invariable, u homóloga aberrante y heteróloga, aunque estas últimas son menos frecuentes (Lai, 1992). La recombinación homóloga aberrante ocurre entre secuencias de virus similares donde los entrecruzamientos no tienen lugar en regiones homólogas en las dos secuencias, sino en regiones diferentes aunque no muy alejadas, y la heteróloga, entre secuencias no relacionadas de RNA. A nivel poblacional, la recombinación puede generar cambios dramáticos en las propiedades biológicas de los virus con importantes consecuencias epidemiológicas, incluyendo la aparición de cepas resistentes o la posibilidad de infectar una nueva gama de huéspedes (Legg y Tresh, 2000; Monci *et al.*, 2002; García-Andrés *et al.*, 2006).

iii) Reordenamiento génico. Este proceso ocurre cuando dos virus de genoma bipartito o multipartito coinfectan una célula y la progenie resultante posee componentes genómicos de ambos parentales. Este fenómeno se ha descrito en poblaciones naturales de virus vegetales (Fraile *et al.*, 1997), y puede desempeñar un papel importante en la evolución viral (Qiu y Moyer, 1999; Miranda *et al.*, 2000). Al igual que la recombinación, los efectos de las reorganizaciones genéticas en los virus pueden ser muy relevantes, pero su papel en la estructura de las poblaciones virales no ha sido analizado con detalle.

5.3. Factores que determinan la estructura genética de las poblaciones virales

Desde la década de los sesenta, se han venido discutiendo principalmente dos escenarios para explicar la evolución de los organismos. En el primero, las variaciones observadas al estudiar una determinada molécula o una población de organismos son el resultado de la acción de la selección natural y, en el segundo, dichas variaciones son la consecuencia de procesos estocásticos asociados a la deriva genética, es decir, son debidas a un muestreo aleatorio de moléculas o genotipos entre generaciones, que se produce en ausencia de selección natural (Kimura *et al.*, 1985). Actualmente se acepta que la distribución de variantes generadas por mutación, recombinación o reordenamiento génico en una población dependerán fundamentalmente de una combinación de los dos procesos evolutivos: la selección y la deriva genética.

5.3.1. Selección natural

La selección natural es considerada como el motor esencial de la evolución desde que fue aceptada por la comunidad científica como la mejor explicación para la evolución de las especies. El concepto de selección natural se basa en que la adecuación de las características fenotípicas de los organismos frente a distintas condiciones ambientales determina la capacidad de los mismos para su supervivencia y reproducción y, mientras el medio permanezca inalterado, los individuos más exitosos (en esos mismos términos, supervivencia y reproducción) son los que se irán imponiendo y aumentando su frecuencia en la población. Si se produce un cambio ambiental, la variación que existe en la población será la materia prima que empleará la selección natural para favorecer a las variantes mejor adaptadas en esa nueva situación ambiental. La acción de la selección puede ser negativa o positiva. La selección es negativa cuando la eficacia de los genotipos mutantes es inferior a la del genotipo silvestre. Este tipo de selección permite explicar la eliminación de las mutaciones que son desventajosas en la población de estudio (Saitou y Nei, 1986; Seibert *et al.*, 1995; Young y de Pamphilis, 2000; Hughes *et al.*, 2003). En contraste, se produce selección positiva cuando la eficacia de los genotipos mutantes es superior a la del silvestre, lo que resulta en la fijación de mutaciones que son ventajosas en la población de estudio y permite explicar la existencia de polimorfismos (Seibert *et al.*, 1995; Wu *et al.*, 1997; Haydon *et al.*, 2001). La selección da lugar a una

disminución de la diversidad en la población viral y puede ocasionar un aumento de la diversidad entre poblaciones si éstas están bajo diferentes presiones selectivas.

En el caso de los virus, la selección puede actuar a distintos niveles. Un primer grupo de presiones selectivas está relacionado con el mantenimiento de algunas características estructurales del virus. Por ejemplo, los aminoácidos de las CPs implicados en el ensamblaje y estabilización de la nucleocápsida están en general muy conservados dentro de grupos virales (Altschuh *et al.*, 1987), y lo mismo ocurre con las estructuras primarias, secundarias o de orden superior del RNA viral que son importantes en la replicación (Bacher *et al.*, 1994). Otros factores de selección están asociados con las plantas huéspedes o con los vectores de transmisión (Power, 2000). En este contexto se ha observado una gran conservación en proteínas virales implicadas en el proceso de transmisión mediado por mosca blanca o por áfidos (Pirone y Blanc, 1996; Flasiński y Cassidy, 1998; Liu *et al.*, 2002). Asimismo, estudios con distintos agentes virales o subvirales han puesto de manifiesto variaciones importantes en las poblaciones como consecuencia de su adaptación al huésped (Liang *et al.*, 2002a; Sacristán *et al.*, 2005; Ayllón *et al.*, 2006; Rico *et al.*, 2006). En ocasiones, los cambios son tan drásticos que las variantes seleccionadas tras pases de un aislado del patógeno en una determinada especie vegetal son incapaces de infectar con eficacia al huésped original del que se obtuvo el aislado (Fagoaga *et al.*, 1995; Liang *et al.*, 2002a; Tan *et al.*, 2005). El efecto que un virus puede tener sobre la población de su planta huésped o sobre el insecto vector dependerá de su virulencia, definida como el efecto del patógeno en disminuir las aptitudes del organismo que parasita. Esta propiedad puede repercutir en la selección (positiva o negativa) de un patógeno y desempeña un papel importante en su evolución (Anderson y May, 1982; May y Nowak, 1995).

5.3.2. Deriva genética

Los virus vegetales pueden alcanzar grandes tamaños poblacionales en las plantas que infectan. Sin embargo, el tamaño efectivo de la población (el número de individuos susceptibles de transmitir sus genes a una nueva generación), puede ser mucho menor debido a que una fracción de la población está compuesta por mutantes que no se multiplicarán. Las poblaciones de muchos organismos pueden no ser lo suficientemente grandes como para asegurar que

cada variante tenga una descendencia en la próxima generación, de modo que algunos factores fortuitos pueden determinar la transmisión de los rasgos genéticos a nuevas generaciones. Este proceso se conoce con el nombre de deriva genética. La pérdida de eficacia biológica de un virus se puede estudiar en sistemas donde se evita la competición entre virus y que constituyen los denominados cuellos de botella genéticos o poblacionales. Estos cuellos de botella conducen a un modo de deriva genética conocida como efecto fundador, que es una situación extrema de deriva genética y tiene lugar cuando una población nueva se inicia con un número pequeño de genotipos seleccionados al azar a partir de la población inicial. Este tipo de efectos tiende a disminuir la diversidad dentro de una población y a incrementar la diversidad entre poblaciones y, en el caso de virus de plantas, puede darse cada vez que una nueva planta huésped es infectada, una nueva especie de planta se convierte en huésped, o una nueva área geográfica se coloniza. Asimismo, el proceso de dispersión sistémica del virus o transmisión mediante vectores también puede conducir a este tipo de situaciones (García-Arenal *et al.*, 1999; French y Stenger, 2003; Sacristán *et al.*, 2003; Monsion *et al.*, 2008).

5.4. Coevolución

La coevolución es un término de la Biología por el que se designa un fenómeno de adaptación evolutiva mutua entre dos o varias especies de seres vivos como resultado de su influencia recíproca por relaciones como la simbiosis, el parasitismo, la polinización o las interacciones entre presa y depredador. La serie de adaptaciones coordinadas que desarrollan les permite poder llevar a cabo correctamente la función que desempeñaban o mantener la relación que tenían. Para que el fenómeno de coevolución ocurra, son necesarios dos requisitos: i) *especificidad*, la evolución de cada carácter de una entidad se debe a presiones selectivas del carácter de la otra y, ii) *reciprocidad y simultaneidad*, ambos caracteres en las dos entidades deben evolucionar conjuntamente y al mismo tiempo (Janzen, 1980).

La coevolución se puede producir también entre moléculas involucradas en un mismo proceso biológico. Un ejemplo de esto podría estar representado por proteínas que intervienen en una determinada ruta metabólica o que se asocian para desempeñar una función concreta. En muchos casos, este tipo de asociaciones conlleva una interacción física entre las moléculas

involucradas. Esto condiciona que cuando se producen cambios en una de ellas, estos cambios sean compensados en la otra proteína, asegurando el correcto funcionamiento del organismo que las codifica. El fenómeno de coevolución se puede producir incluso dentro de una misma molécula, ya sea proteína o ácido nucleico, por la dependencia que puede existir entre distintos residuos para mantener la eficacia biológica de la molécula en cuestión (revisiones recientes en Galtier y Dutheil, 2007; Pazos y Valencia, 2008).

La predicción *a priori* de las regiones o aminoácidos que están implicados en el establecimiento de interacciones entre proteínas o entre dominios de una misma proteína (requeridas para realizar correctamente cualquiera de sus funciones), será de gran ayuda para el investigador que pretende abordar experimentalmente un estudio mecanístico del proceso en el que participa(n). Esto también es aplicable a regiones distintas de un ácido nucleico o de diferentes ácidos nucleicos relacionados funcionalmente. La detección de covariaciones dentro de una misma molécula o entre moléculas distintas ha sido considerada como un signo de posible coevolución y, por lo tanto, la búsqueda de mutaciones correlacionadas (asistida por programas bioinformáticos) ha sido ampliamente utilizada para inferir interacciones funcionales entre diferentes sitios de una molécula o de moléculas distintas (e.j.: Altschuh *et al.*, 1987; Hughes, 1992; Korber *et al.*, 1993; Göbel *et al.*, 1994; Pollock *et al.*, 1999; Larson *et al.*, 2000; Goh y Cohen, 2002; Duntheil *et al.*, 2005; Codoñer *et al.*, 2006; Codoñer y Fares, 2008).

6. VIROSIS DEL GERANIO

El geranio (*Pelargonium* spp.) ocupa una posición importante dentro de la industria floral nacional con un consumo anual de más de 20 millones de plantas. Sin embargo, su producción se ve a menudo afectada por distintos factores entre los que destacan los daños causados por diversos patógenos de tipo viral tales como: el PFBV, el virus del rizado de la hoja del *Pelargonium* (*Pelargonium leaf curl virus*, PLCV, *Tombusvirus*), el TMV, el CMV, el virus del bronceado del tomate (*Tomato spotted wild virus*, TSWV, *Tospovirus*), el virus de las manchas necróticas del *Impatiens* (*Impatiens necrotic spot virus*, INSV, *Tospovirus*), el virus del punteado del geranio zonal (*Pelargonium zonale spot virus*, PZSV, *Ourmiavirus*), el virus de las manchas anulares del tomate (*Tomato ringspot virus*, ToRSV, *Nepovirus*), el virus de las manchas anulares del tabaco (*Tobacco ringspot virus*, TRSV, *Nepovirus*) y el PLPV, entre otros. Aunque

estos patógenos normalmente no ocasionan la muerte de la planta, reducen su crecimiento vegetativo, provocan problemas de enraizamiento y pueden inducir la aparición de manchas o patrones cloróticos/necróticos en las hojas, estrías y roturas del color en los pétalos y deformaciones en las inflorescencias (Stone, 1980; Nameth y Adkins, 1993; Alonso y Borja, 2005). Además de estos perjuicios en la calidad de las plantas, la presencia de estos agentes infecciosos limita las posibilidades de exportación del material producido ya que las normativas a este respecto son cada vez más estrictas. Así pues, las virosis del geranio repercuten de forma muy negativa en la comercialización de esta planta ornamental y a menudo conducen a importantes pérdidas económicas. Esta situación se ha agravado en los últimos años en los que se ha producido un aumento progresivo de las infecciones virales del geranio en Europa occidental propiciado, muy probablemente, por la habitual multiplicación del material por propagación vegetativa y por el hecho de que la mayoría de los virus sólo causan una sintomatología evidente bajo ciertas condiciones ambientales.

6.1. El virus del arabesco del *Pelargonium* (PLPV)

Una prospección realizada recientemente en España ha puesto de manifiesto que el PLPV es el agente de tipo viral más frecuente en geranio con porcentajes de incidencia que oscilan entre el 40 y el 90 % dependiendo del área geográfica examinada (Alonso y Borja, 2005). Este virus puede aparecer sólo o en infecciones mixtas, principalmente junto al PFBV, el CMV, el TMV o el PLCV. Una situación similar es previsible en países de nuestro entorno y, probablemente, en otros más alejados (Stone, 1980; Bouwen y Maat, 1992; Franck y Loebenstein, 1994), aunque no hay estudios detallados a este respecto.

6.1.1. Propiedades biológicas y modos de transmisión

El PLPV fue descrito por primera vez por Stone y Hollings (1971, 1976). La expresión de los síntomas inducidos por el PLPV en geranio se encuentra altamente influenciada por las condiciones externas y varían estacionalmente (Bouwen y Maat, 1992). Las infecciones causadas por el PLPV en *P. peltatum*, *P. glandiflorum* y *P. zonale*, sus huéspedes naturales, son frecuentemente asintomáticas, aunque esta última especie puede presentar manchas cloróticas y patrones de líneas en hojas (Brunt *et al.*, 1996; Alonso *et al.*, 1999) (Fig. 16).

La gama de huéspedes experimentales del PLPV incluye 37 especies distribuidas en 9 familias de plantas, incluyendo a *Beta vulgaris*, *Solanum lycopersicum*, *Ocimum basilicum*, *Phaseolus vulgaris*, *Pisum sativum*, *Solanum tuberosum* o *Vicia faba*, entre otros. La especie utilizada como huésped para el diagnóstico de este virus suele ser *Chenopodium quinoa* y el huésped más empleado en el mantenimiento y propagación del virus es *Nicotiana clevelandii* (Brunt *et al.*, 1996).

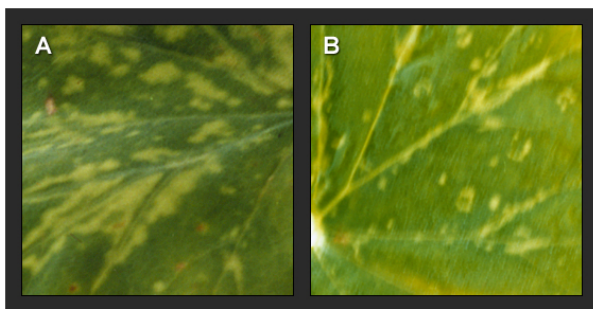


Fig. 16. Síntomas inducidos por el PLPV en hojas de geranio (fotos cedidas por M. Borja del vivero "La Veguilla", Fundación Promiva)

El PLPV se transmite por inoculación mecánica, por injerto y por contacto entre plantas. Hasta el momento no se ha descrito su transmisión por medio de ningún vector biológico (Brunt *et al.*, 1996).

6.1.2. Propiedades moleculares

A pesar de que la enfermedad causada por el PLPV es bien conocida desde hace tiempo, los datos moleculares acerca de este patógeno al inicio de este trabajo eran muy escasos. Concretamente se disponía de la secuencia nucleotídica del gen de la CP de un aislado del virus que había sido depositada en la base de datos (Kinard *et al.*, 1999). Está descrito como un virus muy resistente (punto de inactivación térmica de 85°C), que presenta partículas isométricas de 30 nm y un genoma monopartido de RNA de simple cadena y polaridad positiva de aproximadamente 4 kb (Brunt *et al.*, 1996), características que comparte con miembros de la familia *Tombusviridae* (Lommel *et al.*, 2005), aunque formalmente no ha sido adscrito a ningún grupo viral.

Objetivos

El PLPV es el agente de tipo viral más frecuente en geranio en España y, probablemente, en otros países del entorno. Las propiedades morfológicas y físico-químicas de este virus sugerían que podría tratarse de un nuevo miembro de la familia *Tombusviridae*, aunque la escasez de datos moleculares sobre el mismo no había permitido su adscripción formal a esta familia. Dada la importancia que el PLPV ha ido adquiriendo como patógeno de una planta ornamental ampliamente utilizada, en esta tesis se ha pretendido profundizar en sus características moleculares. Los objetivos concretos que se han planteado son los siguientes:

1.- Determinación de la organización genómica del PLPV, seleccionando un aislado del virus para llevar a cabo la secuenciación completa de su gRNA con el fin de tratar de establecer su posición taxonómica. Este objetivo se complementará con la identificación y caracterización de los RNAs virales que se acumulan en plantas infectadas para obtener datos preliminares acerca de las estrategias de expresión génica del PLPV.

2.- Obtención de un clon infeccioso del PLPV y análisis de la actividad biológica de los transcritos generados *in vitro* en diferentes huéspedes. Se pretende disponer de la herramienta necesaria para poder realizar un estudio detallado de las relaciones estructura-función en el PLPV.

3.- Análisis mutacional de las ORFs identificadas en el genoma del PLPV mediante experimentos de genética reversa con el clon infeccioso. Se persigue identificar ORFs funcionales, obtener información acerca del papel de dichas ORFs en el ciclo biológico del virus y dilucidar los mecanismos de traducción que el virus emplea para expresar todos sus genes.

4.- Estudio de la variabilidad molecular del PLPV. Para obtener datos acerca de las presiones selectivas a las que se encuentra sometido el genoma viral, se partirá de una colección de aislados de distinta procedencia y se secuenciará y comparará una región del genoma que abarca a los genes de las MPs y de la CP así como segmentos flanqueantes incluyendo la 3' UTR completa.

Capítulo I

Arch Virol (2005) 150: 949–965
DOI 10.1007/s00705-004-0464-y

**Archives of
Virology**
Printed in Austria

**Complete nucleotide sequence and genome
organization of Pelargonium line pattern virus
and its relationship with the family
*Tombusviridae***

A. Castaño and C. Hernández

Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC),
Universidad Politécnica de Valencia, Valencia, Spain

Received July 9, 2004; accepted November 10, 2004
Published online January 13, 2005 © Springer-Verlag 2005

Summary. The complete nucleotide sequence of Pelargonium line pattern virus (PLPV) has been determined. The PLPV genomic RNA comprises 3884 nt and contains six open reading frames (ORFs) potentially encoding proteins of 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), and 37 kDa (p37), respectively. The arrangement of these ORFs on the PLPV genome closely resembles that of members of the genus *Carmovirus* in the family *Tombusviridae* and, moreover, most of the putative PLPV gene products showed high identity with proteins of this viral group. However, several striking differences were noticed. Carmoviruses generate two subgenomic RNAs whereas PLPV produces a single one. In addition, only p7 showed similarity with movement proteins of carmoviruses whereas p6 (as p13) has no viral (or other) homologs. This protein might be expressed from a non-canonical start codon or, alternatively, through a -1 frameshift (FS) mechanism. Both, the production of one subgenomic RNA and the likely involvement of a -1 FS for expression of an internal ORF parallel the translation strategies reported for the unique species of the genus *Panicovirus*, belonging also to the family *Tombusviridae*. Overall, the results support the placement of PLPV in this family although its peculiar characteristics preclude its direct assignment to any of the current genera.

The nucleotide sequence reported has been deposited in the GenBank databases under the accession number AY613852.

Introduction

Pelargonium line pattern virus (PLPV) has a broad experimental host range including more than 30 plant species from nine families, although reports on PLPV infections in nature are limited to geranium (*Pelargonium* spp.). This virus may induce the appearance of yellow/green spots and line patterns on the leaves of the affected plants leading to significant quality deterioration [42, 43]. In Spain, a survey done on approximately one thousand geranium samples collected from different regions of the country revealed the presence of PLPV in more than 50% of the tested scions (M. Borja, personal communication). This situation is probably similar to that existing in other countries where this ornamental crop has been largely cultivated, as studies made in the nineties showed a high incidence of PLPV in the Netherlands and Israel [4, 11]. PLPV is transmitted by vegetative propagation and mechanical inoculation, and no vector has been identified.

PLPV virions are icosahedral in shape, about 30 nm in diameter and contain a linear positive sense single stranded RNA of approximately 4 kb [35, 44]. The taxonomic position of PLPV is unclear. It was considered as a tentative member of the genus *Carmovirus* (family *Tombusviridae*) on the basis of its morphological and physico-chemical properties [32], but the last report of the ICTV does not include PLPV within any established viral group or in the list of unassigned viruses [51]. This is most likely due to the poor characterization of PLPV and of other small isometric viruses infecting geranium which, as a consequence, has led to some controversies about their identity and to a confusing nomenclature [4, 18, 19, 27, 41, 43]. This situation underlines the need for additional information about these pathogens to facilitate their discrimination and to shed light on their phylogenetic relationships. A partial sequence corresponding to the PLPV coat protein gene has only recently become available at the sequence databases (Accession No. AY038067). In this work we report the entire primary structure of the PLPV genomic RNA and compare the genomic organization and the deduced amino acid sequences from its ORFs with those of related viruses. In addition, we have characterized the single PLPV subgenomic RNA detected in infected plants to gain an insight into the expression strategies of this virus. The results support the inclusion of PLPV in the family *Tombusviridae* but its assignment to genus level is still not unequivocal, and is discussed.

Material and methods

Virus source

PLPV isolate PV-0193 was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and maintained by mechanical sap inoculation onto the experimental hosts *Chenopodium quinoa* and *Nicotiana clevelandii*.

RNA extraction and analysis

Total RNA preparations of healthy and infected *C. quinoa* and/or *N. clevelandii* leaves were obtained by phenol extraction and lithium precipitation [52]. Double stranded (ds) RNAs were purified from infected tissue by total nucleic acid extraction and subsequent partitioning

on nonionic cellulose (CF11, Whatman) with STE (50 mM Tris-HCl pH 7.2, 100 mM NaCl, 1 mM EDTA) containing 16% ethanol [33]. The dsRNAs were separated in non-denaturing 5% polyacrylamide gels and after ethidium bromide staining, the bands corresponding in size to the PLPV genomic and subgenomic RNAs were eluted and recovered by ethanol precipitation.

For Northern blot analysis, 4 µg of total RNA from healthy and PLPV infected leaves were denatured by glyoxal-dimethyl sulfoxide treatment, electrophoresed in 1% agarose gels, blotted to nylon membranes (Scheicher & Schuell), and cross-linked by UV irradiation. The membranes were hybridized with ³²P-radioactive DNA probes at 42 °C for at least 12 h in the presence of 50% formamide, 5 × SSC and 5% sodium dodecyl sulfate (SDS). The probes, prepared with Ready to go DNA labelling beads (Amersham Biosciences), corresponded to selected PLPV genome regions: a 5' terminal probe which included from nt 1 to nt 181 of the PLPV genomic RNA, and a 3' terminal probe which comprised from nt 3095 to 3884. After hybridization, membranes were washed for 15 min, twice at room temperature in 2 × SSC plus 0.1% SDS and once at 60 °C in 0.1 × SSC plus 0.1% SDS.

cDNA synthesis, cloning and sequencing

A partial PLPV cDNA clone (PL-1030) containing an uncharacterized PLPV cDNA ligated to the pUC9 vector, was provided by the DSMZ plant virus collection. Sequencing of the insert termini allowed the design of two PLPV specific oligonucleotides, CH17 (5'-GAAAATGGCCTTCTACGGGGAC-3'), homologous to nt 2067–2088 of the PLPV genomic RNA, and CH18 (5'-GCCTCCCTGCTGGCATAAACTAATAC-3'), complementary to nt 3578–3603. These primers were used for RT-PCR reactions with Titan One Tube RT-PCR system (Roche) employing total RNA preparations from PLPV infected leaves as templates. The amplified cDNA product of 1536 bp was inserted into the pGEM-T easy vector (Promega) and after determination of its sequence, new oligonucleotides were designed to fully characterize the primary structure of PLPV. Oligonucleotide CH32 (5'-CTGACCCCAAACGCCAAGTGAATG-3'), complementary to nt 2156–2180, was used together with a degenerate primer, able to anneal to carmovirus related polymerase sequences (31), in RT-PCR reactions that yielded a cDNA of approximately 1000 bp. To obtain clones containing the 5' and 3' termini of PLPV genomic RNA, the corresponding dsRNA was denatured with 10 mM methylmercuric hydroxide and polyadenylated using yeast poly(A) polymerase (U.S. Biochemicals). After phenol–chloroform extraction and ethanol precipitation, the polyadenylated RNAs were subjected to reverse transcription (RT) with Superscript II-RT (Invitrogen) using primer GeneRacer Oligo dT (Invitrogen), which has a dT tail of 18 nt at its 3' end and contains the priming sites for the GeneRacer 3' and GeneRacer 3' Nested oligonucleotides (Invitrogen) at 5' end. RT products were PCR amplified with the Expand High Fidelity PCR System (Roche) using primers GeneRacer 3' and GeneRacer 3' Nested in combination with others derived from internal regions: CH43 (5'-TAGCTACTGGTTTCTGGAATTGGAG-3') or CH46 (5'-TGGCCAACCCATGGACGCCAATCCG-3'), complementary to nt 1378–1402 and nt 158–181, respectively, for the 5' terminus, and CH36 (5'-ATAGAGGCGGAGCAACCAGGTAGTGC-3') or CH53 (5'-CCTTAATGAAGAACGGGGGGTAACC-3'), homologous to nt 3548–3573 and nt 3840–3865, respectively, for the 3' terminus.

Complete cDNA clones of the PLPV subgenomic RNA were obtained using the single-primer method for cloning dsRNAs with unknown sequences [17, 54] with minor modifications. Briefly, the purified dsRNA was polyadenylated and reverse transcribed with primer GeneRacer Oligo dT. After treatment with RNase H (Roche), the cDNA complementary strands were annealed for 10 min at 80 °C, 16 h at 65 °C and 3 h at 30 °C and the resulting

hybrid was filled in with T4 DNA polymerase (Roche) and PCR amplified with Expand High Fidelity PCR System (Roche) and primer GeneRacer Oligo dT.

PCR products were separated by electrophoresis in 1% agarose gels or in non-denaturing 5% polyacrilamide gels, depending on the expected size. The DNA fragments were eluted and cloned into the pGEM-T easy vector (Promega) or the plasmid pTZ19R (Fermentas). At least two clones were selected for each DNA fragment and the nucleotide sequence of the inserts was determined with an ABI PRISM DNA sequencer 377 (Perkin-Elmer).

Primer extension assays

The 5' terminal sequence of PLPV genomic RNA and the transcriptional start site of the subgenomic RNA were corroborated by primer extension assays. The substrates for the RT reactions were total RNA preparations from PLPV infected tissue in conjunction with primer CH46 (to map the 5' end of the genomic RNA) or primer CH50 (5'-GCACCGA CAAGATCGCGAGATGTAC-3'), complementary to nt 2386–2410 (to map the 5' end of the subgenomic RNA). RT reactions were performed with Superscript-RT II in the presence of (α -³²P)(dCTP). The RT products were run in sequencing gels together with sequencing ladders obtained with the corresponding primer and proper recombinant plasmids with PLPV derived DNA inserts.

Sequence analysis

Analysis of the coding capacity of the nucleotide sequences was carried out with the TRANSLATE program (available at the ExpASY proteomics server; <http://www.expasy.org/tools/dna.html>). Database searches for nucleotide and protein similarities were done using the programs BLASTN, BLASTP and BLASTX (www.genome.ad.jp, www.ncbi.nlm.nih.gov, www.ebi.ac.uk). Pairwise comparisons of the PLPV nucleotide and amino acid sequences with those of related viruses were performed using the ALIGN program at the Genestream network (<http://www2.igh.cnrs.fr/bin/align-guess.cgi>) [34]. The RNA secondary structure of the PLPV 3' region was predicted at 37 °C using MFOLD version 3.1 [30, 56]. Multiple alignments of protein sequences were generated with the CLUSTAL X program [47]. These alignments were used to conduct phylogenetic analysis employing programs of the MEGA 2.1 package (24). Sequences used for comparison were retrieved from the sequence databases and correspond to the following accession numbers: CarMV (*Carnation mottle virus*; X02986), TCV (*Turnip crinkle virus*, M22445), MNSV (*Melon necrotic spot virus*, M29671), CCFV (*Cardamine chlorotic fleck virus*, L16015), CPMoV (*Cowpea mottle virus*, U20976), SCV (*Saguaro cactus virus*, NC001780), GaMV (*Galinsoga mosaic virus*, Y13463), HCRSV (*Hibiscus chlorotic ringspot virus*, X86448), JINRV (*Japanese iris necrotic ring virus*, NC002187), PSNV (*Pea stem necrosis virus*, NC004995), PFBV (*Pelargonium flower break virus*, AJ514833), PoLV (*Pothos latent virus*, X87115), OCSV (*Oat chlorotic stunt virus*, X83964), CRSV (*Carnation ringspot virus*, L18870), RCNMV (*Red clover necrotic mosaic virus*, J04357), SCNMV (*Sweet clover necrotic mosaic virus*, L07884), MCMV (*Maize chlorotic mottle virus*, X14736), TNV-A (*Tobacco necrosis virus-A*, M33002), TNV-D (*Tobacco necrosis virus-D*, D00942), OLV-1 (*Olive latent virus 1*, X85989), LWSV (*Leek white stripe virus*, X94560), PMV (*Panicum mosaic virus*, U55002), TBSV (*Tomato bushy stunt virus*, M21958) AMCV (*Artichoke mottled crinkle virus*, X62493), CIRV (*Carnation Italian ringspot virus*, X85215), CuNV (*Cucumber necrosis virus*, M25270), CymRSV (*Cymbidium ringspot virus*, X15511), and PNSV (*Pelargonium necrotic spot virus*, AJ607402). The coat protein sequences of PelRSV (*Pelargonium ringspot virus*, AY038068), PCRPV (*Pelargonium chlorotic ring pattern virus*, AY038069), and ELV (*Elderberry latent virus*, AY038066) were also used for phylogenetic analysis.

Results

Nucleotide sequence and coding regions

The complete nucleotide sequence of PLPV and its deduced amino acid sequences have been determined. The PLPV genomic RNA consists of 3884 nt and contains six open reading frames (ORFs) (Fig. 1). The first ORF potentially encodes a protein of 27 kDa (p27). This ORF(p27) begins at AUG (nt 7–9) in a favourable translation context (CAAAAUGGCAC), with an A at position –3 and a G at position +4 (23, 28), and terminates at an UAG codon (nt 727–729). If this amber codon is read-through, the ORF extends to a downstream UAA (nt 2302–2304) giving rise to a protein of 87 kDa. Both p27 and p87 present homology with proteins of members of the family *Tombusviridae* (Table 1) which have been shown

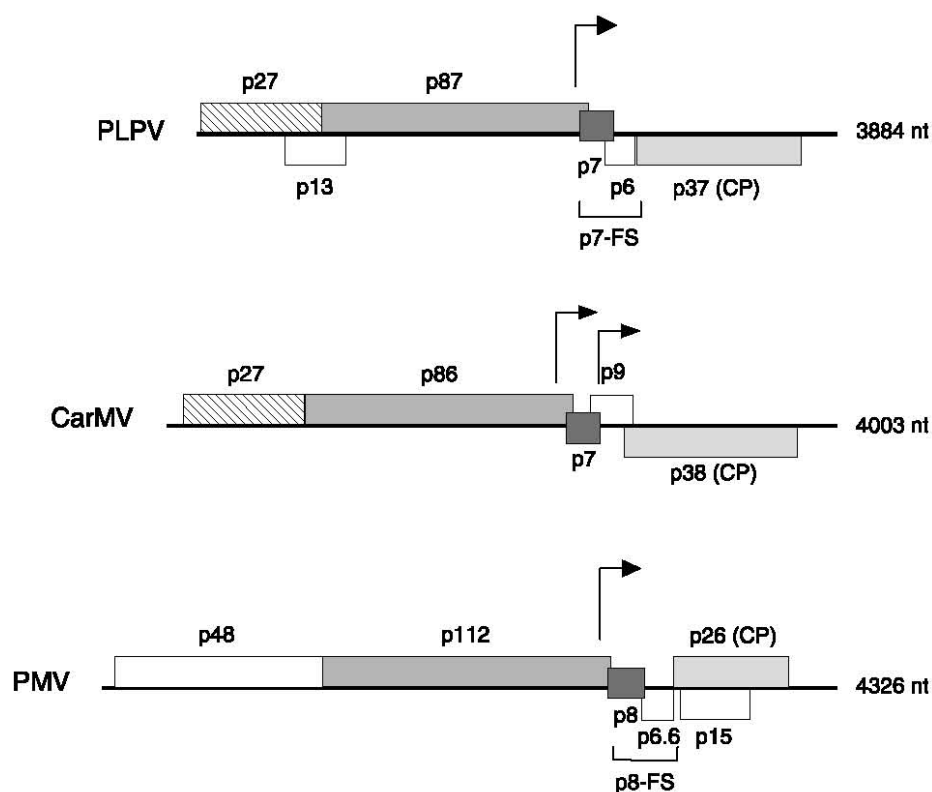


Fig. 1. Comparison between the genome organization of PLPV and those of the type species of the genera *Carmovirus* (*Carnation mottle virus*, CarMV) and *Panicovirus* (*Panicum mosaic virus*, PMV) in the family *Tombusviridae*. The ORFs are indicated by boxes which are similarly shaded when the encoded proteins share significant amino acid identity. The transcription start sites of the corresponding subgenomic RNAs are marked by arrows. Both p6 and p6.6 of PLPV and PMV, respectively, might be translated from non-canonical initiation codons. A –1 frameshift (FS) event could be required for the generation of the putative p7-FS and p8-FS proteins of PLPV and PMV, respectively. These potential frameshift products would have an N-terminus corresponding to p7 (for PLPV) or p8 (for PMV), and a C-terminus corresponding to almost the complete p6 (for PLPV) or the complete p6.6 (for PMV)

Table 1. Sequence identity (%) between the predicted PLPV proteins and those of viruses from the different genera of the family *Tombusviridae*^a

Virus	Proteins				
	p27	p86	RT ^b	p7	CP
Carmoviruses					
CarMV	25.7	43.2	51.8	37.9	31.0
TCV	28.2	41.8	49.2	34.7	26.5
MNSV	24.5	39.8	47.6	27.5	24.6
CCFV	26.8	42.7	50.5	39.4	26.0
CPMoV	24.4	38.9	45.8	27.5	25.9
SCV	28.2	42.3	49.1	46.0	34.9
GaMV	21.5	35.8	42.6	23.2	23.1
HCRSV	28.2	44.0	51.3	38.6	27.6
JINRV	27.9	40.2	46.1	27.6	27.7
PSNV	25.8	39.8	46.2	25.4	20.5
PFBV	32.8	43.4	48.4	35.7	33.5
Aureusviruses					
PoLV	20.0	33.8	40.7	– ^c	24.3
Avenaviruses					
OCSV	21.9	32.0	36.5	18.9	20.6
Dianthoviruses					
CRSV	21.7	26.9	28.8	–	21.6
RCNMV	19.8	28.7	34.2	–	21.4
SCNMV	18.5	28.5	32.1	–	21.3
Machlomoviruses					
MCMV	18.0	33.4	46.2	22.1	18.8
Necroviruses					
TNV-A	26.0	38.0	44.2	19.4	17.5
TNV-D	21.3	33.7	40.0	25.4	18.2
OLV-1	26.4	38.5	44.4	26.7	17.2
LWSV	15.4	31.4	38.9	20.8	16.8
Panicoviruses					
PMV	17.2	32.5	44.0	27.0	16.7
Tombusviruses					
TBSV	19.2	33.3	41.2	–	24.1
AMCV	17.9	32.9	41.6	–	21.3
CIRV	19.0	32.4	41.1	–	18.9
CuNV	20.6	33.9	42.0	–	22.8
CymRSV	19.9	33.0	40.9	–	26.2
PNSV	20.2	33.2	41.2	–	24.5

^aSee the Methods section for viral abbreviations and accession numbers of the sequences used for comparison

^bRead-through portion of the RdRp

^cHomologue is absent

to be involved in replication [1, 14, 55] and, moreover, p87 contains the motifs characteristic of the viral RNA dependent RNA polymerases (RdRp) including the highly conserved GDD box [22]. Another ORF, starting at AUG (nt 530–532, in the context CCGGAUGAAAA) and ending at UAA (nt 911–913), is nested between ORFs (p27) and (p87) and potentially encodes a protein of 13 kDa (p13) that is unique to this virus.

A small centrally located ORF begins at AUG (nt 2274–2276, in the context UCGAAUGGAUA) and stops at UGA (nt 2463–2465), and may encode a protein of about 7 kDa (p7) which exhibits high identity with movement proteins from species of the family *Tombusviridae*, especially with those of the genus *Carmovirus* (Table 1). In this PLPV putative product, a domain adopting a characteristic α -helix, similar to that able to bind RNA in the p7 of CarMV (29), can be predicted (data not shown). An additional small ORF, overlapping ORF(p7), potentially encodes a protein of 6 kDa (p6) which does not show clear similarity with any viral (or other) protein. This ORF might be expressed from a non-canonical start codon, AUU (nt 2441–2443, in the context UGGAAUUUGUG), and/or by a -1 frameshift (FS) mechanism which would lead to the production of a 12 kDa fusion protein (p12 or p7-FS). Supporting the latter possibility, an heptanucleotide is found immediately upstream of the stop codon of ORF(p7) whose sequence (Fig. 2) fits the canonical frameshifting motifs (5, 6). Moreover, a stem-loop is predicted downstream of the putative shifty site (data not shown), a type of structure which seems to favour frameshifting (reviewed in [12]).

Finally, the 3' proximal ORF of PLPV, starting at AUG (nt 2621–2623, in the context GGAAAUGGCGG) and ending at UGA (nt 3635–3637), is predicted to encode a 37 kDa coat protein (p37 or CP), as indicates the high amino acid identity found between this product and proteins involved in encapsidation in the genus *Carmovirus* and in other genera of the family *Tombusviridae* (Table 1). The three different structural domains described in this type of proteins (39) can be also distinguished in the PLPV CP: (i) R, the N-terminal internal domain which

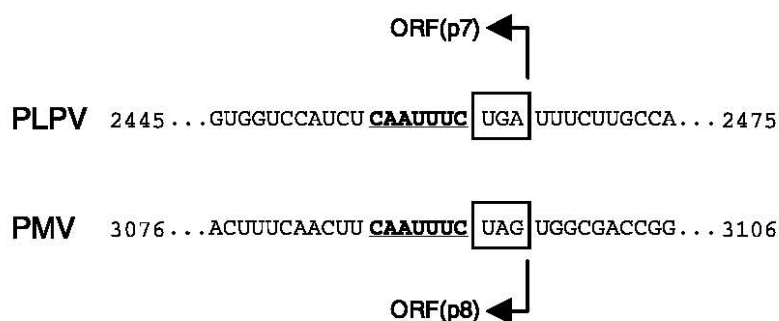


Fig. 2. Comparison of the nucleotide sequences of equivalent genome regions of PLPV and PMV. The heptanucleotide which fits the canonical frameshifting motifs [5] is shown in bold underlined letters. The stop codons of the PLPV ORF(p7) and PMV ORF(p8) are boxed

contains many positively charged residues and interacts with RNA, (ii) S, the shell domain which forms a barrel structure made up of β strands and constitutes the capsid backbone and, (iii) P, the protruding C-terminal domain.

When overlapping cDNA clones were compared, eighteen nucleotide substitutions were detected within coding regions. Only eight of these point mutations resulted in amino acid changes, five affecting the p27 and/or the p87 (two of them non-conservative, substitutions P \rightarrow S and R \rightarrow W at positions 40 and 336, respectively, of p87), two the p13 (both non-conservative, substitutions S \rightarrow W and A \rightarrow D at positions 64 and 86, respectively), and one the CP.

Comparison of nucleotide sequence of the CP gene of the PLPV isolate characterized in this work with that available at the databases showed an overall sequence identity of 93.7%. The corresponding protein sequences differed at nine amino acid positions including two non-conservative changes affecting the S and P domains (substitutions E \rightarrow G and L \rightarrow Q at positions 140 and 296, respectively).

5' and 3' untranslated regions of PLPV genomic RNA

The 5' and 3' terminal sequences of the PLPV genomic RNA were determined by sequencing of cDNA clones derived from these regions as indicated in the Methods section. This approach was complemented by primer extension assays (Fig. 3A) which confirmed that the 5' untranslated region (UTR) of PLPV RNA comprises only 6 nt, to our knowledge, the shortest 5' UTR described so far for the genomic RNA of a plant virus. The closest examples in this respect correspond to subgenomic RNAs encoding viral coat proteins as that of *Brome mosaic virus* (BMV) which has a 5' UTR composed of only 9 nt [3]. It is also worth mentioning that the primer extension assays suggested that the PLPV genomic RNA lacks a 5' terminal cap structure since the cDNA doublet normally associated with a capped RNA was absent [2, 13].

The 3' UTR of PLPV RNA contains 247 nt. As observed in most viral species of the family *Tombusviridae*, the 3' terminal sequence may fold into a stem-loop which is particularly stable in the case of PLPV because of the presence of six contiguous G–C pairs in the stem (Fig. 4). Studies with several representatives of the family *Tombusviridae* have shown that this type of structure is required for the efficient *in vivo* replication of the viral RNAs [7, 9, 10, 46, 50]. In addition, a midrange base-pairing interaction between the 5 nt in the 3' terminus and a sequence located about 50 nt upstream, similar to that described recently as a replication silencer in *Tomato bushy stunt virus* and related viruses [36], can also be predicted for PLPV (Fig. 4).

Identification and characterization of a PLPV subgenomic RNA

As stated above, analysis of the coding capacity of PLPV predicts the synthesis of at least three proteins from downstream ORFs. To determine if subgenomic RNAs are generated during PLPV infection which could serve as mRNAs for the production of these proteins, total RNA from *C. quinoa* leaves infected by PLPV

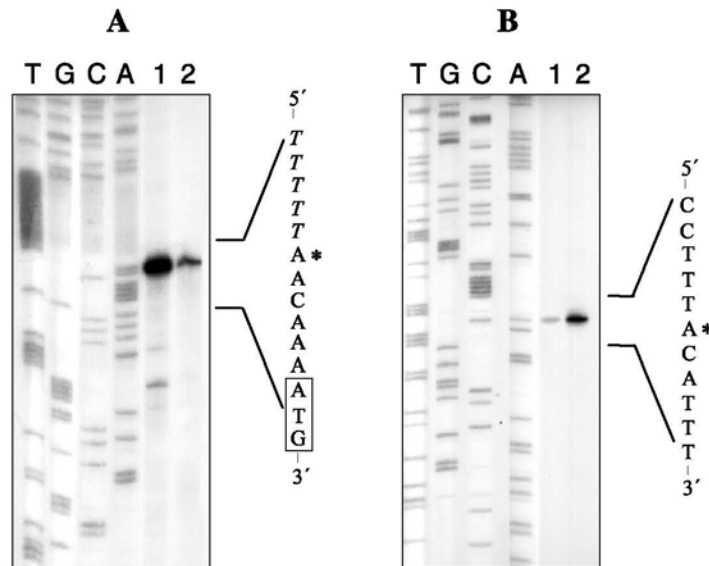


Fig. 3. Determination of the 5' termini of the PLPV genomic (A) and subgenomic (B) RNAs by primer extension assays. Reverse transcriptions (RT) were performed using oligonucleotides complementary to nt 158–181 (A) or nt 2386–2410 (B) of the PLPV genome as primers and two distinct total RNA preparations (1 and 2) from *C. quinoa* PLPV infected leaves as templates. The RT products were run in sequencing gels together with sequencing ladders obtained with the same primers employed in the RT reactions and proper PLPV cDNA clones. The T tail derived from the cDNA cloning procedure is shown in italics. The 5' terminal nucleotides of the genomic (A) and subgenomic (B) RNAs are indicated by asterisks, and the initiation codon of the ORF(p27) is boxed

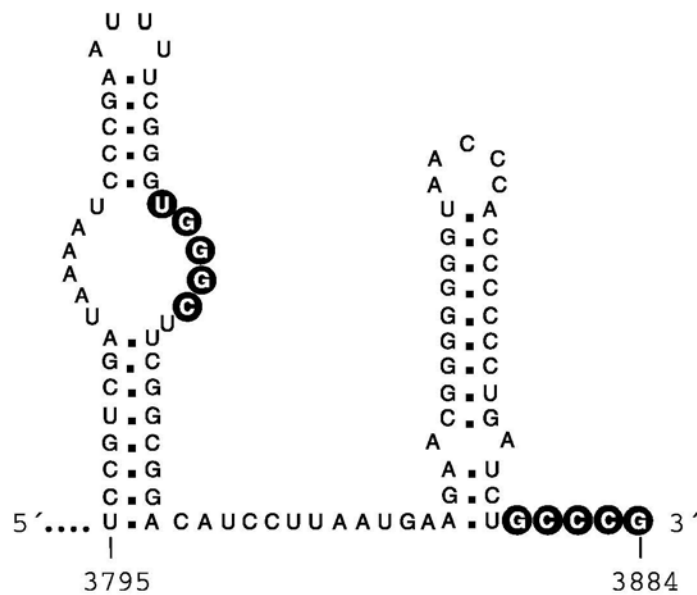


Fig. 4. MFOLD-predicted RNA secondary structure of the 3'-terminal region of PLPV. The nucleotides that could participate in a silencer 3'-terminus interaction are white in black circles. Numbers below correspond to positions in the genomic RNA

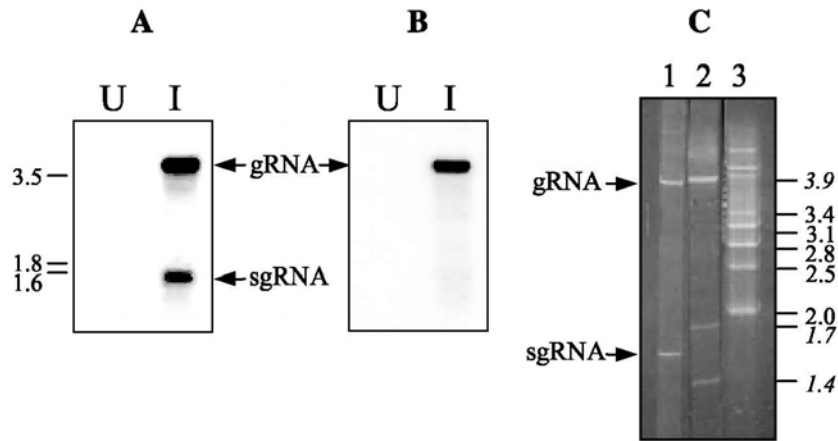


Fig. 5. Analysis of PLPV RNAs. **A** and **B.** Northern blot hybridizations were performed on total RNA preparations from uninfected (*U*) and PLPV-infected (*I*) *C. quinoa* leaves using ^{32}P -labeled probes derived from the 3' (**A**) or 5' (**B**) terminus of the PLPV genomic RNA. The positions of single stranded RNA molecular weight markers (in kb), corresponding to *in vitro* transcripts, are indicated at the left margin in **A**. **C** PAGE analysis in 5% non-denaturing polyacrilamide gels of PLPV dsRNAs (*I*). The dsRNAs of *Pelargonium flower break virus* (PFBV) and the mycoviral-like ones associated to Cherry chlorotic rusty spot disease (Covelli et al. and Coutts et al., in press) have been included for comparison (2 and 3, respectively) with the sizes (in kbp) of some of them shown at the right margin (in italics those corresponding to the dsRNAs of PFBV). In all panels, the bands corresponding to the PLPV genomic (*g*) and subgenomic (*sg*) RNAs are indicated with arrows

were analyzed by Northern blot hybridization with cDNA probes derived from different regions of the genomic RNA. When a 3' terminal probe was used, a strongly hybridizing RNA species of about 1.6 kb was detected in infected tissue in addition to the genomic size RNA (Fig. 5A). Such 1.6 kb RNA corresponded very likely to a subgenomic RNA since a 5' specific probe did not hybridize with it (Fig. 5B).

Analysis of dsRNA extracts from PLPV infected tissue revealed the presence of two dsRNA species of about 4.0 kbp and 1.6 kbp, in agreement with the results of the Northern blot (Fig. 5C). To determine the precise transcription start site of the 1.6 kb subgenomic RNA, the corresponding dsRNA was eluted from the gel, reverse-transcribed and cloned using the single-primer approach described in the Methods section. Sequencing of some of the clones showed that the 5' terminus of the PLPV subgenomic RNA maps 23 nt upstream (position 2251 of the genomic RNA) from the initiation codon of ORF(p7). Primer extension analysis confirmed this result and suggested that such RNA lacks a 5' terminal cap structure as seems to be the case for the genomic RNA (Fig. 3B). The molecular structure of the PLPV subgenomic RNA suggests that is generated for the translation of the small internal ORFs and the CP ORF having, therefore, the potential to encode at least three proteins.

Phylogenetic analysis

Multiple alignments with the sequences of different proteins of PLPV and those from members of the family *Tombusviridae* were generated and used for phylogenetic analysis. Comparisons based on the p27 and p7 homologs yielded tentative phylogenetic trees in which the distinction by genera was not clearly discernible since the viruses showed a tendency to form a continuum rather than to cluster (data not shown) in agreement with recent reports [8, 45]. When the p87 homologs were used to perform the analysis, the resulting tree exhibited a better correlation between grouping and genera (with the same exceptions reported in 45) and revealed a close association of PLPV with carmoviruses (Fig. 6A). In the analysis based on the CP sequences, those corresponding to three additional isometric viruses infecting *Pelargonium* spp., PeIRSV, PCRPV and ELV, were included. The CP tree showed PLPV grouped with most carmoviruses but, remarkably, appeared tightly associated to PeIRSV, PCRPV and ELV, three viral entities whose taxonomic position remains to be defined (see below).

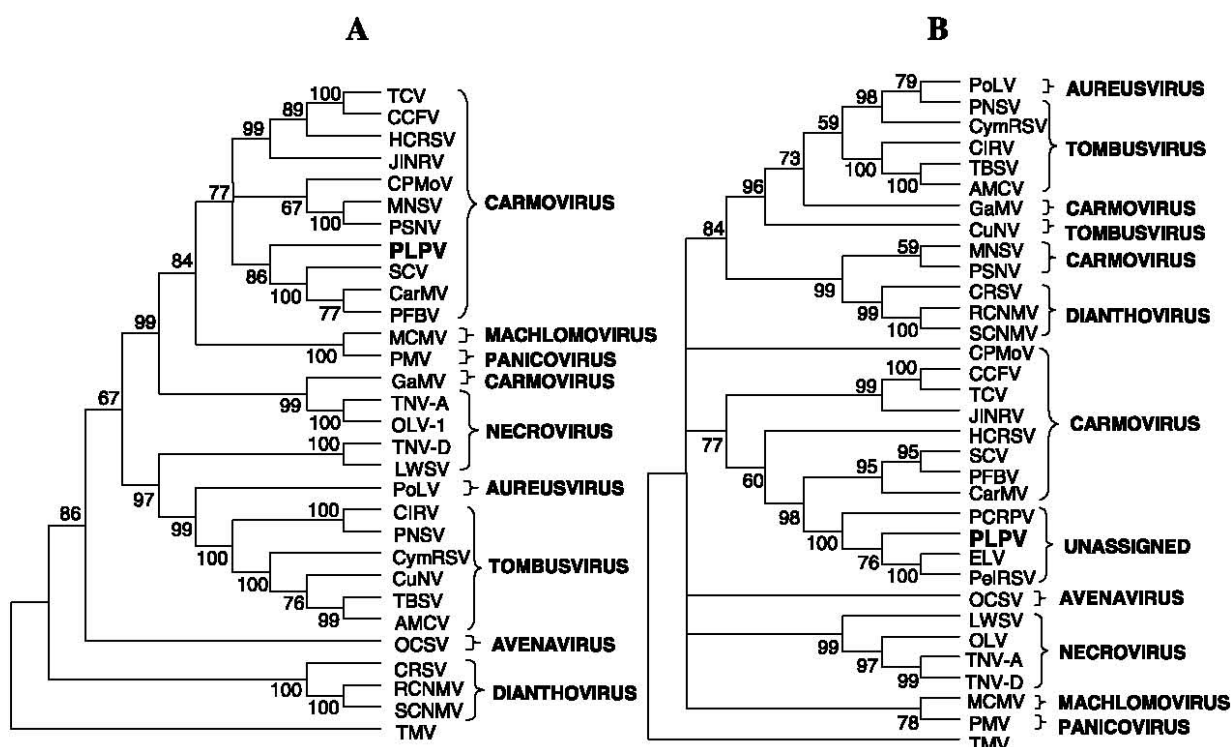


Fig. 6. Phylogenetic trees inferred from RNA dependent RNA polymerase (A) and coat protein (B) sequences of PLPV and members of the family *Tombusviridae*. Virus name abbreviations and accession numbers of the sequences used to construct the tree are indicated in the Methods section. Phylogenetic analysis was conducted with programs included into the MEGA 2.1 package. The numbers at the nodes indicate the number of times out of 10000 trees that this grouping occurred after bootstrapping the data; only values >50% are shown. The unrelated sequences of the corresponding proteins of *Tobacco mosaic virus* (TMV, accession number D38444) were used as outgroups

Discussion

We have determined the complete 3884 nt sequence of the genomic RNA of PLPV. The arrangement of ORFs in this virus resembles that of members of the genus *Carmovirus* and, moreover, phylogenetic analyses have shown that PLPV is closely related to this viral group. However, PLPV presents some distinctive features that prevent its direct assignment to this genus. The central region of the genomic RNA of carmoviruses contains two small out-of-frame ORFs whose translation gives rise to two proteins, p7 (or p8) and p9, involved in cell-to-cell movement [14, 25]. PLPV genome is not predicted to code for a product homologue of the second movement protein of carmoviruses and presents, instead, a small ORF which potentially encodes a polypeptide, p6, with no clear similarity to any viral protein. In addition, only one subgenomic RNA seems to be produced for translation of downstream PLPV ORFs, whereas carmoviruses generate two subgenomic RNAs for expression of the movement proteins and the CP, respectively [39].

The ORF(p13) is also unique to PLPV although its location on the viral genome is similar to that of the novel ORF(p23) of *Hibiscus chlorotic ringspot carmovirus* which encodes a protein indispensable for host specific replication [26]. Database searches have not demonstrated similarities between the putative PLPV p13 and any known protein and, moreover, computer-aided sequence analysis has not revealed any motif in p13 which could provide an insight about its function. Hence, the biological significance of p13, if produced in infected plants, is yet to be determined.

Caution is required when the possible translation events on RNA are deduced only from sequence data but should all proteins potentially encoded by PLPV be produced *in vivo*, the virus must use a variety of mechanisms for expression of its genes. Translation of PLPV ORFs (p27) and (p87) will probably occur from the genomic RNA involving a read-through mechanism in the case of ORF(p87), as demonstrated for related ORFs of members of the family *Tombusviridae* [16, 40, 48, 55]. Indeed, the sequence surrounding the amber stop codon of ORF(p27) is AAA-UAG-GGA, which is consistent with the consensus sequence required for efficient read-through AA(A/G)-UAG-G(G/U)(G/A) [16]. In the case of ORF(p13), an internal ribosomal entry is likely for its translation because the optimal sequence context of the initiation codon of ORF(p27) makes its expression via a leaky scanning mechanism unlikely [12]. The remaining downstream ORFs may be translated from the abundant 1.6 kb subgenomic RNA detected in PLPV infected plants. ORF(p7) would be expressed by conventional ribosomal scanning whereas ORF(p6) might be translated from a non-canonical start codon by leaky scanning and/or by a -1 FS mechanism giving rise to a p12 fusion protein (p7-FS). This also distinguishes PLPV from carmoviruses, whose small internal ORFs present canonical initiation codons and are presumably translated by a leaky scanning mechanism, with the exception of *Melon necrotic spot virus* for which a read-through event has been predicted [38]. While we cannot exclude the production of a subgenomic RNA accumulating to very low amounts for the CP gene's expression,

over exposure of the Northern blots did not reveal such potential RNA species. Thus, a leaky scanning mechanism or an internal ribosomal entry may account for its translation.

Interestingly, the arrangement and predicted expression strategies of the ORFs on the PLPV subgenomic RNA parallel those proposed for the unique subgenomic RNA of *Panicum mosaic virus* (PMV), the sole member of the genus *Panicovirus* in the family *Tombusviridae*, with the exception of an additional ORF overlapping the CP gene within PMV (Fig. 1) [48]. Both subgenomic RNAs have a region upstream of the CP gene devoid of AUG codons except for the one opening the 5' proximal ORF which, moreover, has not an optimal translation context. Furthermore, the proposed “shifty” heptanucleotide of PLPV is identical in sequence and in an equivalent position to one found immediately upstream of the stop codon of the PMV ORF(p8) (Fig. 2), thus reinforcing the hypothesis of the production of a frameshift fusion protein. Attempts to detect the putative PMV p8-FS *in vivo* have failed however [49]. In contrast, *in vitro* translation assays have demonstrated that the non-canonical start codon of PMV ORF(p6.6) is active despite its suboptimal initiation context (CUAGUGG). Subsequent mutagenesis experiments have shown this ORF is required for efficient infection in plants [49]. Despite not share clear sequence similarity, the PMV p6.6 and the potential PLPV p6, also presumably translated from a non-canonical codon in a suboptimal context, do have similar hydrophobicity profiles and, moreover, resemble those of the p9 movement proteins of carmoviruses (data not shown). Only one membrane-spanning domain however is predicted in the PMV p6.6 and the PLPV p6 whereas the carmovirus proteins contain two transmembrane segments [53]. Bioassays with PMV mutants have suggested that the p6.6 plays a role in virus spread [49] and, considering the analogies found between the two proteins, an equivalent function can be anticipated for the PLPV p6.

A series of small isometric viruses naturally restricted to geranium has been described including PFBV, *Pelargonium leaf curl virus* (PLCV), PNSV, PeIRSV, PCRPV, and PLPV. The characterization of these viruses has been mainly based on their biological/physicochemical properties and serological relationships, raising questions as to their taxonomy and status as distinct virus species [4, 18, 19, 27, 41, 43]. This has been further complicated by designating the same virus with different names, for example PLPV, which was formerly known as Pelargonium ring pattern virus [4]. Information on the genomic organization of these pathogens will undoubtedly shed light on this issue. The primary structures of PFBV, a carmovirus, and of PNSV, a proposed new species of the genus *Tombusvirus*, have been recently reported [15, 37]. In addition, some authors have determined the complete nucleotide sequence of other pelargonium viruses such as PCRPV, PeIRSV and also PLPV [20] but have not made the data public. Recently, the genomic organization of PCRPV which potentially encodes five proteins very similar to those involved in replication (p27 and p87), movement (p7 and p9) and encapsidation (p37) in the genus *Carmovirus* has been described [21]. Interestingly, PCRPV resembles PLPV in producing a single subgenomic RNA

for expression of downstream ORFs and in the likely involvement of a -1 FS mechanism for translation of the second small internal ORF. It has been suggested that such traits are also shared by PeIRSV and ELV [21] and, based on these common features and sequence similarities, it has been proposed to place the four viruses in a new genus within the family *Tombusviridae* for which the name *Pelarspovirus* has been advanced [21]. Only the CP sequences of PCRPV, PeIRSV and ELV are available and show high identity with that of PLPV (44, 48 and 50%, respectively). Consistently, the phylogenetic tree based on CP sequences grouped the four pathogens together (Fig. 6). However, the limited sequence information has not allowed us to further analyze similarities and differences between these viruses and, therefore, commenting on the taxonomic proposal done.

To conclude, the results of this study support the classification of PLPV as a member of the family *Tombusviridae*. Although the PLPV genomic organization and the characteristics of its putative proteins indicate a close association with members of the genus *Carmovirus*, the uniqueness in the arrangement of some ORFs and proposed expression strategies hinder the direct assignment of the virus to this genus. The detailed description of the genomes of PCRPV, PeIRSV and ELV, which resemble PLPV, will allow the establishment of more accurate relationships and the performance of phylogenetic analysis based on whole genome sequences as described recently [45]. We are currently attempting to elucidate the PLPV translation mechanisms and to ascertain the role of the novel ORFs found through the viral genome.

Acknowledgements

We thank Drs. R. Flores and V. Pallás for critical reading of the manuscript and suggestions, and Dr. M. Borja for valuable comments in the course of this work. This research was supported by grants AGL2000-0942 and AGL2003-04249 from the Ministerio de Ciencia y Tecnología de España and by grant Grupos 03/064 from the Generalitat Valenciana (to C. H.). A. C. was the recipient of predoctoral fellowships from the Generalitat Valenciana and from CSIC-Fundación Bancaja.

References

1. Andriessen M, Meulewaeter F, Cornelissen M (1995) Expression of *Tobacco necrosis virus* open reading frames 1 and 2 is sufficient for the replication of satellite *Tobacco necrosis virus*. *Virology* 212: 222–224
2. Ahlquist P, Janda M (1984) cDNA cloning and *in vitro* transcription of the complete *Brome mosaic virus* genome. *Mol Cell Biol* 4: 2876–2882
3. Ahlquist P, Luckow V, Kaesberg P (1981) Complete nucleotide sequence of *Brome mosaic virus* RNA 3. *J Mol Biol* 153: 23–38
4. Bouwen I, Maat DZ (1992) *Pelargonium flower-break* and *Pelargonium* line pattern viruses in the Netherlands; purification, antiserum preparation, serological identification, and detection in *pelargonium* by ELISA. *Neth Plant Pathol* 98: 141–156
5. Brierley I (1995) Ribosomal frameshifting on viral RNAs. *J Gen Virol* 76: 1885–1892
6. Brierley I, Jenner AJ, Inglis SC (1992) Mutational analysis of the “slippery sequence” component of a coronavirus ribosomal frameshifting signal. *J Mol Biol* 227: 463–479

7. Bringle DH, Pleij CW, Coutts RH (1999) Mutation analysis of cis-elements in the 3'- and 5'-untranslated regions of satellite *Tobacco necrosis virus* strain C RNA. *Virology* 264: 76–84
8. Cañizares MC, Marcos JF, Pallás V (2001) Molecular variability of twenty-one geographically distinct isolates of *Carnation mottle virus* (CarMV) and phylogenetic relationships within the *Tombusviridae* family. *Arch Virol* 146: 2039–2051
9. Carpenter CD, Simon AE (1998) Analysis of sequences and predicted structures required for viral satellite RNA accumulation by *in vivo* genetic selection. *Nucleic Acids Res* 26: 2426–2432
10. Fabian MR, Na H, Ray D, White KA (2003) 3'-Terminal RNA secondary structures are important for accumulation of Tomato bushy stunt DI RNAs. *Virology* 313: 567–580
11. Franck A, Loebenstein G (1994) Virus and virus-like diseases of pelargonium in Israel. *Acta Hort* 377: 31–39
12. Fütterer J, Hohn T (1996) Translation in plants—rules and exceptions. *Plant Mol Biol* 32: 159–189
13. Goldberg K-B, Modrell B, Hillman BI, Heaton LA, Choi TJ, Jackson AO (1991) Structure of the glycoprotein gene of *Sonchus yellow net virus*, a plant rhabdovirus. *Virology* 185: 32–38
14. Hacker DL, Petty ITD, Wei N, Morris TJ (1992) *Turnip crinkle virus* genes required for RNA replication and virus movement. *Virology* 186: 1–8
15. Heinze C, Wobbe V, Lesemann DE, Zhang DY, Willingmann P, Adam G (2004) Pelargonium necrotic spot virus: a new member of the genus *Tombusvirus*. *Arch Virol* 149: 1527–1539
16. Huang M, Koh D, Weng L, Chang M, Yap Y, Zhang L, Wong S (2000) Complete nucleotide sequence and genome organization of *Hibiscus chlorotic ringspot virus*, a new member of the genus *Carmovirus*: evidence for the presence and expression of two novel open reading frames. *J Virol* 74: 3149–3155
17. Isogai M, Uyeda I, Hataya T (1998) An efficient cloning strategy for viral double-stranded RNAs with unknown sequences. *Ann Phytopathol Soc Jpn* 64: 244–248
18. Jones AT, McGavin WJ, Brunt AA, Phillips S (2000) Elderberry latent virus: its relationship to Pelargonium ringspot virus and its identification as a distinct member of the genus *Carmovirus*, family *Tombusviridae*. *Ann Appl Biol* 136: 147–152
19. Kinard GR, Jordan RL, Hurtt SS (1996) Partial characterization of Pelargonium line pattern and Pelargonium ringspot viruses. *Acta Hort* 432: 148–155
20. Kinard GR, Jordan RL (1998) Genome organization of Pelargonium ringspot and Elderberry latent viruses. *Phytopathology* 88: S48
21. Kinard GR, Jordan R (2002) Genome organization of Pelargonium chlorotic ring pattern virus: further implications for *Tombusviridae* taxonomy. *Acta Hort* 568: 17–27
22. Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* 28: 375–430
23. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283–292
24. Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA 2: molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244–1245
25. Li WZ, Qu F, Morris TJ (1998) Cell-to-cell movement of *Turnip crinkle virus* is controlled by two small open reading frames that function *in trans*. *Virology* 244: 405–416
26. Liang XZ, Lucy AP, Ding SW, Wong SM (2002) The p23 protein of *Hibiscus chlorotic ringspot virus* is indispensable for host-specific replication. *J Virol* 76: 12312–12319

27. Lisa V, Vaira AM, Dellavalle G, Masenga V, Milne RG (1996) Viruses of pelargonium in Italy. *Acta Hort* 432: 108–117
28. Lütke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J* 6: 43–48
29. Marcos JF, Vilar M, Pérez-Payá E, Pallás V (1999) *In vivo* detection, RNA-binding properties and characterization of the RNA-binding domain of the p7 putative movement protein from *Carnation mottle carmovirus* (CarMV). *Virology* 255: 354–365
30. Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288: 911–940
31. Morozov S, Ryabov E, Leiser R, Zavriev S (1995) Use of highly conserved motifs in plant virus polymerases as the tags for specific detection of carmovirus-related RNA-dependent RNA-polymerase genes. *Virology* 207: 312–315
32. Morris T, Carrington J (1988) *Carnation mottle virus* and viruses with similar properties. In: Koenig R (ed). *The plant viruses*, vol 3: Polyhedral virions with monopartite RNA genomes. Plenum Press, New York, pp 73–112
33. Morris TJ, Dodds JA (1979) Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology* 69: 854–858
34. Pearson WR, Wood T, Zhang Z, Miller W (1997) Comparison of DNA sequences with protein sequences. *Genomics* 46: 24–36
35. Plese N, Stefanac Z (1980). Some properties of a distinctive isometric virus from pelargonium. *Acta Hort* 110: 183–190
36. Pogany J, Fabian MR, White KA, Nagy PD (2003) A replication silencer element in a plus-strand RNA virus. *EMBO J* 22: 5602–5611
37. Rico P, Hernández C (2004) Complete nucleotide sequence and genome organization of *Pelargonium flower break virus*. *Arch Virol* 149: 641–651
38. Riviere CJ, Rochon DM (1990) Nucleotide sequence and genomic organization of *Melon necrotic spot virus*. *J Gen Virol* 71: 1887–1896
39. Russo M, Burgyan J, Martelli GP (1994) Molecular Biology of *Tombusviridae*. *Adv Virus Res* 44: 381–428
40. Scholthof KB, Scholthof HB, Jackson AO (1995) The *Tomato bushy stunt virus* replicase proteins are coordinately expressed and membrane associated. *Virology* 208: 365–369
41. Stone OM (1980) Nine viruses isolated from pelargonium in the United Kingdom. *Acta Hort* 110: 177–182
42. Stone OM, Hollings M (1971) Pelargonium. *Rep Glasshouse Crops Res Inst* 1970, pp 153–155
43. Stone OM, Hollings M (1976) Pelargonium. *Rep Glasshouse Crops Res Inst* 1975, pp 119–120
44. Stone OM, Hollings M (1977) Pelargonium. *Rep Glasshouse Crops Res Inst* 1976, pp 121–122
45. Stuart G, Moffett K, Bozarth RF (2004) A whole genome perspective on the phylogeny of the plant virus family *Tombusviridae*. *Arch Virol* 149: 1595–1610
46. Stupina V, Simon AE (1997) Analysis *in vivo* of *Turnip crinkle virus* satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. *Virology* 238: 470–477
47. Thompson JD, Gibson TJ, Plewniak F, Higgins DG (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882

48. Turina M, Maruoka M, Monis J, Jackson AO, Scholthof KB (1998) Nucleotide sequence and infectivity of a full-length cDNA clone of *Panicum mosaic virus*. *Virology* 241: 141–155
49. Turina M, Desvoyes B, Scholthof KB (2000) A gene cluster encoded by *Panicum mosaic virus* is associated with virus movement. *Virology* 266: 120–128
50. Turner RL, Buck KW (1999) Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA 2 of *Red clover necrotic mosaic virus*. *Virology* 253: 115–124
51. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (2000) *Virus Taxonomy*. Seventh Report of the International Committee on Taxonomy of Viruses, Academic Press, San Diego
52. Verwoerd TC, Dekker BMM, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res* 17: 2362
53. Vilar M, Sauri A, Monne M, Marcos JF, von Heijne G, Pérez-Payá E, Mingarro I (2002) Insertion and topology of a plant viral movement protein in the endoplasmic reticulum membrane. *J Biol Chem* 244: 23447–23452
54. Vreede FT, Cloete M, Napier GB, van Dijk AA, Viljoen GJ (1998) Sequence-independent amplification and cloning of large dsRNA virus genome segments by poly(dA)-oligonucleotide ligation. *J Virol Methods* 72: 243–247
55. White KA, Skuzeski JM, Li W, Wei N, Morris TJ (1995) Immunodetection, expression strategy and complementation of *Turnip crinkle virus* p28 and p88 replication components. *Virology* 211: 525–534
56. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406–3415

Author's address: Carmen Hernández, Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC), Universidad Politécnica de Valencia, Avenida de los Naranjos s/n, 46022 Valencia, Spain; e-mail: cahernan@ibmcp.upv.es

Capítulo II

BIOLOGICAL ACTIVITY OF TRANSCRIPTS FROM cDNA OF PELARGONIUM LINE PATTERN VIRUS

A. CASTAÑO, C. HERNÁNDEZ*

Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, Campus Universidad Politécnica de Valencia, Ciudad de la Innovación, Avenida Ingeniero Fausto Elio s/n, 46022 Valencia, Spain

Received June 20, 2007; accepted October 22, 2007

Summary. – A set of cDNAs of Pelargonium line pattern virus (PLPV) was assembled under the control of T7 RNA polymerase promoter and ligated into the plasmid pUC18. Transcripts synthesized *in vitro* from cDNA were infectious on *Chenopodium quinoa* according to locally induced lesions and hybridization assay. The biological activity of the viral transcripts was particularly sensitive to the short 3' terminus extensions, whereas inclusion of the 3 extra bases at the 5' terminus did not substantially affect the infectivity. Inoculation of the transcripts on plants *Nicotiana benthamiana* and *Nicotiana clelandii* give rise to the systemic infection indistinguishable from that established by the parental isolate. This is the first report about the preparation of infectious RNA transcripts from a full-length cDNA clone of PLPV.

Key words: Pelargonium line pattern virus; cDNA; infectious transcripts

PLPV is one of the most frequent viral pathogens in *Pelargonium* spp. worldwide (Stone, 1980; Bouwen and Maat, 1992; Franck and Loebenstein, 1994; Alonso and Borja, 2005). The infections are often asymptomatic, but under certain environmental conditions, the yellow-green spots and line patterns on the leaves may appear what impair their quality and marketability (Nameth and Adkins, 1993). It should be noted that PLPV is proposed as a tentative member of a prospective new genus in the family *Tombusviridae* (Kinard and Jordan, 2002; Castaño and Hernández, 2005) though the virus has not been formally classified into taxonomic scheme yet (Fauquet *et al.*, 2005). This is most likely due to the limited information on PLPV and other related viruses infecting geranium, what has complicated their classification to a viral group so far.

PLPV virions are isometric, about 30 nm in diameter and contain a linear positive-sense single-stranded RNA (ssRNA). The complete nucleotide sequence of PLPV genomic RNA (gRNA) has been recently determined (Castaño and Hernández, 2005). It comprises 3884 nt and contains 6 ORFs flanked by an unusually short untranslated regions of 6 nt at the 5' end and 247 nt at the 3' end. The 6 ORFs potentially encode proteins of 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6) and 37 (p37) K. The ORF arrangement on the PLPV genome closely resembles the one found in the members of genus *Carmovirus*, family *Tombusviridae*. Moreover, the putative PLPV gene products presumably involved in replication (p27 and its readthrough product p87), movement (p7), and encapsidation (p37) show a high identity with proteins of this viral group. However, some unique characteristics distinguish PLPV from the carmoviruses, as the production of only one subgenomic RNA in contrast to the carmoviruses that generate two. Three related small isometric ssRNA viruses, Pelargonium ringspot virus (unclassified), Pelargonium chlorotic ring pattern virus (unclassified), and Elderberry latent virus (a tentative species of the genus *Carmovirus*) share the same characteristics. In advance, it has been suggested that they should be grouped

*Corresponding author. E-mail: cahernan@ibmcp.upv.es; fax: +3496-3877859.

Abbreviations: cDNA = complementary DNA; dsRNA = double-stranded RNA; gRNA = genomic RNA; PLPV = Pelargonium line pattern virus; RT = reverse transcription; ssRNA = single-stranded RNA

together with PLPV into a new genus *Pelarspovirus* within the family *Tombusviridae* (Kinard and Jordan, 2002; Castaño and Hernández, 2005).

The availability of a full-length cDNA prepared from the viral genome capable of providing infectious transcripts has enormously contributed to the studies about genomic functions of plant viruses. In this paper, we report the construction of a full-length cDNA of the PLPV genome under the control of the bacteriophage T7 promoter that served as a template for the synthesis of infectious transcripts *in vitro*. To our knowledge, this is the first report about the preparation of cDNA clone for a tentative member of the prospective genus *Pelarspovirus*. We supposed that preparation of cDNA allowed us to perform studies concerning the replication, gene expression and pathogenesis of this viral group.

PLPV, isolate PV-0193, was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and maintained by a sap inoculation on the experimental host *Chenopodium quinoa*. Total RNA from infected *C. quinoa* leaves was prepared by the phenol extraction and lithium chloride precipitation (Verwoerd *et al.*, 1989) and used as a template for RT-PCR reaction with Superscript II-RT (Invitrogen) and Expand High Fidelity DNA polymerase (Roche). Two overlapping cDNAs were generated, one with primers CH18 (5'-GCCTCCCTGCTGGCATAAATAATAC-3') complementary to nt 3578–3603 of PLPV gRNA and CH61 (5'-GCAAGCTTGTAAATACGACTCACTATAGGGAACA AAATGGCACACTATTTTGG-3'), which contains a *Hind*III site (underlined) fused to a T7 RNA polymerase promoter sequence (in bold) followed by 23 nt of the 5' end of the virus sequence. The second cDNA was generated with primers CH17 (5'-GAAAATGGCCTTCTACGGGGAC-3'), homologous to nt 2067–2088, and CH60 (5'-CCGG ATCCCGGGCAGATCAGGGGGGTGGGTTAC-3'), complementary to the 3' terminus of the viral sequence (nt 3859–3884) with a *Sma*I site (underlined) and a *Bam*HI site (in italics) at the 5' terminus. Ligation of the RT-PCR products through the *Xmn*I site (present in the PLPV sequence at nt 3091–3100) allowed us to generate independent full-length viral cDNAs, which were subsequently cloned into *Hind*III and *Bam*HI sites of the plasmid pUC18. Two recombinant plasmids were prepared and signed as pPLDS-1 and pPLDS-2. The plasmids were digested with *Bam*HI and transcribed using T7 RNA polymerase (Fermentas). No cap analog was included in the transcription reactions as the cap structure is not required for infectivity of viral RNA among the members of the family *Tombusviridae*. After recognition of its promoter, T7 RNA polymerase will start transcription at the first of the 3 terminal G nt in the promoter sequence, thus the transcripts were expected to contain 3 extra G nt at the 5' terminus

compared to the viral RNA. In addition, as the plasmids were linearized with *Bam*HI prior *in vitro* transcription, the transcripts were also predicted to harbor 5 extraviral nt (GGAUC) at the 3' terminus derived from the *Bam*HI target site. The *C. quinoa* plants were mechanically inoculated with RNAs derived from each cDNA clone (approximately 0.7 µg per leaf). The mock-inoculated plant served as a negative control. None of the inoculated leaves developed chlorotic lesions typical of PLPV infections. Likewise, hybridization assays confirmed the absence of the virus in the inoculated leaves. To study more closely the absence of infectivity of these clones, we determined their nucleotide sequences that showed 99% nt identity with the PLPV (Castaño and Hernández, 2005). No obvious lethal nucleotide changes were detected in the construct pPLDS-2. On the other hand, one of the nucleotide substitutions present in the construct pPLDS-1 (C703T) led to the introduction of a premature stop codon in ORF (p27), which might be responsible for the non-viability of the corresponding transcripts.

Two following sets of partial cDNAs were generated by RT-PCR using two different pairs of primers: CH61 and CH64 (5'-CCAGTCTAGACCTCACGCCAATTTTC-3'), complementary to the nt 1663–1687 of the PLPV gRNA, or CH60 and CH63 (5'-CGTGAGGTCTAGACTGGTGA ACAACG-3'), homologous to the nt 1671–1697. These cDNAs covering the 5' and 3' portion of the gRNA, respectively, and containing an *Xba*I site (present at nt 1678–1683 in PLPV genome) in the overlapping region, were used to ready exchange of these regions in the initial non-infectious construct pPLDS-2. Transcripts derived from several of the resulting clones with the 5' portion replaced, induced chlorotic lesions when inoculated on *C. quinoa* leaves. The presence of the virus in lesions was confirmed by the Northern blot hybridization (data not shown). One representative of these clones, signed pPLDS-10 (Fig. 1), was randomly selected for sequencing and was found to differ from pPLDS-2 at six nt positions within ORF (p27) and/or ORF (p87). Which of these nucleotide changes reversed the non-infectivity of pPLDS-2 is currently under investigation. Despite the biological activity displayed by pPLDS-10 derived transcripts, the number of lesions that appeared on the inoculated leaves was considerably lower than that observed with the viral RNA as inoculum (Fig. 1 and data not shown). Therefore, we followed two approaches to test the effect of additional nucleotides at the 5' and/or 3' terminus of the transcripts on their infectivity. The viral full-length cDNA inserted into pPLDS-10 was amplified by PCR and ligated into pUC18-derived plasmid. This plasmid contained the CaMV 35S promoter and the terminator sequence from the nopaline synthase gene in such a way that RNA polymerase II would initiate transcription exactly at the first nucleotide of the viral sequence. *C. quinoa* leaves inoculated with the resulting construct (approximately 2 µg

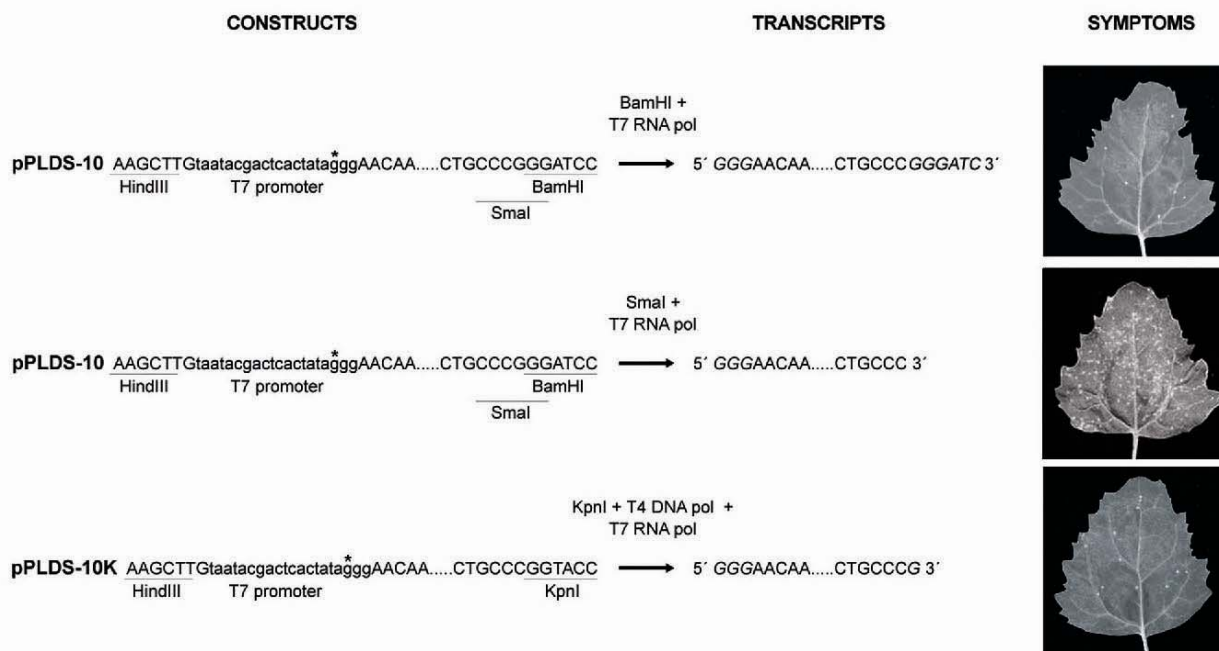


Fig. 1

Schematic representation of the generation of transcripts from constructs pPLDS-10 and pPLDS-10K

The nucleotide sequences and restriction sites are shown at the borders of the full-length viral cDNA. The initiation site for transcription is indicated by asterisk and non-viral nucleotides in the transcripts are shown in italics. Photographs at the right show the lesions induced by the different transcripts on *C. quinoa* leaves.

per leaf) developed lesions similar in number to those induced by transcripts synthesized in vitro from *Bam*HI digested-pPLDS-10, although their appearance was delayed 1–2 days with respect to the transcript-inoculated plants.

In view of the above result, the possible effect of the extra nt present at the 3' terminus on the infectivity of the transcripts was investigated. The viral cDNA inserted in pPLDS-10 was amplified by PCR with primers CH61 and CH89 (5'-ACGTGGTACCGGGCAGATCAGGGGGGTGGGTAC-3'), complementary to the 3' terminus of the viral sequence (nt 3859–3884) with a *Kpn*I site (underlined) at the 5' terminus. The PCR product was digested using *Hind*III and *Kpn*I and ligated to the properly digested pUC18. The resulting recombinant plasmid signed as pPLDS-10K was used as a template to synthesize transcripts after its linearization with *Kpn*I and subsequent polishment with T4 DNA polymerase (Roche). In addition, new transcripts were synthesized from the construct pPLDS-10 linearized with *Sma*I. The number of lesions induced by the transcripts derived from the *Sma*I digested plasmid that were expected to contain a deletion at the 3' terminus was about 10-fold higher than number of lesions induced by the transcripts from the *Bam*HI linearized construct or by the *Kpn*I plus T4

DNA polymerase treated pPLDS-10K template that were predicted to contain the precise 3' end (Fig. 1). This result strongly supported the idea that the G nt previously determined as the 3' terminal residue of the PLPV genome (Castaño and Hernández, 2005) is not in fact present in the viral ssRNA. The replicating form of double-stranded RNA (dsRNA) in several positive RNA viruses carries an extra residue at the 3' terminus of the plus and/or minus strand (Karasev *et al.*, 1995; Galiakparov *et al.*, 1999). Since this kind of dsRNA was the template used for cloning and sequencing of the PLPV 3'-proximal region, our present observations indicated that PLPV as well as the remaining members of the family *Tombusviridae* ended with CCC-OH (White and Nagy, 2004).

In addition, the transcripts derived from the *Sma*I digested pPLDS-10 construct were used to inoculate *N. clelandii* and *N. benthamiana* plants. The inoculated plants did not develop obvious symptoms, but the viral RNA was detected by Northern blot hybridization in both local and systemic leaves like in the control plants inoculated with the parental virus (Fig. 2).

In conclusion, we have generated a full-length cDNA from PLPV and obtained the infectious transcripts. The

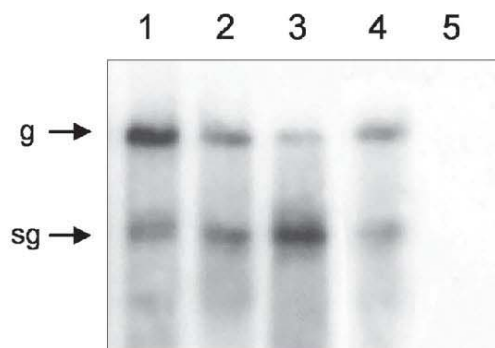


Fig. 2

Northern-blot hybridization of total RNA from plants inoculated with transcripts derived from *Sma*I digested pPLDS-10

RNA samples were prepared from inoculated leaves of *C. quinoa* (2) or from systemic leaves of *N. clelandii* (3) and *N. benthamiana* (4). RNA preparations from *C. quinoa* plants infected with PLPV isolate PV-0193 (1) and from the mock-infected plants (5) were included as positive and negative controls, respectively. The arrows indicate PLPV genomic (g) and subgenomic (sg) RNA.

results have shown a strong effect of additional nucleotides at the 3' terminus of the viral RNA on infectivity of transcripts, whereas three extra residues at the 5' terminus did not substantially interfere with the viral multiplication. Our results were in contrast with those reported for other members of the family *Tombusviridae* as *Pothos latent virus* or *Melon necrotic spot virus*, which infectivity was not compromised by 3' extensions in the viral RNA (Rubino and Russo, 1997; Díaz *et al.*, 2003). To deepen our knowledge of PLPV genomic organization and to clarify the involvement of each viral protein in PLPV basic biology, at this time reverse genetics experiments are being performed using the biologically active cDNA clone.

Acknowledgements. This study was supported by the grants AGL2000-0942 and AGL2003-04249 from the Ministerio de Ciencia y Tecnología de España and by the grants Grupos 03/064 and ACOM06/210 from the Generalitat Valenciana. A.C. was the recipient of predoctoral fellowships from the Generalitat Valenciana and from CSIC-Fundación Bancaja.

References

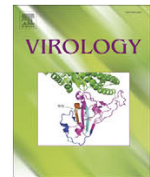
- Alonso M, Borja M (2005): High incidence of Pelargonium line pattern virus infecting asymptomatic Pelargonium spp. in Spain. *Eur. J. Plant Pathol.* **112**, 95–100.
- Bouwen I, Maat DZ (1992): Pelargonium flower-break and Pelargonium line pattern viruses in the Netherlands; purification, antiserum preparation, serological identification, and detection in pelargonium by ELISA. *Neth. Plant Pathol.* **98**, 141–156.
- Castaño A, Hernández C (2005): Complete nucleotide sequence and genome organization of Pelargonium line pattern virus and its relationship with the family Tombusviridae. *Arch. Virol.* **150**, 949–965.
- Díaz JA, Bernal JJ, Moriones E, Aranda MA (2003): Nucleotide sequence and infectious transcripts from a full-length cDNA clone of the carmovirus Melon necrotic spot virus. *Arch. Virol.* **148**, 599–607.
- Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (2005): *Virus Taxonomy. 8th Report of the International Committee on Taxonomy of Viruses*. Elsevier Inc. pp. 907–936.
- Franck A, Loebenstein G (1994): Virus and virus-like diseases of pelargonium in Israel. *Acta Hort.* **377**, 31–39.
- Galiakparov N, Tamme E, Sela E, Gafny R (1999): Infectious RNA transcripts from Grapevine virus A cDNA clone. *Virus Genes* **19**, 235–242.
- Karasev AV, Boyko VP, Gowda S, Nikolaeva OV, Hilf ME, Koonin EV, Niblett CL, Cline K, Gumpf DJ, Lee RF, Garnsey SM, Lewandowski DJ, Dawson WO (1995): Complete sequence of the Citrus tristeza virus RNA genome. *Virology* **208**, 511–520.
- Kinard GR, Jordan R (2002): Genome organization of Pelargonium chlorotic ring pattern virus: further implications for Tombusviridae taxonomy. *Acta Hort.* **568**, 17–27.
- Nameth ST, Adkins ST (1993): Viral diseases. In White JW (Ed.): *Geraniums IV-The Grower's Manual*. Ball Publishing, Geneva, USA, pp. 267–275.
- Rubino L, Russo M (1997): Molecular analysis of the Photos latent virus genome. *J. Gen. Virol.* **78**, 1219–1226.
- Stone OM (1980): Nine viruses isolated from pelargonium in the United Kingdom. *Acta Hort.* **110**, 177–182.
- Verwoerd TC, Dekker BMM, Hoekema A (1989): A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res.* **17**, 2362.
- White KA, Nagy PD (2004): Advances in the molecular biology of tombusviruses: gene expression, genome replication and recombination. *Prog. Nucleic Acid Res. Mol. Biol.* **78**, 187–226.

Capítulo III



Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus

Aurora Castaño, Leticia Ruiz, Carmen Hernández*

Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas-UPV), Campus Universidad Politécnica de Valencia, Avenida de los Naranjos, 46022 Valencia, Spain

ARTICLE INFO

Article history:

Received 25 November 2008
 Returned to author for revision
 13 December 2008
 Accepted 15 January 2009
 Available online 12 February 2009

Keywords:

Pelargonium line pattern virus
 Tombusviridae
 Gene expression
 Virus translation
 Leaky-scanning

ABSTRACT

Pelargonium line pattern virus (PLPV), a proposed member of a prospective genus (*Pelarspovirus*) within family *Tombusviridae*, has a positive-sense, single-stranded genomic RNA. According to previous predictions, it contains six open reading frames (ORFs) potentially encoding proteins of 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), and 37 kDa (p37). Using a variety of techniques we demonstrate that all predicted ORFs are functional, with the exception of (p13) and (p6). We also characterize a previously unidentified ORF which encodes a 9.7 kDa protein (p9.7) that is essential for viral movement. Furthermore, we present evidence that the single subgenomic RNA (sgRNA) produced by the virus directs synthesis of p7, p9.7 and p37. Remarkably, the translation of these totally unrelated proteins is coordinated via leaky-scanning. This mechanism seems to be favoured by the poor translation context of the start codon of ORF(p7), the non-AUG weak initiation codon of ORF(p9.7) and the lack of additional AUG codons in any reading frame preceding ORF(p37). The results also suggest that precise regulation of protein production from the sgRNA is critical for virus viability. Altogether, the data supports the notion that PLPV belongs to a new genus of plant viruses.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Translational control plays a key role in modulating eukaryotic gene expression. Because of the genome size constraints, and to expand the expressible gene content, viruses frequently use distinct mechanisms to modify ribosome behaviour. Such modifications may relieve the restrictions to monocistronic expression that typically characterize the eukaryotic translation system (Gale et al., 2000; Gallie, 1996; Kozak, 2002). Leaky scanning is one of the most common strategies employed by viruses to express overlapping open reading frames (ORFs). It occurs when a proportion of ribosomes scanning from the 5'-end of the mRNA bypass either the 5'-most AUG due to its suboptimal sequence context (absence of purine in position -3 and/or G at position +4) or certain non-AUG triplets that can serve as weak initiation codons (Gordon et al., 1992; Kozak, 1989a; Peabody, 1989), to initiate translation at an alternative start codon downstream. Re-initiation may also take place if ribosomes that have completed translation of an initial ORF, resume scanning and (re)initiate translation at an AUG located farther downstream (Kozak, 2001; Luukkonen et al., 1995). Backward scanning and re-initiation at upstream sites has also been reported for translation of overlapping ORFs. This may occur when the stop codon of the upstream ORF is close to the start codon of the downstream ORF (Kozak, 2001; Peabody et al., 1986; Thomas and Capecchi, 1986). Finally, internal

initiation has been demonstrated for animal and plant viral RNAs. This is mediated by internal ribosome entry site (IRES) elements that facilitate direct recruitment of the pre-initiation complex, bypassing 5'-dependent scanning to access internal start codons (Hellen and Sarnow, 2001; Kneller et al., 2006).

Besides the above mechanisms related to alternative translation initiation, viruses employ strategies to ensure that all their genes are accessible to the eukaryotic protein-synthesizing system. Examples include "recoding" that permits readthrough and frameshifting, or the generation of subgenomic RNAs (sgRNAs) (Fauquet et al., 2005). Another intriguing feature of viral RNAs is that many of them lack a 5'-cap structure and/or a 3'-poly(A) tail. This is the case in more than 80% of plant positive strand RNA viruses. Deciphering how these non-canonical RNAs regulate their expression may provide valuable clues on the diversity and complexity of gene translation mechanisms.

Pelargonium line pattern virus (PLPV) is one of the most prevalent viruses affecting geraniums (*Pelargonium* spp.) (Alonso and Borja, 2005; Bouwen and Maat, 1992; Franck and Loebeinstein, 1994; Stone, 1980). PLPV has a positive-sense, single-stranded genomic RNA (gRNA) of 3883 nt that is packaged into an icosahedral particle of about 30 nm. This gRNA lacks a 5'-cap structure or a 3'-poly(A) tail and, according to *in silico* analyses, contains six ORFs potentially encoding proteins of 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), and 37 kDa (p37) (Castaño and Hernández, 2005, 2007). Such ORFs are flanked by an unusually short untranslated region (UTR) of 6 nt at the 5'-end and by a 247 nt long UTR at the 3'-end. The arrangement of ORFs on the gRNA is similar to that of members of the genus

* Corresponding author. Fax: +34 96 877859.

E-mail address: cahernan@ibmcp.upv.es (C. Hernández).

Carmovirus in the family *Tombusviridae*. With the exception of p13 and p6, which do not show significant similarity to protein sequences in data banks, putative PLPV gene products share considerable sequence identity with proteins involved in replication (p27 and p87, the latter one with the motifs typical of RNA dependent-RNA polymerases), movement (p7) and encapsidation (p37) in this viral group. However, whereas carmoviruses generate two sgRNAs, PLPV produces only one as do three other related small isometric single-stranded RNA viruses, Pelargonium ringspot virus, Pelargonium chlorotic ring pattern virus (PCRPV) and Elderberry latent virus. This common trait has led to the suggestion that the four pathogens should be clustered together into a new genus, *Pelarspovirus*, within the family *Tombusviridae*. Such a proposal is supported by phylogenetic analyses (Castaño and Hernández, 2005; Kinard and Jordan, 2002; Stuart et al., 2006).

Based on sequence data, it was anticipated that if all proteins potentially encoded by PLPV are produced in vivo the virus must use a variety of mechanisms for expression of its genes (Castaño and Hernández, 2005). Thus, translation of ORFs (p27), (p13) and (p87) were proposed to occur from the gRNA involving read-through of the amber stop codon of ORF(p27) in the case of ORF(p87) and possibly an IRES in the case of ORF(p13). The remaining downstream ORFs could be translated from the abundant sgRNA of 1.6 kb detected in PLPV infected plants. ORF(p7), located 5'-proximal in the sgRNA, would be expressed by conventional ribosomal scanning whereas the overlapping ORF(p6) might be translated from a non-AUG initiation codon by leaky scanning and/or by a -1 frameshift mechanism giving rise to

a p12 fusion protein (p7-FS). Internal ribosomal entry or leaky scanning was considered the most plausible mechanisms for translation of the 3'-proximal ORF(p37).

In this work, we have firstly identified the biologically active ORFs of PLPV. Next, we have investigated the mechanisms of gene expression from the gRNA and the sgRNA. We present evidence that the latter directs synthesis of three distinct proteins and, most remarkably, that translation of these proteins is coordinated via leaky-scanning. The results also support that maintenance of protein production ratios from the sgRNA is critical for virus viability. Globally, the data provide further justification for the contention that PLPV belongs to a new genus of plant viruses.

Results

Four out of the six ORFs predicted in PLPV genome are essential for effective infection in local and systemic hosts

As mentioned above, our previous sequence analysis of PLPV genome indicated the presence of six ORFs potentially encoding proteins p27, p13, p87, p7, p6 and p37 (Fig. 1A). To assess their biological significance, a set of mutants with the presumed start codon of each ORF independently modified (Fig. 1B) was generated using as template the PLPV infectious clone pPLDS-10 (Castaño and Hernández, 2007). Transcripts were synthesized from pPLDS-10 and from the individual mutant constructs and used to mechanically inoculate *Chenopodium*

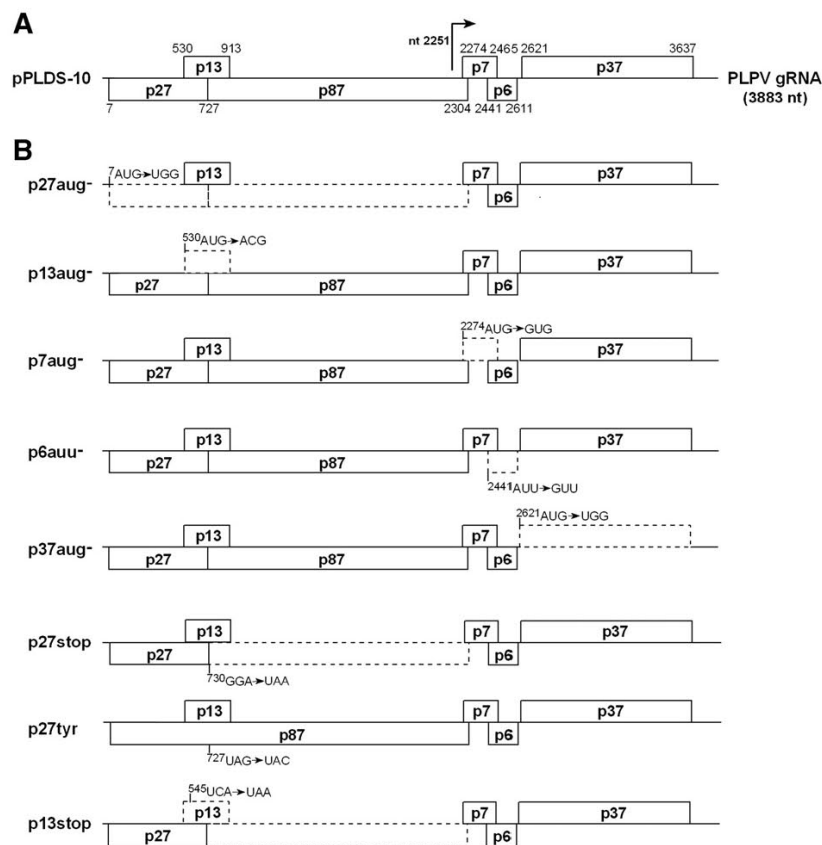


Fig. 1. Location of mutations in PLPV constructs. (A) Previously proposed structure of PLPV gRNA (Castaño and Hernández, 2005, 2007). Open boxes correspond to predicted ORFs. Numbers above and below the boxes represent the nucleotide positions of the codons defining each ORF. The transcription start site of the sgRNA at nt 2251 is marked by an arrow. (B) Schematic representation of PLPV mutants. Nucleotide substitutions were introduced into the full-length cDNA clone pPLDS-10 leading to triplet changes that are depicted for each mutant. In mutants p13aug- and p6auu-, besides altering the presumed start codons of the corresponding ORFs, close downstream triplets have been also modified (⁵³⁵AUG in p13aug- and ²⁴⁶⁵CUG and ²⁴⁶⁵AUU in p6auu-) to ensure abolishment of protein production. In all mutants, dashed boxes correspond to ORFs that are not expected to be translated.

Table 1

Construct	Lesions in <i>C. quinoa</i>	PLPV detection		
		<i>C. quinoa</i> (inoculated leaf)	<i>N. benthamiana</i> (systemic leaf)	<i>N. benthamiana</i> (agroinfiltrated leaf)
pPLDS-10	++	++	++	++
p27aug-	-	-	-	-
p27stop	-	-	-	-
p27tyr	-	-	-	-
p13aug-	++	++	++	NA
p13stop	++	++	++	NA
p7aug-	-	-	-	++
p6auu-	++	++	++	NA
p9.7gug-	-	-	-	++
p37aug-	-	-	-	++

NA: not applicable.

quinoa plants. A similar number of lesions was observed at 7 days post-inoculation on *C. quinoa* leaves challenged either with the wild type (wt) construct, pPLDS-10, or with constructs p13aug- and p6auu-, and Northern blot hybridizations revealed equivalent accumulation levels of viral RNAs in these leaves (results summarized in Table 1). The virus was also detected in local and systemic leaves of *Nicotiana benthamiana* plants 28 days after inoculation with the latter constructs (Table 1) though plants did not develop obvious symptoms in line with previous observations (Castaño and Hernández, 2007). Though the infectivity of mutant p13aug- suggested that ORF(p13) lacked biological significance, its conservation in different isolates of the virus (Castaño, Ruiz and Hernández, unpublished results) prompted us to further explore whether it could play a non-essential or a host-specific role during infection. To this aim, the bioassay of mutant p13aug- was extended to a third host, *N. clevelandii*, and the progress of the infection in this host and also in *N. benthamiana* was weekly monitored by dot-blot hybridization. Mutant p13aug- was able to systemically infect *N. clevelandii* and *N. benthamiana* plants as efficiently as the wt construct, with no significant differences in the time course of virus spread or in viral accumulation levels (data not shown). The change of the presumed ⁵³⁰AUG start codon of ORF(p13) by ACG in mutant

p13aug- (Fig. 1) could still allow production of low levels of p13 considering that triplet ACG may function as a noncanonical start codon. Thus a new mutant was generated, p13stop, with a premature stop codon in ORF(p13) (Fig. 1). This mutant was also mechanically inoculated onto *C. quinoa*, *N. benthamiana* and *N. clevelandii* plants giving rise to infections indistinguishable from those established by the wt construct (Table 1 and data not shown).

In contrast with that observed with constructs p6auu-, p13aug- and p13stop, no lesions appeared on *C. quinoa* leaves inoculated with transcripts derived from mutants p27aug-, p7aug- or p37aug-, and hybridization assays corroborated the absence of PLPV RNAs in these leaves as well as in *N. benthamiana* plants inoculated with the same transcripts (Table 1). Among these mutants, only p27aug- was undetectable in *N. benthamiana* leaves after *Agrobacterium*-mediated viral inoculation (Table 1), an inoculation system that allows measuring the RNA accumulation level in cells independently of the ability of virus cell-to-cell movement (Chiba et al., 2006; Gopinath et al., 2005; Voinnet et al., 2000). The results were in agreement with the anticipated involvement of p27 (and its presumed read-through product, p87) in viral replication and the likely involvement of p7 and p37 in viral movement and encapsidation, respectively (Castaño and Hernández, 2005).

To confirm the lack of expression of the mutagenized ORFs, transcripts derived from the wt and the mutant genomic and/or subgenomic constructs were employed as templates for in vitro translation reactions using wheat germ extracts and including [³⁵S]methionine or [¹⁴C]leucine to label the products. Two proteins with apparent molecular masses of 27 and 87 kDa (Fig. 2A, lane 1) were synthesized from the wt genomic transcript. Mutation of the ⁷AUG codon of ORF(p27) in construct p27aug-, led to abolishment of p27 and p87 production from the genomic RNA (Fig. 2A, lane 2), supporting that both proteins share the initiation codon and that p87 is the readthrough product of p27. Moreover, in vitro translation of two additional mutants, p27stop and p27tyr (Fig. 1B), both non-infectious in plants and unable to accumulate in *N. benthamiana* leaves after agroinfiltration (Table 1), confirmed this point: no p87 was produced from the former mutant whereas this protein, and not

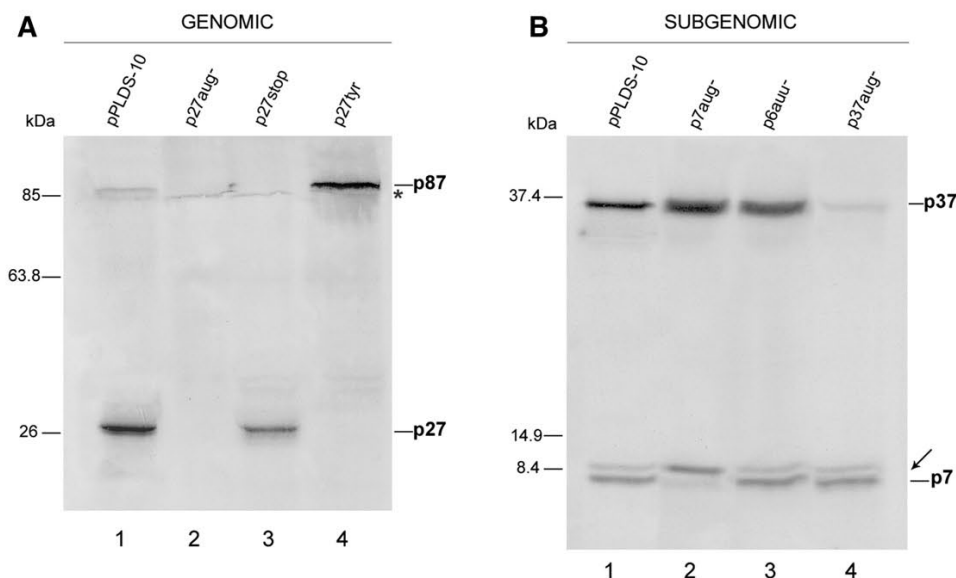


Fig. 2. In vitro translation assay of PLPV transcripts derived from genomic (A) and subgenomic (B) wt and mutant constructs. The constructs, whose description is provided in the text and in Fig. 1, are indicated above the lanes. Translation products labeled with [³⁵S]methionine (A) or [¹⁴C]leucine (B) were separated on SDS-12% polyacrylamide gels and visualized by autoradiography. Positions of protein molecular weight markers (in kDa) are depicted on the left side of the panels and PLPV proteins are indicated on the right side of each blot. The asterisk in (A) marks the position of a background band and the arrow in (B) marks a protein whose identification was subsequently done (see text and Fig. 4A).

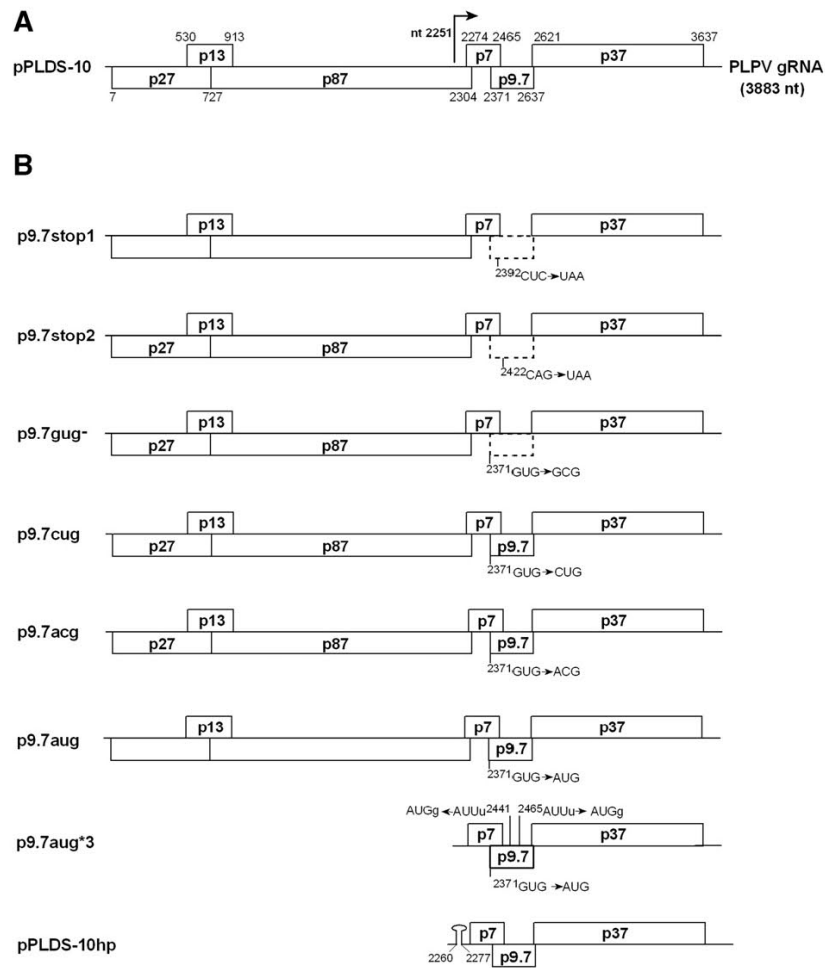


Fig. 3. Location of mutations in PLPV constructs. (A) Structure of PLPV genome proposed after functional analysis of ORFs (this work). Open boxes correspond to biologically active ORFs. The transcription start site of the sgRNA at nt 2251 is marked by an arrow. (B) Schematic representation of PLPV mutants. Nucleotide substitutions introduced into the full-length cDNA clone pPLDS-10 are depicted for each mutant. Mutants p9.7aug*3 and pPLDS-10hp were generated only as subgenomic constructs. PLPV positions flanking the hairpin structure engineered in mutant pPLDS-10hp are indicated. Other details as in Fig. 1.

p27, was synthesized from the latter mutant (Fig. 2A, lanes 3 and 4). Production of the putative p13 was not detected from in vitro translation of the wt gRNA (Fig. 2A, lane 1) which, together with the outcome of bioassays with mutants p13aug- and p13stop (Table 1), suggested that ORF(p13) is not expressed in vivo.

Expression of the two small ORFs (p6) and (p7) as well as of ORF (p37) was expected to occur from the PLPV sgRNA and three proteins with estimated molecular masses reasonably consistent with this expectation were observed from in vitro translation of wt subgenomic transcripts (Fig. 2B, lane 1). Though the translation patterns were the same irrespectively of using [³⁵S]methionine or [¹⁴C]leucine to label the in vitro synthesized proteins, the latter one permitted better detection of the intermediate size product. This may be because its leucine content is higher than that of methionine. Alteration of ²⁶²¹AUG in mutant p37aug- strongly diminished production of the bigger protein allowing its assignment to ORF(p37) (Fig. 2B, lane 4). The low p37 production still detected from p37aug- sgRNA was likely the result of translation from a noncanonical AAG start codon located six nt downstream of ²⁶²¹AUG. However, mutation of ²²⁷⁴AUG in p7aug- sgRNA revealed that the smaller detected protein was the product of ORF(p7) (Fig. 2B, lane 2) whereas modification of ²⁴⁴¹AUU in p6auu- sgRNA did not alter the pattern of translation products

(Fig. 2B, lane 3). Moreover, no evidence for synthesis of a frameshift product of ORF(p7) was obtained from in vitro translation experiments. These results, together with those from the bioassay in plants of mutant p6auu- (Table 1), indicated that the initially predicted ORF (p6) was unlikely to be biologically active and that, instead, another missed ORF could be relevant for viral infection as related viruses produce two small proteins to assist viral movement (Lommel et al., 2005).

A newly identified ORF with a non-AUG start codon encodes a protein of 9.7 kDa indispensable for virus cell-to-cell movement

Upon re-examination of PLPV sequence an ORF initially ignored, (positions 2371 to 2637), was identified. This putative ORF, which overlaps ORF(p7), lacked AUG codons and could give rise to a protein with a molecular mass up to 9.7 kDa considering triplet ²³⁷¹GUG as the initiation codon. Interestingly, this protein showed high sequence identity (44.4%) with one potentially encoded by PCRVP, another proposed member of the prospective genus *Pelarspovirus*, suggesting it may have functionality in vivo.

In order to assess whether the protein with a size slightly higher than that of p7 detected from in vitro translation of wt sgRNA could be

assigned to the newly identified ORF, two additional mutants were generated by converting two presumed triplets of the corresponding reading frame, 2392 CUC and 2422 CAG, into stop codons (constructs p9.7stop1 and p9.7stop2, respectively; Fig. 3). In vitro translation of subgenomic transcripts derived from these two mutants did not yield the protein migrating immediately above p7 (Fig. 4A, lanes 2 and 3) indicating that it was the product of the predicted ORF and that its translation was initiated upstream triplet 2392 CUC. This result pointed to triplet 2371 GUG or, alternatively, 2383 AGG, as the start codon of the reading frame. Conversion of triplet 2371 GUG into GCG in mutant p9.7gug⁻ (Fig. 3) abolished the production of the protein from the corresponding subgenomic transcript (Fig. 4A, lane 4), implying that its translation starts at the first triplet of the open reading phase deviating from AUG at one position as the vast majority of non-AUG initiation codons (Kozak, 1991; Touriol et al., 2003). Mutant p9.7gug⁻ did not induce lesions when mechanically inoculated on *C. quinoa*

leaves and hybridization assays confirmed the absence of an effective infection (Table 1). This mutant was also unable to infect *N. benthamiana* plants but accumulated at wt levels in *N. benthamiana* leaves after agroinoculation (Table 1). Globally, the results showed that the newly predicted ORF is expressed in vivo and that the resulting product of 9.7 kDa (p9.7) is most likely involved in virus movement.

Translation of internal ORFs from PLPV sgRNA is most likely regulated by leaky scanning

As shown above, in vitro translation assays clearly indicated that PLPV sgRNA directs the synthesis of three distinct proteins, p7, p9.7 and p37, being the translation of latter one very efficient despite the 3'-proximal position of the corresponding ORF (Fig. 2B, lane 1). Indeed, taking into account the relative intensity of the labeled protein

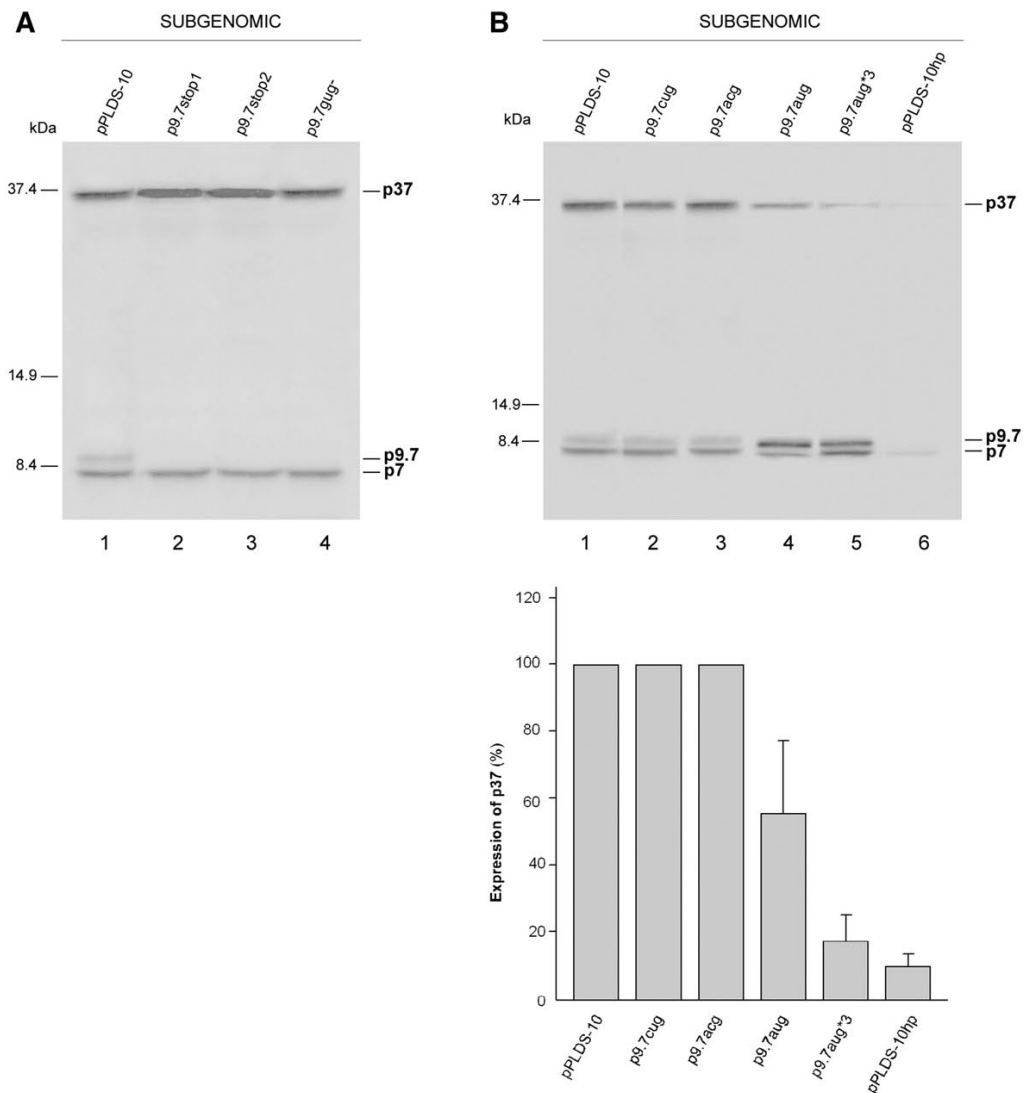


Fig. 4. In vitro translation assay of PLPV transcripts derived from wt and two different sets (A and B) of mutant subgenomic constructs. The constructs, whose description is provided in the text and in Fig. 3, are indicated above the lanes. Translation products labeled with [14 C]leucine were separated on SDS-12% polyacrylamide gels and visualized by autoradiography. A graphic representation of levels of p37 expressed in vitro is shown below the autoradiogram of panel B. The expression of p37 from pPLDS-10 is represented as 100% and the effect of mutations on the translation of ORF(p37) from other templates are represented as percentages of the p37 expression from pPLDS-10. The percentage is shown as the mean and standard error of the mean. Other details as in Fig. 2.

bands and the distinct leucine content of the proteins, the efficiencies of p9.7 and p37 gene translation were estimated to be about 10% and 50%, respectively, of that of the p7 gene. The possibility that p37 would be expressed from a monocistronic degradation product of sgRNA instead of the full-length molecule was ruled out as we did not detect any outstanding degradation products during the in vitro translation reactions (data not shown). The tricistronic arrangement of PLPV sgRNA has few parallels in the plant (or eukaryotic) world, and suggests that different translational mechanisms may account for expression of internal ORFs which, moreover, should be tightly regulated in order to achieve proper ratios of the encoded proteins. The mechanism involved in ORF(p9.7) expression could already be inferred from the in vitro translation results with p7aug- sgRNA which showed that mutation of the ²²⁷⁴AUG start codon of ORF(p7) to GUG increased p9.7 production by at least 2-fold (Fig. 2B, lane 2). This observation strongly supported that translation of ORF(p9.7) is accomplished by leaky scanning which would be favoured by the poor translation context (cgaAUGg) of the start codon of ORF(p7).

Three alternative mechanisms were considered to be the most plausible for expression of the p37 gene. First, translation of p37 could occur by direct internal initiation because of the presence of an IRES as described for different plant viruses including *Hibiscus chlorotic ringspot virus* (HCRSV), a member of the genus *Carmovirus* whose CP may be translated from genomic and subgenomic RNAs through such type of element (Koh et al., 2003, 2006). Second, ribosomes may resume scanning after translating the initial short ORF(p7) by a termination–reinitiation mechanism. Alternatively, ribosomes may be able to scan backward and reinitiate at the once-bypassed upstream AUG codon of the p37 gene after reaching the stop codon of the p9.7 gene, as both codons are separated by only 11 nt. Last, p37 synthesis may involve leaky scanning.

IRES have been proven to drive protein synthesis of a second cistron when placed outside of its genetic context (Hellen and Sarnow, 2001; Martínez-Salas, 1999). Insertion of PLPV sgRNA regions preceding the start codon of the p37 gene into HIS-GFP bicistronic constructs did not promote translation of the GFP gene (data not shown), arguing against the involvement of an IRES in p37 translation. In concordance with these results, removal of distinct segments preceding the start codon of ORF(p37) in PLPV sgRNA did not substantially affect production of p37 from the corresponding subgenomic transcripts (data not shown).

On the other hand, the 63 codon ORF(p7) is considerably larger than what would be expected for an ORF that directs translation reinitiation (Meijer and Thomas, 2002) and thus, improbably provides access to the ORF(p37) start codon by “escorting” some ribosomes past the ORF(p9.7) start site. In line with this view, abolishment of p7 production in mutant p7aug- did not negatively affect p37 translation, indicating that ribosome scanning is unlikely to be resumed after translation of the first ORF to reinitiate translation at the downstream ORF(p37) start codon (Fig. 2B, lane 2). Ribosomal backwards scanning after reaching the stop codon of the p9.7 gene was also discarded since neither elimination nor premature termination of ORF(p9.7) in mutants p9.7gug-, p9.7stop1 and p9.7stop2 had detrimental effects on p37 production (Fig. 4A). The leaky scanning hypothesis was then explored. It was reasoned that this mechanism could be favoured by the poor translation context initiating ORF(p7), and by the non-AUG weak initiation codon of ORF(p9.7). Another remarkable feature of PLPV sgRNA is that the region preceding ORF(p37) is devoid of AUG codons in any reading frame except for that initiating ORF(p7), which would facilitate scanning toward a downstream AUG. To test this hypothesis, a new set of mutants was generated by converting the GUG start codon of ORF(p9.7) to AUG, CUG or ACG (mutants p9.7aug, p9.7cug and p9.7acg, respectively; Fig. 3B). As expected, the CUG and the ACG codons directed translation initiation with an efficiency similar to that of the wt GUG (Fig. 4B, lanes 2 and 3), whereas a 4-fold enhanced p9.7 production was observed with the AUG codon (Fig. 4B,

lane 4). This enhancement correlated with a 2.5-fold reduction in p37 synthesis supporting leaky scanning as the mechanism that directs translation of p37. In line with this possibility, introduction of two additional AUGs (in good contexts) within the region preceding the start codon of ORF(p37) in mutant p9.7aug*3 (Fig. 3B), led to a stronger reduction (10-fold) in p37 production (Fig. 4B). It is noteworthy that the AUG codon at position 2441 engineered in mutant p9.7aug*3, which is in frame with that engineered at position 2465 but out of frame with respect to that opening ORF(p9.7), could potentially function as start codon of an ORF ending at ²⁶⁰⁹UAG. Translation of such ORF would lead to a polypeptide with predicted molecular mass of 6.2 kDa. Though no clear evidence for expression of this polypeptide was obtained, a reinforcement of the p7 band was observed with construct p9.7aug*3 (Fig. 4B, lane 5) which suggested co-migration of p7 with an additional translation product, a point which could not be unequivocally confirmed.

Finally, the presence of strong secondary structures in the leader sequence of mRNAs usually reduces the translation efficiency of the downstream ORF(s) as they interfere with assembly of the pre-initiation complex and/or impede conventional scanning of 40S subunits (Kozak, 1989b). Introduction of a small and stable hairpin in the leader sequence of PLPV sgRNA, 6 nt down from the 5'-end, in mutant pPLDS-10hp (Fig. 3B), severely diminished production of the three sgRNA encoded proteins (Fig. 4B, lane 5), clearly showing that translation of p37, as that of p7 and p9.7, is 5'-dependent as expected for a leaky scanning mechanism. Our results taken as a whole strongly suggested coordinate expression of p7, p9.7 and p37 genes from PLPV sgRNA via leaky scanning.

Alteration of production ratios of PLPV sgRNA encoded proteins strongly affects viral infectivity

To determine whether modification of the production ratios of sgRNA encoded proteins could affect virus infectivity, mutants p9.7cug, p9.7acg and p9.7aug were bioassayed in *C. quinoa*. Mutant p9.7cug, whose sgRNA directed protein translation in fashion similar to that of wt molecule (Fig. 4B, lane 2), gave rise to infections indistinguishable from that established by construct pPLDS-10, as assessed by the appearance of lesions on the inoculated leaves and by Northern blot hybridization (data not shown). Conversely, no infection was detected after inoculation with mutant p9.7aug which showed a strong bias in protein ratios in in vitro translation assays with the corresponding sgRNA (Fig. 4B, lane 4). The engineered nucleotide substitution in the latter mutant (²³⁷¹G → A; Fig. 3) caused a change (S33N) in the amino acid sequence encoded by the overlapping ORF(p7) which could be responsible of the inability of the mutant to infect. However, mutant p9.7acg carries nucleotide substitutions (²³⁷¹GU → AC; Fig. 3) that also lead to the change S33N in p7 but allow p9.7 translation initiation at an efficiency equivalent to that of the wt start codon (Fig. 4B, lane 3), and this mutant was infectious when inoculated on *C. quinoa* though the appearance of lesions was delayed 3–4 days with respect to the wt. Sequence analysis of total RNA from infected leaves 2 weeks post-inoculation, revealed that the mutated codon (²³⁷¹ACG) was maintained and did not revert to the wild-type GUG, indicating that the change S33N in p7 is not deleterious and that the lack of infectivity of mutant p9.7aug is likely due to the unbalance in protein production.

Discussion

In vitro transcripts from mutagenized PLPV cDNA clones have allowed assignment of functional significance to ORFs on the PLPV genome. Based on amino acid sequence comparisons with proteins of known function it was hypothesized that the translation products of ORFs (p27) and (p87) would be involved in viral RNA replication, and those of ORFs (p7) and (p37) in viral movement and encapsidation,

respectively (Castaño and Hernández, 2005). The results of this work are consistent with these anticipations and have provided information on two additional, previously predicted ORFs (p13 and p6). These newly identified ORFs potentially encode for proteins with no noticeable similarity to protein sequences in data banks. Mutations affecting ORF(p13) had no obvious effects on viral infection in three distinct experimental hosts and, moreover, the expected translation product was not detected in *in vitro* translation assays suggesting that this ORF is not expressed *in vivo*. Evidence for expression of p6 was also not obtained. Instead, another protein, p9.7, was found to be essential for virus spread indicating that the double gene block typically involved in movement of carmo-like viruses (Hull, 2002; Li et al., 1998; Navarro et al., 2006) must be formed by ORF(p7) and the newly identified ORF(p9.7), both overlapping and lying in the middle of the genome. In agreement with this view, p9.7 has a hydrophobicity profile which closely resembles those of the larger movement proteins of carmoviruses (Navarro et al., 2006) despite it shows scarce sequence similarity with them (data not shown). Remarkably, this PLPV protein shares notable sequence identity (44.4%, as indicated above) with a 9.4 kDa protein potentially encoded by a centrally located ORF of PCRPV. The initiation codon of such ORF would also correspond to a non-AUG triplet (Kinard and Jordan, 2002) further extending the resemblances among PLPV and PCRPV, both tentative members of the prospective genus *Pelarspovirus*.

Bioassay of mutant p37aug- showed that abolishment of CP production has a drastic effect on local and, consequently, also on systemic PLPV movement. The requirement of CP for cell-to-cell and systemic movement varies among distinct members of the *Tombusviridae*. Thus, CP expression is dispensable for cell-to-cell movement of species of the genera *Aureusvirus* (Reade et al., 2003; Rubino and Russo, 1997), *Necrovirus* (Molnar et al., 1997), and *Dianthovirus* (Vaewhongs and Lommel, 1995). Conversely, CP-deficient mutants of *Panicum mosaic virus* (PMV), in the genus *Panicovirus* (Turina et al., 2000), or *Melon necrotic spot virus*, in the genus *Carmovirus* (Genovés et al., 2006), failed to move from cell to cell. In the case of *Turnip crinkle virus*, another member of the genus *Carmovirus*, CP is required for cell-to-cell movement in *N. benthamiana* but not in *Arabidopsis*, though is essential for systemic movement in both hosts (Cohen et al., 2000; Hacker et al., 1992; Heaton et al., 1991). These observations contrast with those obtained with members of the closely related genus *Tombusvirus* which can move both cell to cell and systemically in the absence of CP (Russo et al., 1994; Sit et al., 1995). Whether the CP of PLPV has a direct role in movement by forming part of the complex that spreads the virus in the plant or contributes indirectly to the establishment and maintenance of the infection remains to be ascertained. Such indirect contribution could result from the ability of the protein to inhibit RNA silencing-mediated degradation of viral RNA as several related carmoviral CPs have been reported to suppress of RNA silencing (Genovés et al., 2006; Martínez-Turiño and Hernández, 2009; Meng et al., 2006; Qu et al., 2003; Thomas et al., 2003). Alternatively, PLPV CP could function as an accessory factor protecting the genome or promoting *in vivo* translation and/or stability of movement protein(s) as suggested for the CP of PMV (Turina et al., 2000).

Besides identifying functional ORFs in the PLPV genome, valuable insights into the mechanisms involved in their translation have been obtained. Expression of the distinct PLPV ORFs seems to be achieved by a variety of mechanisms of translation which, have to be cap-independent as this structure is absent from the 5' end of PLPV RNAs (Castaño and Hernández, 2005), as occurs in most, if not all, members of the *Tombusviridae* (Lommel et al., 2005). *In vitro* translation assays have confirmed the expression of the replicase proteins, p27 and its readthrough product p87, from the gRNA and, moreover, have corroborated that the 6 nt long leader sequence (to our knowledge, the shortest one reported for a natural mRNA) is able to direct efficient translation even from an uncapped molecule. The

layout of *in vitro* translation assays have also indicated that the sgRNA is polycistronic directing the production of three distinct proteins, p7, p9.7 and p37. Though full characterization of the mechanism(s) implicated in sgRNA gene expression will require additional work, the results of the present study strongly suggest that translation of the internal ORFs is accomplished by leaky scanning. A peculiarity of the sgRNA sequence further supports this view: the 370 nt region preceding the start codon of ORF(p37) contains a single AUG codon. This start codon occurs in an unfavourable context and opens ORF(p7). In all other regions of the viral genome, AUG codons are about as frequent as expected (i.e., one every 64 nt for an incompletely unbiased sequence). The lack of additional AUG codons, even in nontranslated reading phases, seems to indicate that potential translation start codons are under strong negative selection in this segment of the genome. In this scenario, it can be postulated that this particular sequence has evolved to allow a fraction of the scanning ribosomal complexes that have reached the start codon in a suboptimal translation context of ORF(p7) to proceed toward ORF(p9.7). The weak initiation signal provided by the non-canonical start codon of ORF(p9.7) would favour that part of the scanning complexes proceeds further toward ORF(p37). According to this model, introduction of additional upstream initiation signals would be expected to reduce downstream translation as indeed was observed with subgenomic transcripts from mutant p9.7aug and, particularly, p9.7aug*3 (Fig. 4B). Conversely, weakening or elimination of upstream initiation signals should increase downstream translation as observed with p7aug- which showed enhanced p9.7 production (Fig. 2B). However, we could not detect significant enhancement of p37 production with p9.7gug- sgRNA transcripts (Fig. 4A) which could be justified by the low efficiency at which p9.7 is translated, indicating that the presence of ORF(p9.7) only slightly reduces the number of ribosomes that reach ORF(p37) and initiate there. The model also implies that translation of all sgRNA ORFs must be 5'-dependent and, consistently, a strong negative effect on p7, p9.7 and p37 production was observed when a stable hairpin was introduced in the leader sequence of the sgRNA (Fig. 4B).

Leaky scanning is a mechanism frequently used to translate different proteins from viral RNAs (reviewed by Ryabova et al., 2006). This mechanism usually leads to the production of two polypeptides either with a common C-terminal region, when being translated from the same ORF, or non related, if the start codons are located in different phases (Belsham and Lomonosoff, 1991; Dinesh-Kumar and Miller, 1993; Herzog et al., 1995; Shivaprasad et al., 2005; Simón-Buela et al., 1997; Verchot et al., 1998; Weiland and Dreher, 1989). The most remarkable case of leaky scanning described to date refers to the production of three totally unrelated proteins from ORFs I, II and III of a single RNA of the pararetrovirus *Rice tungro bacilliform virus* (RTBV; Fütterer et al., 1997). Two additional viral RNAs that give rise to three proteins via leaky scanning have also been reported though two of them corresponded to C-terminal nested polypeptides (Latorre et al., 1998; Suzuki et al., 1996). Translation of PLPV sgRNA provides a notable new example of the use of leaky scanning for synthesis of three distinct proteins from three separate, out of frame, translation initiation sites. Leaky scanning in PLPV sgRNA parallels that of RTBV RNA which is favoured by the non-AUG codon that opens ORFI and the poor translational context of the AUG start codon of ORFII allowing a significant fraction of ribosomes to reach the further downstream ORFIII AUG (Fütterer et al., 1997). Moreover, a strong bias against AUG codons is also observed in the region preceding ORFIII, as ORFs I and II lack internal AUGs in any reading frame, a feature conserved in related badnaviruses which suggests that the leaky scanning mechanism is employed by two distinct genera of bacilliform pararetroviruses (Pooggin et al., 1999). In this context, it is worth mentioning that PCRPV sgRNA is predicted to resemble PLPV sgRNA in ORF arrangement, in type and translational contexts of ORF start

codons and in the absence of AUGs in the region preceding the CP gene excepting for that opening the corresponding ORF(p7) (Kinard and Jordan, 2002), indicating that the mechanism of leaky scanning proposed for PLPV might be shared by other members of the prospective genus *Pelarspovirus*.

The lack of infectivity of mutant p9.7aug contrasts with the good infectivity levels of mutants p9.7cug and p9.7acg and strongly suggests that preservation of the ratios of the proteins produced from the sgRNA is a requisite for virus viability. Moreover, these results show that the wt GUG start codon of ORF(p9.7) can be functionally replaced by CUG or ACG but not by a canonical AUG start codon. The utilization of weakly initiating non-AUG codon for maintenance of correct ratios of proteins has been reported for both cellular and viral mRNAs (Kozak, 1991; Touriol et al., 2003), including some from plant viruses belonging to different taxonomic groups (Fütterer et al., 1996; Shirako, 1998). Within the family *Tombusviridae*, HCRSV is the only carmovirus which has been shown to utilize a non-AUG initiation codon for translation of a virus-specific protein which seems to affect symptom expression (Koh et al., 2006; Zhou et al., 2006). A second example of non-AUG start codon usage in the family *Tombusviridae* corresponds to PMV, which encodes a small movement protein whose translation initiation occurs at AUU (Turina et al., 1998). As PLPV, PMV produces a unique, polycistronic sgRNA from which four proteins are expressed, and it has been suggested that leaky scanning may account for translation of the two 5'-proximal genes whereas translation of the two 3'-terminal genes may involve a combination of leaky scanning with as yet unidentified translation strategies (Batten et al., 2006; Turina et al., 2000). Though the data presented here strongly support that polycistronic translation of PLPV sgRNA is mainly accomplished by leaky scanning, the contribution of other mechanisms cannot completely be ruled out. Another open question refers to the strategies PLPV RNAs adopt to be efficiently translated in the absence of both a 5' cap and a poly(A) tail. To compensate for the lack of these structures at the RNA termini, members of the families *Tombusviridae* and *Luteoviridae* contain sequences in the 3' region of their genomes which function as translational enhancers (TEs) (reviewed by Kneller et al., 2006). PLPV has also been predicted to possess a 3' TE (Fabian and White, 2006). Investigation on this issue may help to further understand the regulation complexity of PLPV gene expression and the network of molecular interactions involved.

Materials and methods

Plasmid construction

Plasmid pPLDS-10, which contains a wild-type (wt) full-length PLPV cDNA inserted into pUC18 downstream from a T7 RNA polymerase promoter, has been previously described (Castaño and Hernández, 2007). This plasmid was used as template to generate a set of mutant constructs bearing nucleotide substitutions that either altered the presumed start codons of PLPV ORFs or led to the incorporation of premature stop codons in a given ORF. The mutations were introduced by PCR with the Quick Change Site-Directed Mutagenesis kit (Stratagene) and proper specific oligonucleotide pairs. Three mutant constructs carrying internal deletions were also generated. To this aim, plasmid pPLDS-10 was used as template for PCR reactions with *Pfu Ultra* DNA polymerase (Stratagene) and different pairs of 5'-phosphorylated primers that were complementary and homologous to distinct regions preceding ORF(p37) in PLPV sgRNA. PCR products (comprising the cloning vector, pUC18, fused to a PLPV 5' region at one side and to a PLPV 3' region at the other side) were eluted after agarose electrophoresis, self-ligated and used for transformation. The deleted regions in the different constructs comprised nt 2496–2601, 2296–2376 and 2378–2601 of PLPV sequence. Each mutant construct was verified by DNA sequencing with an ABI PRISM DNA sequencer 377 (Perkin-Elmer). The type and

position of the modifications introduced in each construct is indicated in the figures depicting the mutants. Unless indicated, the point mutations introduced in a particular ORF did not affect the amino acid sequence encoded by overlapping ORFs.

A region embracing the cDNA sequence of the PLPV sgRNA was PCR amplified from pPLDS-10 and the mutagenized clones with the Expand High Fidelity PCR System (Roche) using primer CH94 (5'-GCAAGCTT**GTAATACGACTCACTATAGGG**ACATTGAAGTTGGCGCTATC-3'), which contains a HindIII site (underlined) fused to a T7 RNA polymerase promoter sequence (in bold) followed by 21 nt of the 5' end of the sgRNA (position 2251–2271 in the gRNA), and CH60 (5'-CCGGATCCCGGGCAGATCAGGGGGTGGGTTAC-3'), complementary to the 3' terminus of the viral sequence (nt 3859–3883) with a SmaI site (underlined) and a BamHI site (in italics) at the 5' end. PCR products were run in 1% agarose gels, eluted, digested with HindIII and BamHI and subsequently cloned into the corresponding restriction sites of pUC18. An inverted repeat (5'-GGGGGGGCATGCCCCCCC-3') promoting the formation of a small hairpin into the leader sequence of the PLPV sgRNA was introduced in a subgenomic construct with Quick Change Site-Directed Mutagenesis kit (Stratagene) and proper specific oligonucleotide pairs.

A plasmid containing the full-length PLPV cDNA of pPLDS-10 under the control of the CaMV 35S promoter (Castaño and Hernández, 2007) was used as template to introduce selected mutations by PCR with the Quick Change Site-Directed Mutagenesis kit (Stratagene) and proper specific oligonucleotide pairs. The wt and the mutant full-length PLPV cDNAs fused to the CaMV 35S promoter were cut out with suitable restriction enzymes from the corresponding constructs, ligated at its 3' side with the terminator sequence of the *Solanum tuberosum* proteinase inhibitor II gene, and cloned into the binary Ti vector pMOG800 (Knoester et al., 1998).

To generate bicistronic constructs, a plasmid containing the HIS and the GFP genes in tandem, pHIS-GFP, was made. To this aim, the HIS and the GFP genes were PCR amplified with High Fidelity PCR System (Roche) from proper plasmids using specific oligonucleotides. The oligonucleotide homologous to the 5'-terminus of the HIS gene harboured a BamHI site at the 5'-end, and the one complementary to the 3'-terminus contained a SphI site at the 5' end. The GFP-specific oligonucleotides led to the incorporation of SphI/BamHI and PstI sites at the 5'- and 3'-ends, respectively, of the corresponding PCR product. The amplified HIS and GFP genes were ligated through the SphI site and then cloned into BamHI and PstI sites of pBluescript KS (+) (Stratagene). Additional bicistronic constructs were generated from pHIS-GFP by inserting PCR amplified PLPV sequences (nt 2283–2370 and 2283–2495), preceding ORF(p37) in the sgRNA, at the SphI and BamHI sites present between the HIS and GFP genes.

Mechanical and *Agrobacterium*-mediated inoculation of plants

For mechanical inoculation of plants, uncapped RNA transcripts were synthesized in vitro from pUC18 derived full-length PLPV clones with T7 RNA polymerase (Fermentas) following digestion of plasmids with SmaI. The resulting transcripts were gently rubbed onto carborundum-dusted leaves of *C. quinoa*, *N. benthamiana* or *N. clevelandii* plants (three leaves per plant employing approximately 1 µg of transcript per leaf). Plants were maintained under greenhouse conditions (16 h days at 24 °C, 8 h nights at 20 °C) and leaf samples were harvested 7 (in the case of *C. quinoa*) or 28 (in the case of *N. benthamiana* and *N. clevelandii*) days post-inoculation.

For *Agrobacterium*-mediated inoculation of plants, the Ti plasmid constructs were introduced into *A. tumefaciens* strain C58C1 by heat shock. Cultures of *A. tumefaciens* harbouring the different Ti plasmid constructs were infiltrated at an OD₆₆₀ of 0.5 on the abaxial side of *N. benthamiana* leaves using a 5 ml needleless syringe. The infiltrated plants were kept under greenhouse conditions and leaf samples were taken 3–4 days after infiltration.

RNA extraction and analysis

Total RNA preparations of healthy and PLPV infected *C. quinoa*, *N. benthamiana* or *N. clevelandii* leaves were obtained by phenol extraction and lithium precipitation (Verwoerd et al., 1989). For Northern blot analysis, 4 µg of total RNA were denatured by glyoxal-dimethyl sulfoxide treatment, electrophoresed in 1% agarose gels, blotted to nylon membranes (Hybond N+, Amersham), and hybridized as described previously (Castaño and Hernández, 2005) with a ³²P-radioactive DNA probe encompassing nt 3095 to 3883 of the PLPV gRNA. For dot-blot analysis, leaf samples were squeezed in 20 volumes (w/v) of extraction buffer (50 mM sodium citrate, pH 8.5), and the homogenates were clarified by centrifugation at 6000 rpm for 5 min. The clarified extracts were spotted (4 µl) onto positively charged nylon membranes (Roche). Hybridization with a digoxigenin-labeled PLPV riboprobe was performed as described previously (Ivars et al., 2004).

In vitro translation assays

Uncapped transcripts were synthesized in vitro from PLPV genomic and subgenomic constructs as indicated above or from bicistronic constructs linearized with XhoI. The RNA transcripts were used for in vitro translation experiments with Wheat Germ Extract Plus following manufacturer recommendations (Promega). [¹⁴C] leucine or [³⁵S]methionine was included in the reactions and the translation products were treated with 2% sodium dodecyl sulphate (SDS) and separated by electrophoresis through 12% polyacrylamide gels in Tris-glycine buffer (Sambrook and Russell, 2001). The gels were subsequently fixed in 25% isopropanol and 10% acetic acid for 30 min, incubated with Amersham Amplify™ Fluorographic Reagent (GE Healthcare), dried and exposed to X-ray film or, alternatively, scanned with a bioimage analyzer (Fuji BAS1500) for quantitation of signals.

Acknowledgments

This research was supported by grants AGL2003-04249 (MCYT) and BFU2006-11230 (MEC) and by grant ACOM06/210 from the Generalitat Valenciana (Spain) to C.H.. A.C. was a recipient of predoctoral fellowships from the Generalitat Valenciana and from the CSIC-Fundación Bancaja and L.R. received a postdoctoral contract from the Juan de la Cierva program of MEC. We are indebted to Olga Fernández-Miragall for construct pHIS-GFP and to Dolores Arocas and Isabella Avellaneda for excellent technical assistance. We also thank Dr. Campbell W. Gourlay (University of Kent) for generously reviewing the English grammar.

References

Alonso, M., Borja, M., 2005. High incidence of Pelargonium line pattern virus infecting asymptomatic *Pelargonium* spp. in Spain. *Eur. J. Plant Pathol.* 112, 95–100.

Batten, J.S., Desvoyes, B., Yamamura, Y., Scholthof, K.B., 2006. A translational enhancer element on the 3'-proximal end of the Panicum mosaic virus genome. *FEBS Lett.* 580, 2591–2597.

Belsham, G.J., Lomonosoff, G.P., 1991. The mechanism of translation of cowpea mosaic virus middle component RNA: no evidence for internal initiation from experiments in an animal cell transient expression system. *J. Gen. Virol.* 72, 3109–3113.

Bouwen, I., Maat, D.Z., 1992. Pelargonium flower-break and Pelargonium line pattern viruses in the Netherlands; purification, antiserum preparation, serological identification, and detection in pelargonium by ELISA. *Neth. Plant Pathol.* 98, 141–156.

Castaño, A., Hernández, C., 2005. Complete nucleotide sequence and genome organization of Pelargonium line pattern virus and its relationship with the family *Tombusviridae*. *Arch. Virol.* 150, 949–965.

Castaño, A., Hernández, C., 2007. Biological activity of transcripts from cDNA of Pelargonium line pattern virus. *Acta Virol.* 51, 271–274.

Chiba, M., Reed, J.C., Prokhnovsky, A.I., Chapman, E.J., Mawassi, M., Koonin, E.V., Carrington, J.C., Dolja, V.V., 2006. Diverse suppressors of RNA silencing enhance agroinfection by a viral replicon. *Virology* 346, 7–14.

Cohen, Y., Gisel, A., Zambryski, P.C., 2000. Cell-to-cell and systemic movement of recombinant green fluorescent protein-tagged turnip crinkle viruses. *Virology* 273, 258–266.

Dinesh-Kumar, S.P., Miller, W.A., 1993. Control of start codon choice on a plant viral RNA encoding over overlapping genes. *Plant Cell* 5, 679–692.

Fabian, M.R., White, K.A., 2006. Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA* 12, 1304–1314.

Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A., 2005. *Virus Taxonomy*. Eighth Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.

Franck, A., Loebenstein, G., 1994. Virus and virus-like diseases of pelargonium in Israel. *Acta Hort.* 377, 31–39.

Fütterer, J., Potrykus, I., Bao, Y., Li, L., Burns, T.M., Hull, R., Hohn, T., 1996. Position-dependent ATT initiation during plant pararetrovirus rice tungro bacilliform virus translation. *J. Virol.* 70, 2999–3010.

Fütterer, J., Rothnie, H.M., Hohn, T., Potrykus, I., 1997. Rice tungro bacilliform virus open reading frames II and III are translated from polycistronic pregenomic RNA by leaky scanning. *J. Virol.* 71, 7984–7989.

Gale Jr, M., Tan, S.L., Katze, M.G., 2000. Translational control of viral gene expression in eukaryotes. *Microbiol. Mol. Biol. Rev.* 64, 239–280.

Gallie, D.R., 1996. Translational control of cellular and viral mRNAs. *Plant Mol. Biol.* 32, 145–158.

Genovés, A., Navarro, J.A., Pallás, V., 2006. Functional analysis of the five melon necrotic spot virus genome-encoded proteins. *J. Gen. Virol.* 87, 2371–2380.

Gopinath, K., Dragnea, B., Kao, C., 2005. Interaction between Brome mosaic virus proteins and RNAs: effects on RNA replication, protein expression, and RNA stability. *J. Virol.* 79, 14222–14234.

Gordon, K., Fütterer, J., Hohn, T., 1992. Efficient initiation of translation at non-AUG triplets in plant cells. *Plant J.* 2, 809–813.

Hacker, D.L., Petty, I.T.D., Wei, N., Morris, T.J., 1992. Turnip crinkle virus genes required for RNA replication and virus movement. *Virology* 186, 1–8.

Heaton, L.A., Lee, T.C., Wei, N., Morris, T.J., 1991. Point Mutations in the Turnip crinkle virus capsid protein affect the symptoms expressed by *Nicotiana benthamiana*. *Virology* 183, 143–150.

Hellen, C.U., Sarnow, P., 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* 15, 1593–1612.

Herzog, E., Guillely, H., Fritsch, C., 1995. Translation of the second gene of peanut clump virus RNA2 occurs by leaky scanning in vitro. *Virology* 208, 215–225.

Hull, R., 2002. Virus movement through the plant and effects on plant metabolism. *Matthews' Plant Virology*, Fourth edition. Academic Press, San Diego, pp. 373–436.

Ivars, P., Alonso, M., Borja, M., Hernández, C., 2004. Development of a non-radioactive dot-blot hybridisation assay for the detection of Pelargonium flower break virus and Pelargonium line pattern virus. *Eur. J. Plant Pathol.* 110, 275–283.

Kinard, G.R., Jordan, R., 2002. Genome organization of Pelargonium chlorotic ring pattern virus: further implications for *Tombusviridae* taxonomy. *Acta Hort.* 568, 17–27.

Kneller, E.L., Rakotondrafara, A.M., Allen Miller, W., 2006. Cap-independent translation of plant viral RNAs. *Virus Res.* 119, 63–75.

Knoester, M., van Loon, L.C., van den Heuvel, J., Hennig, J., Bol, J.F., Linthorst, H.J.M., 1998. Ethylene-insensitive tobacco lacks nonhost resistance against soil-borne fungi. *Proc. Natl. Acad. Sci. U. S. A.* 95, 1933–1937.

Koh, D.C., Wong, S.M., Liu, D.X., 2003. Synergism of the 3'-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein. *J. Biol. Chem.* 278, 20565–20573.

Koh, D.C., Wang, X., Wong, S.M., Liu, D.X., 2006. Translation initiation at an upstream CUG codon regulates the expression of *Hibiscus chlorotic ringspot virus* coat protein. *Virus Res.* 122, 35–44.

Kozak, M., 1989a. Context effects and inefficient initiation at non-AUG codons in eukaryotic cell free translation systems. *Mol. Cell Biol.* 9, 5073–5080.

Kozak, M., 1989b. Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs. *Mol. Cell Biol.* 9, 5134–5142.

Kozak, M., 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* 266, 19867–19870.

Kozak, M., 2001. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res.* 29, 5226–5232.

Kozak, M., 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1–34.

Latorre, P., Kolakofsky, D., Curran, J., 1998. Sendai virus Y proteins are initiated by a ribosomal shunt. *Mol. Cell Biol.* 18, 5021–5031.

Li, W.-Z., Qu, F., Morris, T.J., 1998. Cell-to-cell movement of turnip crinkle virus is controlled by two small open reading frames that function in trans. *Virology* 244, 405–416.

Lommel, S.A., Martelli, G.P., Rubino, L., Russo, M., 2005. Family *Tombusviridae*. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Eighth Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, USA, pp. 907–936.

Luukkonen, B.G.M., Tan, W., Schwartz, S., 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J. Virol.* 69, 4086–4094.

Martínez-Salas, E., 1999. Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.* 10, 458–464.

Martínez-Turiño, S., Hernández, C., 2009. Inhibition of RNA silencing by the coat protein of Pelargonium flower break virus: distinctions from closely related suppressors. *J. Gen. Virol.* 90, 519–525.

Meijer, H.A., Thomas, A.A., 2002. Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem. J.* 367, 1–11.

- Meng, C., Chen, J., Peng, J., Wong, S.M., 2006. Host-induced avirulence of *Hibiscus chlorotic ringspot virus* mutant correlates with reduced gene-silencing suppression activity. *J. Gen. Virol.* 87, 451–459.
- Molnar, A., Havelda, Z., Dalmay, T., Szutorisz, H., Burgyan, J., 1997. Complete nucleotide sequence of tobacco necrosis virus strain DH and genes required for RNA replication and virus movement. *J. Gen. Virol.* 78, 1235–1239.
- Navarro, J.A., Genovés, A., Climent, J., Saurí, A., Martínez-Gil, L., Mingarro, I., Pallás, V., 2006. RNA-binding properties and membrane insertion of Melon necrotic spot virus (MNSV) double gene block movement proteins. *Virology* 356, 57–67.
- Peabody, D.S., 1989. Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* 264, 5031–5035.
- Peabody, D.S., Subramani, S., Berg, P., 1986. Effect of upstream reading frames on translation efficiency in simian virus 40 recombinants. *Mol. Cell. Biol.* 6, 2704–2711.
- Pooggin, M.M., Fütterer, J., Skryabin, K.G., Hohn, T., 1999. A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. *J. Gen. Virol.* 80, 2217–2228.
- Qu, F., Ren, T., Morris, T.J., 2003. The coat protein of Turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. *J. Virol.* 77, 511–522.
- Reade, R., Miller, J., Robbins, M., Xiang, Y., Rochon, D., 2003. Molecular analysis of the cucumber leaf spot virus genome. *Virus Res.* 91, 171–179.
- Ryabova, L.A., Pooggin, M.M., Hohn, T., 2006. Translation reinitiation and leaky scanning in plant viruses. *Virus Res.* 119, 52–62.
- Rubino, L., Russo, L., 1997. Molecular analysis of the pothos latent virus genome. *J. Gen. Virol.* 78, 1219–1226.
- Russo, M., Burgyan, J., Martelli, G., 1994. Molecular biology of *Tombusviridae*. *Adv. Virus Res.* 44, 381–428.
- Sambrook, J., Russell, D., 2001. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbour, Laboratory Press, New York.
- Shirako, Y., 1998. Non-AUG translation initiation of a plant RNA virus: a forty-amino-acid extension is added to the N terminus of the soil-borne *Wheat mosaic virus* capsid protein. *J. Virol.* 72, 1677–1682.
- Shivaprasad, P.V., Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T., Pooggin, M.M., 2005. Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. *J. Virol.* 79, 8149–8163.
- Simón-Buela, L., Guo, H.S., García, J.A., 1997. Cap-independent leaky scanning as the mechanism of translation initiation of a plant viral genomic RNA. *J. Gen. Virol.* 78, 2691–2699.
- Sit, T.L., Johnston, J.C., Ter-Borg, M.G., Frison, E., McLean, M.A., Rochon, D., 1995. Mutational analysis of the cucumber necrosis virus coat protein gene. *Virology* 206, 38–48.
- Stone, O.M., 1980. Nine viruses isolated from pelargonium in the United Kingdom. *Acta Hort.* 110, 177–182.
- Stuart, G.W., Moffett, P.K., Bozarth, R.F., 2006. A comprehensive open reading frame phylogenetic analysis of isometric positive strand ssRNA plant viruses. *Arch. Virol.* 151, 1159–1177.
- Suzuki, N., Sugawara, M., Nuss, D.L., Matsuura, Y., 1996. Polycistronic (tri- or bicistronic) phytoreoviral segments translatable in both plant and insect cells. *J. Virol.* 70, 8155–8159.
- Thomas, K.K., Capecci, M.R., 1986. Introduction of homologous DNA sequences into mammalian cell lines induces mutations in the cognate gene. *Nature* 324, 34–38.
- Thomas, C.L., Leh, V., Lederer, C., Maule, A.J., 2003. Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. *Virology* 306, 33–41.
- Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A.C., Vagner, S., 2003. Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol. Cell* 95, 169–178.
- Turina, M., Maruoka, M., Monis, J., Jackson, A.O., Scholthof, K.B., 1998. Nucleotide sequence and infectivity of a full-length cDNA clone of *Panicum mosaic virus*. *Virology* 241, 141–155.
- Turina, M., Desvoyes, B., Scholthof, K.B., 2000. A gene cluster encoded by *Panicum mosaic virus* is associated with virus movement. *Virology* 266, 120–128.
- Vaewhongs, A.A., Lommel, S.A., 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. *Virology* 212, 607–613.
- Verchot, J., Angell, S.M., Baulcombe, D., 1998. In vivo translation of the triple gene block of potato virus X requires two subgenomic RNAs. *J. Virol.* 72, 8316–8320.
- Verwoerd, T.C., Dekker, B.M.M., Hoekema, A., 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res.* 17, 2362.
- Voinnet, O., Lederer, C., Baulcombe, D.C., 2000. A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. *Cell* 103, 157–167.
- Weiland, J., Dreher, T.W., 1989. Infectious TYMV RNA from cloned cDNA: effects in vitro and in vivo of point substitutions in the initiation codons of two extensively overlapping ORFs. *Nucleic Acids Res.* 17, 4675–4687.
- Zhou, T., Fan, Z.F., Li, H.F., Wong, S.M., 2006. *Hibiscus chlorotic ringspot virus* p27 and its isoforms affect symptom expression and potentiate virus movement in kenaf (*Hibiscus cannabinus* L.). *Mol. Plant-Microb. Interact.* 19, 948–957.

Capítulo IV

Genetic structure and evolution of Pelargonium line pattern virus populations

Aurora Castaño, Leticia Ruiz, Santiago F. Elena, Carmen Hernández*

Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas-UPV), Avenida de los Naranjos, 46022 Valencia, Spain

To be submitted

Abstract

The genomic structure of Pelargonium line pattern virus (PLPV), a tentative member of a proposed new genus within the family *Tombusviridae*, has been recently determined. However, little is known about the genetic variability and population structure of this pathogen. Here, we have investigated the heterogeneity of PLPV isolates from different origins by sequence analysis of a 1817 nt fragment encompassing the movement (p7 and p9.7) and coat (CP) protein genes as well as flanking segments including the complete 3' untranslated region. The results of the analyses have revealed that distinct protein domains are under different selective constraints and that maintenance of certain foldings in RNA regulatory regions might be an important factor limiting the heterogeneity. Covarying amino acids within and between proteins have been detected suggesting specific intra- and inter-protein interactions during the infectious cycle. We also report evidences of an extremely rapid evolution of PLPV after being transferred to a new plant species. Moreover, the evolutionary trajectory after host switch was perfectly reproducible, highlighting the strong constraints that the host imposes to the genetic structure of PLPV populations.

INTRODUCTION

RNA viruses are known to generate high levels of genetic variation that allow them to evolve rapidly facilitating their successful adaptation to new environments and the expansion of their host ranges. The low fidelity of the viral encoded RNA dependent-RNA polymerases (RdRps), that may lack proofreading functions (Agol, 2006; Castro *et al.*, 2005; Holland *et al.*, 1982, Steinhauer and Holland, 1986), is the underlying source for most variation. Indeed, the mutation rates of RNA viruses (i.e., the number of nucleotide misincorporations per site per round of replication) are 10^4 to 10^5 higher than those of their DNA hosts (Domingo and Holland, 1997; Drake and Holland, 1999).

In addition to the error rate of viral RdRps, other factors play a major role in structuring genetic diversity, including selection and genetic bottlenecks as those that occur during both systemic infection (French and Stenger, 2003; Li and Roossinck, 2004; Sacristán *et al.*, 2003) and horizontal transmission by vectors (Ali *et al.*, 2006; Betancourt *et al.*, 2008). These other factors may lead to a considerable genetic stability as it has been reported for many different plant RNA viruses that appear more genetically stable than their animal counterparts (García-Arenal *et al.*, 2001, 2003). This could be due to a combination of intrinsically lower rates of mutation (Malpica *et al.*, 2002) and a reduced fixation rate of advantageous non-synonymous mutations because of weaker immune selection (García-Arenal *et al.*, 2001). The identification and manipulation of factors that regulate the composition of the viral populations may offer a new set of tools to predict or control emerging diseases.

Pelargonium line pattern virus (PLPV) is a major geranium (*Pelargonium* spp.) pathogen in Spain and probably worldwide (Alonso and Borja, 2005; Bouwen and Maat, 1992; Franck and Loebenstein, 1994; Stone, 1980). The frequent symptomless condition of PLPV infections (Alonso and Borja, 2005) compromise regulatory inspections and might have contributed to the spread of the virus. The factors that influence the appearance of symptoms, characterized by yellow-green spots and line patterns on the leaves, remain unclear but they are likely a combination of the viral isolate, the environmental conditions and the geranium cultivar. Though natural infections by PLPV seem to be restricted to *Pelargonium* spp., some experimental hosts as *Chenopodium quinoa* or distinct *Nicotiana* spp. become easily infected after mechanical inoculation of the virus in the greenhouse.

PLPV virions are isometric in shape and hold a single stranded RNA molecule. Cloning and sequencing of genomic RNA (gRNA) together with reverse genetic experiments have recently allowed determination of the genome organization of PLPV (Castaño and Hernández, 2005, 2007; Castaño *et al.*, 2009). The gRNA comprises 3883 nt and contains five open reading frames (ORFs) flanked by an unusually short untranslated region (UTR) of 6 nt at the 5' end and by a 246 nt long UTR at the 3' end. The two 5'-proximal ORFs encode proteins essential for replication, p27 and its read-through product p87 (the viral RdRp). Two small overlapping ORFs, located at the middle of the genome, encode proteins involved in viral movement (p7 and p9.7), while the 3'-proximal ORF encodes the coat protein (p37 or CP). The two replication proteins are translated directly from the gRNA whereas the movement and encapsidation proteins are translated from the unique PLPV subgenomic RNA (sgRNA) of 1.6 kb detected in infected tissue (Castaño *et al.*, 2009).

PLPV has not been formally classified into taxonomic scheme yet. It was formerly considered as a tentative member of the genus *Carmovirus* but recent results supported its inclusion into a prospective new genus (*Pelarspovirus*) in the family *Tombusviridae* (Castaño and Hernández, 2005; Castaño *et al.*, 2009; Stuart *et al.*, 2006). Other tentative species of the prospective genus would be Pelargonium ringspot virus (PeIRSV), Pelargonium chlorotic ring pattern virus (PCRPV) and Elderberry latent virus (ELV), that, as PLPV, produce only one sgRNA (Kinard and Jordan, 2002) in contrast with typical carmoviruses that generate two (Lommel *et al.*, 2005).

The extent of PLPV variability is unknown as, so far, only the complete sequence of a German isolate, that from which the genomic organization of the virus was deduced (Castaño and Hernández, 2005, 2007; Castaño *et al.*, 2009), and a partial sequence of an American isolate, corresponding to the CP gene (Accession No. AY038067), have been reported. Sequence information for other tentative members of the proposed genus *Pelarspovirus* is even scarcer. Indeed, only the primary structure of PCRPV genome has been fully determined while just the CP sequences of PeIRSV and ELV are available.

In this work, we have studied the genetic variability among 10 PLPV isolates recovered from naturally infected geranium plants which were collected in four countries at different times. We have obtained data that have allowed to infer selection constraints acting on PLPV genome and

that suggest geographical segregation of PLPV sequences. Additionally, covariation analysis has permitted to draw a putative network of specific protein interactions likely required to accomplish the different steps of the infection cycle. The combination of linked amino acid substitutions in p9.7 and CP in different isolates pointed to host-driven covariation of these proteins, a possibility that was supported by the evolutionary trajectory of a PLPV population founded with a single variant and subjected to serial transfer in *Chenopodium quinoa*. The results have additionally shown that PLPV evolves very rapidly in the experimental host, and that besides the correlated amino acid mutations in p9.7 and in CP, more than 100 nucleotide substitutions in the 1817 nt region might be the result of adaptation to *C. quinoa*. This elevated percentage of host-selected positions are fixed after only 14 days of viral multiplication, suggesting that the mutation rate of PLPV may be particularly high.

MATERIALS AND METHODS

Viral isolates

Field PLPV isolates were obtained from geranium plants collected at distinct geographical locations over four years (2000-2004). Sap from the original plants was used to pass the virus into the experimental host *C. quinoa* by mechanical inoculation and the viral population was recovered from this infected material. The isolates were designed with the first letters of the country of origin followed by a number to distinguish isolates from the same country (Table 1). PLPV isolate PV-0193, obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and characterized previously (Castaño and Hernández, 2005, 2007), was included in the sequence analyses for comparison purposes.

Reverse transcription, PCR amplification, cloning and sequencing

Total RNA preparations were obtained from infected *C. quinoa* leaves by phenol extraction and lithium precipitation (Verwoerd *et al.*, 1989) and used as templates for reverse transcription (RT) reactions with Superscript II-RT (Invitrogen) and primer CH60 (5'-CCGGATCCCGGGCAGATCAGGGGGGTGGGTTAC-3'), complementary to the 3' terminus of the viral sequence (nt 3859–3883) with a *Sma*I site (underlined) and a *Bam*HI site (in italics) at

the 5' terminus. RT products were PCR amplified with the Expand High Fidelity PCR System (Roche) and primers CH60 and CH17 (5'-GAAAATGGCCTTCTACGGGGAC-3'), homologous to nt 2067-2088 of the PLPV genome. After an initial denaturation step at 94 °C for 2 min, PCR was performed for 35 cycles each of 30 sec at 95°C, 30 sec at 60°C and 3 min at 68°C, followed by an extension step of 10 min at 68 °C. The resulting RT-PCR products were separated by electrophoresis in 1% agarose gels, eluted and cloned into the pGEM-T easy vector (Promega) or the plasmid pTZ19R (Fermentas). Two clones for each DNA fragment were selected for sequencing.

For construction of a full-length infectious cDNA clone derived from isolate SPA4 (Table 1), total RNA from *C. quinoa* leaves infected with such isolate was used as template for RT-PCR reactions employing primers CH241 (5'- TGAATATCCATTCTATAGCGCCAACTTC-3'), complementary to nt 2256-2283 of PLPV gRNA, and CH61 (5'- GCAAGCTTGTAATACGACTCACTATAGGGAACAAAATGGCACACTATTTTGG-3'), which contains a HindIII site (underlined) fused to a T7 RNA polymerase promoter sequence (in bold) followed by 23 nt of the 5' end of the virus sequence. The resulting cDNA, representing a 5'-terminal portion of the PLPV gRNA, was gel-purified and cloned into the plasmid pTZ19R. One clone was randomly selected and the insert was excised out by digestion with HindIII and AgeI and fused to an individual DNA fragment obtained with primers CH60 and CH17 (see above), which was recovered from the corresponding recombinant plasmid by digestion with AgeI and *Bam*HI. Ligation of the 5'- and 3'- regions through the AgeI site (present in the PLPV sequence at nt 2130-2135) yielded a full-length viral cDNA which was subsequently cloned into HindIII and *Bam*HI sites of pUC18. The resulting construct was named pPLSPA4 and the infectivity of the cloned sequence was corroborated as previously described (Castaño and Hernández, 2007). The nucleotide sequence of all viral cDNAs was determined with an ABI PRISM DNA sequencer 377 (Perkin-Elmer).

Serial virus transfer

Construct pPLSPA4 was digested with *Sma*I and transcribed using T7 RNA polymerase (Fermentas) without including cap analog in the transcription reactions. The *in vitro* transcripts were used to mechanically inoculate *C. quinoa* plants as previously reported (Castaño and

Hernández, 2007). The leaves infected by transcript inoculation constituted the starting material (passage zero) for serial transfer of PLPV. Subsequent passages (1 through 4) were done at 7-day interval using total RNA preparations from the infected plant material as inocula. Plants were maintained under greenhouse conditions (16 h days at 24 °C, 8 h nights at 20 °C) after inoculation. Total RNA preparations from infected leaves of passages 1, 2 and 4 were used as templates for RT-PCR amplification with primers CH60 and CH17. The resulting cDNAs were cloned and sequenced as indicated above.

Sequence analysis

Multiple sequence alignments were constructed using MUSCLE (Edgar, 2004). For coding regions, translated amino acid sequences were first aligned and used as guide to built protein-coding nucleotide sequence alignments by concatenating codons using PAL2NAL (Suyama *et al.*, 2006). The best-fitting model of nucleotide substitution was identified by MODELTEST (Posada and Crandall, 1998) as the general reversible GTR + Γ 4 model, with the frequency of each substitution type and the gamma distribution of among-site rate variation with four rate categories estimated from the empirical data. Recombination was ruled out as a potential confounding factor by using GARD (Kosakovsky *et al.*, 2006) and RDP (Martin *et al.*, 2005) methods. A minimum evolution tree was constructed using the above model of nucleotide substitution and its statistical significance was evaluated by bootstrap (upon 10000 pseudoreplicates) using MEGA4 (Tamura *et al.*, 2007). MEGA4 was also used for computing within- and among-population nucleotide diversities (standard errors were computed by the bootstrap method based on 1000 pseudoreplicates) as well as for performing Tajima's relative rates test (Tajima F, 1993). RNA secondary structures of the 3' UTR and of the putative subgenomic promoter of PLPV were predicted using MFOLD version 3.1 (Mathews *et al.*, 1999; Zuker, 2003).

Identification of adaptive evolution in PLPV genomes. It is generally assumed that synonymous substitutions accumulate neutrally, at rate dS per synonymous site, because they have no effect on the amino acid composition of proteins and, henceforth, may not affect proteins folding and functioning. In contrast, nonsynonymous substitutions, occurring at rate dN per nonsynonymous site, involve amino acid replacements and are more likely to affect, for bad

or for good, the folding and function of proteins. The intensity of selection, ω , can thus be evaluated as the ratio dN/dS. Values $\omega < 1$, $= 1$ or > 1 indicate, respectively, purifying selection, neutral evolution, and directional selection, respectively (Sharp, 1997). Here, we have used the several maximum likelihood Bayesian methods available in the HYPHY packaged (Kosakovsky and Frost, 2005) as implemented in the www.datamonkey.org server. Each ORF was separately analyzed. Only sites identified by at least half of the six methods available will be reported.

Molecular covariation within and among proteins. Selection not necessarily acts on single amino acids but it may operate on groups of amino acids in a concerted manner. This being the case, changes in one amino acid should appear associated to changes in the other members of the interaction group. Two different methods were used to evaluate the presence of coevolving amino acids within any given protein. First the Bayesian graphical model implemented in SPIDERMONKEY (Poon *et al.*, 2007) and available online in the datamonkey server. Second, the mutual information content (MIC) approach described in Codoñer *et al.* (2006). Significance P-values were computed, based on a million permutations, as the fraction of shuffles with a MIC value greater than or equal to the observed value. To minimize the number of false positives, the FDR method was applied (Benjamini and Hochberg, 1995). Only sites predicted to covary by both methods will be reported. Intermolecular covariation was only evaluated using the second methodology.

RESULTS

Genetic diversity in coding and non-coding regions of PLPV genome

The primary structure of a total of 18 PLPV cDNAs, 1817 nt in length and derived from nine isolates (Table 1), was determined, and the resulting sequences were combined with the two additional ones previously characterized from a German isolate (Castaño and Hernández, 2005, 2007), producing a total dataset of 20 sequences which encompassed part of the RdRp gene, the complete p7, p9.7 and CP genes as well as the 3' UTR.

Genetic distances between each pair of sequences ranged from 0.001 to 0.081. The maximum values were found between pairs of sequences from isolates sampled during different years at different countries (e.g.: pair SPA3-USA1 or SPA3-ITA4), and the minimal values were

detected for pairs of sequences of a given isolate, as they could differ by just 1 nt (e.g., sequences from isolate ITA2).

Table 1. PLPV isolates used in this work

Isolate ^a	Country	Year(s)	Original Host	Accesion number
SPA0	Spain	2000	<i>P. zonale</i>	EU852912, EU852913
SPA1	Spain	2000	<i>P. zonale</i>	EU852914, EU852915
SPA3	Spain	2000	<i>P. zonale</i>	EU852916, EU852917
SPA4	Spain	2000	<i>P. zonale</i>	EU849616, EU852918
SPA7	Spain	2000	<i>P. zonale</i>	EU852919, EU852920
SPA8	Spain	2004	<i>P. zonale</i>	EU852921, EU852922
USA1	USA	2006	<i>P. peltatum</i>	EU852923, EU852924
ITA2	Italy	2004	<i>P. zonale</i>	EU852925, EU852926
ITA4	Italy	2004	<i>P. zonale</i>	EU852927, EU852928
PV-0193	Germany	1990s	<i>P. peltatum</i>	AY613852, EU835946

^a PV-0193 has been previously characterized by Castaño and Hernández (2005; 2007)

The nucleotide diversity for the whole population was 0.055 ± 0.007 (Table 2), which was similar to nucleotide diversity estimates of populations of other plant viruses (García-Arenal *et al.*, 2001). Nucleotide diversity calculated independently for each of the coding and non-coding regions included in the analysis ranged from 0.035 ± 0.008 to 0.062 ± 0.009 , with the highest diversity corresponding to the CP gene and the lowest to the 3' UTR (Table 2).

Table 2. Nucleotide diversity \pm SEM (based on 1000 bootstrap replicates). Isolates from Spain, Italy, USA and Germany have been considered as subpopulations

	Entire population π_T	Mean within subpopulations π_S	Mean among subpopulations $\pi_{ST} = \pi_T - \pi_S$	Coefficient of differentiation $N_{ST} = \delta_{ST} / \pi_T$
RdRp	0.041 ± 0.008	0.010 ± 0.003	0.031 ± 0.007	0.757 ± 0.052
p7	0.054 ± 0.010	0.010 ± 0.003	0.044 ± 0.009	0.813 ± 0.036
p9.7	0.052 ± 0.008	0.011 ± 0.002	0.041 ± 0.007	0.789 ± 0.026
CP	0.062 ± 0.009	0.015 ± 0.002	0.047 ± 0.007	0.765 ± 0.017
3' UTR	0.035 ± 0.008	0.009 ± 0.002	0.026 ± 0.006	0.736 ± 0.046
Complete	0.055 ± 0.007	0.013 ± 0.002	0.042 ± 0.005	0.769 ± 0.012

Remarkably, the sequence of two genomic segments were strictly conserved in all isolates, one encompassing nt 3642-3707 and corresponding to a 5'-portion of the 3' UTR, and the other

comprising nt 2240-2279 and matching the leader sequence of the PLPV sgRNA plus short flanking stretches (Castaño and Hernández, 2005). As observed in other viral systems (Liang *et al.*, 2002; Mansky and Temin, 1995; Rico *et al.*, 2006; Schneider and Roossinck, 2001; Vartanian *et al.*, 1997), transition mutations (77.35%) were much more frequent than transversions (22.65%).

Nucleotide diversity was also estimated between and within PLPV subpopulations, considering a subpopulation as the group of isolates that were originally collected from a given country (Germany, Spain, USA and Italy). Between subpopulation diversity values considering either the complete 1817 nt region, individual ORFs or the 3' UTR were greater than within subpopulation diversity values (Table 2), suggesting that there is differentiation of population according to the country from which the isolates were taken.

To gain a better insight into the relationships between all PLPV isolates, a minimum evolution phylogenetic tree with the nucleotide sequences included in the study was constructed. The results revealed two major groups of PLPV sequences: group I included all

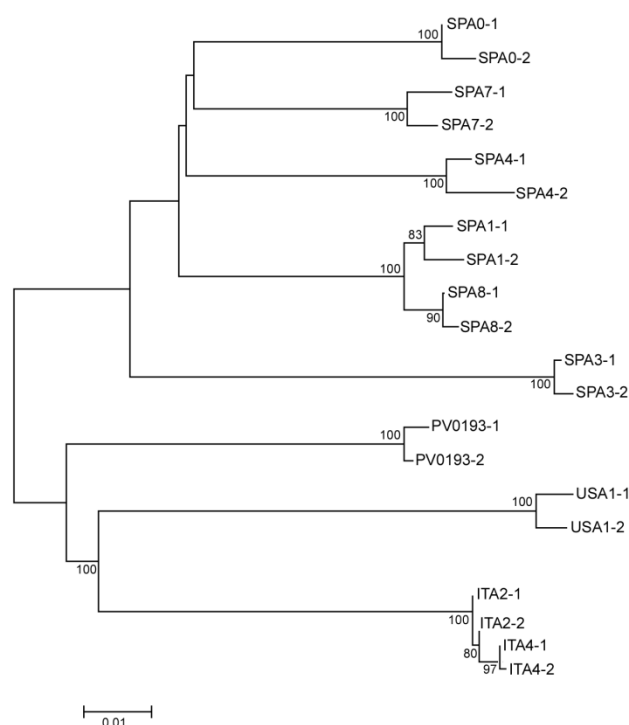


Fig. 1. Minimum evolution unrooted phylogenetic tree inferred from nucleotide sequences derived from 10 isolates of PLPV (see Table 1). Phylogenetic analysis was conducted with programs included in the MEGA4 package. The numbers at the nodes are bootstrap support values based on 10,000 pseudoreplicates; only values >50% are shown.

the partial RdRp, four sites were detected under negative selection (T711, L732, N760, I763). In the case of p7, one site was detected to be under negative selection (S23) and another one under positive selection (S5). Curiously, the latter site overlapped with one of those found under negative selection in the RdRp (N760).

Concerning p9.7, 10 amino acid sites were found to be under negative selection and 27 sites were found in the case of CP. Such sites were mainly concentrated in the central and N-terminal region of p9.7 and CP, respectively (Fig. 2 and 3 and data not shown).

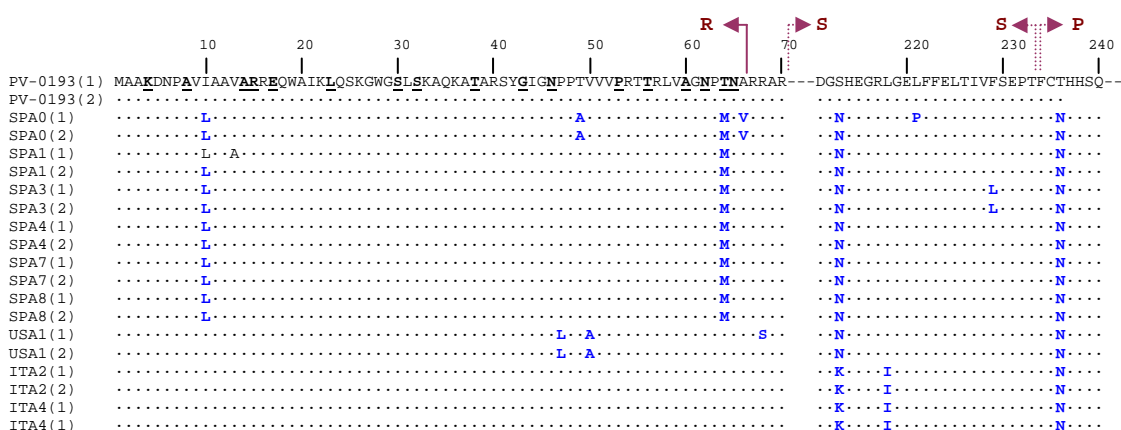


Fig. 3. Alignment of partial amino acid sequences of CP from different isolates of PLPV. The reference sequence derived from isolate PV-0193 (Castaño and Hernández, 2007) is shown at the top of the figure. The domains of the protein at which the depicted residues belong are indicated at the top. Other details as in Fig. 2.

Folding conservation of potential RNA regulatory regions

The PLPV genomic region under study (nt 2067-3883) contains at least two segments that are presumed to play a key role in regulation of viral replication/transcription: the 3' UTR and the promoter for synthesis of the PLPV sgRNA (in the minus strand). Though the latter one has not been experimentally defined, it is expected to embrace a stretch of about 100 nt preceding the initiation site of PLPV sgRNA (nt 2251; Castaño and Hernández, 2005), in line with that reported for subgenomic promoters of related viruses (Li and Wong, 2006; Wang and Simon, 1997; Wang *et al.*, 1999). Such promoters may fold into hairpin-like structure and this type of conformation seems to be critical for the mechanism of transcription of sgRNAs (Li and Wong, 2006; Wang *et al.*, 1999). Secondary structure predictions showed that the putative PLPV subgenomic promoter might also adopt a hairpin-like conformation with a small lateral branch

(Fig. 4A). Interestingly, the sequence variation detected in this segment when comparing isolates, essentially maintained the predicted folding since most mutations were located in single stranded regions or, when affecting double stranded regions, they were compensatory or located at the base of loops or stems (Fig. 4A).

Concerning the 3' UTR, it is expected to contain structural elements critical for viral replication and, most probably, also for translation on the basis of that found in other *Tombusviridae* (Batten *et al.*, 2006; Bringloe *et al.*, 1999; Carpenter and Simon, 1998; Fabian and White, 2006; Fabian *et al.*, 2003; Pogany *et al.*, 2003; Qu and Morris, 2000; Turner and Buck, 1999; Wang and Wong, 2004). The key role of the 3' UTR during the infectious cycle

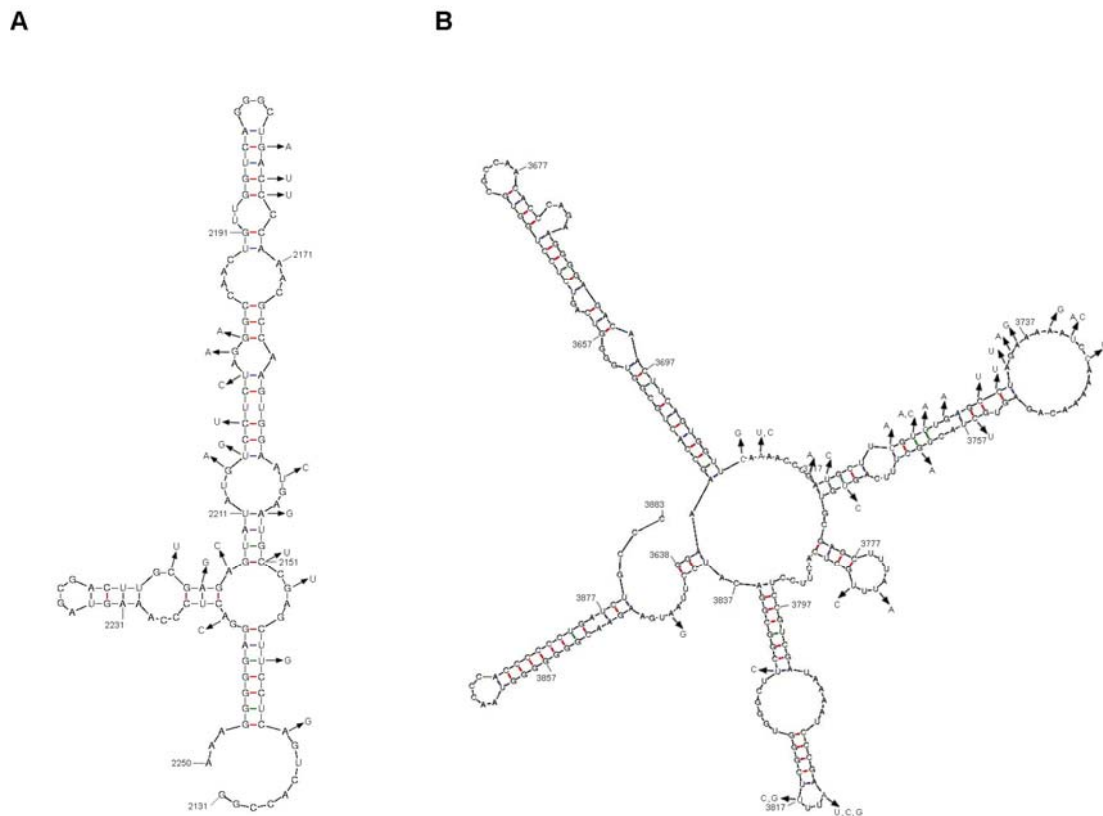


Fig. 4. MFOLD-predicted RNA secondary structures of the putative subgenomic promoter (A) and the 3' UTR (B) of PLPV. The distribution of polymorphic positions is indicated on most stable folding of the reference sequence corresponding to isolate PV-0193 (Castaño and Hernández, 2007). Numbers denote positions in the PLPV gRNA. The minus strand is shown in (A).

implies that strong constraints may operate on the region to preserve its functionality that will likely depend on certain primary, secondary and/or tertiary RNA structures. In line with this view, the 3' UTR showed a value of genetic diversity that, remarkably, was lower than those calculated for ORFs and, moreover, the nucleotide sequence of a 5'-proximal segment of this region was strictly conserved in all isolates as indicated above. *In silico* analysis showed that the 3' UTR may fold into a series of stem-loops that was essentially maintained in the different variants (Fig. 4B). Remarkably, a 5'-proximal stem-loop (Fig. 4B) was formed by the conserved segment whereas the 3'-adjacent stem-loop concentrated most of the heterogeneity found in the non-coding region (Fig. 4B). Collectively, the results suggested that conservation of specific conformations in regulatory regions confer selective advantages to the viral RNA.

Variability in PLPV proteins

The heterogeneity was unevenly distributed in the PLPV proteins. The C-terminal portion of the RdRp inferred from the amplified genomic region was not taken in consideration to analyze variability distribution as it represented only ~1/6 of the complete replication molecule. In the case of p7, 6 out of the 10 polymorphic positions detected were located in the first third of the protein (N-terminal 20 amino acids) despite the corresponding coding sequence overlapped in part with that of RdRp gene. An amino acid replacement mapped at the putative RNA binding domain of p7 (V30I) but they did not affect basic residues which have been found essential for RNA binding in other carmoviruses (Marcos *et al.*, 1999; Navarro *et al.*, 2006). The variability of p9.7 also concentrated at the N-proximal half of the molecule as 12 out of the 17 polymorphic positions were detected within the N-terminal 40 amino acid residues of the protein (Fig. 2), although most of the corresponding coding sequence overlaps with p7 gene. Concerning the CP, the percentage of polymorphic positions was considerable higher in the P domain (18.09%) than in R (10.95%) or S (13.12 %) domains. A stretch within the R domain was absolutely conserved in all isolates (from A14 to N46) probably because structural and/or functional constraints. Supporting the existence of such constraints, a high proportion of the sites detected under negative selection were located in this stretch (Fig. 3).

Next, as an additional test for the effect of selection, we analyzed the possible existence of covariation groups within and between proteins (Fig. 5). Firstly, we focused on covariations

within-proteins. Regarding the partial RdRp, three amino acid residues showed significant covariation, S704N-E713A-S745A. In the case of p7, two covariation groups were detected S8T-V30I and S11I-L43I, whereas for p9.7 three covariation groups were observed, Y3C-V6A, S16L-S39L and N24S-G88R (Fig. 5). Interestingly, the covariation that affected amino acids at positions 16 and 39 distinguished sequences from isolate PV-0193, which showed the combination S16, S39, from those of the remaining isolates, that exhibited L residues at both positions. Up to eight covariation groups were detected for the CP, prominent among which was I10L-T64M, that differentiated the Spanish sequences (bearing the combination L10, M64) from those with other geographical origins (combination I10, T64). The analysis was extended to detect covarying positions between proteins. Remarkably, amino acid residues of p7 covaried with amino acid residues of the other three proteins included in the study. Thus, the p7 covariation S11I-L43I was significantly linked to RdRp substitution N760S and this linkage distinguished Italian isolates (with the combination RdRp S760, p7 I11, I43) from the remaining

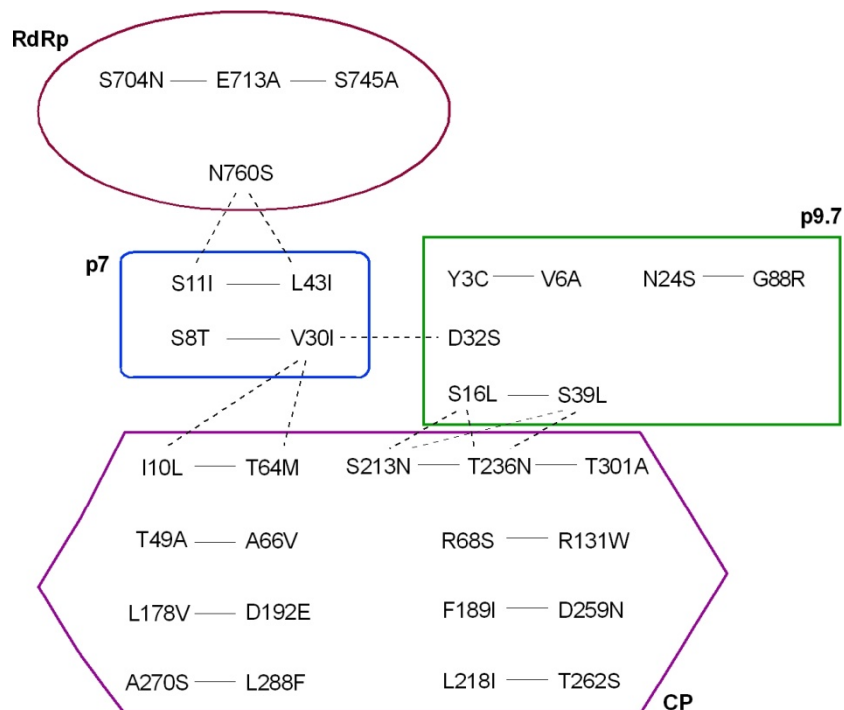


Fig. 5. Covariations within and between p7, p9.7, CP and the partial RdRp. The residues covarying within a given protein are connected by solid lines and those covarying among proteins are connected by dashed lines.

ones (with the combination RdRp N760, p7 S11, L43). In addition, the amino acid replacement V30I in p7 was linked to the amino acid replacement D32S in p9.7 and to the covarying group I10L-T64M in CP, and the corresponding combinations (p7 I30, p9.7 S32, CP L10, M64 versus p7 V30, p9.7 D32, CP I10, T64) segregated the Spanish isolate from the rest, further highlighting geographical distinctions between isolates. Finally, the covarying group S16L-S39L of p9.7 was linked to the amino acid substitution T236N of CP according to the programs employed, though visual inspection of alignments allowed to detect a strict association also with S213N/K (Fig. 3). Indeed, the combination of S16, S39 in p9.7 and S213, T236 in CP was specific for the PV-0193 sequences whereas the remaining isolates showed the combination L16, L39 in p9.7 and N/K213, N236 in CP (Fig. 2 and 3).

Rapid evolution of PLPV in the *C. quinoa*: identification of covarying amino acids in movement and coat proteins consistently selected in the experimental host

The above analyses of intra- and inter-population diversity, phylogenetic relationships and covariation groups supported clustering of isolates by geographical origin rather than by host species or collection times. Grouping by symptomatology was also discarded as, according to our observations, some of the Spanish isolates included in the study induced symptoms but not others. Though according to this initial study on PLPV variability, host species did not seem to define viral populations, the host origin of isolate PV-0193 should be taken with caution. It was recovered at the nineties from a *P. peltatum* plant but it has been maintained by serial transfer in *C. quinoa*, and the segregation of the PV-0193 into a subgroup could be the result of host adaptation, a factor that may result in a decrease or a bias in diversity within a viral population (García-Arenal *et al.*, 1999; Gibbs *et al.*, 1999).

To explore the possible effect of the host on the genetic composition of isolate PV-0193, an infectious full-length cDNA clone derived from the Spanish isolate SPA4 (Table 1), was generated by fusing a T7 RNA polymerase promoter at the 5' end of the viral cDNA. The viral transcripts synthesized *in vitro* from this construct, named pPLSPA4, were used to inoculate *C. quinoa* plants and the resulting viral population was serially passaged in the same host. The PLPV genomic region under study (nucleotides 2067-3883) was RT-PCR amplified from viral progenies of passages 1, 2 and 4, and the resulting DNAs were subsequently cloned and

sequenced. It should be noted that the nucleotide sequence identity among SPA4 and PV-0193 variants in the analyzed region was ~93-94% (i.e., they differ at more than 100 nucleotide sites in the 1817 nt segment). Ten clones of each passage were randomly selected and subjected to digestion with the restriction enzymes HindIII and Scal. The recognition sites of these enzymes were used as probable markers for SPA4-like and PV0193-like sequences as SPA4 isolate bears a HindIII site at nt 2684-2689 and lacks a Scal site while PV0193 isolate lacks a HindIII site and contains a Scal site at nt 2518-2523. The results of the restriction analysis revealed that, whereas the 10 clones from passage 0 had the same restriction pattern as the funding SPA4 sequence, 8 out of the 10 clones of passage 1 had lost the HindIII site of the parental sequence and had acquired the Scal site characteristic of PV-0193 sequences. Moreover, all clones selected from passages 2 and 4 displayed the latter restriction pattern (Table 3). The two remaining clones of passage 1 yielded different restriction profiles: both clones retained the HindIII site of the parental sequence but only one lacked the Scal site. The results strongly suggested that after two passages the viral population founded with an SPA4 variant was genetically indistinguishable from that composing isolate PV-0193, and remained stable in further passages. Remarkably, the results also suggested that one passage was sufficient for almost complete displacement of the parental sequence by variants generated de novo and selected in *C. quinoa*, as about 80 % of the progeny clones of the first passage seemed to correspond to PV0193-like sequences. Sequencing of two or three different clones from each passage confirmed that advanced by the restriction analysis: clones lacking a HindIII site and containing a Scal site were almost identical to those characterized previously from isolate PV-0193 (99.0-99.99 % nucleotide sequence identity which was equivalent to PV-0193 intrapopulation diversity), whereas the nucleotide sequence of the clone of passage 1 harboring a HindIII site and lacking a Scal site was 99.78 % identical to the SPA4 parental variant (Table 3). Determination of the nucleotide sequence of the clone of passage 1 bearing both a HindIII site and a Scal site corroborated the apparent chimeric nature of the molecule: the region encompassing positions 2067 to 2770 shared 99.71 and 94.33 % sequence identity with SPA4 and PV-0193 variants, respectively, whereas the region comprising positions 2771 to 3883 was 93.17 % and 99.73 % identical to SPA4 and PV-0193 sequences, respectively. Remarkably, the p9.7 and the CP encoded by all progeny variants characterized from passages 2 and 4 (and

most from passage 1) presented the combination of covarying residues that, as indicated above, is specific of the PV-0193 isolate: S16, S39 in p9.7 and S213, T236 in CP (Tabla 3). Collectively, analysis of the heterogeneity of PLPV isolates together with monitorization of the molecular evolution of a population initiated with a single viral variant strongly supports that: (i) PLPV is able to evolve very rapidly and reproducibly under new selective pressures as those imposed by the experimental host *C. quinoa* and, (ii) an observed covariation among the movement protein p9.7 and the CP is likely driven by the host.

Table 3. Analysis of PLPV sequence variations in a 1817 nt region (positions 2067-2883) through serial passages in *C. quinoa*. The serial transfer experiment was initiated with a SPA4 variant

Passage number	Clones ^a	Restriction pattern		Nucleotide sequence ^b	Amino acid residues at positions ^c			
		HindIII	ScaI		p9.7		CP	
					16	39	213	236
	SPA4	+	-	SPA4 type	L	L	N	N
	1/10	+	-	SPA4 type	L	L	N	N
1	1/10	+	+	Chimera SPA4/PV-0193	L	L	S	T
	8/10	-	+	PV-0193 type	S	S	S	T
2	10/10	-	+	PV-0193 type	S	S	S	T
4	10/10	-	+	PV-0193 type	S	S	S	T
	PV-0193	-	+	PV-0193 type	S	S	S	T

^a SPA4 and PV-0193 correspond to any of the sequence variants characterized from isolates SPA4 and PV-0193 (Castaño and Hernández, 2005; 2007), respectively

^b SPA4 type: $\geq 99.8\%$ sequence identity with any SPA4 variant; PV-0193: $\geq 99.8\%$ sequence identity with any PV-0193 variant; Chimera SPA4/PV-0193: $\geq 99.8\%$ sequence identity with any SPA4 variant in the region encompassing nt 2067 to 2770 and $\geq 99.8\%$ sequence identity with any PV-0193 variant in the region encompassing nt 2771 to 3883. A total of seven clones were fully sequenced: three from passage 1 (representatives of the three restriction patterns found in this passage), and two from passages 2 and 4 (yielding the unique restriction pattern found in these passages)

^c Only positions involved in a site-specific amino acid covariation are shown

DISCUSSION

In this work, the extent and structure of genetic diversity in PLPV have been explored by sequence analysis of a 1817 nt fragment (representing about 50% of the complete viral genome) of 10 viral isolates sampled from four distinct geographical areas. The results have shown that the populations of PLPV are genetically heterogeneous matching the quasispecies concept that suggests that the population of a virus, even in a single replicating population, is essentially a collection of variants varying around a consensus sequence (Domingo and

Holland, 1997; Eigen, 1993). Despite the sample size is small, our phylogenetic analysis has revealed significant clustering of isolates by country of origin, suggesting that although importation of new material may allow PLPV to cross geographical boundaries, such gene flow is not sufficient to eradicate geographical structure.

The predominant evolutionary pressure on coding regions was that of purifying (negative) selection with a mean ω of 0.148, suggesting that conservation of the amino acid sequence encoded by the corresponding ORFs constitutes an impediment to variation. The portion of the RdRp included in the analysis did not comprise any of the eight motifs conserved in the RdRps (Koonin and Dolja, 1991; Koonin and Dolja, 1993) but, nevertheless, only sites under negative selection were detected. Regarding p7, the unique negatively selected site was located in the putative RNA-binding motif of the protein (Marcos *et al.*, 1999; Navarro *et al.*, 2006), but it did not correspond to any of the basic residues that are presumably critical for RNA-binding capability, suggesting that selection is probably acting on the maintenance of the right secondary/tertiary conformation of the RNA-binding motif. In the case of p9.7, 9 out of the 10 negatively selected sites concentrated in the central part of the molecule (codons 29 to 54) which essentially matched the region that connects the two hydrophobic domains that, according to that reported for related proteins (Vilar *et al.*, 2002; Saurí *et al.*, 2005; Navarro *et al.*, 2006), must be involved in membrane association. Concerning the CP, almost 2/3 of the negatively selected sites (17/27) were located within the N-terminal 66 residues which constitute the R domain, though only two of them corresponded to basic residues that are likely critical for RNA-binding capability suggesting that, as in the case of p7, selection is probably acting on the preservation of the proper conformation of the RNA-binding motif. Moreover, the S and P domains showed identical number of negatively selected sites (5 each) despite the general trend to conservation of the former one in *Tombusviridae* (Lommel *et al.*, 2005).

On the other hand, the variability patterns found in regulatory regions of the viral RNA, such as the putative subgenomic promoter or the 3' UTR, support the existence of structural constraints that prevent the loss of their functionality. In the case of the subgenomic promoter, besides its predicted role in transcriptional regulation, it completely overlaps the 3'-portion of the RdRp gene (in the minus strand) and thus the same stretch is expected to have a dual function as coding and as regulatory sequence, which should considerably restrict heterogeneity.

Accordingly, the mean of nucleotide diversity in this region was lower than in other coding regions (Table 1) and, moreover, the nucleotide substitutions did not disrupt the predicted hairpin-like structure that is presumably required for the promoter function, suggesting that conservation of this conformation significantly influences the profile of naturally occurring mutations. Regarding the 3' UTR, the variability data support that maintenance of a specific folding composed by a series of stem-loops might limit its sequence heterogeneity. Different *Tombusviridae* have been reported to contain in this region *cis* elements critical for viral replication, such as promoters and repressors for minus strand synthesis (Carpenter and Simon, 1998; Na and White, 2006; Pogany *et al.*, 2003; Stupina and Simon, 1997; Zhang *et al.*, 2004a, 2004b), and others relevant for gene expression, such as translational enhancers that promote cap independent translation since the viruses of this family are characterized by non-blocked 5'-termini (reviewed by Kneller *et al.*, 2006). An element of this type has been proposed to be present in the 3' UTR of PLPV RNAs (Fabian and White, 2006) which, remarkably, would be embedded in the 5'-proximal stem-loop that is strictly conserved in all isolates (Fig. 4A) providing indirect evidence for its functional significance.

Correlated amino acid mutation analysis has been widely used to infer functional interactions between different sites in a protein or between distinct proteins (e.g.: Altschuh *et al.*, 1987; Codoñer *et al.*, 2006; Hoffman *et al.*, 2003; Larson *et al.*, 2000; Thomas *et al.*, 1996). The study of PLPV genetic variability have allowed identification of groups of amino acids that covary both within and between PLPV proteins, revealing a putative network of interactions that is likely needed for maintenance of proper protein foldings and for driving the viral RNA from replication to cell-to-cell/systemic translocation. We found especially remarkable a covariation between proteins p9.7 and CP that distinguished PV-0193 isolate, largely propagated in *C. quinoa*, from the remaining ones. Analysis of the molecular evolution of a unique PLPV variant subjected to serial passages in the experimental host has provided solid evidence that the observed p9.7/CP covariation is determined by the host. Both p9.7 and CP are indispensable for virus cell-to-cell movement (Castaño *et al.*, 2009) and, though the exact mechanism that account for inter-cellular transport of carmo-like viruses is not yet known, it is not unlikely to require a physical interaction among CP and movement proteins as reported for other plant viruses (Akamatsu *et al.*, 2007; Kim *et al.*, 2004; Liu *et al.*, 2001; Sánchez-Navarro and Bol, 2001; Sánchez-Navarro

et al., 2006; Takeda *et al.*, 2004). Serial transfer experiments have also revealed that PLPV evolves very rapidly in the new host. Moreover, the evolutionary trajectory of PLPV in *C. quinoa* was extremely reproducible as the population evolved from a Spanish variant collected in 2000 from *P. zonale* was essentially identical to that evolved from a (uncharacterized) German isolate collected in the nineties from *P. peltatum*, indicating that the constraints imposed by the host creates strong bias for “convergent evolution” (Doolittle, 1994). The course of the fixation of mutations in the PLPV molecule was puzzling since the chimeric SPA4/PV-0193 molecule recovered from passage 1 suggested that the 3'-region adapted first to *C. quinoa* and that the nucleotide/amino acid sequences initially selected likely conditioned rapid fixation of changes in the rest of the viral molecule(s) presumably to optimize fitness. Whether selection operates on nucleotides rather than on amino acids remains to be ascertained.

The evolutionary dynamics of plant viruses may have major implications on their ability to emerge in new host species. The extraordinarily rapid accumulation of mutations of PLPV observed in this work parallels our previous results with a related pathogen, *Pelargonium flower break virus* (PFBV) (Rico *et al.*, 2006), and differs from that reported for distinct unrelated viruses (Hébrard *et al.*, 2006; Rebenstorf *et al.*, 2006; Schneider and Roossinck, 2000, 2001). We suggested that viruses with RdRps of supergroup II, as PFBV and PLPV, could show mutation rates higher than those of viruses with RdRps belonging to supergroups I and III (Koonin and Dolja, 1993), which would facilitate their quicker adaptation to changing environments (Rico *et al.*, 2006). Though little is known about the generation time for plant viruses, a medium generation time of one hour can be considered (Schneider and Roossinck, 2001; Wu *et al.*, 1994). After one passage in *C. quinoa* (i.e., 14 days of viral multiplication), there should be about 336 generations, and a theoretical mutation frequency of 3.4×10^{-2} considering an error rate of 10^{-4} , which is the mean estimation previously reported for riboviruses (reviewed by Duffy *et al.*, 2008). The ratio of total mutations/bases sequenced after the first passage would be approximately 1,025/36,340, considering the 20 clones checked in passages 0 and 1, which allows calculating a mutation frequency of about 2.8×10^{-2} . This value is clearly higher than the mutation frequencies estimated for other (plant) RNA viruses not belonging to supergroup II (Fargette *et al.*, 2008; García-Arenal *et al.*, 2001, Schneider and Roossinck, 2001) and, moreover, it almost perfectly match the theoretical mutation frequency

(3.4×10^{-2}) considering a mutation rate of 10^{-4} per round of replication. As the mutation frequency will be a sum of mutation rate, natural selection and genetic drift, it can be anticipated that the mutation rate of PLPV is extraordinarily high, further supporting the possibility of RdRps of supergroup II being endowed with unusual low level of fidelity. Interestingly, the mutation rate can depend on the plant species in which the virus multiplies as has been recently shown for *Cucumber mosaic virus* (Pita *et al.*, 2007), an observation that may have major implications in viral disease progression. In this scenario, it will be interesting to study the evolutionary trajectory of PLPV in other plant species in order to assess if the phenomenon of quick genetic change of the viral population is general for all hosts or is specifically associated with *C. quinoa*.

ACKNOWLEDGEMENTS

We are grateful to Dr. J. van der Meij (Ball Flora Plant, Chicago) for providing PLPV isolate from USA, and to Dr. M. Borja (Fundación Promiva, Madrid) for the Spanish isolates and for valuable comments in the course of this work. We thank to Dolores Arocas and Isabella Avellaneda for excellent technical assistance. This research was supported by grants AGL2003-04249 (MCyT), BFU2006-11230 (MEC) and ACOM06/210 (Generalitat Valenciana) to C.H and by grant BFU2006-14819-C02-01/BMC to SFE. A.C. was the recipient of predoctoral fellowships from the Generalitat Valenciana and from CSIC-Fundación Bancaja and L.R. received a postdoctoral contract from the Juan de la Cierva program of MEC.

REFERENCES

- Agol VI.** 2006. Molecular mechanisms of poliovirus variation and evolution, pp. 211–259. In E. Domingo (ed.), *Quasispecies: concepts and implications for virology*. Current topics in microbiology and immunology, vol. 299.. Springer, Heidelberg, Germany.
- Akamatsu N, Takeda A, Kishimoto M, Kaido M, Okuno T, Mise K.** 2007. Phosphorylation and interaction of the movement and coat proteins of *Brome mosaic virus* in infected barley protoplasts. *Arch. Virol.* **152**: 2087-2093.
- Ali A, Li H, Schneider ML, Sherman DJ, Grey S, Smith D, Roossinck MJ.** 2006. Analysis of genetic bottlenecks during horizontal transmission of *Cucumber mosaic virus*. *J. Virol.* **80**: 8345-8350.
- Altschuh D, Lesk AM, Bloomer AC, Klug A.** 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to *Tobacco mosaic virus*. *J. Mol. Biol.* **193**: 693–707.
- Alonso M, Borja M.** 2005. High incidence of Pelargonium line pattern virus infecting asymptomatic Pelargonium spp. in Spain. *Eur. J. Plant Pathol.* **112**: 95-100.
- Batten JS, Desvoyes B, Yamamura Y, Scholthof KB.** 2006. A translational enhancer element on the 3'-proximal end of the *Panicum mosaic virus* genome. *FEBS Lett.* **580**: 2591-2597.
- Benjamini Y, Hochberg Y.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **57**: 289-300
- Betancourt M, Fereres A, Fraile A, García-Arenal F.** 2008. Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. *J. Virol.* **82**: 12416-12421.
- Bouwen I, Maat DZ.** 1992. Pelargonium flower-break and Pelargonium line pattern viruses in the Netherlands; purification, antiserum preparation, serological identification, and detection in *Pelargonium* by ELISA. *Neth. Plant Pathol.* **98**: 141-156.
- Bringloe DH, Pleij CW, Coutts RH.** 1999. Mutation analysis of cis-elements in the 3'- and 5'-untranslated regions of satellite *Tobacco necrosis virus* strain C RNA. *Virology* **264**: 76-84.
- Carpenter CD, Simon AE.** 1998. Analysis of sequences and predicted structures required for viral satellite RNA accumulation by in vivo genetic selection. *Nucleic Acids Res.* **26**: 2426-2432
- Castaño A, Hernández C.** 2005. Complete nucleotide sequence and genome organization of Pelargonium line pattern virus and its relationship with the family *Tombusviridae*. *Arch. Virol.* **150**: 949-965.
- Castaño A, Hernández C.** 2007. Biological activity of transcripts from cDNA of Pelargonium line pattern virus. *Acta Virol.* **51**: 271-274.

Castaño A, Ruiz L, Hernández C. 2009. Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus. *Virology* (in press).

Castro C, Arnold JJ, Cameron CE. 2005. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res.* **107**: 141–149.

Codoñer FM, Fares MA, Elena SF. 2006. Adaptative covariation between the coat and the movement proteins of *Prunus necrotic ringspot virus*. *J. Virol.* **80**: 5833-5840.

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

Doolittle RF. 1994. Convergent evolution: the need to be explicit. *Trends Biochem. Sci.* **19**: 15-18.

Drake JW, Holland JJ. 1999. Mutation rates among lytic RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**: 13910-13913.

Duffy S, Shackelton LA, Holmes EC. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**: 267-276.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797

Eigen M. 1993. The origin of genetic information: Virus as models. *Gene* **135**: 37-47.

Fabian MR, Na H, Ray D, White KA. 2003. 3'-Terminal RNA secondary structures are important for accumulation of *Tomato bushy stunt virus* DI RNAs. *Virology* **313**: 567-580.

Fabian MR, White KA. 2006. Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA* **12**: 1304-1314.

Fargette D, Pinel A, Rakotomalala M, Sangu E, Traoré O, Sérémé D, Sorho F, Issaka S, Hébrard E, Séré Y, Kanyeka Z, Konaté G. 2008. *Rice yellow mottle virus*, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J. Virol.* **82**: 3584-3589.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

Franck A, Loebenstein G. 1994. Virus and virus-like diseases of pelargonium in Israel. *Acta Hort.* **377**: 31-39.

French R, Stenger DC. 2003. Evolution of *Wheat streak mosaic virus*: dynamics of population growth within plants may explain limited variation. *Annu. Rev. Phytopathol.* **41**: 199-214.

García-Arenal F, Fraile A, Malpica JM. 1999. Genetic variability and evolution, p. 143-159. In G. L. Mandahar (ed.). *Molecular biology of plant viruses*. Kluwer, London.

García-Arenal F, Fraile A, Malpica JM. 2001. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* **39**: 157–186.

García-Arenal F, Fraile A, Malpica JM. 2003. Variation and evolution of plant virus populations. *Int. Microbiol.* **6**: 225–232.

Gibbs A, Keese P, Gibbs M, Garcia-Arenal F. 1999. Plant virus evolution: past, present and future, p 263-285. In E. Domingo, R. Webster and J. Holland (ed.), *Origin and Evolution of Viruses*. Academic Press, San Diego.

Hébrard E, Pinel-Galzi A, Bersoult A, Siré C, Fargette D. 2006. Emergence of a resistance-breaking isolate of *Rice yellow mottle virus* during serial inoculations is due to a single substitution in the genome-linked viral protein VPg. *J. Gen. Virol.* **87**: 1369-1373

Hoffman NG, Schiffer CA, Swanstrom R. 2003. Covariation of amino acid positions in HIV-1 protease. *Virology* **314**: 536–548.

Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, Van de Pol S. 1982. Rapid evolution of RNA genomes. *Science* **215**:1577–1585.

Kim SH, Kalinina NO, Andreev I, Ryabov EV, Fitzgerald AG, Taliensky ME, Palukaitis P. 2004. The C-terminal 33 amino acids of the *Cucumber mosaic virus* 3a protein affect virus movement, RNA binding and inhibition of infection and translation. *J. Gen. Virol.* **85**: 221–230.

Kinard GR, Jordan R. 2002. Genome organization of Pelargonium chlorotic ring pattern virus: further implications for *Tombusviridae* taxonomy. *Acta Hort.* **568**: 17-27.

Kneller EL, Rakotondrafara AM, Miller WA. 2006. Cap-independent translation of plant viral RNAs. *Virus Res.* **119**: 63-75.

Koonin EV, Dolja VV. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* **72**: 2197-2206.

Koonin EV, Dolja VV. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28**: 375-430.

Kosakovsky Pond SL, Frost SDW. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**: 1208-1222

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. 2006. Automated phylogenetic detection of recombination using genetic algorithm. *Mol. Biol. Evol.* **23**: 1891-1901

Larson SM, Di Nardo AA, Davidson AR. 2000. Analysis of covariation in an SH3 domain sequence alignment: applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions. *J. Mol. Biol.* **303**: 433–446.

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

Li W, Wong SM. 2006. Analyses of subgenomic promoters of *Hibiscus chlorotic ringspot virus* and demonstration of 5'-untranslated region and 3'-terminal sequences functioning as subgenomic promoters. *J. Virol.* **80**: 3395-3405.

Liang XZ, Lee BTK, Wong SM. 2002. Covariation in the capsid protein of *Hibiscus chlorotic ringspot virus* induced by serial passing in a host that restricts movement leads to avirulence in its systemic host. *J. Virol.* **76**: 12320-12324.

Liu H, Boulton MI, Oparka KJ, Davies JW. 2001. Interaction of the movement and coat proteins of *Maize streak virus*: implications for the transport of viral DNA. *J. Gen. Virol.* **82**: 35-44.

Lommel SA, Martelli GP, Rubino L, Russo M. 2005. Family *Tombusviridae*, pp. 907-936. In Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U., and L. A. Ball (Eds.). *Eighth Report of the International Committee on Taxonomy of Viruses*, Academic Press, San Diego, USA.

Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, García-Arenal F. 2002. The rate and character of spontaneous mutations in an RNA virus. *Genetics* **162**: 1505-1511.

Mansky LM, Temin HM. 1995. Lower *in vivo* mutation rate of *Human immunodeficiency virus* type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**: 5087-5094.

Marcos JF Vilar M, Pérez-Payá E, Pallás V. 1999. *In vivo* detection, RNA-binding properties and characterization of the RNA-binding domain of the p7 putative movement protein from *Carnation mottle carmovirus* (CarMV). *Virology* **255**: 354-365.

Martin DP, Williamson C, Posada D. 2005. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**: 260-262.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**: 911-940.

Na H, White KA. 2006. Structure and prevalence of replication silencer-3' terminus RNA interactions in *Tombusviridae*. *Virology* **345**: 305-316.

Navarro JA, Genovés A, Climent J, Saurí A, Martínez-Gil L, Mingarro I, Pallás V. 2006. RNA-binding properties and membrane insertion of *Melon necrotic spot virus* (MNSV) double gene block movement proteins. *Virology* **356**: 57–67.

Pita JS, de Miranda JR, Schneider WL, Roossinck MJ. 2007. Environment determines fidelity for an RNA virus replicase. *J. Virol.* **81**: 9072-9077.

Pogany J, Fabian MR, White KA, Nagy PD. 2003. A replication silencer element in a plus-strand RNA virus. *EMBO J.* **22**: 5602-5611.

Pond SL, Frost SDW. 2005. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**: 1208-1222.

Poon AFY, Lewis FI, Kosakowsky Pond SL, Frost SDW. 2007. An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. *PLoS Comp. Biol.* **3**: e231

Posada D, Crandall K. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817-81

Qu F, Morris TJ. 2000. Cap-independent translational enhancement of *Turnip crinkle virus* genomic and subgenomic RNAs. *J. Virol.* **74**: 1085-1093.

Rebenstorf K, Candresse T, Dulucq MJ, Buttner C, Obermeier C. 2006. Host species-dependent population structure of a pollen-borne plant virus, *Cherry leaf roll virus*. *J. Virol.* **80**: 2453-2462.

Rico P, Ivars P, Elena SF, Hernández C. 2006. Insights into the selective pressures restricting *Pelargonium flower break virus* genome variability and evidence of host adaptation. *J. Virol.* **80**: 8124-8132.

Sacristán S, Malpica JM, Fraile A, García-Arenal F. 2003. Estimation of population bottlenecks during systemic movement of *Tobacco mosaic virus* in tobacco plants. *J. Virol.* **77**: 9906-9911.

Sánchez-Navarro JA, Bol JF. 2001. Role of the *Alfalfa mosaic virus* movement protein and coat protein in virus transport. *Mol. Plant Microbe Interact.* **14**: 1051-1062.

Sánchez-Navarro JA, Herranz MC, Pallás V. 2006. Cell-to-cell movement of *Alfalfa mosaic virus* can be mediated by the movement proteins of ilar-, bromo-, cucumo-, tobamo- and comoviruses and does not require virion formation. *Virology* **346**: 66-73.

Saurí A, Saksena S, Salgado J, Jhonson AE, Mingarro I. 2005. Double-spanning plant viral movement protein integration into the endoplasmic reticulum membrane is signal recognition particle-dependent, translocon-mediated, and concerted. *J. Biol. Chem.* **280**: 25907-25912.

Schneider WL, Roossinck MJ. 2000. Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. *J. Virol.* **74**: 3130-3134.

Schneider WL, Roossinck MJ. 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J. Virol.* **75**: 6566-6571.

- Sharp PM.** 1997. In search of molecular Darwinism. *Nature* **385**: 111-112
- Steinhauer DA, Holland JJ.** 1986. Direct method for quantification of extreme polymerase error frequencies at selected single base sites in viral RNA. *J. Virol.* **57**:219–228.
- Stone OM.** 1980. Nine viruses isolated from pelargonium in the United Kingdom. *Acta Hortic.* **110**: 177-182.
- Stuart GW, Moffett PK, Bozarth RF.** 2006. A comprehensive open reading frame phylogenetic analysis of isometric positive strand ssRNA plant viruses. *Arch. Virol.* **151**: 1159-1177.
- Stupina V, Simon AE.** 1997. Analysis *in vivo* of *Turnip crinkle virus* satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. *Virology* **238**: 470-477.
- Suyama M, Torrents D, Bork P.** 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucl. Acids Res.* **34**: 609-612.
- Tajima F.** 1993. Simple methods for testing molecular clock hypothesis. *Genetics* **135**: 599-607
- Takeda A, Kaido M, Okuno T, Mise K.** 2004. The C-terminus of the movement protein of *Brome mosaic virus* controls the requirement for coat protein in cell-to-cell movement and plays a role in long distance movement. *J. Gen. Virol.* **85**: 1751–1761.
- Tamura K, Dudley J, Nei M, Kumar S.** 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**: 1596-1599
- Thomas DJ, Casari G, Sander C.** 1996. The prediction of protein contacts from multiple sequence alignments. *Protein Eng.* **9**: 941–948.
- Turner RL, Buck KW.** 1999. Mutational analysis of *cis*-acting sequences in the 3'- and 5'-untranslated regions of RNA 2 of *Red clover necrotic mosaic virus*. *Virology* **253**: 115-124.
- Vartanian JP, Plikat U, Henry M, Mahieux R, Guillemot L, Meyerhans A, Wain-Hobson S.** 1997. HIV genetic variation is directed and restricted by DNA precursor availability. *J. Mol. Biol.* **270**: 139-151.
- Verwoerd TC, Dekker BMM, Hoekema A.** 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res.* **17**: 2362.
- Vilar M, Sauri A, Monné M, Marcos JF, von Heijne G, Pérez-Payá E, Mingarro I.** 2002. Insertion and topology of a plant viral movement protein in the endoplasmic reticulum membrane. *J. Biol. Chem.* **277**: 23447-23452.
- Wang S, Carpenter CD, Simon AE.** 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for *Turnip crinkle virus*. *Virology* **253**: 327-336.

Wang J, Simon AE. 1997. Analysis of the two subgenomic RNA promoters for *Turnip crinkle virus* in vivo and in vitro. *Virology* **232**: 174-186.

Wang HH, Wong SM. 2004. Significance of the 3'-terminal region in minus strand RNA synthesis of *Hibiscus chlorotic ringspot virus*. *J. Gen. Virol.* **85**: 1763-1776.

Zhang G, Zhang J, Simon AE. 2004a. Repression and derepression of minus-strand synthesis in a plus-strand RNA virus replicon. *J. Virol.* **78**: 7619-33.

Zhang J, Stuntz RM, Simon AE. 2004b. Analysis of a viral replication repressor: sequence requirements for a large symmetrical internal loop. *Virology* **15**: 90-102.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406-3415.

Discusión General

Secuencia nucleotídica y organización genómica del PLPV

El geranio se encuentra a menudo afectado por enfermedades de etiología viral que repercuten negativamente en la producción y en la calidad de esta planta ornamental. En los últimos años, las infecciones causadas por el virus del arabesco del Pelargonium (Pelargonium line pattern virus, PLPV) han alcanzado una importancia notable, con porcentajes de incidencia muy altos en nuestro país y, probablemente, a nivel mundial (Stone, 1980; Bouwen y Maat, 1992; Franck y Loebenstein, 1994; Alonso *et al.*, 1999; Alonso y Borja, 2005). Se trata de un virus con partículas isométricas y un genoma de ssRNA+ cuyas infecciones naturales parecen restringidas a especies del género *Pelargonium*, aunque experimentalmente puede infectar a especies muy diversas. Dada la escasez de datos sobre el PLPV y la importancia que el virus está adquiriendo como patógeno, en esta tesis hemos pretendido profundizar en sus características biológicas y, sobre todo, moleculares.

El primer objetivo abordado en este trabajo ha sido determinar la estructura primaria del PLPV y analizar los RNAs virales que se acumulan en las plantas infectadas. Partiendo de un aislado alemán del virus (denominado PV-0193), se secuenció completamente el gRNA. Los resultados indicaron que esta molécula estaba constituida por 3884 nt y, mediante análisis *in silico*, inicialmente identificamos seis ORFs que potencialmente codificaban proteínas de 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), y 37 kDa (p37). Estas ORFs estaban flanqueadas por una 5' UTR inusualmente corta, con sólo 6 nt, y una 3' UTR de 246 nt. Tanto la organización de las ORFs en el RNA viral como la mayoría de sus productos potenciales se asemejaban mucho a los implicados en replicación (p27 y p87), movimiento (p7) y encapsidación (p37) de miembros del género *Carmovirus* (familia *Tombusviridae*). A pesar de las semejanzas que el PLPV compartía con componentes de este grupo viral, este patógeno presentaba algunas características que impedían su asignación directa a este género. La región central del genoma de los carmovirus contiene dos pequeñas ORFs en distintas fases de lectura (con excepción del MNSV) y con codones de iniciación canónicos que codifican las MPs (Hacker *et al.*, 1992; Li *et al.*, 1998; Navarro *et al.*, 2006), mientras que de las dos proteínas potencialmente codificadas por la región central del genoma del PLPV, sólo la p7 mostraba homología significativa con MPs de carmovirus. Una búsqueda en bases de datos de proteínas no reveló

la existencia de homólogos claros de la putativa p6 ni tampoco de la proteína p13 potencialmente codificadas por el virus. Además, los carmovirus producen dos sgRNAs para la expresión de las MPs y de la CP, respectivamente (Russo *et al.*, 1994), mientras que un análisis Northern indicó que el PLPV producía un solo sgRNA de 1.6 kb. Las características estructurales de esta molécula permitieron sugerir que posiblemente estaba implicada en la traducción de la p7, de la p6 y de la p37 del virus. Es interesante también señalar que los resultados de los experimentos de extensión del cebador llevados a cabo para cartografiar y/o confirmar el extremo 5' del gRNA y del sgRNA del PLPV eran consistentes con los esperados para moléculas sin cap, en línea con lo descrito para miembros de la familia *Tombusviridae* (Lommel *et al.*, 2005).

Además de asignar posibles funciones a las proteínas potencialmente codificadas por el PLPV mediante análisis comparativo con proteínas de función conocida, también se especuló acerca de los posibles mecanismos de expresión génica empleados por el virus. Esta especulación se basó en la organización genómica del mismo y en ciertas características de su secuencia. Las proteínas p27 y p87 presumiblemente implicadas en replicación viral, y la p13, de función desconocida, probablemente eran traducidas a partir del gRNA, la p27 mediante un mecanismo convencional de rastreo ribosomal y la p87 por un mecanismo de lectura a través del codón de terminación de la ORF (p27), como se ha descrito para la mayoría de miembros de la familia *Tombusviridae* (Scholthof *et al.*, 1995; White *et al.*, 1995; Turina *et al.*, 1998; Huang *et al.*, 2000). Un mecanismo de iniciación interna de la traducción fue considerado como plausible para la producción de la p13, ya que la involucración de un mecanismo de escape al proceso de rastreo ribosomal era bastante improbable dado el contexto óptimo que flanquea al codón del inicio de la ORF (p27) (Futterer y Hohn, 1996). El resto de ORFs, localizadas aguas abajo en el genoma, se traducirían a partir del único sgRNA detectado en las plantas infectadas que, además, se acumula a niveles elevados. La proteína p7, presumiblemente implicada en movimiento, debía ser traducida por un mecanismo de rastreo ribosomal convencional, mientras que la p6, de función desconocida, podría ser traducida a partir de un codón de iniciación no canónico por un mecanismo de escape al proceso de rastreo del RNA y/o alternativamente, mediante un mecanismo de corrimiento de la pauta de lectura de -1 nt (-1 FS). Este corrimiento se produciría inmediatamente aguas arriba del codón de parada de la

ORF (p7), dando lugar a una proteína de 12 kDa (p12 o p7-FS). Esta posibilidad estaba apoyada por dos rasgos estructurales del RNA: la presencia de un heptanucleótido, flanqueando en 5' al codón de parada de la ORF (p7), con una secuencia que se ajusta al motivo característico de fenómenos de corrimiento de pauta de lectura, y la posible formación de una estructura de tipo horquilla aguas abajo del mencionado codón. La presencia de esta clase de estructuras parece facilitar la parada de los ribosomas propiciando su deslizamiento posterior. Si efectivamente el virus empleaba un mecanismo de -1 FS para expresar una de sus ORFs internas, esto constituiría otro rasgo distintivo del PLPV frente a los carmovirus, cuyas ORFs internas, además de presentar codones de inicio canónicos, son presumiblemente traducidas por un mecanismo de escape al proceso de rastreo ribosomal. Si bien no podíamos excluir que un sgRNA adicional se produjera en muy bajas concentraciones (o en determinados estadios del ciclo infeccioso) para la expresión de la ORF (p37), los análisis Northern no revelaron la presencia de esta molécula hipotética. Por esta razón, se consideró que la ORF (p37) podría expresarse por un mecanismo de iniciación interna o de escape al proceso de rastreo del RNA.

La organización de ORFs y las estrategias de expresión génica predichas para el sgRNA del PLPV eran muy similares a las descritas para el único sgRNA del PMV (*Panicovirus*), aunque este último además de presentar tres ORFs que codifican las proteínas p8, p6.6 y p38, contiene una ORF adicional dentro del gen de la CP (Turina *et al.*, 1998). También en el caso del sgRNA del PMV se había sugerido la posible generación de una proteína p8-FS como consecuencia de un fenómeno de corrimiento en la pauta de lectura y, curiosamente, la secuencia de siete nucleótidos necesaria para que se produzca este fenómeno en el PMV era idéntica a la encontrada aguas arriba del codón de parada de la ORF (p7) del PLPV, lo que reforzaba la hipótesis de la implicación de un mecanismo de este tipo para la producción de la hipotética p7-FS. Sin embargo, en el caso del PMV la supuesta proteína p8-FS no ha sido detectada *in vivo* y, más aún, ensayos de traducción *in vitro* han demostrado que el codón de inicio no canónico de la ORF p6.6 es funcional, a pesar de que el contexto que lo flanquea no es óptimo (Turina *et al.*, 2000). Bioensayos realizados con mutantes del PMV han demostrado que la proteína p6.6 es necesaria para una infección eficiente y que, además, está involucrada en el movimiento del virus en la planta (Turina *et al.*, 2000). A pesar de que las secuencias de

esta proteína del PMV y del potencial producto p6 del PLPV eran considerablemente distintas, sus perfiles de hidrofobicidad eran similares. Teniendo en cuenta las analogías entre las dos proteínas, se podía anticipar una función equivalente a la de la p6.6 del PMV para la potencial p6 del PLPV.

Además del PLPV, se habían descrito un conjunto de pequeños virus isométricos cuyas infecciones naturales también parecen restringidas a especies del género *Pelargonium*, entre los que se encuentran el virus de las manchas anulares de *Pelargonium* (*Pelargonium ringspot virus*, PeIRSV) y el virus del anillo clorótico de *Pelargonium* (*Pelargonium chlorotic ring pattern virus*, PCRPV). El PCRPV codifica potencialmente 5 proteínas muy similares a las implicadas en replicación (p27 y p87), movimiento (p7 y p9) y encapsidación (p37) en el género *Carmovirus* (Kinard y Jordan, 2002), pero, al igual que el PLPV, genera un único sgRNA presumiblemente implicado en la expresión de las ORFs situadas en posición interna y 3'-proximal en el gRNA. Asimismo, podría utilizar un mecanismo -1 FS para la traducción de la ORF (p9) o, alternativamente, esta ORF podría ser traducida a partir de un codón no-AUG, de forma similar a lo predicho para la ORF (p6) del PLPV. Algunos datos indican que estas características podrían ser compartidas por el PeIRSV y el virus latente de la baya del saúco (*Elderberry latent virus*, ELV) (Kinard *et al.*, 1996; Jones *et al.*, 2000; Kinard y Jordan, 2002), aunque la secuencia completa de los gRNAs de estos patógenos no está disponible. Sin embargo, sobre la base de la producción de un solo sgRNA, de estrategias de expresión génica posiblemente comunes y de ciertas similitudes de secuencia, se propuso incluir a los cuatro virus en un nuevo género dentro de la familia *Tombusviridae* denominado *Pelarspovirus* (Kinard y Jordan, 2002). Un árbol filogenético elaborado con las secuencias de la CP de los cuatro patógenos permitió apoyar esta propuesta (Fig. 6, capítulo I), al igual que análisis filogenéticos posteriores llevados a cabo con las secuencias completas del PLPV y del PCRPV y de los componentes de la familia *Tombusviridae* (Stuart *et al.*, 2006).

Generación de un clon infeccioso del PLPV

Las peculiaridades genómicas del PLPV (y su posible adscripción a un nuevo género) merecían sin duda ser estudiadas en mayor profundidad. Por ello el siguiente objetivo que se

abordó fue la obtención de un clon infeccioso del virus para poder llevar a cabo experimentos de genética reversa. Este tipo de aproximación ha contribuido enormemente al estudio de funciones genómicas en virus de plantas (Boyer y Haenni, 1994), y nos podía permitir contrastar la validez de las predicciones realizadas, mediante análisis *in silico*, acerca del papel de las distintas ORFs en el ciclo biológico del PLPV y de los mecanismos de expresión génica utilizados por este agente infeccioso.

Dado que los intentos de conseguir la amplificación de un cDNA de longitud completa en una sola reacción de RT-PCR fueron infructuosos, probablemente por impedimentos estructurales del propio RNA y/o cDNA o por problemas derivados de la baja procesividad de la polimerasa utilizada, se generaron dos poblaciones de cDNAs solapantes que abarcaban la mitad 5' (fusionada al promotor de la polimerasa del fago T7) y la mitad 3' del genoma viral, respectivamente. Tras la digestión con enzimas de restricción apropiados, estas dos poblaciones de cDNAs fueron ligadas entre sí y los productos resultantes se fusionaron al vector de clonación pUC18. Con esta estrategia, similar a la descrita como "clonaje poblacional" (Yu y Wong, 1998), se intentó maximizar la probabilidad de obtener un clon biológicamente activo al permitir la síntesis de una serie de clones diferentes que podían ser bioensayados simultáneamente aunque de forma independiente.

Mediante esta estrategia se seleccionaron dos clones distintos de longitud completa, denominados pPLDS-1 y pPLDS-2, que fueron linearizados con el enzima de restricción *Bam*HI. Los transcritos generados *in vitro* a partir de estos moldes se utilizaron para inocular hojas del huésped experimental *C. quinoa*. Puesto que estos transcritos no resultaron infecciosos, se secuenciaron completamente los dos clones de partida para intentar averiguar a qué se debía su falta de actividad biológica. En el caso del clon pPLDS-1, se detectó una sustitución nucleotídica que conducía a la introducción de un codón de parada prematuro en la ORF (p27). Esta mutación, que debía ser la responsable de la incapacidad del clon para establecer una infección eficiente, probablemente fue incorporada durante el proceso de transcripción reversa o de amplificación por PCR del cDNA viral. En el caso del clon pPLDS-2, no se observó en su secuencia ningún cambio nucleotídico que pudiera explicar de forma obvia su naturaleza no infecciosa.

Los transcritos derivados de varios de los clones resultantes de sustituir la región 5' del clon pPLDS-2 por nuevos cDNAs, indujeron lesiones cloróticas cuando se inocularon sobre hojas de *C. quinoa*, comprobándose la presencia del PLPV mediante hibridación molecular. Se seleccionó al azar uno de estos clones infecciosos, denominado pPLDS-10, y su secuenciación reveló que presentaba sólo 2 sustituciones nucleotídicas con respecto al clon pPLDS-2; estas sustituciones estaban localizadas dentro de la ORF (p27) y/o ORF (p87) y daban lugar a dos cambios en las correspondientes secuencias aminoacídicas, L40P y W351S (posiciones en p27/p87 y p87, respectivamente). Experimentos de mutagénesis dirigida realizados posteriormente en el laboratorio han puesto de manifiesto que la sustitución nucleotídica que da lugar al cambio de una S por W en el residuo 351 de la RdRp es la responsable de la falta de infectividad del clon pPLDS-2, aunque por el momento se desconoce qué propiedad particular de la proteína está afectada por este cambio.

A pesar de la actividad biológica mostrada por los transcritos derivados del clon pPLDS-10, el número de lesiones que apareció en las hojas inoculadas fue considerablemente menor que el observado utilizando el RNA viral silvestre como inóculo. Estos transcritos eran presumiblemente portadores de nucleótidos “no virales” en los extremos, concretamente, 3 Gs en 5', incorporados durante la transcripción a partir del promotor de la RNA polimerasa del fago T7, y 5 nt más en 3' (GGAUC), resultantes de la utilización del sitio *Bam*HI para la linearización del plásmido previa a la transcripción. Para averiguar si estos nucleótidos adicionales eran responsables de la reducción en la infectividad de los transcritos, se llevaron a cabo dos aproximaciones. Por un lado, se fusionó el cDNA viral recuperado del clon infeccioso pPLDS-10 directamente al promotor 35S del CaMV y al terminador del gen de la nopalina sintasa, y la construcción resultante se inoculó sobre las hojas de *C. quinoa*. Estas hojas desarrollaron un número de lesiones similar al observado tras inoculación con transcritos generados a partir del clon pPLDS-10 linearizado con *Bam*HI, aunque la aparición de los síntomas se retrasó 1-2 días. Este resultado sugería que los nucleótidos adicionales presentes en el extremo 5' de los transcritos no eran responsables de la baja infectividad de los mismos. Por otro lado, el clon pPLDS-10 fue linearizado, además de con *Bam*HI, con *Kpn*I o *Sma*I. En teoría, estas dos digestiones permitían obtener transcritos con el extremo 3' preciso del virus o con 1 nt menos, respectivamente. Sorprendentemente, los transcritos sintetizados tras la linearización con *Sma*I

indujeron un número de lesiones alrededor de 10 veces superior a los observados con los otros transcritos. Este resultado apoyaba firmemente la idea de que el residuo G asignado inicialmente al extremo 3' del genoma del PLPV, no está realmente presente en el ssRNA viral. Los intermediarios replicativos de doble cadena de ciertos virus de ssRNA+, presentan un residuo adicional en el extremo 3' de la cadena positiva o negativa (Karasev *et al.*, 1995; Galiakparov *et al.*, 1999) y, probablemente, a esto se debía nuestra asignación incorrecta del residuo terminal, ya que los dsRNAs del PLPV se habían utilizado como moldes para la clonación y secuenciación de la región 3'- proximal del virus. Conjuntamente, los resultados indicaron que el genoma del PLPV, al igual que los del resto de miembros de la familia *Tombusviridae* termina con la secuencia CCC-OH (White y Nagy, 2004) y, por lo tanto, la linearización del clon infeccioso con *Sma*I dejaba el extremo preciso del virus, lo que explicaba el aumento del número de lesiones en hojas de *C. quinoa*. Esta observación contrasta con lo que ha sido descrito para otros miembros de la familia *Tombusviridae*, como el virus latente del Photos (*Photos latent virus*, PoLV, *Aureusvirus*) o el MNSV (*Carmovirus*) (Rubino y Russo, 1997; Díaz *et al.*, 2003), cuya infectividad no parece estar afectada por extensiones en el extremo 3' del RNA viral. El clon infeccioso del PLPV que habíamos generado era además capaz de infectar sistémicamente otros huéspedes experimentales como *N. clevelandii* o *N. benthamiana*, aunque de forma asintomática, reproduciendo las características biológicas del virus parental.

Identificación de las ORFs del PLPV funcionales *in vivo*

El análisis *in silico* del genoma del PLPV había aportado pistas acerca de la posible función de las proteínas potencialmente codificadas por algunas de sus ORFs y había permitido hipotetizar sobre las estrategias de expresión génica empleadas por el virus. Para identificar qué ORFs de las predichas inicialmente eran biológicamente activas y cuál era su papel en el ciclo infeccioso del patógeno, se generó una colección de mutantes a partir del clon infeccioso. Estos mutantes fueron utilizados como moldes en reacciones de traducción *in vitro* y bioensayados en diferentes huéspedes experimentales. Los resultados de estos experimentos corroboraron la expresión y las funciones anticipadas para las proteínas p27 y p87

(efectivamente implicadas en replicación) así como para las proteínas p7 y p37 (implicadas en movimiento y movimiento/encapsidación, respectivamente). Sin embargo, la expresión de las otras dos ORFs predichas inicialmente, ORFs (p13) y (p6), no pudo ser detectada *in vitro* y, más aún, los bioensayos mostraron que no eran requeridas *in vivo*, ya que mutantes con estas ORFs inactivadas eran capaces de establecer infecciones indistinguibles de las establecidas por el virus silvestre en tres huéspedes distintos, indicando que las mencionadas ORFs probablemente carecen de significado biológico. En su lugar se identificó una nueva ORF, localizada en la región central del genoma, que se traduce a partir de un codón débil de iniciación (GUG) y que da lugar a una proteína de tamaño 9.7 kDa implicada en el movimiento del virus. A pesar de que presenta un perfil de hidrofobicidad muy similar al de proteínas de movimiento del género *Carmovirus* (Navarro *et al.*, 2006), su homología de secuencia con dichas proteínas es escasa. Curiosamente, la p9.7 del PLPV sí muestra una notable identidad de secuencia (44,4%) con otra de tamaño 9.4 kDa, potencialmente codificada por una ORF del PCRPV también ubicada en la región central del genoma y cuyo codón de inicio es, asimismo, un triplete no-AUG (Kinard y Jordan, 2002). Esta observación amplía las semejanzas entre estos dos virus, ambos miembros provisionales del género propuesto *Pelarspovirus*.

La CP del PLPV por su parte, desempeña un papel fundamental en el movimiento local y, por lo tanto, también sistémico del virus. El requerimiento de la CP para el movimiento célula a célula y sistémico varía entre los distintos miembros de la familia *Tombusviridae*. Así, la CP no es necesaria para el movimiento local del virus de las manchas de las hojas del pepino (*Cucumber leaf spot virus*, CLSV) y del PoLV, ambos en el género *Aureusvirus* (Rubino y Russo, 1997; Reade *et al.*, 2003), del virus de la necrosis del tabaco (*Tobacco necrosis virus*, TNV, *Necrovirus*) (Molnar *et al.*, 1997), y del RCNMV (*Dianthovirus*) (Vaewhongs y Lommel, 1995). En este último sin embargo, la CP es estrictamente necesaria para el movimiento a larga distancia, excepto en el huésped *N. benthamiana* a bajas temperaturas (Xiong *et al.*, 1993). Por el contrario, la CP del PMV (*Panicovirus*) (Turina *et al.*, 2000) y del MNSV (*Carmovirus*) (Genovés *et al.*, 2006) es requerida para el movimiento célula a célula. En el caso del TCV (*Carmovirus*), la CP es indispensable para el movimiento célula a célula en *N. benthamiana*, pero no en *Arabidopsis*, aunque es esencial para la dispersión sistémica en ambos huéspedes (Heaton *et al.*, 1991; Hacker *et al.*, 1992; Cohen *et al.*, 2000). Estas

observaciones contrastan con lo descrito para miembros del género *Tombusvirus*, que pueden moverse célula a célula o sistémicamente en ausencia de CP (Russo *et al.*, 1994; Sit *et al.*, 1995). Si la CP del PLPV tiene un papel directo en el movimiento del virus formando parte del complejo que permite su propagación en la planta o contribuye indirectamente al inicio y mantenimiento de la infección, está por determinar. Esta contribución indirecta podría ser el resultado de la capacidad de la proteína para inhibir la degradación del RNA viral mediada por la respuesta del huésped frente al virus basada en silenciamiento del RNA, ya que CPs relacionadas de carmovirus actúan como supresores de este tipo de mecanismo defensivo (Qu *et al.*, 2003; Thomas *et al.*, 2003; Genovés *et al.*, 2006; Meng *et al.*, 2006; Martínez-Turiño y Hernández, 2009). De manera alternativa, la CP del PLPV podría funcionar como un factor accesorio protegiendo el genoma o promoviendo la traducción *in vivo* y/o estabilizando a las MPs, como se ha propuesto para la CP del PMV (Turina *et al.*, 2000).

Mecanismos de expresión génica del PLPV

Una vez identificadas las ORFs del PLPV funcionales *in vivo* y determinada la función que desempeñan en el ciclo biológico del patógeno, nos planteamos obtener información sobre las estrategias que utiliza el virus para la expresión de todos sus genes. Los resultados indicaron que el PLPV emplea mecanismos de traducción variados que permiten que los distintos genes sean accesibles a la maquinaria de síntesis proteica del huésped. Además, estos procesos traduccionales deben ser independientes de cap ya que los RNAs del PLPV, como los de los otros miembros de la familia *Tombusviridae* (Lommel *et al.*, 2005), carecen de esta estructura en su extremo 5'. Ensayos de traducción *in vitro* confirmaron que: i) el gRNA dirige la síntesis de la proteína p27 y de su producto *readthrough* p87 de manera eficiente, a pesar de que este RNA presenta una región líder de sólo 6 nt (la más corta descrita para un mRNA natural) y, como se ha indicado arriba, carece de estructura cap, y, ii) el sgRNA actúa como mRNA para la producción de las tres proteínas restantes, p7, p9.7 y p37, mediante procesos de escape al proceso de rastreo ribosomal. Estos procesos parecen estar favorecidos por el contexto subóptimo que flanquea al codón de inicio del gen p7, por el codón débil de inicio (GUG) del gen p9.7, y por la ausencia de codones AUG adicionales en cualquier pauta de lectura en la

región de 370 nt que precede al codón de inicio del gen p37. De hecho, en cualquier otra región del RNA viral los tripletes AUG son tan frecuentes como se puede esperar de forma estadística, es decir, aproximadamente uno cada 64 nt. Esta observación sugiere que los codones AUG están bajo fuerte selección negativa en el segmento que precede al gen p37 en el sgRNA del PLPV. En este contexto, se puede postular que esta secuencia ha evolucionado para permitir que una fracción de los ribosomas que han alcanzado el codón de inicio del gen p7, flanqueado por un contexto subóptimo, no empiecen la síntesis proteica y continúen el rastreo hasta el gen p9.7. La débil señal de inicio proporcionada por el codón GUG de este gen probablemente favorece que una fracción de los ribosomas continúen el rastreo hasta alcanzar el gen p37. De acuerdo con este modelo, la introducción de codones AUG adicionales aguas arriba del gen p37, debe reducir la producción de la proteína que codifica, como de hecho se observó con los mutantes p9.7aug y p9.7aug*3, que incorporan uno y tres codones AUG, respectivamente, siendo la reducción más drástica (superior al 80%) con el último mutante, como cabía esperar. Del mismo modo, la eliminación del codón de inicio AUG del gen p7, conlleva un aumento en la producción de la p9.7. Sin embargo, no hemos podido detectar un aumento significativo en la producción de p37 al eliminar el codón de inicio del gen p9.7, probablemente por la baja eficiencia con que el codón GUG promueve el comienzo de la traducción, lo que indica que la presencia de esta ORF sólo reduce ligeramente el número de ribosomas que llegan a la ORF (p37). El modelo también implica que la traducción de todos los genes del sgRNA es dependiente del extremo 5', como efectivamente se comprobó mediante la introducción de una horquilla estable en la secuencia líder del sgRNA que tuvo un acusado efecto negativo sobre la producción de las proteínas p7, de la p9.7 y de la p37.

El mecanismo de escape al proceso de rastreo ribosomal es con frecuencia responsable de la traducción de diferentes proteínas a partir de RNAs virales (revisado por Ryabova *et al.*, 2006). Este mecanismo por lo general conduce a la síntesis de dos polipéptidos con una misma región C-terminal, si son traducidos a partir de una misma ORF, o de dos polipéptidos independientes, cuando las ORFs se encuentran en diferentes pautas de lectura (Weiland y Dreher, 1989; Belsham y Lomonosoff, 1991; Dinesh-Kumar y Miller, 1993; Herzog *et al.*, 1995; Simón-Buela *et al.*, 1997; Verchot *et al.*, 1998; Shivaprasad *et al.*, 2005). El caso más notable de este tipo de mecanismo descrito previamente afecta al RTBV cuyo pgRNA dirige, mediante

escape al proceso de rastreo ribosomal al igual que el sgRNA del PLPV, la síntesis de tres proteínas independientes codificadas por las ORFs I, II y III del virus (Fütterer *et al.*, 1997). Como en el caso del PLPV, una de las ORFs implicadas en el mecanismo de escape al proceso de rastreo del RTBV es traducida a partir de un codón no-AUG (ORFI) y otra presenta un codón de inicio flanqueado por un contexto subóptimo (ORFII). Además, la región de alrededor de 1 kb que precede al codón de inicio de la ORFIII en el pgRNA del RTBV también está desprovista de tripletes AUG en cualquier pauta de lectura, aparte del que inicia a la ORFII que, como se ha indicado anteriormente, se encuentra en un contexto traduccional pobre. Estas peculiaridades compartidas por los RNAs del RTBV y del PLPV probablemente permiten que una importante fracción de los ribosomas que inicia el rastreo desde el extremo 5' pueda alcanzar la tercera ORF. El sesgo en contra de codones AUG previos a la ORFIII es una característica no sólo del RTBV, sino que está conservada en badnavirus relacionados, lo que sugiere que la estrategia de escape al proceso de rastreo ribosomal es empleada por miembros de dos géneros, *Tungrovirus* y *Badnavirus* (Pooggin *et al.*, 1999). En este contexto, es interesante mencionar que el sgRNA del PCRPV se asemeja al sgRNA del PLPV en cuanto a número y organización de ORFs, contextos que flanquean sus codones de inicio y ausencia de codones AUG aguas arriba del gen de la CP, con excepción del codón de inicio del gen p7 (Kinard y Jordan, 2002), lo que indica que el mecanismo de escape al proceso de rastreo del RNA propuesto para el PLPV podría ser también empleado por otros miembros del género propuesto *Pelarspovirus*.

Los buenos niveles de infectividad de los mutantes p9.7cug y p9.7acg del PLPV (con reemplazamientos del codón de inicio GUG del gen p9.7 por otros codones de inicio no-AUG) contrastan con la falta de infectividad del mutante p9.7aug (con reemplazamiento del codón de inicio GUG del gen p9.7 por el canónico AUG), y sugieren que la preservación de las proporciones de las proteínas producidas a partir del sgRNA es un requisito indispensable para la viabilidad del virus. Estos resultados indican además que otros codones de inicio no-AUG pueden sustituir funcionalmente al GUG del gen silvestre, pero no el AUG convencional. La utilización de codones de inicio débiles para regular las proporciones de las proteínas producidas ha sido descrita tanto para mRNAs celulares como virales (Kozak, 1991; Touriol *et al.*, 2003). Entre estos últimos se incluyen algunos virus de plantas pertenecientes a diferentes

grupos taxonómicos (Fütterer *et al.*, 1996; Shirako, 1998). Dentro de la familia *Tombusviridae*, el HCRSV es el único carmovirus donde se ha demostrado la utilización de un codón de inicio no-AUG; este codón está presente en un gen que codifica una proteína específica que parece influir en la expresión de los síntomas (Koh *et al.*, 2006; Zhou *et al.*, 2006). Un segundo ejemplo corresponde al PMV que utiliza un codón débil de iniciación para la síntesis de una proteína de pequeño tamaño, p6.6, implicada en movimiento (Turina *et al.*, 1998), tal como se ha indicado con anterioridad en esta discusión. El único sgRNA producido por el PMV es, como el del PLPV, policistrónico y dirige la traducción de cuatro proteínas distintas. Se ha sugerido que un mecanismo de escape al proceso de rastreo ribosomal sería responsable de la expresión de los dos genes 5'-proximales mientras que la traducción de los dos genes 3'-proximales podría tener lugar mediante una combinación del mecanismo de escape al proceso de rastreo del RNA y alguna estrategia adicional no identificada (Turina *et al.*, 2000; Batten *et al.*, 2006a).

Aunque los datos presentados aquí apoyan firmemente que el mecanismo principalmente involucrado en la traducción de las tres ORFs del sgRNA del PLPV es el de escape al proceso de rastreo ribosomal, no se puede descartar completamente la contribución de otro tipo de mecanismos. Otra cuestión abierta concierne a la forma en que los RNAs del PLPV reclutan al ribosoma pese a carecer de cap y de cola poliA. Miembros de las familias *Tombusviridae* y *Luteoviridae*, que también carecen de este tipo de estructuras en sus extremos, contienen secuencias en la 3' UTR que funcionan como estimuladores de la traducción (TEs) (revisado por Kneller *et al.*, 2006). En el PLPV también se ha postulado la existencia de un TE basándose en un análisis de complementariedad entre la 3' UTR del virus y las 5' UTRs del gRNA y del sgRNA (Fabian y White, 2006). La investigación sobre este tema puede ayudar a comprender mejor la complejidad de la regulación de la expresión génica en el PLPV y la red de interacciones moleculares involucradas.

Estudio de la variabilidad molecular del PLPV

Puesto que la información de secuencia disponible sobre el PLPV se limitaba a la del aislado de origen alemán a partir del cual habíamos determinado la organización genómica del patógeno (aislado PV-0193), y a la del gen CP de otro aislado, de procedencia incierta, que

estaba depositada en la base de datos, consideramos de interés caracterizar aislados adicionales del virus como una primera aproximación para estimar su grado de diversidad genética. Para ello se partió de una colección de aislados provenientes de plantas de *P. zonale* o de *P. peltatum* recolectadas en tres países diferentes (España, Italia y Estados Unidos) en años distintos. La región genómica analizada fue muy amplia, concretamente de 1817 nt lo que representa aproximadamente la mitad del gRNA viral. Esto ha posibilitado la obtención de datos acerca de tres regiones codificantes completas (los genes de las MPs, p7 y p9.7, y de la CP) así como de la 3' UTR y de un segmento del gen de la RdRp. Los resultados del análisis de estas secuencias y su comparación con las del aislado alemán inicial PV-0193, han proporcionado información relevante acerca de la estructura genética de las poblaciones del PLPV como se discute a continuación.

Los cambios nucleotídicos detectados entre las secuencias de un mismo aislado han confirmado que el PLPV, como otros ribovirus, se propaga como una cuasiespecie, siendo la variabilidad nucleotídica dentro de cada aislado muy limitada y similar a la estimada para otros virus de RNA (Fields y Winter, 1981; Arias *et al.*, 1986; Steinhauer *et al.*, 1989; Hedges *et al.*, 1999; Albiach-Martí *et al.*, 2000). A pesar de que el tamaño de la muestra es relativamente pequeña (20 secuencias en total), los análisis filogenéticos han puesto de manifiesto una agrupación significativa de los aislados por país de origen, indicando que el flujo genético derivado de la importación de nuevo material vegetal no es suficiente para erradicar la estructura geográfica de las poblaciones del PLPV.

Del análisis de las secuencias del PLPV también se deduce que el genoma del virus se encuentra bajo fuertes presiones selectivas. Dichas presiones han sido evaluadas en las regiones codificantes mediante la comparación de las tasas de sustituciones sinónimas y no sinónimas utilizando métodos estadísticos apropiados. Los resultados indican que estas regiones se encuentran fundamentalmente bajo selección negativa y han permitido identificar aminoácidos específicos que podrían ser diana de este tipo de selección. Concretamente en la RdRp, se han detectado cuatro sitios bajo selección negativa a pesar de que la porción incluida en el análisis no comprende ninguno de los ocho motivos conservados en la RdRps (Koonin y Dolja, 1991; Koonin y Dolja, 1993). En la proteína p7, el único aminoácido identificado bajo selección negativa se localiza en el supuesto motivo de unión a RNA de la molécula (Marcos *et*

al., 1999, Navarro *et al.*, 2006), pero no corresponde a ninguno de los residuos básicos que presumiblemente son los más críticos en dicha unión. Esto sugiere que la selección probablemente actúe preservando la conformación activa del mencionado motivo. En la proteína p9.7, nueve de los diez sitios que se encuentran bajo selección negativa se concentran en la parte central de la molécula. La función concreta de esta región se desconoce pero corresponde al segmento que conecta los dos dominios hidrofóbicos supuestamente implicados en la asociación de la proteína a membranas, como se ha descrito para proteínas relacionadas (Vilar *et al.*, 2002; Sauri *et al.*, 2005, Navarro *et al.*, 2006), lo que sugiere que este segmento también debe tener determinados requerimientos estructurales para mantener la topología correcta. Por último, en la CP, casi 2/3 de los sitios seleccionados negativamente se localizan en la parte N-terminal (dominio R), aunque sólo dos de ellos corresponden a residuos básicos probablemente críticos en la función de unión a RNA de esta región de la proteína (Lommel *et al.*, 2005). Al igual que se ha sugerido para la p7, la selección podría actuar preservando la conformación de la CP que facilite su unión al RNA.

Además de las presiones selectivas en regiones codificantes, la conservación de la estructura primaria y/o secundaria de regiones reguladoras del RNA viral, tales como la 3' UTR o el presunto promotor subgenómico, parece también restringir la heterogeneidad del PLPV. Esta observación es consistente con el papel crítico que deben desempeñar dichas regiones durante la multiplicación del patógeno, ya que su función estará determinada básicamente por sus características estructurales. En el caso del presunto promotor subgenómico hay que señalar que, además de estar involucrado en la regulación de la transcripción, se encuentra imbricado dentro de la porción 3'-terminal del gen de la RdRp y, por lo tanto, su actividad dual como secuencia codificante y reguladora debe restringir considerablemente su variabilidad. En consistencia con esta previsión, la diversidad nucleotídica media de esta región es más baja que la de otras regiones codificantes y, más aún, la mayoría de sustituciones nucleotídicas detectadas en esta parte del genoma no perturban la estructura de tipo horquilla que presumiblemente es requerida para la actividad promotora, tal como se ha descrito para virus relacionados (Wang *et al.*, 1999; Li y Wong, 2006). Esta observación sugiere que la conservación de la mencionada conformación determina en gran medida el perfil de mutaciones naturales. De forma similar, la distribución de variabilidad en la 3' UTR indica que el

mantenimiento de un plegamiento específico compuesto de una serie de estructuras de tipo horquilla, similar al propuesto para otros *Tombusviridae* (Fabian *et al.*, 2003; Wang y Wong, 2004; Zhang y Simon, 2005; Na y White, 2006), es un factor limitante de la heterogeneidad en esta región. Al igual que en otros virus de la familia, la 3' UTR del PLPV debe contener elementos en *cis* esenciales para la replicación tales como el promotor mínimo requerido para el reconocimiento del RNA viral por la RdRp o elementos represores de la síntesis de cadenas negativas (Stupina y Simón, 1997; Carpenter y Simon, 1998; Pogany *et al.*, 2003; Zhang *et al.*, 2004a, 2004b; Na y White, 2006). Asimismo, en esta región se han descrito TEs que promueven la expresión génica por un mecanismo independiente de cap (revisado por Kneller *et al.*, 2006). Como se ha indicado anteriormente, en el caso del PLPV también se ha propuesto la presencia de un TE en su 3' UTR (Fabian y White, 2006). Una observación interesante a este respecto es la conservación estricta tanto de la estructura primaria como secundaria de este presunto TE en todos los aislados incluidos en este estudio lo que, indirectamente, apoya su funcionalidad *in vivo*.

Otro dato relevante derivado del análisis de variabilidad del PLPV se refiere a la identificación de una red potencial de interacciones específicas dentro y entre proteínas del virus, probablemente requerida para mantener la conformación correcta de dichas proteínas y conducir al RNA viral desde el proceso de replicación hasta su diseminación en planta. Esta presunta red se ha inferido a partir de un análisis covariacional, una aproximación que ha sido ampliamente utilizada para deducir interacciones funcionales entre sitios distintos de una proteína y/o entre diferentes proteínas (Altschuh *et al.*, 1987; Thomas *et al.*, 1996; Larson *et al.*, 2000; Hoffman *et al.*, 2003; Codoñer *et al.*, 2006). Una de las covariaciones detectadas entre moléculas del PLPV afecta a las proteínas p9.7 y CP y distingue al aislado PV-0193, propagado repetidamente en el huésped experimental *C. quinoa*, del resto de aislados. El análisis de la evolución molecular de una única variante del PLPV sometida a pases seriados en dicho huésped experimental ha proporcionado pruebas sólidas de que dicha covariación está determinada por el huésped. Como se ha concluido de los experimentos de genética reversa realizados en el curso de este trabajo, ambas proteínas son indispensables para el movimiento célula a célula del virus y, aunque el mecanismo preciso del transporte intercelular del virus (o de virus relacionados) no se conoce todavía, no es improbable que requiera de una

interacción física entre la(s) MP(s) y la CP, como parece ocurrir en otros virus de plantas (Liu *et al.*, 2001; Sánchez-Navarro y Bol, 2001; Kim *et al.*, 2004; Takeda *et al.*, 2004; Sánchez-Navarro *et al.*, 2006; Akamatsu *et al.*, 2007).

El análisis de la evolución molecular de una variante del virus sometida a pases seriados en *C. quinoa* también ha puesto de manifiesto la capacidad de este patógeno para evolucionar de forma extraordinariamente rápida cuando es transferido a una nueva especie vegetal. Además, la trayectoria evolutiva del PLPV en *C. quinoa* es totalmente reproducible ya que la población derivada de la variante con la que iniciamos los pases seriados, proveniente de un aislado español (SPA4) recuperado de una planta de *P. zonale* recolectada en el año 2000, es esencialmente idéntica a la derivada de un aislado alemán procedente de una planta de *P. peltatum* recolectada en la década de los 90. Aunque no tenemos datos acerca de la composición genética inicial de dicho aislado, es muy probable que difiriera de forma notable del aislado español SPA4, de modo que los datos indican que el huésped introduce un sesgo muy marcado en la población viral dando lugar a fenómenos de evolución convergente (Doolittle, 1994). A pesar de que se desconocen las interacciones precisas que se establecen entre el virus y la especie *C. quinoa*, es interesante destacar que las mutaciones específicas de huésped parecen fijarse en primer lugar en la región 3' del genoma viral, como sugiere la secuencia de una de las variantes de la población recuperada tras el primer pase que posee la región 5' del aislado parental y la región 3' característica de los aislados adaptados a *C. quinoa*. Los cambios fijados inicialmente en esta región 3' probablemente condicionen la fijación posterior de mutaciones en el resto de la(s) molécula(s) viral(es), presumiblemente para optimizar la eficacia biológica del patógeno en el nuevo huésped.

La rápida acumulación de mutaciones detectada en el PLPV en este trabajo es similar a la observada con anterioridad en otro patógeno relacionado, el PFBV (*Carmovirus*) (Rico *et al.*, 2006), y difiere de lo descrito para otros virus más distantes desde un punto de vista taxonómico (Schneider y Roossinck, 2000, 2001; Hébrard *et al.*, 2006; Rebenstorf *et al.*, 2006; Fargette *et al.*, 2008). Estos otros virus codifican RdRps pertenecientes a los supergrupos I y III, a diferencia del PLPV y del PFBV cuyas RdRps corresponden al supergrupo II (Koonin y Dolja, 1993). Una posibilidad ya apuntada en el trabajo de Rico *et al.* (2006), es que los virus del supergrupo II presenten tasas de mutación más altas que los de los otros supergrupos

facilitando así su rápida adaptación a un entorno cambiante. Aunque se sabe poco acerca del tiempo de generación de virus de plantas, algunos estudios permiten estimar un tiempo medio de generación de una hora (Wu *et al.*, 1994; Schneider y Roossinck, 2001). De acuerdo con esta estimación, deberían haberse producido 336 generaciones del virus después del primer pase en *C. quinoa* (es decir, después de 14 días de multiplicación viral). Puesto que la tasa de mutación estimada para ribovirus, incluyendo a virus de plantas aunque hay pocos datos sobre estos últimos (Malpica *et al.*, 2002), se sitúa en torno a 10^{-4} por ronda de replicación (revisado por Duffy *et al.*, 2008), la frecuencia de mutación teórica sería de $3,4 \times 10^{-2}$ tras el primer pase. Si consideramos los 20 clones que se han caracterizado tras dicho pase (10 en pase 0 y 10 en pase 1), el número total de mutaciones/número total de bases secuenciadas sería de $2,8 \times 10^{-2}$. Este valor es claramente superior a la frecuencia de mutación descrita para otros virus de RNA de plantas no pertenecientes al supergrupo II (García-Arenal *et al.*, 2001, Schneider y Roossinck, 2001; Fargette *et al.*, 2008) y, además, es casi igual a la frecuencia de mutación teórica ($3,4 \times 10^{-2}$) considerando una tasa de mutación de 10^{-4} . Puesto que la frecuencia de mutación no sólo depende de la tasa de mutación, sino que está también fuertemente influenciada por la selección y la deriva genética (que, como se ha mencionado en la Introducción general, tienden a disminuir la variabilidad en las poblaciones virales), se puede anticipar que la tasa de mutación del PLPV es inusualmente elevada, una característica que puede tener implicaciones muy relevantes en la progresión de las infecciones causadas por este patógeno y por otros relacionados. En este contexto, es interesante señalar que la capacidad para evolucionar rápidamente podría estar incrementada por la ocurrencia de fenómenos de recombinación, muy frecuente en algunos miembros de los géneros *Tombusvirus* y *Carmovirus*. De hecho, en las infecciones de estos virus en muchos casos se generan DIs por recombinación (Hillman *et al.*, 1987; Burgyan *et al.*, 1989; Cascone *et al.*, 1993; Finnen y Rochon, 1993; Rubino *et al.*, 1995; White y Morris, 1999) y, además, distintos trabajos han permitido detectar un número importante de genomas virales recombinantes tanto en protoplastos como en plantas transgénicas (Cascone *et al.*, 1990, 1993; White y Morris, 1994, 1995; Borja *et al.*, 1999; Desvoyes y Scholthof, 2002). Se ha especulado que los altos niveles de recombinación de este tipo de virus podrían ser consecuencia de la ausencia de dominios helicasa en sus replicasas, lo que podría tener efectos muy importantes sobre la

procesividad de estos enzimas y, por tanto, sobre la probabilidad de que cambien de molde durante la síntesis de la cadena complementaria (Borja *et al.*, 1999). La ausencia de dominios helicasa es un rasgo general de miembros de la familia *Tombusviridae*, de modo que, aunque en este trabajo no hemos detectado recombinación en el PLPV (datos no mostrados), la contribución de este tipo de fenómenos en la adaptación del PLPV (o de virus afines) a determinados ambientes podría ser relevante.

Por último, cabe mencionar que recientemente se ha demostrado que la tasa de mutación de un virus de plantas, concretamente del CMV, puede depender de la especie vegetal en la que se multiplica (Pita *et al.*, 2007). Teniendo en cuenta esta observación, será interesante estudiar la trayectoria evolutiva de PLPV en distintas especies vegetales con el fin de evaluar si el fenómeno del rápido cambio genético de la población viral está específicamente asociado a *C. quinoa* o es general para todos los huéspedes.

Conclusiones

1. La determinación de la estructura primaria completa del gRNA del PLPV, con un tamaño de casi 4 kb, ha permitido confirmar la inclusión de este virus dentro de la familia *Tombusviridae*. Mediante análisis *in silico* inicialmente se identificaron seis ORFs que potencialmente codificaban proteínas de 27 (p27), 13 (p13), 87 (p87), 7 (p7), 6 (p6), y 37 kDa (p37) y estaban flanqueadas por una 5' UTR inusualmente corta, con sólo 6 nt, y una 3' UTR de 246 nt. La organización de las ORFs en el RNA viral y la mayoría de sus productos potenciales eran muy similares a los de miembros del género *Carmovirus*; sin embargo, las estrategias de expresión génica predichas para el PLPV se asemejaban mucho a las descritas para el único miembro del género *Panicovirus*, el PMV, impidiendo su asignación directa a alguno de los géneros ya establecidos dentro de la familia *Tombusviridae*.
2. El análisis de los RNAs del PLPV que se acumulan en plantas infectadas ha revelado la presencia del gRNA y de un solo sgRNA de 1.6 kb. Los resultados de experimentos de extensión del cebador para la caracterización del extremo 5' de ambas moléculas son consistentes con lo esperado para moléculas sin cap, ajustándose a lo descrito para el resto de miembros de la familia *Tombusviridae*.
3. La particularidad del PLPV de producir un solo sgRNA parece ser compartida por otros pequeños virus isométricos relacionados entre los que se encuentran el PeIRSV, el PCRPV y el ELV. Sobre la base de esta característica común y de similitudes de secuencia, se propuso incluir a los cuatro virus en un nuevo género dentro de la familia *Tombusviridae* denominado *Pelarspovirus*, una propuesta apoyada por análisis filogenéticos llevados a cabo tras la determinación de la estructura primaria del PLPV.
4. Se ha generado un clon de cDNA del PLPV a partir del cual se pueden sintetizar *in vitro* transcritos que son infecciosos cuando se inoculan mecánicamente en los huéspedes experimentales *C. quinoa*, *N. clelandii* y *N. benthamiana*. Este clon ha sido una herramienta fundamental para determinar las ORFs del PLPV que son funcionales *in vivo* y para analizar las estrategias de expresión génica utilizadas por este patógeno.

5. Las ORFs (p27), (p87), (p7) y (p37) predichas inicialmente son expresadas por el virus y resultan indispensables para una infección efectiva por su implicación en replicación (p27 y p87), movimiento (p7) y movimiento/encapsidación (p37). Sin embargo, las ORFs (p13) y (p6) carecen de relevancia biológica y, en su lugar, se ha identificado una nueva ORF, localizada en la región central del genoma, que se traduce a partir de un codón no-AUG de iniciación (GUG) y que da lugar a una proteína de 9.7 kDa (p9.7) implicada en el movimiento del virus. Esta proteína muestra una notable identidad de secuencia (44,4%) con otra de tamaño 9.4 kDa, potencialmente codificada por una ORF del PCRPV también ubicada en la región central del genoma y cuyo codón de inicio es, asimismo, un triplete no-AUG, lo que amplía las semejanzas entre estos dos virus, ambos miembros provisionales del género propuesto *Pelarspovirus*.

6. El gRNA del PLPV dirige la síntesis de las proteínas p27 y p87 de manera eficiente (la última mediante un mecanismo de lectura a través del codón de parada débil del gen p27), pese a carecer de una estructura cap en su extremo 5' y presentar una secuencia líder extremadamente corta; el sgRNA actúa como mRNA para la producción de las tres proteínas restantes, p7, p9.7 y p37, mediante un mecanismo de escape al proceso de rastreo ribosomal o *leaky-scanning*. Estos procesos parecen estar favorecidos por el contexto subóptimo del codón de inicio del gen p7, por el codón débil de inicio (GUG) del gen p9.7, y por la ausencia de codones AUG adicionales en cualquier pauta de lectura en la región que precede al codón de inicio del gen p37 en el sgRNA.

7. El estudio de la variabilidad de la mitad 3' del gRNA del PLPV ha puesto de manifiesto un alto grado de conservación en las secuencias aminoacídicas de determinados dominios de las proteínas y la preservación del plegamiento de regiones reguladoras del RNA viral. Análisis filogenéticos con las secuencias de aislados de distintas procedencia sugieren una agrupación de los aislados por país de origen indicando que el flujo genético derivado de la importación de material no es suficiente para erradicar la estructuración geográfica de las poblaciones del PLPV.

8. El estudio de covariaciones dentro y entre proteínas del PLPV ha permitido identificar una supuesta red de interacciones específicas probablemente requerida para mantener la conformación correcta en las proteínas virales y para conducir al RNA viral desde el proceso de replicación hasta su diseminación en la planta.

9. Pases seriados de una variante natural del PLPV en el huésped experimental *C. quinoa*, han mostrado que este patógeno es capaz de evolucionar de forma muy rápida y reproducible cuando es transferido a una nueva especie vegetal, poniendo de manifiesto el papel crítico que desempeñan las presiones selectivas ejercidas por el huésped en la estructura genética de las poblaciones del virus. La rápida acumulación de sustituciones nucleotídicas en *C. quinoa* también sugiere que el virus presenta una tasa de mutación inusualmente elevada, lo que puede tener implicaciones relevantes en el progreso de las infecciones causadas por el PLPV.

Bibliografía

- Adkins S, Hammond J, Gera A, Maroon-Lango CJ, Sobolev I, Harness A, Zeidan M, Spiegel S.** 2006. Biological and molecular characterization of a novel carmovirus isolated from *Angelonia*. *Phytopathology* **96**: 460-467.
- Agol VI.** 2006. Molecular mechanisms of poliovirus variation and evolution. In Domingo E (ed.), Quasispecies: concepts and implications for virology. *Current topics in Microbiology and Immunology* vol. 299. pp. 211–259. Springer, Heidelberg, Germany.
- Akamatsu N, Takeda A, Kishimoto M, Kaido M, Okuno T, Mise K.** 2007. Phosphorylation and interaction of the movement and coat proteins of *Brome mosaic virus* in infected barley protoplasts. *Arch. Virol.* **152**: 2087-2093.
- Albiach-Martí MR, Navassi M, Gowda S, Satyanarayana T, Hilf ME, Shanker S, Almira EC, Vives MC, López C, Guerri J, Flores R, Moreno P, Garnsey SM, Dawson WO.** 2000. Sequences of *Citrus tristeza virus* separated in time and space are essentially identical. *J. Virol.* **74**: 6856-6865.
- Alejska M, Figlerowicz M, Malinowska N, Urbanowicz A, Figlerowicz M.** 2005a. A universal BMV-based RNA recombination system-how to search for general rules in RNA recombination. *Nucleic Acids Res.* **33**: 105-119.
- Alejska M, Kurzyńska-Kokorniak A, Broda M, Kierzek R, Figlerowicz M.** 2001. How RNA viruses exchange their genetic material. *Acta Biochim. Pol.* **48**: 391-407.
- Alejska M, Malinowska N, Urbanowicz A, Figlerowicz M.** 2005b. Two types of non-homologous RNA recombination in *Brome mosaic virus*. *Acta Biochim. Pol.* **52**: 833-844.
- Alicai T, Fendy NS, Gibson RW, Adipala E, Vetten HJ, Foster GD, Seal SE.** 1999. Occurrence of two serotypes of *Sweet potato chlorotic stunt virus* in East Africa and their associated differences in coat protein and HSP70 homologue gene sequences. *Plant Pathol.* **78**: 718-726.
- Alonso M, Borja M.** 2005. High incidence of Pelargonium line pattern virus infecting asymptomatic *Pelargonium* spp. in Spain. *Eur. J. Plant Pathol.* **112**: 95-100.
- Alonso M, Gisbert E, Fraga M, Borja M.** 1999. Sintomatología de los virus que infectan geranio. Libro de resúmenes del VI Congreso Nacional de Virología, p. 136.
- Althaus CL, Bonhoeffer S.** 2005. Stochastic interplay between mutation and recombination during the acquisition of drug resistance mutations in *Human immunodeficiency virus* type 1. *J. Virol.* **79**: 13572-13578.
- Altschuh D, Lesk AM, Bloomer AC, Klug A.** 1987. Correlation of co-ordinated aminoacid substitutions with function in viruses related to *Tobacco mosaic virus*. *J. Mol. Biol.* **193**: 693-707.

- Anderson RM, May RM.** 1982. Coevolution of host and parasites. *Parasitology* **85**: 411-426.
- Aparicio F, Vilar M, Perez-Payá, E, Pallás V.** 2003. The coat protein of *Prunus necrotic ringspot virus* specifically binds to and regulates the conformation of its genomic RNA. *Virology* **313**: 213-223.
- Arias CF, López S, Espejo RT.** 1986. Heterogeneity in base sequence among different DNA clones containing equivalent sequence of rotavirus double-stranded RNA. *J. Virol.* **57**: 1207-1209.
- Ashoub A, Rohde W, Prufer D.** 1998. In planta transcription of a second subgenomic RNA increases the complexity of the subgroup 2 luteovirus genome. *Nucleic Acids Res.* **26**: 420-426.
- Asurmendi S, Berg RH, Koo JC, Beachy RN.** 2004. Coat protein regulates formation of replication complexes during *Tobacco mosaic virus* infection. *Proc. Natl. Acad. Sci. USA* **101**: 1415-1420.
- Ayllón MA, Rubio L, Sentandreu V, Moya A, Guerri J, Moreno P.** 2006. Variations in two gene sequences of *Citrus tristeza virus* after host passage. *Virus Genes* **32**: 119-128.
- Bacher JW, Warkentin D, Ramsdell D, Hancock F.** 1994. Selection versus recombination: What is maintaining identity in the 3' termini of *Blueberry leaf mottle nepovirus* RNA1 and RNA2? *J. Gen. Virol.* **75**: 2133-2137.
- Balmori E, Gilmer D, Richards K, Guilley H, Jonard G.** 1993. Mapping the promoter for subgenomic RNA synthesis on *Beet necrotic yellow vein virus* RNA 3. *Biochimie* **75**: 517-521.
- Baranov PV, Gesteland RF, Atkins JF.** 2002. Recoding: Translational bifurcations in gene expression. *Gene* **286**: 187-201.
- Basso J, Dallaire P, Charest PJ, Devantier Y, Laliberte JF.** 1994. Evidence for an internal ribosome entry site within the 5' non-translated region of *Turnip mosaic potyvirus* RNA. *J. Gen. Virol.* **75**: 3157-3165.
- Batten JS, Desvoyes B, Yamamura Y, Scholthof KB.** 2006a. A translational enhancer element on the 3'-proximal end of the *Panicum mosaic virus* genome. *FEBS Lett.* **580**: 2591-2597.
- Batten JS, Turina M, Scholthof KBG.** 2006b. *Panicovirus* accumulation is governed by two membrane-associated proteins with a newly identified conserved motif that contributes to pathogenicity. *Virol. J.* **3**: 12-24.
- Beauchemin C, Boutet N, Laliberté JF.** 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 iso4E in planta. *J. Virol.* **81**: 775-782.

- Belsham GJ, Lomonossoff GP.** 1991. The mechanism of translation of *Cowpea mosaic virus* middle component RNA: no evidence for internal initiation from experiments in an animal cell transient expression system. *J. Gen. Virol.* **72**: 3109-3113.
- Bol JF.** 2005. Replication of alfamo- and ilarviruses: Role of the coat protein. *Annu. Rev. Phytopathol.* **43**: 39-62.
- Bonneville JM, Sanfaçon H, Fütterer J, Hohn T.** 1989. Posttranscriptional trans-activation in *Cauliflower mosaic virus*. *Cell* **59**: 1135-1143.
- Borja M, Rubio T, Scholthof HB, Jackson AO.** 1999. Restoration of wild-type virus by double recombination of tombusvirus mutants with a host transgene. *Mol. Plant Microbe Interact.* **12**: 153-162.
- Bouwen I, Maat D.** 1992. *Pelargonium flower break* and *Pelargonium line pattern* viruses in the Netherlands: purification, antiserum preparation, serological identification and detection in *Pelargonium* by ELISA. *Neth. J. Plant Pathol.* **98**: 141-156.
- Boyer JC, Haenni AL.** 1994. Infectious transcripts and cDNA clones of RNA viruses. *Virology* **198**: 415-426.
- Boyko V, Ferralli J, Ashby J, Schellenbaum P, Heinlein M.** 2000. Function of microtubules in intercellular transport of plant virus RNA. *Nat. Cell. Biol.* **2**: 826-32.
- Braut V, Miller WA.** 1992. Translational frameshifting mediated by a viral sequence in plant cells. *Proc. Natl. Acad. Sci. USA* **89**: 2262-2266.
- Braut V, Perigon S, Reinbold C, Erdinger M, Scheidecker D, Herrbach E, Richards K, Ziegler-Graff V.** 2005. The polerovirus minor capsid protein determines vector specificity and intestinal tropism in the aphid. *J. Virol.* **79**: 9685-9693.
- Braut V, Van den Heuvel JFJM, Verbeek M, Ziegler-Graff V, Reutenauer A, Herrbach E, Garaud JC, Guilley H, Richards K, Jonard G.** 1995. Aphid transmission of *Beet western yellows luteovirus* requires the minor capsid read-through protein P74. *EMBO J.* **14**: 650-659.
- Brill LM, Nunn RS, Kahn TW, Yeager M, Beachy RN.** 2000. Recombinant *Tobacco mosaic virus* movement protein is an RNA-binding, alpha-helical membrane protein. *Proc. Natl. Acad. Sci. USA* **97**: 7112-7117.
- Brostrom CO, Brostrom MA.** 1998. Regulation of translational initiation during cellular responses to stress. *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 79-125.

- Brown CM, Dinesh-Kumar SP, Miller WA.** 1996. Local and distant sequences are required for efficient read-through of the *Barley yellow dwarf virus*-PAV coat protein gene stop codon. *J. Virol.* **70**: 5884–5892.
- Browning KS.** 2004. Plant translation initiation factors: it is not easy to be green. *Biochem. Soc. Trans.* **32**: 589-591.
- Brunt AA, Crabtree K, Dallwitz MJ, Gibbs AJ, Watson L, Zurcher EJ.** 1996. *Pelargonium flower break virus* and *Pelargonium line pattern virus*. Plant Viruses Online: Descriptions and lists from the vide database.
- Burgyan J, Rubino L, Russo M.** 1989. Defective interfering RNA molecule in *Cymbidium ringspot virus* infections. *J. Gen. Virol.* **70**: 235-239.
- Calkhoven CF, Müller C, Leutz A.** 2000. Translational control of C/EBPa and C/EBPb isoform expression. *Genes Dev.* **14**: 1920-1932.
- Cañizares MC, Navas-Castillo J, Moriones E.** 2008. Multiple suppressors of RNA silencing encoded by both genomic RNAs of the crinivirus, *Tomato chlorosis virus*. *Virology* **379**: 168-174.
- Carpenter CD, Simon AE.** 1998. Analysis of sequences and predicted structures required for viral satellite RNA accumulation by *in vivo* genetic selection. *Nucleic Acids Res.* **26**: 2426-2432.
- Carrington JC, Heaton LA, Zuidema D, Hillman BI, Morris TJ.** 1989. The genome structure of *Turnip crinkle virus*. *Virology* **170**: 219-226.
- Carrington JC, Morris TJ.** 1985. Characterization of the cell-free translation products of *Carnation mottle virus* genomic and subgenomic RNAs. *Virology* **144**: 1-10.
- Carrington JC, Morris TJ, Stockley PG, Harrison SL.** 1987. Structure and assembly of *Turnip crinkle virus*. IV. Analysis of the coat protein gene and implications of the subunit primary structure. *J. Mol. Biol.* **194**: 265-276.
- Cascone PJ, Carpenter CD, Li XH, Simon AE.** 1990. Recombination between satellite RNAs of *Turnip crinkle virus*. *EMBO J.* **9**: 1709-1715.
- Cascone PJ, Haydar TF, Simon AE.** 1993. Sequences and structures required for recombination between virus-associated RNAs. *Science* **260**: 801-805.
- Castro C, Arnold JJ, Cameron CE.** 2005. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res.* **107**: 141–149.

- Cevik B, Lee RF, Niblett CL.** 2008. *In vivo* and *in vitro* expression analysis of the RNA-dependent RNA polymerase of *Citrus tristeza virus*. *Arch. Virol.* **153**: 315–321.
- Chamorro M, Parkin N, Varmus HE.** 1992. An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc. Natl. Acad. Sci. USA* **89**: 713-717.
- Chappell SA, Dresios J, Edelman GM, Mauro VP.** 2006. Ribosomal shunting mediated by a translational enhancer element that base pairs to 18S rRNA. *Proc. Natl. Acad. Sci. USA* **103**: 9488-9493.
- Chare ER, Holmes EC.** 2006. A phylogenetic survey of recombination frequency in plant RNA viruses. *Arch. Virol.* **151**: 933-946.
- Chung BY, Miller WA, Atkins JF, Firth AE.** 2008. An overlapping essential gene in the *Potyviridae*. *Proc. Natl. Acad. Sci. USA* **105**: 5897-5902.
- Citovsky V, Knorr D, Schuster G, Zambryski P.** 1990. The P30 movement protein of *Tobacco mosaic virus* is a single-strand nucleic acid-binding protein. *Cell* **60**: 637-647.
- Ciuffreda P, Rubino L, Russo M.** 1998. Molecular cloning and complete nucleotide sequence of *Galinsoga mosaic virus* genomic RNA. *Arch. Virol.* **143**: 173-180.
- Codoñer FM, Fares MA, Elena SF.** 2006. Adaptive covariation between the coat and the movement proteins of *Prunus necrotic ringspot virus*. *J. Virol.* **80**: 5833-5840.
- Codoñer FM, Fares M.** 2008. Why should we care about molecular coevolution?. *Evol. Bioinform.* **4**: 29-38.
- Cohen Y, Qu F, Gisel A, Morris TJ, Zambryski PC.** 2000. Nuclear localization of *Turnip crinkle virus* movement protein p8. *Virology* **273**: 276-285.
- Crawford KM, Zambryski PC.** 2001. Non-targeted and targeted protein movement through plasmodesmata in leaves in different development and physiological states. *Plant Physiol.* **125**: 1802-1812.
- De Filippis VR, Villareal LP.** 1999. Virus evolution. *Virology* **1**: 353-370.
- Desvoyes B, Scholthof KB.** 2000. RNA: protein interactions associated with satellites of *Panicum mosaic virus*. *FEBS Lett.* **485**: 25-28.
- Desvoyes B, Scholthof KB.** 2002. Host-dependent recombination of a *Tomato bushy stunt virus* coat protein mutant yields truncated capsid subunits that form virus-like complexes which benefit systemic spread. *Virology* **304**: 434-442.

- Di R, Dinesh-Kumar SP, Miller WA.** 1993. Translational frameshifting by *Barley yellow dwarf virus* RNA (PAV serotype) in *Escherichia coli* and in eukaryotic cell-free extracts. *Mol. Plant-Microbe Interact.* **6**: 444-452.
- Díaz JA, Bernal JJ, Moriones E, Aranda MA.** 2003. Nucleotide sequence and infectious transcripts from a full-length cDNA clone of the carmovirus *Melon necrotic spot virus*. *Arch. Virol.* **148**: 599-607.
- Díez J, Marcos JF, Pallás V.** 1999. Characterization and *in vitro* translation analysis of *Pelargonium flower break virus*. *Arch. Virol.* **144**: 1627-1637.
- Dinesh-Kumar SP, Braut V, Miller WA.** 1992. Precise mapping and *in vitro* translation of a trifunctional subgenomic RNA of *Barley yellow dwarf virus*. *Virology* **187**: 711-722.
- Dinesh-Kumar SP, Miller WA.** 1993. Control of start codon choice on a plant viral RNA encoding overlapping genes. *Plant Cell* **5**: 679-692.
- Ding SW, Li WX, Symons RH.** 1995. A novel naturally occurring hybrid gene encoded by a plant RNA virus facilitates long distance movement. *EMBO J.* **14**: 5762-5772.
- Dinkova TD, Zepeda H, Martínez-Salas E, Martínez LM, Nieto-Sotelo J, de Jiménez ES.** 2005. Cap-independent translation of maize Hsp101. *Plant J.* **41**: 722-731.
- Doan DN, Lee KC, Laurinmaki P, Butcher S, Wong SM, Dokland T.** 2003. Three-dimensional reconstruction of *Hibiscus chlorotic ringspot virus*. *J. Struct. Biol.* **144**: 253-261.
- Domingo E.** 1994. Virus en evolución. Una aproximación a la dinámica poblacional de los virus y al origen de nuevas enfermedades víricas. Eudema, Madrid.
- Domingo E, Escarmís C, Sevilla N, Moya A, Elena SF, Quer J, et al.** 1996. Basic concepts in RNA virus evolution. *FASEB J.* **10**: 859-864.
- Domingo E, Holland JJ.** 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**: 151-178.
- Domingo E, Martínez-Salas E, Sobrino F, de la Torre JC, Portela A, Ortiz J, et al.** 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance-a review. *Gene* **40**: 1-8.
- Dominguez DI, Ryabova LA, Pooggin MM, Schmidt-Puchta W, Fütterer J, Hohn T.** 1998. Ribosome shunting in *Cauliflower mosaic virus*. Identification of an essential and sufficient structural element. *J. Biol. Chem.* **273**: 3669-3678.

- Donald RG, Lawrence DM, Jackson AO.** 1997. The *Barley stripe mosaic virus* 58-kilodalton beta (b) protein is a multifunctional RNA binding protein. *J. Virol.* **71**: 1538-1546.
- Doolittle RF.** 1994. Convergent evolution: the need to be explicit. *Trends Biochem. Sci.* **19**: 15-18.
- Dorokhov YL, Ivanov PA, Komarova TV, Skulachev MV, Atabekov JG.** 2006. An internal ribosome entry site located upstream of the crucifer-infecting tobamovirus coat protein (CP) gene can be used for CP synthesis *in vivo*. *J. Gen. Virol.* **87**: 2693-2697.
- Dorokhov YL, Skulachev MV, Ivanov PA, Zvereva SD, Tjulkina LG, Merits A, Gleba YY, Hohn TJ, Atabekov, JG.** 2002. Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. *Proc. Natl. Acad. Sci. USA* **99**: 5301–5306.
- Drake JW.** 1993. Role of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90**: 4171-4175.
- Drake JW, Holland JJ.** 1999. Mutation ranges among RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**: 13910-13913.
- Dreher TW, Miller WA.** 2006. Translational control in positive strand RNA plant viruses. *Virology* **344**: 185-197.
- Dreher TW, Uhlenbeck OC, Browning KS.** 1999. Quantitative assessment of *EF-1alpha*. GTP binding to aminoacyl-tRNAs, aminoacyl-viral RNA, and tRNA shows close correspondence to the RNA binding properties of EF-Tu. *J. Biol. Chem.* **274**: 666–672.
- Dugeon G, Urcuqui-Inchima S, Milner M, Kadare G, Valle RPC, Voyatzakis A, Haenni A-L, Schirawski J.** 1999. The strategies of plant virus gene expression: models of economy. *Plant Sci.* **148**: 77-88.
- Duffy S, Shackelton LA, Holmes EC.** 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**: 267-276.
- Duncan R, Hershey JWB.** 1989. Protein synthesis and protein phosphorylation during heat stress, recovery and adaptation. *J. Cell Biol.* **109**: 1467-1481.
- Duntheil J, Pupko T, Jean-Marie A, Galtier N.** 2005. A model-based approach for detecting coevolving positions in a molecule. *Mol. Biol. Evol.* **22**: 1919-1928.
- Eigen M.** 1996. On the nature of quasispecies. *Trends Microbiol.* **4**: 212-214.
- Eigen M, Biebricher CK.** 1988. Sequence space and quasispecies distribution. In Domingo E, Ahlquist P, Holland J. (eds). *RNA Genetics* vol. 3, pp. 211-245. CRC Press, Boca Raton, Florida, USA.

- Eigen M, Schuster P.** 1981. Comments on "growth of a hypercycle" by King. *Biosystems* **13**: 235.
- Eigen M, Gardiner W, Schuster P, Winkler-Oswatitsch R.** 1981. The origin of genetic information. *Sci. Am.* **244**: 88-92.
- Fabian MR, Na H, Ray D, White KA.** 2003. 3'-Terminal RNA secondary structures are important for accumulation of *Tomato bushy stunt virus* DI RNAs. *Virology* **313**: 567-580.
- Fabian MR, White KA.** 2004. 5'-3' RNA-RNA interaction facilitates cap- and poly(A) tail-independent translation of *Tomato bushy stunt virus* mRNA: a potential common mechanism for *Tombusviridae*. *J. Biol. Chem.* **279**: 28862-28872.
- Fabian MR, White KA.** 2006. Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA* **12**: 1304-1314.
- Fagoaga C, Semancik JS, Durán-Vila N.** 1995. A *Citrus exocortis viroid* variant from broad bean (*Vicia faba* L.): infectivity and pathogenesis. *J. Gen. Virol.* **76**: 2271-2277.
- Fargette D, Pinel A, Rakotomalala M, Sangu E, Traoré O, Sérémé D, Sorho F, Issaka S, Hébrard E, Séré Y, Kanyeka Z, Konaté G.** 2008. *Rice yellow mottle virus*, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J. Virol.* **82**: 3584-3589.
- Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA.** 2005. Virus taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, USA.
- Fields S, Winter G.** 1981. Nucleotide sequence heterogeneity and sequence rearrangements in *Influenza virus* cDNA. *Gene* **15**: 207-214.
- Finnen RL, Rochon DM.** 1993. Sequence and structure of defective interfering RNAs associated with *Cucumber necrosis virus* infections. *J. Gen. Virol.* **74**: 1715-1720.
- Flasinski S, Cassidy BG.** 1998. Potyvirus aphid transmission requires helper component and homologous coat proteins for maximal efficiency. *Arch. Virol.* **143**: 2159-2172.
- Fraile A, Alonso-Prados JL, Aranda MA, Bernal JJ, Malpica JM, García-Arenal F.** 1997. Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. *J. Virol.* **71**: 934-940.
- Fraile A, Malpica JM, Aranda MA, Rodríguez-Cerezo E, García-Arenal F.** 1996. Genetic diversity in *Tobacco mild green mosaic tobamovirus* infecting the wild plant *Nicotiana glauca*. *Virology* **233**: 148-155.

- Franck A, Loebenstein G.** 1994. Virus and virus-like diseases of Pelargonium in Israel. *Acta Hort.* **377**: 31-39.
- Fraser CS, Doudna JA.** 2007. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat. Rev. Microbiol.* **5**: 29-38.
- French R, Stenger DC.** 2003. Evolution of *Wheat streak mosaic virus*: dynamics of population growth within plants may explain limited variation. *Annu. Rev. Phytopathol.* **41**: 199-214.
- Fujita M, Mise K, Kajiura Y, Dohi K, Furusawa I.** 1998. Nucleic acid-binding properties and subcellular localization of the 3a protein of *Brome mosaic bromovirus*. *J. Gen. Virol.* **79**: 1273-1280.
- Fütterer J, Gordon K, Sanfaçon H, Bonneville JM, Hohn T.** 1990. Positive and negative control of translation by the leader sequence of *Cauliflower mosaic virus* pregenomic 35S RNA. *EMBO J.* **9**: 1697-1707.
- Fütterer J, Hohn T.** 1991. Translation of a polycistronic mRNA in the presence of the *Cauliflower mosaic virus* transactivator protein. *EMBO J.* **10**: 3887-3896.
- Fütterer J, Hohn T.** 1996. Translation in plants—rules and exceptions. *Plant Mol. Biol.* **32**: 159-189.
- Fütterer J, Kiss-László Z, Hohn T.** 1993. Nonlinear ribosome migration on *Cauliflower mosaic virus* pregenomic 35S RNA. *Cell* **73**: 789-802.
- Fütterer J, Potrykus I, Bao Y, Li L, Burns TM, Hull R, Hohn T.** 1996. Position-dependent ATT initiation during plant pararetrovirus *Rice tungro bacilliform virus* translation. *J. Virol.* **70**: 2999-3010.
- Fütterer J, Rothnie HM, Hohn T, Potrykus I.** 1997. *Rice tungro bacilliform virus* open reading frames II and III are translated from polycistronic pregenomic RNA by leaky-scanning. *J. Virol.* **71**: 7984-7989.
- Gaba A, Wang Z, Krishnamoorthy T, Hinnebusch AG, Sachs MS.** 2001. Physical evidence for distinct mechanisms of translational control by upstream open reading frames. *EMBO J.* **20**: 6453-6463.
- Gale M Jr, Tan SL, Katze MG.** 2000. Translational control of viral gene expression in eukaryotes. *Microbiol. Mol. Biol. Rev.* **64**: 239-280.
- Galiakparov N, Tanne E, Sela I, Gafny R.** 1999. Infectious RNA transcripts from *Grapevine virus A* cDNA clone. *Virus Genes* **19**: 235-242.
- Gallie DR.** 1996. Translational control of cellular and viral mRNAs. *Plant Mol. Biol.* **32**: 145-158.
- Gallie DR.** 1998. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* **216**: 1–11.

- Gallie DR.** 2001. Cap-independent translation conferred by the 5' leader of *Tobacco etch virus* is eukaryotic initiation factor 4G dependent. *J. Virol.* **75**: 12141-12152.
- Gallie DR, Kovayashi M.** 1994. The role of the 3'- untranslated region of non-polyadenylated plant viral mRNAs in regulating translational efficiency. *Gene* **142**: 159-165.
- Gallie DR, Walbot V.** 1990. RNA pseudoknot domain of *Tobacco mosaic virus* can functionally substitute for a poly(A) tail in plant and animal cells. *Genes Dev.* **4**: 1149-1157.
- Galtier N, Dutheil J.** 2007. Coevolution within and between genes. *Genome Dyn.* **3**: 1-12.
- Gao Z, Johansen E, Eysers S, Thomas CL, Noel Ellis TH, Maule AJ.** 2004. The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *Plant J.* **40**: 376-385.
- García-Andrés S, Monci F, Navas-Castillo J, Moriones E.** 2006. *Begomovirus* genetic diversity in the native plant reservoir *Solanum nigrum*: evidence for the presence of a new virus species of recombinant nature. *Virology* **350**: 433-442.
- García-Andrés S, Tomás DM, Sánchez-Campos S, Navas-Castillo J, Moriones E.** 2007. Frequent occurrence of recombinants in mixed infections of *Tomato yellow leaf curl* disease-associated begomoviruses. *Virology* **365**: 210-219.
- García-Arenal F, Fraile A, Malpica JM.** 1999. Genetic variability and evolution. In Mandahar GL (ed.). *Molecular biology of plant viruses*, pp. 143-159. Kluwer, London.
- García-Arenal F, Fraile A, Malpica JM.** 2001. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* **39**: 157-186.
- García-Arenal F, Fraile A, Malpica JM.** 2003. Variation and evolution of plant virus populations. *Int. Microbiol.* **6**: 225-232.
- García-Castillo S, Sánchez-Pina A, Pallás V.** 2003. Spatio-temporal analysis of the RNAs, coat and movement (p7) proteins of *Carnation mottle virus* in *Chenopodium quinoa* plants. *J. Gen. Virol.* **84**: 745-749.
- Genovés A, Navarro JA, Pallás V.** 2006. Functional analysis of the five *Melon necrotic spot virus* genome-encoded proteins. *J. Gen. Virol.* **87**: 2371-2380.
- Giedroc DP, Theimer CA, Nixon PL.** 2000. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.* **298**: 167-185.

- Giedroc DP, Cornish PV.** 2009. Frameshifting RNA pseudoknots: Structure and mechanism. *Virus Res.* **139**: 193-208.
- Göbel U, Sander C, Schneider R, Valencia A.** 1994. Correlated mutations and residue contacts in proteins. *Proteins* **18**: 309–317.
- Goh ChS, Cohen FE.** 2002. Coevolutionary analysis reveals insights into protein-protein interactions. *J. Mol. Biol.* **324**: 177–192.
- Gordon K, Fütterer J, Hohn T.** 1992. Efficient initiation of translation at non-AUG triplets in plant cells. *Plant J.* **2**: 809-813.
- Graber TE, Holcik M.** 2007. Cap-independent regulation of gene expression in apoptosis. *Mol. Biosyst.* **3**: 825-834.
- Gramstat A, Prufer D, Rohde W.** 1994. The nucleic acid-binding zinc finger protein of *Potato virus M* is translated by internal initiation as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage. *Nucleic Acids Res.* **22**: 3911–3917.
- Grant CM, Miller PF, Hinnebusch AG.** 1995. Sequences 5' of the first upstream open reading frame in GCN4 mRNA are required for efficient translational reinitiation. *Nucleic Acids Res.* **23**: 3980-3988.
- Guilley H, Carrington JC, Balázs E, Jonard G, Richards K, Morris TJ.** 1985. Nucleotide sequence and genome organization of *Carnation mottle virus* RNA. *Nucleic Acids Res.* **13**: 6663-6677.
- Guo L, Allen E, Miller WA.** 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Mol. Cell.* **7**: 1103–1109.
- Hacker DL, Petty IT, Wei N, Morris TJ.** 1992. *Turnip crinkle virus* genes required for RNA replication and virus movement. *Virology* **186**: 1-8.
- Hann LE, Webb AC, Cai JM, Gehrke L.** 1997. Identification of a competitive translation determinant in the 3' untranslated region of *Alfalfa mosaic virus* coat protein mRNA. *Mol. Cell Biol.* **17**: 2005-2013.
- Harbison SA, Davies JW, Wilson TMA.** 1985. Expression of high molecular weight polypeptides by *Carnation mottle virus* RNA. *J. Gen. Virol.* **66**: 2597-2604.
- Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D.** 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* **6**: 1099-1108.
- Havelda Z, Dalmay T, Burgyan J.** 1995. Localization of cis-acting sequences essential for *Cymbidium ringspot tobamovirus* defective interfering RNA replication. *J. Gen. Virol.* **76**: 2311-2316.

- Haydon DT, Bastos AD, Knowles NJ, Samuel AR.** 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **157**: 7-15.
- Heaton LA, Lee TC, Wei N, Morris TJ.** 1991. Point mutations in the *Turnip crinkle virus* capsid protein affect the symptoms expressed by *Nicotiana benthamiana*. *Virology* **183**: 143-150.
- Hébrard E, Pinel-Galzi A, Bersoult A, Siré C, Fargette D.** 2006. Emergence of a resistance-breaking isolate of *Rice yellow mottle virus* during serial inoculations is due to a single substitution in the genome-linked viral protein VPg. *J. Gen. Virol.* **87**: 1369-1373.
- Hedges JF, Balasuyira UB, Yimoney PJ, Mccollum WH, Maclachlan NJ.** 1999. Genetic divergence with emergence of novel phenotypic variants of *Equine arteritis virus* during persistent infection of stallions. *J. Virol.* **73**: 3672-3681.
- Heinlein M, Epel BL, Padgett HS, Beachy RN.** 1995. Interaction of tobamovirus movement proteins with the plant cytoskeleton. *Science* **270**: 1983-1985.
- Hellen CU, Sarnow P.** 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **15**: 1593-1612.
- Hemmings-Mieszczak M, Hohn T, Preiss T.** 2000. Termination and peptide release at the upstream open reading frame are required for downstream translation on synthetic shunt-competent mRNA leaders. *Mol. Cell Biol.* **20**: 6212-6223.
- Herranz MC, Pallás V.** 2004. RNA-binding properties and mapping of the RNA-binding domain from the movement protein of *Prunus necrotic ringspot virus*. *J. Gen. Virol.* **85**: 761-768.
- Herranz MC, Sánchez-Navarro JA, Saurí A, Mingarro I, Pallás V.** 2005. Mutational analysis of the RNA-binding domain of the *Prunus Necrotic ringspot virus* (PNRSV) movement protein reveals its requirement for cell-to-cell movement. *Virology* **339**: 31-41.
- Herzog E, Guilley H, Fritsch C.** 1995. Translation of the second gene of *Peanut clump virus* RNA 2 occurs by leaky scanning *in vitro*. *Virology* **208**: 215-225.
- Hillman BI, Carrington JC, Morris TJ.** 1987. A defective interfering RNA that contains a mosaic of a plant virus genome. *Cell* **51**: 427-433.
- Hoffman NG, Schiffer CA, Swanstrom R.** 2003. Covariation of amino acid positions in HIV-1 protease. *Virology* **314**: 536-548.

- Holcik M, Sonenberg N.** 2005. Translational control in stress and apoptosis. *Nat. Rev. Mol. Cell Biol.* **6**: 318-327.
- Holland JJ, Domingo E, de la Torre JC, Steinhauer DA.** 1990. Mutation frequencies at defined single codon sites in *Vesicular stomatitis virus* and poliovirus can be increased only slightly by chemical mutagenesis. *J. Virol.* **64**: 3960-3962.
- Holness CL, Lomonossoff GP, Evans D, Maule AJ.** 1989. Identification of the initiation codons for translation of *Cowpea mosaic virus* middle component RNA using site-directed mutagenesis of an infectious cDNA clone. *Virology* **172**: 311-320.
- Huang M, Koh DC, Weng LJ, Chang ML, Yap YK, Zhang L, Wong SM.** 2000. Complete nucleotide sequence and genome organization of *Hibiscus chlorotic ringspot virus*, a new member of the genus *Carmovirus*: evidence for the presence and expression of two novel open reading frames. *J. Virol.* **74**: 3149-3155.
- Hughes AL.** 1992. Coevolution of the vertebrate integrin alpha- and beta-chain genes. *Mol. Biol. Evol.* **9**: 216-234.
- Hughes AL, Packer B, Welch R, Bergen AW, Chanock SJ, Yeager M.** 2003. Widespread purifying selection at polymorphic sites in human protein-coding loci. *Proc. Natl. Acad. Sci. USA* **100**: 15754-15757.
- Hunter T, Jackson R, Zimmern D.** 1983. Multiple proteins and subgenomic mRNAs may be derived from a single open reading frame on *Tobacco mosaic virus* RNA. *Nucleic Acids Res.* **11**: 801-821.
- Iacono M, Mignone F, Pesole G.** 2005. uAUG and uORFs in human and rodent 5' untranslated mRNAs. *Gene* **349**: 97-105.
- Ivanov PA, Karpova OV, Skulachev MV, Tomashevskaya OL, Rodionova NP, Dorokhov YuL, Atabekov JG.** 1997. A tobamovirus genome that contains an internal ribosome entry site functional *in vitro*. *Virology* **26**: 32-43.
- Iwakawa HO, Kaido M, Mise K, Okuno T.** 2007. *Cis-acting* core RNA elements required for negative-strand RNA synthesis and cap-independent translation are separated in the 3'-untranslated region of *Red clover necrotic mosaic virus* RNA1. *Virology* **369**: 168-181.
- Jaag HM, Kawchuk L, Rohde W, Fischer R, Emans N, Pruffer D.** 2003. An unusual internal ribosomal entry site of inverted symmetry directs expression of a *Potato leafroll polerovirus* replication-associated protein. *Proc. Natl. Acad. Sci. USA* **100**: 8939-8944.

- Jacobson A.** 1996. Poly(A) metabolism and translation: the closed-loop model. In Hershey JWB, Mathews MB, Sonenberg N (eds). *Translational control*, pp. 451–480. Plainview, N.Y. Cold Spring Harbor Laboratory Press.
- Jang SK.** 2006. Internal initiation: IRES elements of picornaviruses and *Hepatitis C virus*. *Virus Res.* **119**: 2-15.
- Jang SK, Krausslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E.** 1988. A segment of the 5' nontranslated region of *Encephalomyocarditis virus* RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* **62**: 2636–2643.
- Janzen DH.** 1980. When is it coevolution? *Evolution* **34**: 611-612.
- Johnston JC, Rochon DM.** 1995. Deletion analysis of the promoter for the *Cucumber necrosis virus* 0.9-kb subgenomic RNA. *Virology* **214**: 100-109.
- Jones AT, McGavin WJ, Brunt AA, Phillips S.** 2000. Elderberry latent virus: its relationship to Pelargonium ringspot virus and its identification as a distinct member of the genus *Carmovirus*, family *Tombusviridae*. *Ann. Appl. Biol.* **136**: 147-152.
- Joshi S, Pleij CW, Haenni AL, Chapeville F, Bosch L.** 1983. Properties of the *Tobacco mosaic virus* intermediate length RNA-2 and its translation. *Virology* **127**: 100–111.
- Joshi CP, Zhou H, Huang X, Chiang VL.** 1997. Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* **35**: 993-1001.
- Karasev AV, Boyko VP, Gowda S, Nikolaeva OV, Hilf ME, Koonin EV, Niblett CL, Cline K, Gumpf DJ, Lee RF, et al.** 1995. Complete sequence of the *Citrus tristeza virus* RNA genome. *Virology* **208**: 511–520.
- Karasev AV, Kashina AS, Gelfand VI, Dolja VV.** 1992. HSP70-Related 65 kDa protein of *Beet yellows closterovirus* is a microtubule-binding protein. *FEBS Lett.* **304**: 12-14.
- Kelly L, Gerlach WL, Waterhouse PM.** 1994. Characterization of the subgenomic RNAs of an Australian isolate of *Barley yellow dwarf luteovirus*. *Virology* **202**: 565-573.
- Khan MA, Miyoshi H, Ray S, Natsuaki T, Suehiro N, Goss DJ.** 2006. Interaction of genome-linked protein (VPg) of *Turnip mosaic virus* with wheat germ translation initiation factors eIFiso4E and eIFiso4F. *J. Biol. Chem.* **281**: 28002-28010.

- Khan MA, Yumak H, Gallie DR, Goss DJ.** 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.
- Kim JW, Bozarth RF.** 1992. Mapping and sequence analysis of the capsid protein gene of *Cowpea mottle virus*. *Interviol.* **33**: 135-47.
- Kim SH, Kalinina NO, Andreev I, Ryabov EV, Fitzgerald AG, Taliensky ME, Palukaitis P.** 2004. The C-terminal 33 amino acids of the *Cucumber mosaic virus* 3a protein affect virus movement, RNA binding and inhibition of infection and translation. *J. Gen. Virol.* **85**: 221-230.
- Kim KH, Lommel SA.** 1994. Identification and analysis of the site of -1 frameshifting in *Red clover necrotic mosaic virus*. *Virology* **200**: 574-582.
- Kimura M.** 1985. The Neutral Theory of Molecur Evolution. *New Sci.* **11**: 41-46.
- Kinard G, Guaragna MA, Jordan R.** 1999. Characterization and genome organization of four unique viruses infecting geranium and their assignment to a new provisional genus in the *Tombusviridae* family. Libro de resúmenes del V Congreso Internacional de Virología, p. 415.
- Kinard GR, Jordan RL.** 2002. Genome organization of Pelargonium chlorotic ring pattern virus: further implications for *Tombusviridae* taxonomy. *Acta Hortic.* **568**: 17-27.
- Kinard GR, Jordan RL, Hurtt SS.** 1996. Partial characterization of Pelargonium line pattern and Pelargonium ringspot viruses. *Acta Hortic.* **432**: 148-155.
- Kneller EL, Rakotondrafara AM, Miller WA.** 2006. Cap-independent translation of plant viral RNAs. *Virus Res.* **119**: 63-75.
- Koev G, Miller WA.** 2000. A positive-strand RNA virus with very different subgenomic RNA promoters. *J. Virol.* **74**: 5988-5996.
- Koev G, Mohan BR, Miller WA.** 1999. Primary and secondary structural elements required for synthesis of *Barley yellow dwarf virus* subgenomic RNA1. *J. Virol.* **73**: 2876-2885.
- Koh DCY, Liu DX, Wong SM.** 2002. A six-nucleotide segment within the 3' untranslated region of *Hibiscus chlorotic ringspot virus* plays an essential role in translational enhancement. *J. Virol.* **76**: 1144-1153.
- Koh DC, Wang X, Wong SM, Liu DX.** 2006. Translation initiation at an upstream CUG codon regulates the expression of *Hibiscus chlorotic ringspot virus* coat protein. *Virus Res.* **122**: 35-44.
- Koh DC, Wong SM, Liu DX.** 2003. Synergism of the 3'-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein. *J. Biol. Chem.* **278**: 20565-20573.

- Koonin EV, Dolja VV.** 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* **72**: 2197-2206.
- Koonin EV, Dolja VV.** 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28**: 375-430.
- Korber BTM, Farber RM, Wolpert DH, Lapedes AS.** 1993. Covariation of mutations in the V3 loop of *Human immunodeficiency virus* type 1 envelope protein: An information theoretic analysis. *Proc. Natl. Acad. Sci. USA* **90**: 7176-7180.
- Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**: 283-292.
- Kozak, M.** 1989. Context effects and inefficient initiation at non-AUG codons in eukaryotic cell free translation systems. *Mol. Cell Biol.* **9**: 5134-5142.
- Kozak M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**: 19867-19870.
- Kozak, M.** 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**: 1-34.
- Krab IM, Caldwell C, Gallie DR, Bol JF.** 2005. Coat protein enhances translational efficiency of *Alfalfa mosaic virus* RNAs and interacts with the eIF4G component of initiation factor eIF4F. *J. Gen. Virol.* **86**: 1841-1849.
- Krczal G, Albouy J, Damy I, Kusiak C, Deogratias J, Moreau J, Berkelman B, Wohanka W.** 1995. Transmission of *Pelargonium flower break virus* by irrigation systems and by thrips. *Plant Dis.* **79**: 163-166.
- Laakso MM, Heaton LA.** 1993. Asp-Asn substitutions in the putative calcium-binding site of the *Turnip crinkle virus* coat protein affect virus movement in plants. *Virology* **197**: 774-777.
- Lai MMC.** 1992. RNA recombination in animal and plant viruses. *Microbiol. Rev.* **56**: 61-79.
- Larson SM, Di Nardo AA, Davidson AR.** 2000. Analysis of covariation in an SH3 domain sequence alignment: applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions. *J. Mol. Biol.* **303**: 433-446.
- Latorre P, Kolakofsky D, Curran J.** 1998. *Sendai virus* Y proteins are initiated by a ribosomal shunt. *Mol. Cell Biol.* **18**: 5021-5031.
- Leathers V, Tanguay R, Kobayashi M, Gallie DR.** 1993. A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. *Mol. Cell Biol.* **13**: 5331-5347.

- Lee KC, Lim D, Wong SM, Dokland T.** 2003. Purification, crystallization and X-ray analysis of *Hibiscus chlorotic ringspot virus*. *Acta Crystallogr. D. Biol. Crystallogr.* **59**: 1481-1483.
- Legg J, Tresh J.** 2000. *Cassava mosaic virus* disease in East Africa: a dynamic disease in a changing environment. *Virus Res.* **71**: 135-149.
- Levis C, Astier-Manifacier S.** 1993. The 5' untranslated region of PVY RNA, even located in an internal position, enables initiation of translation. *Virus Genes* **7**: 367-379.
- Lewis SM, Holcik M.** 2008. For IRES *trans*-acting factors, it is all about location. *Oncogene* **27**: 1033-1035.
- Li WZ, Qu F, Morris TJ.** 1998. Cell-to-cell movement of *Turnip crinkle virus* is controlled by two small open reading frames that function in trans. *Virology* **244**: 405-416.
- Li W, Wong SM.** 2006. Analyses of subgenomic promoters of *Hibiscus chlorotic ringspot virus* and demonstration of 5'-untranslated region and 3'-terminal sequences functioning as subgenomic promoters. *J. Virol.* **80**: 3395-3405.
- Liang XZ, Lee BTK, Wong SM.** 2002a. Covariation in the capsid protein of *Hibiscus chlorotic ringspot virus* induced by serial passing in a host that restricts movement leads to avirulence in its systemic host. *J. Virol.* **76**: 12320-12324.
- Liang XZ, Lucy AP, Ding SW, Wong SM.** 2002b. The p23 protein of *Hibiscus chlorotic ringspot virus* is indispensable for host-specific replication. *J. Virol.* **76**: 12312-12319.
- Lin B, Heaton LA.** 1999. Mutational analyses of the putative calcium binding site and hinge of the *Turnip crinkle virus* coat protein. *Virology* **259**: 34-42.
- Lin HX, White KA.** 2004. A complex network of RNA-RNA interactions control subgenomic mRNA transcription in a tombusvirus. *EMBO J.* **18**: 3365-3374.
- Lin HX, Xu W, White KA.** 2007. A multicomponent RNA-based control system regulates subgenomic mRNA transcription in a tombusvirus. *J. Virol.* **81**: 2429-2439.
- Liu H, Boulton MI, Oparka KJ, Davies JW.** 2001. Interaction of the movement and coat proteins of *Maize streak virus*: implications for the transport of viral DNA. *J. Gen. Virol.* **82**: 35-44.
- Liu L, Grainger J, Cañizares MC, Angell SM, Lomonossoff GP.** 2004. *Cowpea mosaic virus* RNA-1 acts as an amplicon whose effects can be counteracted by a RNA-2-encoded suppressor of silencing. *Virology* **323**: 37-48.

- Liu S, He X, Park G, Josefsson C, Perry KL.** 2002. A conserved capsid protein surface domain of *Cucumber mosaic virus* is essential for efficient aphid vector transmission. *J. Virol.* **76**: 9756-9762.
- Lommel SA, Martelli GP, Rubino L, Russo M.** 2005. Family *Tombusviridae*. In Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA. (eds.). *Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 907-936. Academic Press, San Diego, USA.
- Lu R, Folimonov A, Shintaku M, Li WX, Falk BW, Dawson WO, Ding SW.** 2004. Three distinct suppressors of RNA silencing encoded by a 20-kb viral RNA genome. *Proc. Natl. Acad. Sci. USA* **101**: 15742-15747.
- Lu PD, Harding HP, Ron D.** 2004. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J. Cell Biol.* **167**: 27-33.
- Lukaszewicz M, Feuermann M, Jerouville B, Stas A, Boutry M.** 2000. *In vivo* evaluation of the context sequence of the translation initiation codon in plants. *Plant Sci.* **154**: 89-98.
- Luukkonen BGM, Tan W, Schwartz S.** 1995. Efficiency or reinitiation of translation on *Human immunodeficiency virus* type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J. Virol.* **69**: 4086-4094.
- Mahajan S, Dolja VV, Carrington JC.** 1996. Roles of the sequence encoding *Tobacco etch virus* capsid protein in genome amplification: requirements for the translation process and a *cis*-active element. *J. Virol.* **70**: 4370-4379.
- Makinen K, Naess V, Tamm T, Truve E, Aaspollu A, Saarma M.** 1995. The putative replicase of *Cocksfoot mottle sobemovirus* is translated as a part of the polyprotein by -1 ribosomal frameshift. *Virology* **207**: 566-571.
- Makino DL, Day J, Larson SB, McPherson A.** 2006. Investigation of RNA structure in *satellite Panicum mosaic virus*. *Virology* **351**: 420-431.
- Malik PS, Kumar V, Bagewadi B, Mukherjee SK.** 2005. Interaction between coat protein and replication initiation protein of *Mung bean yellow mosaic India virus* might lead to control of viral DNA replication. *Virology* **337**: 273-283.
- Malpica JM, Fraile A, Moreno I, Obies C, Drake JW, García-Arenal F.** 2002. The rate and character of spontaneous mutation in an RNA virus. *Genetics* **162**: 1505-1511.
- Marco CF, Aranda MA.** 2005. Genetic diversity of a natural population of *Cucurbit yellow stunting disorder virus*. *J. Gen. Virol.* **86**: 815-822.

- Marcos JF, Vilar M, Pérez-Payá E, Pallás V.** 1999. *In vivo* detection, RNA-binding properties and characterization of the RNA-binding domain of the p7 putative movement protein from *Carnation mottle carmovirus* (CarMV). *Virology* **255**: 354-365.
- Mardanova ES, Zamchuk LA, Skulachev MV, Ravin NV.** 2008. The 5' untranslated region of the maize alcohol dehydrogenase gene contains an internal ribosome entry site. *Gene* **420**: 11-16.
- Marsh LE, Pogue GP, Hall TC.** 1989. Similarities among plant virus (+) and (-) RNA termini imply a common ancestry with promoters of eucaryotic tRNAs. *Virology* **172**: 415-427.
- Martelli GP.** 1994. Epidemiology and management of plant virus diseases. In Maden M, Raccach B, Tresh J (eds.). Springer-Verlag, Berlin.
- Martínez-Gil L, Saurí A, Vilar M, Pallás V, Mingarro I.** 2007. Membrane insertion and topology of the p7B movement protein of *Melon necrotic spot virus* (MNSV). *Virology* **367**: 348-357.
- Martínez-Salas E.** 1999. Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.* **10**: 458-464.
- Martínez-Salas E, Pacheco A, Serrano P, Fernández N.** 2008. New insights into internal ribosome entry site elements relevant for viral gene expression. *J. Gen. Virol.* **89**: 611-626.
- Martínez-Turiño S, Hernández C.** 2009. Inhibition of RNA silencing by the coat protein of *Pelargonium flower break virus*: distinctions from closely related suppressors. *J. Gen. Virol.* **90**: 519-525.
- Matsuda D, Dreher TW.** 2004. The tRNA-like structure of *Turnip yellow mosaic virus* RNA is a 3'-translational enhancer. *Virology* **321**: 36-46.
- Mauro VP, Chappell SA, Dresios J.** 2007. Analysis of ribosomal shunting during translation initiation in eukaryotic mRNAs. *Methods Enzymol.* **429**: 323-354.
- May RM, Nowak MA.** 1995. Coinfection and the evolution of parasite virulence. *Proc. R. Soc. London Ser. B.* **261**: 209-215.
- McBurney SP, Ross TM.** 2008. Viral sequence diversity: challenges for AIDS vaccine designs. *Expert Rev. Vaccines* **7**: 1405-1417.
- McLean BG, Zupan J, Zambryski PC.** 1995. *Tobacco mosaic virus* movement protein associates with the cytoskeleton in tobacco cells. *Plant Cell* **7**: 2101-2114.
- Meijer HA, Thomas AA.** 2002. Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem. J.* **367**: 1-11.

- Meng C, Chen J, Peng J, Wong SM.** 2006. Host-induced avirulence of *Hibiscus chlorotic ringspot virus* mutant correlates with reduced gene-silencing suppression activity. *J. Gen. Virol.* **87**: 451-459.
- Merrick WC.** 2004. Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332**: 1-11.
- Meulewaeter F, van Lipzig R, Gultyaev AP, Pleij CW, Van Damme D, Cornelissen M, van Eldik G.** 2004. Conservation of RNA structures enables TNV and BYDV 5' and 3' elements to cooperate synergistically in cap-independent translation. *Nucleic Acids Res.* **32**: 1721-1730.
- Miller EA, Dinesh-Kumar SP, Paul CP.** 1995. *Luteovirus* gene expression. *Crit. Rev. Plant Sci.* **14**: 179-211.
- Miller WA, Koev G.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. *Virology* **273**: 1-8.
- Miranda GJ, Azzam O, Shirako Y.** 2000. Comparison of nucleotide sequences between northern and southern Philippine isolates of *Rice grassy stunt virus* indicates occurrence of natural genetic reassortment. *Virology* **266**: 26-32.
- Mizumoto H, Iwakawa HO, Kaido M, Mise K, Okuno T.** 2006. Cap-independent translation mechanism of *Red clover necrotic mosaic virus* RNA2 differs from that of RNA1 and is linked to RNA replication. *J. Virol.* **80**: 3781-3791.
- Mizumoto H, Tatsuta M, Kaido M, Mise K, Okuno T.** 2003. Cap-independent translational enhancement by the 3' untranslated region of *Red clover necrotic mosaic virus* RNA1. *J. Virol.* **77**: 12113-12121.
- Mohan BR, Dinesh-Kumar SP, Miller WA.** 1995. Genes and *cis*-acting sequences involved in replication of *Barley yellow dwarf virus*-PAV RNA. *Virology* **212**: 186-195.
- Molnar A, Havelda Z, Dalmay T, Szutorisz H, Burgyan J.** 1997. Complete nucleotide sequence of *Tobacco necrosis virus* strain DH and genes required for RNA replication and virus movement. *J. Gen. Virol.* **78**: 1235-1239.
- Monci F, Sánchez-Campos S, Navas-Castillo J, Moriones E.** 2002. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and *Tomato yellow leaf curl virus* exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Virology* **303**: 317-326.
- Monsion B, Froissart R, Michalakakis Y, Blanc S.** 2008. Large bottleneck size in *Cauliflower mosaic virus* populations during host plant colonization. *PLoS Pathog.* **4**: e1000174.

- Morris TJ.** 1991. Classification and Nomenclature of Viruses. In Francki R, Fauquet C, Knudson D, Brown F (eds.). *Fifth Report of the International Committee on Taxonomy of Viruses*. *Arch. Virol. Suppl.* 2, pp. 303-305. Springer-Verlag, Vienna.
- Morris DR, Geballe AP.** 2000. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **20**: 8635-8642.
- Na H, White KA.** 2006. Structure and prevalence of replication silencer-3' terminus RNA interactions in *Tombusviridae*. *Virology* **345**: 305-316.
- Nagy PD, Bujarski JJ.** 1993. Targeting the site of RNA-RNA recombination in *Brome mosaic virus* with antisense sequences. *Proc. Natl. Acad. Sci. USA* **90**: 6390-6394.
- Nagy PD, Ogiela C, Bujarski JJ.** 1999a. Mapping sequences active in homologous RNA recombination in *Brome mosaic virus*: prediction of recombination hot spots. *Virology* **254**: 92-104.
- Nagy PD, Pogany J, Simon AE.** 1999b. RNA elements required for RNA recombination function as replication enhancers *in vitro* and *in vivo* in a plus-strand RNA virus. *EMBO J.* **18**: 5653-5665.
- Nagy PD, Pogany J, Simon AE.** 2001. *In vivo* and *in vitro* characterization of an RNA replication enhancer in a satellite RNA associated with *Turnip crinkle virus*. *Virology* **288**: 315-324.
- Nameth ST, Adkins ST.** 1993. Viral diseases. In White JW (eds.). *Geraniums IV-The Grower's Manual*, pp. 267-275. Ball Publishing, Geneva, USA
- Namy O, Hatin I, Rousset JP.** 2001. Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep.* **2**: 787-793.
- Navarro JA, Genovés A, Climent J, Saurí A, Martínez-Gil L, Mingarro I, Pallás V.** 2006. RNA-binding properties and membrane insertion of *Melon necrotic spot virus* (MNSV) double gene block movement proteins. *Virology* **356**: 57-67.
- Navarro B, Rubino L, Russo M.** 2004. Expression of the *Cymbidium ringspot virus* 33 kDa protein in *Saccharomyces cerevisiae* and molecular dissection of the peroxisomal targeting signal. *J. Virol.* **78**: 4744-4752.
- Navas-Castillo J, Sánchez-Campos S, Noris E, Louro D, Accotto GP, Moriones E.** 2000. Natural recombination between *Tomato yellow leaf curl virus*-I_s and *Tomato leaf curl virus*. *J. Gen. Virol.* **81**: 2797-2801.
- Neeleman L, Linthorst HJ, Bol JF.** 2004. Efficient translation of alfamovirus RNAs requires the binding of coat protein dimers to the 3' termini of the viral RNAs. *J. Gen. Virol.* **85**: 231-240.

- Neeleman L, Olsthoorn RC, Linthorst HJ, Bol JF.** 2001. Translation of a nonpolyadenylated viral RNA is enhanced by binding of viral coat protein or polyadenylation of the RNA. *Proc. Natl. Acad. Sci. USA* **98**: 14286–14291.
- Niepel M, Gallie DR.** 1999. Identification and characterization of the functional elements within the *Tobacco etch virus* 5' leader required for cap-independent translation. *J. Virol.* **73**: 9080–9088.
- Noueiry AO, Lucas WJ, Gilbertson RL.** 1994. Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell* **76**: 925-932.
- Olsthoorn RC, Bruyere A, Dzianott A, Bujarsji JJ.** 2002. RNA recombination in *Brome mosaic virus*: effects of strand-specific stem-loop insert. *J. Virol.* **76**: 12654-12662.
- Omarov RT, Qi D, Scholthof KB.** 2005. The capsid protein of *satellite Panicum mosaic virus* contributes to systemic invasion and interacts with its helper virus. *J. Virol.* **79**: 9756-9764.
- Oster SK, Wu B, White KA.** 1998. Uncoupled expression of p33 and p92 permits amplification of *Tomato bushy stunt virus* RNAs. *J. Virol.* **72**: 5845–5851.
- Panavas T, Hawkins CM, Panaviene Z, Nagy PD.** 2005. The role of the p33: p33/p92 interaction domain in RNA replication and intracellular localization of p33 and p92 proteins of *Cucumber necrosis virus*. *Virology* **338**: 81-95.
- Panavas T, Nagy PD.** 2005. Mechanism of stimulation of plus-strand synthesis by an RNA replication enhancer in a tombusvirus. *J. Virol.* **79**: 9777-9785.
- Panavas T, Pogany J, Nagy PD.** 2002. Analysis of minimal promoter sequences for plus-strand synthesis by the *Cucumber necrosis virus* RNA-dependent RNA polymerase. *Virology* **296**: 263-274.
- Park HS, Himmelbach A, Browning KS, Hohn T, Ryabova LA.** 2001. A plant viral “reinitiation” factor interacts with the host translational machinery. *Cell* **106**: 723-733.
- Pazos F, Valencia A.** 2008. Protein co-evolution, co-adaptation and interactions. *EMBO J.* **27**: 2648-2655.
- Peabody DS.** 1989. Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* **264**: 5031–5035.
- Pelczyk M, Obrepalska-Stepłowska A, Pospieszny H.** 2006. Subviral molecules of RNA associated with plant ss(+)RNA viruses. *Postepy Biochem.* **52**: 212-221.
- Pelletier J, Sonenberg N.** 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**: 320–325.

- Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI, Hellen CU.** 2001. Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. USA* **98**: 7029–7036.
- Pirone TD, Blanc S.** 1996. Helper-dependent vector transmission of plant viruses. *Annu. Rev. Phytopathol.* **34**: 227-247.
- Pita JS, de Miranda JR, Schneider WL, Roossinck MJ.** 2007. Environment determines fidelity for an RNA virus replicase. *J. Virol.* **81**: 9072-9077.
- Pogany J, Fabian MR, White KA, Nagy PD.** 2003. A replication silencer element in a plus-strand RNA virus. *EMBO J.* **22**: 5602-5611.
- Pollock DD, Taylor WR, Goldman N.** 1999. Coevolving protein residues: maximum likelihood identification and relationship to structure. *J. Mol. Biol.* **287**: 187-198.
- Pooggin MM, Fütterer J, Skryabin KG, Hohn T.** 1999. A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. *J. Gen. Virol.* **80**: 2217-2228.
- Pooggin MM, Hohn T, Fütterer J.** 2000. Role of a short open reading frame in ribosome shunt on the *Cauliflower mosaic virus* RNA leader. *J. Biol. Chem.* **275**: 17288–17296.
- Pooggin MM, Ryabova LA, He X, Fütterer J, Hohn T.** 2006. Mechanism of ribosome shunting in *Rice tungro bacilliform pararetrovirus*. *RNA* **12**: 841-850.
- Power AG.** 2000. Insect transmission of plant viruses: a constraint on virus variability. *Curr. Opin. Plant Biol.* **3**: 336-340.
- Prlijc J, Veljkovic N, Veljkovic V.** 2004. Recombination property of the HIV-1 gp 120 gene. *Int. Rev. Immunol.* **23**: 447-454.
- Qi D, Omarov RT, Scholthof KB.** 2008. The complex subcellular distribution of satellite *Panicum mosaic virus* capsid protein reflects its multifunctional role during infection. *Virology* **376**: 154-164.
- Qi D, Scholthof KB.** 2008. Multiple activities associated with the capsid protein of satellite *Panicum mosaic virus* are controlled separately by the N- and C-terminal regions. *Mol. Plant Microbe Interact.* **21**: 613-621.
- Qiu WP, Moyer W.** 1999. *Tomato spotted wilt tospovirus* adapts to the TSWV N gene-derived resistance by genome reassortment. *Phytopathology* **89**: 575-582.
- Qiu W, Scholthof HB.** 2001. Effects of inactivation of the coat protein and movement genes of *Tomato bushy stunt virus* on early accumulation of genomic and subgenomic RNAs. *J. Gen. Virol.* **82**: 3107-3114.

- Qu F, Morris TJ.** 2000. Cap-independent translational enhancement of *Turnip crinkle virus* genomic and subgenomic RNAs. *J. Virol.* **74**: 1085–1093.
- Qu F, Ren T, Morris TJ.** 2003. The coat protein of *Turnip crinkle virus* suppresses posttranscriptional gene silencing at an early initiation step. *J. Virol.* **77**: 511-522.
- Rajendran KS, Nagy PD.** 2003. Characterization of the RNA-binding domains in the replicase proteins of *Tomato bushy stunt virus*. *J. Virol.* **77**: 9244-9258.
- Rajendran KS, Nagy PD.** 2004. Interaction between the replicase proteins of *Tomato bushy stunt virus* *in vitro* and *in vivo*. *Virology* **326**: 250-261.
- Rajendran KS, Nagy PD.** 2006. Kinetics and functional studies on interaction between the replicase proteins of *Tomato bushy stunt virus*: requirement of p33:p92 interaction for replicase assembly. *Virology* **345**: 270-279.
- Rao AL, Cooper B.** 2006. Capsid protein gene and the type of host plant differentially modulate cell-to-cell movement of *Cowpea chlorotic mottle virus*. *Virus Genes* **32**: 219-227.
- Rao AL, Grantham GL.** 1996. Molecular studies on bromovirus capsid protein. II. Functional analysis of the amino-terminal arginine-rich motif and its role in encapsidation, movement, and pathology. *Virology* **226**: 294-305.
- Ray D, White KA.** 1999. Enhancer-like properties of an RNA element that modulates *Tombusvirus* RNA accumulation. *Virology* **256**: 162-171.
- Ray S, Yumak H, Domashevskiy A, Khan MA, Gallie DR, Goss DJ.** 2006. *Tobacco etch virus* mRNA preferentially binds wheat germ eukaryotic initiation factor (eIF) 4G rather than eIFiso4G. *J. Biol. Chem.* **281**: 35826-35834.
- Reade R, Miller J, Robbins M, Xiang Y, Rochon D.** 2003. Molecular analysis of the *Cucumber leaf spot virus* genome. *Virus Res.* **91**: 171-179.
- Reavy B, Arif M, Cowan GH, Torrance L.** 1998. Association of sequences in the coat protein/readthrough domain of *Potato mop-top virus* with transmission by *Spongospora subterranea*. *J. Gen. Virol.* **79**: 2343–2347.
- Rebenstorf K, Candresse T, Dulucq MJ, Buttner C, Obermeier C.** 2006. Host species-dependent population structure of a pollen-borne plant virus, *Cherry leaf roll virus*. *J. Virol.* **80**: 2453-2462.

- Revers F, Le Gall O, Candresse T, Le Romancer M, Dunez J.** 1996. Frequent occurrence of recombinant potyvirus isolates. *J. Gen. Virol.* **77**: 1953-1965.
- Rico P, Hernández C.** 2004. Complete nucleotide sequence and genome organization of *Pelargonium flower break virus*. *Arch. Virol.* **149**: 641-651.
- Rico P, Hernández C.** 2006. Infectivity of *in vitro* transcripts from a full-length cDNA clone of *Pelargonium flower break virus* in an experimental and a natural host. *J. Plant Pathol.* **88**: 103-106.
- Rico P, Ivars P, Elena SF, Hernández C.** 2006. Insights into the selective pressures restricting *Pelargonium flower break virus* genome variability: evidence for host adaptation. *J. Virol.* **80**: 8124-8132.
- Riechmann JL, Lain S, Garcia JA.** 1991. Identification of the initiation codon of *Plum pox potyvirus* genomic RNA. *Virology* **185**: 544-552.
- Riis B, Rattan SIS, Clark BFC, Merrick WC.** 1990. Eukaryotic protein elongation factors. *Trends Biochem. Sci.* **15**: 420-424.
- Riviere CJ, Rochon DM.** 1990. Nucleotide sequence and genomic organization of *Melon necrotic spot virus*. *J. Gen. Virol.* **71**: 1887-1896.
- Roberts IM, Boevink P, Roberts AG, Sauer N, Reichel C, Oparka KJ.** 2001. Dynamic changes in the frequency and architecture of plasmodesmata during the sink-source transition in tobacco leaves. *Protoplasma* **218**: 31-44.
- Robertson NL, Coté F, Paré C, Leblanc E, Bergeron MG, Leclerc D.** 2007. Complete nucleotide sequence of Nootka lupine vein-clearing virus. *Virus Genes* **35**: 807-814.
- Rochon DM, Johnston JC.** 1991. Infectious transcripts from cloned *Cucumber necrosis virus* cDNA. Evidence for a bifunctional subgenomic RNA. *Virology* **181**: 656-665.
- Rodríguez-Cerezo E, Moya A, García-Arenal F.** 1989. Variability and evolution of the plant RNA virus *Pepper mild mottle virus*. *J. Virol.* **63**: 2198-2203.
- Rogers GW Jr, Edelman GM, Mauro VP.** 2004. Differential utilization of upstream AUGs in the beta-secretase mRNA suggests that a shunting mechanism regulates translation. *Proc. Natl. Acad. Sci. USA* **101**: 2794-2799.
- Roossinck MJ.** 1997. Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* **35**: 191-209.
- Rothnie HM, Chapdelaine Y, Hohn T.** 1994. Pararetroviruses and retroviruses: a comparative review of viral structure and expression strategies. *Adv. Virus Res.* **44**: 1-67.

- Rubino L, Russo M.** 1997. Molecular analysis of the *Phothos latent virus* genome. *J. Gen. Virol.* **78**: 1219-1226.
- Rubino L, Russo M, Martelli GP.** 1995. Sequence analysis of *Pothos latent virus* genomic RNA. *J. Gen. Virol.* **76**: 2835-2839.
- Rubio L, Abou-Jawdah Y, Lin HX, Falk BW.** 2001. Geographically distant isolates of the *Crinivirus cucurbit yellow stunting disorder virus* show very low genetic diversity in the coat protein gene. *J. Gen. Virol.* **82**: 929-933.
- Rubio L, Soong J, Kao J, Falk W.** 1999. Geographic distribution and molecular variation of isolates of three whitefly-borne closteroviruses of cucurbits: Lettuce infectious yellows virus, *Cucurbit yellow stunting disorder virus* and *Beet pseudo-yellows virus*. *Phytopathology* **89**: 707-711.
- Russo M, Burgyan J, Martelli G.** 1994. Molecular biology of *Tombusviridae*. *Adv. Virus Res.* **44**: 381-428.
- Ryabov EV, Generozov EV, Kendall TL, Lommel SA, Zavriev SK.** 1994. Nucleotide sequence of *Carnation ringspot dianthovirus* RNA-1. *J. Gen. Virol.* **75**: 243-247.
- Ryabov EV, Robinson DJ, Taliensky ME.** 1999. A plant virus-encoded protein facilitates long-distance movement of heterologous viral RNA. *Proc. Natl. Acad. Sci. USA* **96**: 1212-1217.
- Ryabova LA, Pooggin MM, Dominguez DI, Hohn T.** 2000. Continuous and discontinuous ribosome scanning on the *Cauliflower mosaic virus* 35 S RNA leader is controlled by short open reading frames. *J. Biol. Chem.* **275**: 37278-37284.
- Ryabova LA, Pooggin MM, Hohn T.** 2006. Translation reinitiation and leaky scanning in plant viruses. *Virus Res.* **119**: 52-62.
- Sachs AB, Sarnow P, Hentze MW.** 1997. Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89**: 831-838.
- Sacristán S, Fraile A, Malpica JM, García-Arenal F.** 2005. An analysis of host adaptation and its relationship with virulence in *Cucumber necrosis virus*. *Phytopathology* **95**: 827-833.
- Sacristán S, Malpica JM, Fraile A, García-Arenal F.** 2003. Estimation of population bottlenecks during systemic movement of *Tobacco mosaic virus* in tobacco plants. *J. Virol.* **77**: 9906-9911.
- Saitou N, Nei M.** 1986. Polymorphism and evolution of *Influenza A virus* genes. *Mol. Biol. Evol.* **3**: 57-74.
- Sánchez-Navarro JA, Bol JF.** 2001. Role of the *Alfalfa mosaic virus* movement protein and coat protein in virus transport. *Mol. Plant Microbe Interact.* **14**: 1051-1062.

- Sánchez-Navarro JA, Herranz MC, Pallás, V.** 2006. Cell-to-cell movement of *Alfalfa mosaic virus* can be mediated by the movement proteins of Ilar-, bromo-, cucumo-, tobamo- and comoviruses and does not require virion formation. *Virology* **346**: 66-73.
- Sareila O, Hohkuri M, Wahlroos T, Susi P.** 2004. Role of viral movement and coat proteins and RNA in phloem-dependent movement and phloem unloading of tobamoviruses. *J. Phytopathol.* **152**: 622-629.
- Sarrazin S, Starck J, Gonnet C, Doubeikovski A, Melet F, Morle F.** 2000. Negative and translation termination-dependent positive control of Fli-1 protein synthesis by conserved overlapping 5' upstream open reading frames in Fli-1 mRNA. *Mol. Cell. Biol.* **20**: 2959-2969.
- Sato M, Nakahara K, Yoshii M, Ishikawa M, Uyeda I.** 2005. Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of *Arabidopsis thaliana* by potyviruses. *FEBS Lett.* **579**: 1167-1171.
- Sauri A, Saksena S, Salgado J, Jhonson AE, Mingarro I.** 2005. Double-spanning plant viral movement protein integration into the endoplasmic reticulum membrane is signal recognition particle-dependent, translocon-mediated, and concerted. *J. Biol. Chem.* **280**: 25907-25912.
- Schmitt C, Balmori E, Jonard G, Richards KE, Guilley H.** 1992. *In vitro* mutagenesis of biologically active transcripts of *Beet necrotic yellow vein virus* RNA 2: evidence that a domain of the 75-kDa readthrough protein is important for efficient virus assembly. *Proc. Natl. Acad. Sci. USA* **89**: 5715-5719.
- Schneider WL, Roossinck MJ.** 2000. Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. *J. Virol.* **74**: 3130-3134.
- Schneider WL, Roossinck MJ.** 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J. Virol.* **75**: 6566-6571.
- Scholthof KB, Scholthof HB, Jackson AO.** 1995. The *Tomato bushy stunt virus* replicase proteins are coordinately expressed and membrane associated. *Virology* **208**: 365-369.
- Schultze M, Hohn T, Jiricny J.** 1990. The reverse transcriptase gene of *Cauliflower mosaic virus* is translated separately from the capsid gene. *EMBO J.* **9**: 1177-1185.
- Seibert SA, Howell CY, Hughes MK, Hughes AL.** 1995. Natural selection on the *gag*, *pol*, and *env* genes of *Human immunodeficiency virus 1* (HIV-1). *Mol. Biol. Evol.* **12**: 803-813.
- Sen N, Cao F, Tavis JE.** 2004. Translation of duck *Hepatitis B virus* reverse transcriptase by ribosomal shunting. *J. Virol.* **78**: 11751-11757.

- Shapka N, Stork J, Nagy PD.** 2005. Phosphorylation of the p33 replication protein of *Cucumber necrosis tobravirus* adjacent to the RNA binding site affects viral RNA replication. *Virology* **343**: 65-78.
- Shen R, Miller WA.** 2004. The 3' untranslated region of *Tobacco necrosis virus* RNA contains a *Barley yellow dwarf virus*-like cap-independent translation element. *J. Virol.* **78**: 100-111.
- Shen LX, Tinoco I.** 1995. The structure of an RNA pseudoknot that causes efficient frameshifting in *Mouse mammary tumor virus*. *J. Mol. Biol.* **247**: 963-978.
- Shirako Y.** 1998. Non-AUG translation initiation of a plant RNA virus: a forty-amino acid extension is added to the N terminus of the soil-borne *Wheat mosaic virus* capsid protein. *J. Virol.* **72**: 1677-1682.
- Shivaprasad PV, Akbergenov R, Trinks D, Rajeswaran R, Veluthambi K, Hohn T, Pooggin MM.** 2005. Promoters, transcripts, and regulatory proteins of *Mungbean yellow mosaic geminivirus*. *J. Virol.* **79**: 8149-8163.
- Simon AE, Bujarski JJ.** 1994. RNA-RNA recombination and evolution in virus-infected plant. *Ann. Rev. Phytopathol.* **32**: 337-362.
- Simón-Buela L, Guo HS, García JA.** 1997. Cap-independent leaky scanning as the mechanism of translation initiation of a plant viral genomic RNA. *J. Gen. Virol.* **78**: 2691-2699.
- Simon A, Howell S.** 1986. The virulent satellite RNA of *Turnip crinkle virus* has a major domain homologous to the 3'-end of the helper virus genome. *EMBO J.* **5**: 3423-3428.
- Sit TL, Johnston JC, Ter-Borg MG, Frison E, McLean MA, Rochon D.** 1995. Mutational analysis of the *Cucumber necrosis virus* coat protein gene. *Virology* **206**: 38-48.
- Sivakumaran K, Hacker DL.** 1998. The 105-kDa polyprotein of *Southern bean mosaic virus* is translated by scanning ribosomes. *Virology* **246**: 34-44.
- Skotnicki ML, Mackenzie AM, Torronen M, Gibbs AJ.** 1993. The genomic sequence of *Cardamine chlorotic fleck carmovirus*. *J. Gen. Virol.* **74**: 1933-1937.
- Skulachev MV, Ivanov PA, Karpova OV, Korpela T, Rodionova NP, Dorokhov YL, Atabekov JG.** 1999. Internal initiation of translation directed by the 5'-untranslated region of the tobamovirus subgenomic RNA I_2 . *Virology* **263**: 139-154.
- Skuzeski JM, Nichols LM, Gesteland RF, Atkins JF.** 1991. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *J. Mol. Biol.* **218**: 365-373.

Song C, Simon AE. 1994. RNA-dependent RNA polymerase from plants infected with *Turnip crinkle virus* can transcribe (+) and (-) strands of virus associated RNAs. *Proc. Natl. Acad. Sci. USA* **91**: 8792-8796.

Song C, Simon AE. 1995. Requirement of a 3'-terminal stem-loop in *in vitro* transcription by an RNA-dependent RNA polymerase. *J. Mol. Biol.* **254**: 6-14.

Soto MJ, Chen LF, Seo YS, Gilbertson RL. 2005. Identification of regions of the *Beet mild curly top virus* (family *Geminiviridae*) capsid protein involved in systemic infection, virion formation and leafhopper transmission. *Virology* **341**: 257-270.

Spriggs KA, Stoneley M, Bushell M, Willis AE. 2008. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol. Cell* **100**: 27-38.

Stavolone L, Ragozzino A, Hohn T. 2003. Characterization of *Cestrum yellow leaf curling virus*: A new member of the family *Caulimoviridae*. *J. Gen. Virol.* **84**: 3459-3464.

Steinhauer DA, de la Torre JA, Meier E, Holland JJ. 1989. Extreme heterogeneity in populations of *Vesicular stomatitis virus*. *J. Virol.* **63**: 2072-2080.

Stone OM. 1980. Nine viruses isolated from Pelargonium in United Kingdom. *Acta Hortic.* **110**: 177-182.

Stone OM, Hollings M. 1971. Annual Report Glasshouse Crops Research Institute for 1970, Littlehampton, UK, pp. 153-155.

Stone OM, Hollings M. 1976. Annual Report Glasshouse Crops Research Institute for 1975, Littlehampton, UK, pp. 119-120.

Stork J, Panaviene Z, Nagy PD. 2005. Inhibition of *in vitro* RNA binding and replicase activity by phosphorylation of the p33 replication protein of *Cucumber necrosis tobusvirus*. *Virology* **343**: 79-92.

Stuart GW, Moffett PK, Bozarth RF. 2006. A comprehensive open reading frame phylogenetic analysis of isometric positive strand ssRNA plant viruses. *Arch. Virol.* **151**: 1159-1177.

Stupina VA, Meskauskas A, McCormack JC, Yingling YG, Shapiro BA, Dinman JD, Simon AE. 2008. The 3' proximal translational enhancer of *Turnip crinkle virus* binds to 60S ribosomal subunits. *RNA* **14**: 2379-2393.

Stupina V, Simon AE. 1997. Analysis *in vivo* of *Turnip crinkle virus* satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. *Virology* **238**: 470-477.

Suzuki S, Hase S, Takahashi H, Ikegami M. 2002. The genome organization of *Pea stem necrosis virus* and its assignment to the genus *Carmovirus*. *Intervirology* **45**: 160-163.

- Tacke E, Prüfer D, Reinold S, Schmitz J, Salamini F, Rohde W.** 1990. Characterization of a *Potato leafroll luteovirus* subgenomic RNA: differential expression by internal translation initiation and AUG suppression. *J. Gen. Virol.* **71**: 2265-2272.
- Takamatsu N, Watanabe Y, Meshi T, Okada Y.** 1990. Mutational analysis of the pseudoknot region in the 3' noncoding region of *Tobacco mosaic virus* RNA. *J. Virol.* **64**: 3686-3693.
- Takeda A, Kaido M, Okuno T, Mise K.** 2004. The C-terminus of the movement protein of *Brome mosaic virus* controls the requirement for coat protein in cell-to-cell movement and plays a role in long distance movement. *J. Gen. Virol.* **85**: 1751-1761.
- Takemoto Y, Kanehira T, Shinoraha M, Yamashita S, Hibi T.** 2000. The nucleotide sequence and genome organization of *Japanese iris necrotic ring virus*, a new species in the genus *Carmovirus*. *Arch. Virol.* **145**: 651-657.
- Taliansky ME, García-Arenal F.** 1995. Role of *Cucumovirus* capsid protein in long-distance movement within the infected-plant. *J. Virol.* **69**: 916-922.
- Tamada T, Schmitt C, Saito M, Guilley H, Richards K, Jonard G.** 1996. High resolution analysis of the readthrough domain of *Beet necrotic yellow vein virus* readthrough protein: a KTER motif is important for efficient transmission of the virus by *Polymyxa betae*. *J. Gen. Virol.* **77**: 1359-1367.
- Tamada T, Kusume T.** 1991. Evidence that the 75K readthrough protein of *Beet necrotic yellow vein virus* RNA-2 is essential for transmission by the fungus *Polymyxa betae*. *J. Gen. Virol.* **72**: 1497-1504.
- Tan Z, Gibbs AJ, Tomitaka Y, Sánchez F, Ponz F, Ohshima K.** 2005. Mutations in *Turnip mosaic virus* genomes that have adapted to *Raphanus sativus*. *J. Gen. Virol.* **86**: 501-510.
- Tanguay RL, Gallie DR.** 1996. Isolation and characterization of the 102-kilodalton RNA-binding protein that binds to the 5' and 3' translational enhancers of *Tobacco mosaic virus* RNA. *J. Biol. Chem.* **271**: 14316-14322.
- Ten Dam EB, Pleij CWA, Bosch L.** 1990. RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes* **4**: 121-136.
- Ten Dam EB, Verlaan PWG, Pleij CWA.** 1995. Analysis of the role of the pseudoknot component in the SRV-1 gag-pro ribosomal frameshift signal – Loop lengths and stability of the stem regions. *RNA* **1**: 146-154.

- Thivierge K, Nicaise V, Dufresne PJ, Cotton S, Laliberte JF, Le Gall O, Fortin MG.** 2005. Plant virus RNAs. Coordinated recruitment of conserved host functions by (+) ssRNA viruses during early infection events. *Plant Physiol.* **138**: 1822–1827.
- Thomas DJ, Casari G, Sander C.** 1996. The prediction of protein contacts from multiple sequence alignments. *Protein Eng.* **9**: 941–948.
- Thomas CL, Leh V, Lederer C, Maule AJ.** 2003. Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. *Virology* **306**: 33-41.
- Timm J, Roggendorf M.** 2007. Sequence diversity of *Hepatitis C virus*: implications for immune control and therapy. *World J. Gastroenterol.* **13**: 4808-4817.
- Touriol C, Bornes S, Bonnal S, Audigier S, Prats H, Prats AC, Vagner S.** 2003. Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol. Cell.* **95**: 169-178.
- Treder K, Kneller EL, Allen EM, Wang Z, Browning KS, Miller WA.** 2008. The 3' cap-independent translation element of *Barley yellow dwarf virus* binds eIF4F via the eIF4G subunit to initiate translation. *RNA* **14**: 134-147.
- Truniger V, Nieto C, González-Ibeas D, Aranda M.** 2008. Mechanism of plant eIF4E-mediated resistance against a *Carmovirus* (*Tombusviridae*): cap-independent translation of a viral RNA controlled *in cis* by an (a)virulence determinant. *Plant J.* **56**: 716-727.
- Turina M, Desvoyes B, Scholthof KB.** 2000. A gene cluster encoded by *Panicum mosaic virus* is associated with virus movement. *Virology* **266**: 120-128.
- Turina M, Maruoka M, Monis J, Jackson AO, Scholthof KB.** 1998. Nucleotide sequence and infectivity of a full-length cDNA clone of *Panicum mosaic virus*. *Virology* **241**: 141-155.
- Turner RL, Buck KW.** 1999. Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA 2 of *Red clover necrotic mosaic virus*. *Virology* **253**: 115-124.
- Turner KA, Sit TL, Callaway AS, Allen NS, Lommel SA.** 2004. *Red clover necrotic mosaic virus* replication proteins accumulate at the endoplasmic reticulum. *Virology* **320**: 276-290.
- Vaewhongs AA, Lommel SA.** 1995. Virion formation is required for the long distance movement of *Red clover necrotic mosaic virus* in movement protein transgenic plants. *Virology* **212**: 607-613.
- Vaquero C, Liao YC, Nahring J, Fisher R.** 1997. Mapping of the RNA-binding domain of the *Cucumber mosaic virus* movement protein. *J. Gen. Virol.* **78**: 2095-2099.

- van der Vossen EA, Notenboom T, Bol JF. 1995. Characterization of sequences controlling the synthesis of *Alfalfa mosaic virus* subgenomic RNA *in vivo*. *Virology* **212**: 663-672.
- van Regenmortel MH, Mahy BW. 2004. Emerging issues in virus taxonomy. *Emerg. Infect. Dis.* **10**: 8-13.
- Veidt I, Lot H, Leiser M, Scheidecker D, Guilley H, et al. 1988. Nucleotide sequence of *Beet western yellows virus* RNA. *Nucleic Acids Res.* **16**: 9917-9932.
- Verchot J, Angell SM, Baulcombe DC. 1998. *In vivo* translation of the triple gene block of *Potato virus X* requires two subgenomic mRNAs. *J. Virol.* **72**: 8316-8320.
- Vilar M, Esteve V, Pallás V, Marcos JF, Pérez-Payá E. 2001. Structural properties of *Carnation mottle virus* p7 movement protein and its RNA-binding domain. *J. Biol. Chem.* **276**: 18122-18129.
- Vilar M, Sauri A, Monné M, Marcos JF, von Heijne G, Pérez-Payá E, Mingarro I. 2002. Insertion and topology of a plant viral movement protein in the endoplasmic reticulum membrane. *J. Biol. Chem.* **277**: 23447-23452.
- Vives MC, Galipienso L, Navarro L, Moreno P, Guerri J. 2002. Characterization of two kinds of subgenomic RNAs produced by *Citrus leaf blotch virus*. *Virology* **295**: 328-336.
- Wada Y, Tanaka H, Yamashita E, Kubo C, Ichiki-Uehara T, Nakazono-Nagaoka E, Omura T, Tsukihara T. 2008. The structure of *Melon necrotic spot virus* determined at 2.8 Å resolution. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **64**: 8-13.
- Wang S, Browning KS, Miller WA. 1997. A viral sequence in the 3'-untranslated region mimics a 5' cap in facilitating translation of uncapped mRNA. *EMBO J.* **16**: 4107-4116.
- Wang S, Carpenter CD, Simon AE. 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for *Turnip crinkle virus*. *Virology* **253**: 327-336.
- Wang S, Miller WA. 1995. A sequence located 4.5 to 5 kilobases from the 5' end of the *Barley yellow dwarf virus* (PAV) genome strongly stimulates translation of uncapped RNA. *J. Biol. Chem.* **270**: 13446-13452.
- Wang S, Mortazavi L, White KA. 2008. Higher-order RNA structural requirements and small-molecule induction of tombusvirus subgenomic mRNA transcription. *J. Virol.* **82**: 3864-3871.
- Wang XQ, Rothnagel JA. 2004. 5'-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. *Nucleic Acids Res.* **32**: 1382-1391.

- Wang J, Simon AE.** 1997. Analysis of the two subgenomic RNA promoters for *Turnip crinkle virus* *in vivo* and *in vitro*. *Virology* **232**: 174-186.
- Wang JL, Simon AE.** 1999. Symptom attenuation by a satellite RNA *in vivo* is dependent on reduced levels of virus coat protein. *Virology* **259**: 234-245.
- Wang HH, Wong SM.** 2004. Significance of the 3'-terminal region in minus strand RNA synthesis of *Hibiscus chlorotic ringspot virus*. *J. Gen. Virol.* **85**: 1763-1776.
- Weber-Lotfi F, Dietrich A, Russo M, Rubino L.** 2002. Mitochondrial targeting and membrane anchoring of a viral replicase in plant and yeast cells. *J. Virol.* **76**: 10485-10496.
- Weiland JJ, Dreher TW.** 1989. Infectious TYMV RNA from cloned cDNA: effects *in vitro* and *in vivo* of point substitutions in the initiation codons of two extensively overlapping ORFs. *Nucleic Acids Res.* **17**: 4675-4687.
- Weng Z, Xiong Z.** 1997. Genome organization and gene expression of *Saguaro cactus carmovirus*. *J. Gen. Virol.* **78**: 525-534.
- White KA.** 2002. The premature termination model: a possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. *Virology* **304**: 147-154.
- White KA, Morris TJ.** 1994. Recombination between defective tobusvirus RNAs generates functional hybrid genomes. *Proc. Natl. Acad. Sci. USA* **91**: 3642-3646.
- White KA, Morris TJ.** 1995. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA* **1**: 1029-1040.
- White KA, Morris TJ.** 1999. Defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses. *Curr. Top. Microbiol. Immunol.* **239**: 1-17.
- White KA, Nagy PD.** 2004. Advances in the molecular biology of tobusviruses: gene expression, genome replication, and recombination. *Prog. Nucleic Acids Res. Mol. Biol.* **78**: 187-226.
- White KA, Skuzeski JM, Li W, Wei N, Morris J.** 1995. Immunodetection, expression strategy and complementation of *Turnip crinkle virus* p28 and p88 replication components. *Virology* **211**: 525-534.
- Wobbe KK, Akgoz M, Dempsey DM, Klessig DF.** 1998. A single aminoacid change in *Turnip crinkle virus* movement protein p8 affects RNA binding and virulence on *Arabidopsis thaliana*. *J. Virol.* **72**: 6247-6250.
- Wolf S, Deom CM, Beachy RN, Lucas WJ.** 1989. Movement protein of *Tobacco mosaic virus* modified plasmodesmatal size exclusion limit. *Science* **246**: 377-379.

- Wong SM, Koh DC, Liu D.** 2008. Identification of plant virus IRES. *Methods Mol. Biol.* **451**: 125-133.
- Worobey M, Holmes EC.** 1999. Evolutionary aspects of recombination in RNA viruses. *J. Gen. Virol.* **80**: 2535-2543.
- Wu W, Goodman M, Lomax MI, Grossman LI.** 1997. Molecular evolution of cytochrome c oxidase subunit IV: evidence for positive selection in simian primates. *J. Mol. Evol.* **44**: 477-491.
- Wu B, White KA.** 1999. A primary determinant of cap-independent translation is located in the 3'-proximal region of the *Tomato bushy stunt virus* genome. *J. Virol.* **73**: 8982-8988.
- Xiong Z, Kim KH, Giesman-Cookmayer D, Lommel SA.** 1993. The roles of the *Red clover necrotic virus* capsid and cell movement proteins in systemic infection. *Virology* **192**: 27-32.
- Xu W, White KA.** 2008. Subgenomic mRNA transcription in an aureusvirus: down-regulation of transcription and evolution of regulatory RNA elements. *Virology* **371**: 430-438.
- Yang LJ, Hidaka M, Sonoda J, Masaki H, Uozumi T.** 1997. Mutational analysis of the *Potato virus Y* 5' untranslated region for alteration in translational enhancement in tobacco protoplasts. *Biosci. Biotechnol. Biochem.* **61**: 2131-2133.
- You XJ, Kim JW, Stuart GW, Bozarth RF.** 1995. The nucleotide sequence of *Cowpea mottle virus* and its assignment to the genus *Carmovirus*. *J. Gen. Virol.* **76**: 2841-2845.
- Young ND, de Pamphilis CW.** 2000. Purifying selection detected in the plastid gene matK and flanking ribozyme regions within a group II intron of nonphotosynthetic plants. *Mol. Biol. Evol.* **17**: 1933-1941.
- Yu HH, Wong SM.** 1998. Synthesis of biologically active cDNA clones of *Cymbidium mosaic potexvirus* using a population cloning strategy. *Arch. Virol.* **143**: 1617-1620.
- Yueh A, Schneider RJ.** 2000. Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. *Genes Dev.* **14**: 414-421.
- Zaccomer B, Haenni A-L, Macaya G.** 1995. The remarkable variety of plant virus genomes. *J. Gen. Virol.* **76**: 231-247.
- Zavriev SK, Hickey CM, Lommel SA.** 1996. Mapping of the *Red clover necrotic mosaic virus* subgenomic RNA. *Virology* **216**: 407-410.
- Zeenko V, Gallie DR.** 2005. Cap-independent translation of *Tobacco etch virus* is conferred by an RNA pseudoknot in the 5'-leader. *J. Biol. Chem.* **280**: 26813-26824.

- Zhang J, Simon AE.** 2005. Importance of sequence and structural elements within a viral replication repressor. *Virology* **333**: 301-315.
- Zhang J, Stuntz RM, Simon AE.** 2004a. Analysis of a viral replication repressor: sequence requirements for a large symmetrical internal loop. *Virology* **15**: 90-102.
- Zhang G, Zhang J, Simon AE.** 2004b. Repression and derepression of minus-strand synthesis in a plus-strand RNA virus replicon. *J. Virol.* **78**: 7619-7633.
- Zhou T, Fan ZF, Li HF, Wong SM.** 2006. *Hibiscus chlorotic ringspot virus* p27 and its isoforms affect symptom expression and potentiate virus movement in kenaf (*Hibiscus cannabinus* L.). *Mol. Plant-Microbe Interact.* **19**: 948-957.
- Zhou H, Jackson AO.** 1996. Expression of the *Barley stripe mosaic virus* RNA beta "triple gene block". *Virology* **216**: 367-379.
- Zhou W, Song W.** 2006. Leaky scanning and reinitiation regulate BACE1 gene expression. *Mol. Cell Biol.* **26**: 3353-3364.
- Zhouraleva G, Frolova L, Le Goff X, Le Guellec R, Ingevechtomov S, Kisselev L, Philippe M.** 1995. Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *EMBO J.* **14**: 4065-4072.

