



**INSTITUTO POTOSINO DE INVESTIGACIÓN
CIENTÍFICA Y TECNOLÓGICA, A.C.**

**POSGRADO EN CIENCIAS EN BIOLOGÍA
MOLECULAR**

**“BASES MOLECULARES DE LA INCOMPATIBILIDAD
REPLICATIVA DE DOS CEPAS DE *Rhynchosia mosaic*
Sinaloa virus”**

OMAYRA CITLALLI BOLAÑOS MARTÍNEZ

Para obtener el grado de
Maestra en Ciencias en Biología Molecular

Director de Tesis
Dr. Gerardo Rafael Argüello Astorga

San Luis Potosí, S.L.P., Noviembre del 2014



Constancia de aprobación de la tesis

La tesis "***Bases moleculares de la incompatibilidad replicativa de dos cepas de "Rhynchosia mosaic Sinaloa virus"***" presentada para obtener el Grado de Maestra en Ciencias en Biología Molecular fue elaborada por **Omayra Citlalli Bolaños Martínez** y aprobada el **veintiocho de noviembre del dos mil catorce** por los suscritos, designados por el Colegio de Profesores de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Dr. Gerardo Rafael Argüello Astorga
Director de la tesis

Dr. Ángel Gabriel Alpuche Solís
Miembro del Comité Tutorial

Dr. J. Sergio Casas Flores
Miembro del Comité Tutorial



Créditos Institucionales

Esta tesis fue elaborada en el Laboratorio de Biología Molecular de Plantas de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. Gerardo Rafael Argüello Astorga.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (No. de registro: 490394) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.

El proyecto de investigación descrito en esta Tesis fue financiado con recursos otorgados al Dr. Ramón G. Guevara González por el CONACYT (Proyecto interinstitucional FORDECYT/193512/2012 “Sistema Regional de Investigación, Innovación y Desarrollo en Biotecnología, Farmacología y Salud”).



Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 123 del Libro Primero de Actas de Exámenes de Grado del Programa de Maestría en Ciencias en Biología Molecular está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 28 días del mes de noviembre del año 2014, se reunió a las 16:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

Dr. J. Sergio Casas Flores	Presidente	IPICYT
Dr. Ángel Gabriel Alpuche Solís	Secretario	IPICYT
Dr. Gerardo Rafael Argüello Astorga	Sinodal	IPICYT

a fin de efectuar el examen, que para obtener el Grado de:

MAESTRA EN CIENCIAS EN BIOLOGÍA MOLECULAR

sustentó la C.

Omayra Citlalli Bolaños Martínez

sobre la Tesis intitulada:

Bases moleculares de la incompatibilidad replicativa de dos cepas de "Rhynchosia mosaic Sinaloa virus"

que se desarrolló bajo la dirección de

Dr. Gerardo Rafael Argüello Astorga

El Jurado, después de deliberar, determinó

APROBARLA

Dándose por terminado el acto a las 17:35 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición de la interesada y para los fines que a la misma convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 28 días del mes de noviembre de 2014.

Mtra. Ivonne Lizette Cuevas Vélez
Jefa del Departamento del Posgrado

Dr. Marcial Bonilla Marín
Secretario Académico



Dedicatorias

A mi madre, la Sra. Rosa Mirna Martínez Ortiz, que a pesar de la distancia siempre está al pendiente de mí. Por apoyarme en cada decisión que he tomado, por todo su esfuerzo y sacrificio para brindarme siempre lo mejor y por enseñarme con sus actos el significado de las palabras fortaleza, ingenio y responsabilidad.

A mis hermanos César Tonatiuh Bolaños Martínez y Mauricio Huitzilopochtli Bolaños Martínez, en especial a mi bro Mauri por siempre aligerar mis días con su alegría, ocurrencias y carrillas.

A mi tío Gonzalo Martínez Ortiz (Chito), ser humano bondadoso que este año partió físicamente, pero su recuerdo permanece en nosotros. Porque un hombre nace con libertad, pero esa libertad se destruye cuando el mismo hombre se la impide.

Agradecimientos

Al Dr. Gerardo Rafael Argüello Astorga por darme la oportunidad de pertenecer a su grupo de trabajo y confiar en mí para la realización de este proyecto.

Al Dr. Ángel Alpuche Solís y al Dr. J. Sergio Casas Flores por su valiosa ayuda en la revisión de manuscritos y de esta tesis.

Al Dr. Guillermo Pastor Palacios, por su amistad, paciencia y la enseñanza de diversas técnicas de biología molecular, sin su ayuda este trabajo no hubiera sido el mismo.

A Salvador Ambriz Granados por su disposición y ayuda en diversas técnicas en varias etapas del proyecto. A la Sra. Rosy Cruz por abastecerme oportunamente del material necesario para trabajar.

Al CONACYT e IPICYT por el apoyo económico y la oportunidad de cursar mis estudios de Maestría.

A mis amigas Gloria, Byanca, Magnolia, Cesia y Maritere porque sin su compañía y amistad mi estancia en San Luis Potosí no hubiera sido agradable y divertida.

A mis amigos entrañables, Angélica, Cindy, Chichi, Martha, Paola y Soraya, a mi prima Kenia, que a pesar de la distancia siempre están para apoyarme.

A Kevin por estar presente desde el inicio de esta etapa, por darme su amor y ayudarme a ser un mejor ser humano.

Contenido

Constancia de Aprobación de la Tesis	<u>ii</u>
Créditos Institucionales	iii
Acta de Exámen	iv
Dedicatorias	v
Agradecimientos	vi
Lista de Tablas	viii
Lista de Figuras	ix
Anexos	x
Resumen	1
Abstract	2
Introducción	4
Métodos	15
Resultados	21
Discusión	39
Conclusión	
¡Error! Marcador no definido.	
Referencias	44
Material	suplementario
¡Error! Marcador no definido.	

Lista de Tablas

Tabla 1. - Virus identificados en las plantas analizadas, linaje al que pertenecen e imágenes de las plantas hospederas	28
--	-----------

Lista de Figuras

- Figura 1.-** Replicación por Círculo Rodante (RCR). **6**
- Figura 2.-** Representación esquemática de los motivos conservados en las especies virales pertenecientes al clado SLCV. **14**
- Figura 3.-** Procesamiento de 4 muestras vegetales. ¡Error! Marcador no definido.
- Figura 4.-** Relaciones filogenéticas de los tres virus aislados.¡Error! Marcador no definido.
- Figura 5.-** Alineamiento de secuencias de aminoácidos de las regiones comunes de RhMSinV-GU, PepGMV y las dos nuevas cepas aisladas.¡Error! Marcador no definido.
- Figura 6.-** Modelado por homología de las proteínas Rep de dos de los begomovirus aislados. ¡Error! Marcador no definido.
- Figura 7.-** Ensayo por Southern blot de la autorreplicación de los componentes A de RhMSinV-CO y RhMSinV-GU en protoplastos de células de *Nicotiana tabacum*. **36**
- Figura 8. -** Southern blot de la replicación del componente B de RhMSinV-GU y PepGMV en protoplástos de células de *Nicotiana tabacum*. **38**
- Figura 9. –** Inducción de síntomas en plantas de *Nicotiana benthamiana* infectadas con los virus PepGMV, RhMSinV-GU y RhMSinV-CO. **42**

Anexos

Tabla S1.- Oligonucleótidos universales utilizados para amplificar DNA-A y DNA-B de begomovirus y oligonucleótidos diferenciales diseñados como iniciadores para amplificar genomas virales pertenecientes a la especie RhMSinV **49**

Figura S1.- Análisis de los polimorfismos de fragmentos de restricción (RFLP) **50**

Preparación de protoplastos de células NT1 **51**

Figura S2.- Generación de protoplastos **53**

Tabla S2.- Nombres, acrónimos y números de acceso en GenBank de las secuencias de los geminivirus utilizados en el análisis filogenético **54**

Componentes virales aislados de las muestras Manzanillo 1-10 **55**

Resumen

Los begomovirus tienen un genoma compuesto por una o dos moléculas circulares de DNA de cadena sencilla (ssDNA) de ~2.6 kb, que se replican por el mecanismo de círculo rodante (RCR) en el núcleo de la célula vegetal. La proteína iniciadora de la replicación (Rep) codificada por estos virus pertenece a la superfamilia de endonucleasas HUH, las cuales tienen un papel importante en la RCR de plásmidos, bacteriófagos, y virus circulares de ssDNA de organismos eucariotas, en la transferencia de plásmidos durante la conjugación bacteriana, y en distintos tipos de transposición. En el 2006 se aisló de una planta de *Rhynchosia mínima* colectada en Guasave, Sinaloa, una nueva especie de begomovirus que se denominó *Rhynchosia mosaic Sinaloa virus* (RhMSinV). En el presente trabajo aislamos tres begomovirus diferentes de una planta de *R. minima* colectada en Manzanillo, Colima, y su secuenciación reveló que uno de esos virus es una nueva cepa de RhMSinV (<92% de identidad). Este nuevo virus RhMSinV-CO exhibe en el origen de replicación cuatro presuntos sitios de unión de Rep (iterones) cuya secuencia (TGGAGTC) es diferente a la de los elementos homólogos del aislado original de Guasave (TGGAGGGA), y a una variante de este último aislado también en este estudio, que denominamos RhMSinV-GUco. El alineamiento de las proteínas Rep codificadas por las dos cepas reveló una diferencia en el residuo aminoácido 71 (N-D), que fue identificado en trabajos previos de nuestro laboratorio como un potencial determinante de especificidad (SPD) de la proteína Rep. El modelado de la estructura tridimensional del dominio endonucleolítico de las proteínas Rep de RhMSinV-CO y -GU reveló que sus presuntos SPDs (los residuos 9, 11, 71 y 73) exponen sus grupos laterales hacia la superficie de una mini-hoja beta (β 1- β 5) que se considera crítica para el reconocimiento y unión de Rep a su DNA cognado. Experimentos de replicación transitoria en protoplastos, y de inoculación por biobalística de plantas de *Nicotiana benthamiana*, demostraron que RhMSinV-GU y RhMSinV-CO son incompatibles en replicación, como se predijo a partir de las diferencias en sus potenciales determinantes de especificidad (iterones y Rep-SPDs). En contraste, el DNA-A de la cepa RhMSinV-CO y el DNA-B de un begomovirus filogenéticamente distante, *Pepper golden*

mosaic virus (PepGMV) que presenta iterones y Rep-SPDs similares, formaron pseudo-recombinantes viables que indujeron síntomas severos en las plantas bombardeadas. Tomados en conjunto, los datos del presente trabajo dan sustento muy sólido a la hipótesis de que los determinantes de especificidad de Rep se localizan en las hebras beta-1 y beta-5 de esa proteína viral.

Palabras clave: Begomovirus, Replicación, Pseudorecombinación.

Abstract

The begomoviruses have genomes composed by one or two circular molecules of single stranded DNA (ssDNA), ~2,6 kb in length, which are replicated by a rolling-circle (RC) mechanism in the nucleus of the plant cell. The replication initiation protein (Rep) encoded by the begomoviruses belongs to the HUH endonuclease superfamily, and is essential for the replication process. Those proteins play a key role in the RCR of plasmids and bacteriophages, in plasmid transfer, in the replication of several eukaryotic viruses, and in various types of transposition. In 2006 we isolated from a leguminous weed (*Rhynchosia minima*) collected in Guasave, Sinaloa, a new begomoviral species named *Rhynchosia mosaic Sinaloa virus* (RhMSinV). In this work three different begomoviruses were isolated from a single *R. minima* plant collected in Colima, and their sequencing revealed that two of those viruses are novel strains of RhMSinV (<92% identity). One of these exhibited four Rep-binding sites (iterons) similar to those of the original isolate (i.e., TGGAGGA), whereas the second strain displayed iterons with a TGGAGTC core sequence. Sequence alignments of the two Rep proteins showed a difference in the aa residue at position 71 (N/D), which is one of the five aa residues which were predicted to function as major determinants of geminivirus Rep DNA-binding specificity (SPDs). Modeling of the 3-D structure of the Rep endonuclease domain revealed that the putative SPDs (i.e. 9, 11, 71 y 73) point their side chains toward the exposed surface of the small beta-sheet (s1-s5) element, that is apparently

critical for high-affinity DNA-binding of Rep proteins. Experiments of transient replication in protoplasts and plant inoculation by biobalistics revealed that RhMSinV-GU and RhMSinV-CO are incompatible in replication. In contrast, biolistic coinoculation of *N. benthamiana* plants with RhMSinV-CO DNA-A and the genomic component B of *Pepper golden mosaic virus* (PepGMV), a distant relative harboring similar replication specificity determinants, induced severe symptoms 15 dpi, thus demonstrating the formation of infectious pseudo-recombinants. Taken together, our data provide strong support to the hypothesis that specific amino acid residues sited at beta-strands 1 and 5 of the RhMSinV Rep protein are actually virus-specific replication determinants.

Keywords: Begomovirus, Replication, Pseudorecombination

Introducción

Familia *Geminiviridae*

Los geminivirus son patógenos de plantas ampliamente distribuidos en la biósfera. Estos virus son transmitidos por insectos fitófagos pertenecientes al orden Hemiptera e infectan una gran cantidad de plantas monocotiledóneas y dicotiledóneas. Se consideran uno de los grupos de patógenos más destructivos de cultivos importantes como frijol, algodón, yuca, chile, remolacha, papa y tomate (Rojas et al., 2005). Algunas de estas especies agrícolas proveen alimento básico para la subsistencia de poblaciones rurales en los países tropicales y subtropicales. La lucha contra las epidemias agrícolas causadas por los geminivirus es una de las tareas más importantes para la estabilidad económica y social en estas áreas (Moffat, 1999).

Los viriones de los geminivirus presentan una morfología que semeja dos icosaedros incompletos unidos por una de sus caras y un tamaño aproximado de 22 x 38 nm. Ésta estructura característica se ha conservado entre los virus de los diferentes géneros, por ejemplo en *Maize streak virus* (MSV), un mastrevirus, y en *African cassava mosaic virus* (ACMV), un begomovirus. La estructura del virión de estos virus ha sido resuelta por crio-microscopía electrónica (cryo-EM de sus siglas en inglés) y se logró la reconstrucción de imagen tridimensional (Zhang et al., 2001; Bottcher et al., 2004).

El genoma de los geminivirus puede estar formado por una o dos moléculas de DNA de cadena sencilla, con un tamaño que oscila entre los 2,500 y 3,000

nucleótidos. El genoma geminiviral codifica de 5 a 8 proteínas implicadas en la replicación viral, el movimiento, la transmisión y la patogénesis (Jeske, 2009).

La infección por geminivirus inicia cuando el DNA de cadena sencilla es liberado de los viriones dentro de la célula vegetal, los cuales son introducidos por el insecto vector al tejido vascular cuando se alimenta de la planta (Hanley–Bowdoin *et al.*, 2013). La replicación por círculo rodante (RCR) es el mecanismo por el cual los geminivirus replican su genoma (Figura1); el primer paso del proceso es la síntesis de la cadena complementaria de DNA, partiendo del componente genómico de cadena sencilla; este paso es catalizado con la participación de los factores celulares del hospedero. En esta etapa son generados los intermediarios de DNA de cadena doble, los cuales son necesarios para la transcripción de los genes virales por la RNA polimerasa II del hospedero, permitiendo así la producción de los transcritos que serán traducidos a proteínas por la maquinaria de síntesis de la célula. La proteína iniciadora de la replicación (Rep). Rep inicia la replicación por círculo rodante mediante la introducción de un corte (“nick”) en la cadena positiva (+) del DNA de cadena doble (dsDNA), dentro de una secuencia nonanucleotídica altamente conservada en los geminivirus. Este evento genera un 3´-OH libre que sirve como “primer” (iniciador) para la síntesis del ssDNA, lo cual conduce al desplazamiento de la hebra parental.

El DNA de cadena sencilla generado durante el proceso de RCR puede tener tres destinos alternativos: 1) ser convertido a DNA de cadena doble para reincorporarse al ciclo de replicación; 2) ser encapsidado por la proteína CP para formar viriones; 3) ser transportado fuera del núcleo y luego a una célula vecina a

través de los plasmodesmos, con la ayuda de las proteínas de movimiento virales (Fondong, 2013; Hanley-Bowdoin et al 2013)

Actualmente, la familia *Geminiviridae* se divide en siete géneros, sobre la base de diferencias en su organización genómica, el rango de plantas que infecta y el insecto que actúa como su vector natural: *Begomovirus*, *Mastrevirus*, *Topocuvirus*, *Curtovirus* y más recientemente, *Becurtovirus*, *Turncurtovirus* y *Eragrovirus* (Adams et al., 2013; Varsani et al., 2014)

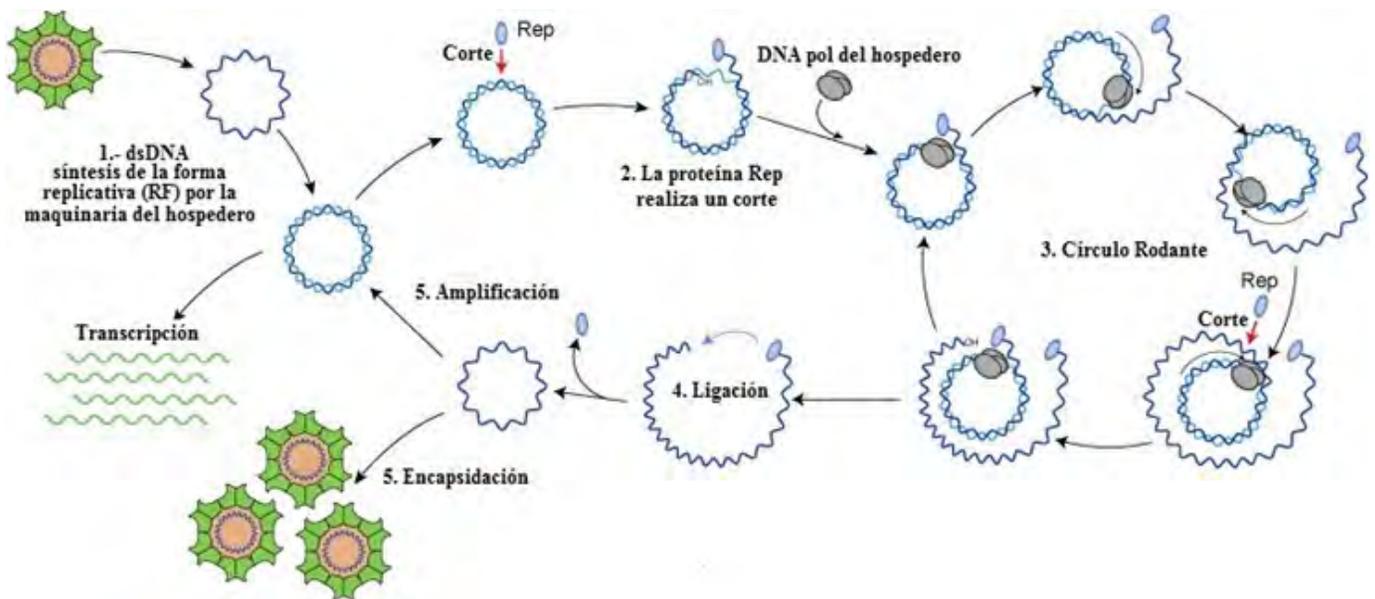


Figura 1.- Replicación por Círculo Rodante (RCR). 1) Formación del intermediario de doble cadena de DNA (Forma Replicativa); 2) Corte endonucleolítico del genoma mediado por la proteína Rep; éste corte genera el 3'-OH libre utilizado como primer para la síntesis de la cadena de DNA, usando la cadena complementaria como molde; 3) Mecanismo de círculo rodante; 4) Ligación del DNA de cadena sencilla liberado durante la RCR; 5) Encapsidación del genoma viral ó formación del intermediario replicativo utilizado para reentrar al ciclo de replicación. Imagen tomada y modificada de Viralzone [http://viralzone.expasy.org/all_by_species/1941.html]

Género *Begomovirus*

El género *Begomovirus* es el más diversificado, con 288 especies descritas, y comprende el 88.6% de las especies pertenecientes a la familia *Geminiviridae* [<http://www.ictvonline.org/virusTaxonomy.asp>]. Los begomovirus están presentes en todas las zonas tropicales y subtropicales del planeta e infectan a una amplia gama de plantas dicotiledóneas. Son transmitidos por un sólo insecto vector, la mosquita blanca (*Bemisia tabaci*), que actualmente se considera un complejo de varias especies morfológicamente indistinguibles (De Barro et al, 2011)

Desde el punto de vista filogenético, los begomovirus pueden ser subdivididos en dos grupos: begomovirus del Viejo Mundo (VM), con genomas mono y bipartitas, distribuidos en Europa, Asia, África y Australia; y los del Nuevo Mundo (NM), con genomas bipartitas exclusivamente, y que se encuentran en todas las regiones del continente americano cuyo clima favorece el desarrollo de poblaciones del insecto vector (Rybicki *et al.*, 1994; Padidam *et al.*, 1997).

Organización genómica de los begomovirus

En los begomovirus bipartitas (ej. todos los del NM) la información genética se encuentra contenida en dos componentes designados como DNA-A y DNA-B. El componente genómico "A" codifica las proteínas necesarias para la replicación del material genético, el control de la expresión de genes, la encapsidación y la evasión del sistema de defensa de la planta. Por ejemplo el gen *AV1* codifica a la proteína de la cápside, CP; el gen *AC1* a la proteína iniciadora de la replicación, Rep; el gen *AC2* al activador transcripcional, TrAP; el gen *AC3* a un enhancer de

la replicación, REn; y *AC4* a un supresor del silenciamiento post-transcripcional. Por su parte, el componente genómico B codifica sólo dos proteínas, las cuales están implicadas en el movimiento intra e intercelular del DNA: BV1 o NSP, que es una proteína de transporte bidireccional núcleo-citoplasma, y BC1 o MP, la proteína de movimiento intercelular (Sanderfoot y Lazarowitz, 1996; Fondong, 2013).

Los dos componentes contienen una región intergénica (IR), la cual contiene promotores en ambas direcciones, los cuales controlan el inicio de la transcripción de los genes codificantes de las proteínas virales. En unos pocos casos las señales de terminación de la transcripción se encuentran localizadas en una IR muy pequeña, localizada entre los extremos 3' de los genes CP y AL3, pero en la mayoría de los begomovirus no existe ninguna IR adicional. La región intergénica mayor de los geminivirus también alberga el origen de replicación del virus.

Los dos componentes genómicos comparten una región de aproximadamente 200 pb (pares de bases), que presentan una identidad en secuencia mayor al 85%; y por ello es llamada la Región Común (CR). Este segmento genómico incluye una secuencia de ~30 nt con el potencial de formar una estructura secundaria de tipo tallo-asa, la cual está conservada entre los geminivirus. En esa conformación, el asa posee una secuencia de nueve nucleótidos invariables (5'-TAATATTAC-3') que es el sustrato endonucleolítico de Rep, y constituye el origen de replicación del DNA viral en sentido estricto.

En la región común también se localizan los sitios de unión específica de la proteína Rep, los cuales son secuencias repetidas, de 6 a 10 nucleótidos de longitud, y son denominados "iterones". Estos elementos presentan una secuencia

nucleotídica específica y difieren aún entre virus emparentados. La CR mantiene la integridad del genoma bipartita, al asegurar que la replicación de ambos componentes sea coordinada por la proteína Rep codificada por el DNA-A (Argüello-Astorga *et al.*, 1994; Orozco *et al.*, 1998; Fontes *et al.*, 1994).

Proteína Asociada a la Replicación

La proteína Rep codificada por el ORF AC1 es esencial para iniciar la replicación del genoma viral mediante RCR. El gen *AC1* está bajo el control de uno de los promotores ubicado en la región intergénica. Rep es una proteína multifuncional, ya que además de participar en el inicio y término de la RCR viral, posee actividad de helicasa, y la capacidad de interactuar con varias proteínas de la célula huésped, principalmente algunas que forman parte de la maquinaria de replicación, como el antígeno nuclear de células en proliferación (PCNA), el factor de replicación C (RFC), y la proteína de replicación A (RPA), entre otras (Hanley-Bowdoin *et al.*, 2013). Además de secuestrar proteínas importantes para la replicación, Rep es capaz de reprogramar el ciclo celular mediante su interacción con la proteína relacionada a retinoblastoma (RBR). El complejo Rep-RBR activa la expresión de genes necesarios para la replicación, con lo que modifica la progresión natural del ciclo celular, para favorecer la síntesis del DNA viral (Orozco *et al.*, 1997; Kong *et al.*, 2000; Gutiérrez, 2000; Bagewadi *et al.*, 2004; Hanley-Bowdoin *et al.*, 2004; Argüello-Astorga *et al.*, 2004).

Al inicio de la replicación por círculo rodante, la proteína Rep se une a secuencias específicas, denominados iterones, presentes en varias copias en el origen de replicación (Argüello-Astorga y Ruiz-Medrano, 2001; Londoño *et al.*,

2010). El dominio de unión al DNA de Rep, se localiza en su extremo amino terminal. Este dominio contiene tres motivos conservados característicos de proteínas iniciadoras de la RCR. El motivo-I está involucrado en el proceso de unión del complejo Rep al DNA; mientras que el motivo-II es un sitio de unión a un ión metálico que puede participar en la conformación de la proteína y en el corte al DNA. El motivo-III es la parte central del dominio catalítico de la endonucleasa; este motivo incluye un residuo de tirosina (Y) esencial para la actividad endonucleotídica de Rep, que genera un “nick” entre la timina y adenina en el extremo 3’ de la secuencia invariante 5’ TAATATTAC 3’ (Ilyina y Koonin, 1992; Koonin & Ilyina, 1992; Jeske, 2009). Recientemente, se ha identificado un motivo conservado denominado “Geminivirus Rep Sequence” (GRS por sus siglas en inglés) que ha mostrado ser también necesaria para el inicio de la replicación del virus (Nash *et al.*, 2011).

Begomovirus del clado SLCV

Es posible distinguir a los begomovirus nativos del continente Americano de los originarios del Viejo Mundo por al menos cuatro características moleculares,: la ausencia del gen AV2 (“pre-coat”); un diferente número y arreglo de iterones (Arguello-Astorga *et al.*, 1994); la conservación del motivo PWRsMaGT en el extremo N-terminal de la proteína de la cápside (CP) el cual está ausente en los begomovirus del VM (Ha *et al.*, 2006, 2008); y la existencia de un motivo conservado (VRFATDKS) en la proteína REn (AC3), que no se encuentra en

ningún begomovirus fuera del continente americano (Mauricio-Castillo et al., 2014)..

No obstante las diferencias antes mencionadas, diversas líneas de investigación indican que los begomovirus nativos de América evolucionaron de begomovirus del Viejo Mundo (Rybicki, 1994; Briddon *et al.*, 2010). Un estudio comparativo de los genomas begomovirales en el que se utilizaron varios métodos de reconstrucción filogenética y programas basados en relojes moleculares, condujo a la hipótesis de que los begomovirus del NM evolucionaron de virus del VM que llegaron al continente americano hace 30-35 millones de años, cuando existía un puente terrestre entre América del Norte y Eurasia, y un clima cálido que favorecía la existencia de una zona continua de vegetación subtropical entre los continentes (Lefeuvre *et al.*, 2011).

De acuerdo con ese hipotético escenario, los begomovirus ancestrales del NM se dispersaron desde Alaska al resto de América del Norte en el Oligoceno y, después de la emergencia del istmo de Panamá (~4 MYA), también a América del Sur, divergiendo evolutivamente en varias ramas correspondientes a los diversos linajes secundarios o clados que existen hoy en día a lo largo del continente.

Uno de los linajes más atípicos de begomovirus del NM es el denominado “clado del *Squash leaf curl virus* (SLCV)”, el cual incluye especies virales distribuidas desde el sur de Estados Unidos hasta Argentina. Dos características principales son las que distinguen a estos virus de otros begomovirus de América: 1) Los sitios de unión a la proteína Rep (Iterones) ubicados en el origen de replicación son más numerosos, poseen una secuencia nuclear (“core”) precedida por un consenso diferente, y exhiben un arreglo único entre los geminivirus

(Argüello-Astorga *et al.*, 1994; Argüello y Ruiz, 2001) y 2) El dominio amino terminal de sus proteínas Rep (segmento 1-160) muestra baja identidad en secuencia (<50%) con proteínas codificadas por begomovirus típicos del NM (Argüello y Ruiz, 2001; Gregorio-Jorge *et al.*, 2010) y carecen de varios motivos aminoácidos característicos de otros virus del NM y todos los del Viejo Mundo.

En un reciente análisis comparativo entre las proteínas de los begomovirus pertenecientes al clado del SLCV (ej., RhMSinV) con las proteínas homólogas de otros begomovirus nativos de América y del VM, se identificaron varios motivos que están conservados específicamente en los miembros del clado SLCV. Estos motivos diferenciales se encontraron en la proteína AC4 y en el extremo N-terminal de Rep, y no existen en las proteínas codificadas por ningún otro begomovirus conocido. Para determinar el posible origen del segmento genómico atípico (de aprox. 600 pb) de los miembros del clado mencionado por un evento de transferencia horizontal, el análisis comparado se extendió para incluir las proteínas Rep de todos los miembros de la familia *Geminiviridae*. Este análisis reveló que los únicos geminivirus que codifican proteínas Rep con los motivos atípicos del clado SLCV son dos especies del género *Curtovirus* denominadas *Horseradish curly top virus* (HrCTV) y *Spinach severe curly top virus* (SpiSCTV), pero otros curtovirus no los presentan (Figura 2) (Torres-Herrera *et al.*, manuscrito en preparación).

En el 2010, Gregorio-Jorge *et al.*, reportaron el hallazgo de una nueva cepa de *Euphorbia mosaic virus* (EuMV) aislada de plantas de Chile colectadas en Jalisco. Esta nueva cepa presenta dos diferencias importantes con el prototipo de la especie, EuMV-YP (aislado en Yucatán): 1) sus cinco iterones difieren en el cuarto

nucleótido (TGGAGTCC en vez de TGGIGTCC); y 2) su proteína Rep exhibe un residuo aa diferente en la posición 11 (Q en vez de A). Las diferencias en los presuntos determinantes de replicación en *cis* y en *trans* de estos virus llevó a los autores a realizar experimentos de intercambio de los componentes genómicos de EuMV-Jal y EuMV-YP; los resultados de esos experimentos mostraron inequívocamente que las dos cepas eran incompatibles en replicación, lo que proporcionó sólido apoyo a la hipótesis de que el residuo 11 de la proteína Rep de los virus del clado SLCV está, efectivamente, involucrado en el reconocimiento del nucleótido 4 del iterón, como se había postulado varios años atrás (Argüello-Astorga y Ruiz-Medrano, 2001).

En el presente trabajo describimos el aislamiento y caracterización molecular de dos nuevas cepas begomovirales, una de las cuales pertenece a la especie RhMSinV, del clado SLCV. Esta nueva cepa, denominada RhMSinV-CO, exhibe iterones distintos al aislado original de Guasave (RhMSinV-GU), y de una variante del mismo (RhMSinV-GUco) aislada de la misma planta en la que se encontró la nueva cepa. Esta última codifica además una proteína Rep que difiere de la de la otra cepa en el residuo 71, que es uno de los presuntos determinantes de especificidad de unión al DNA de esa proteína (Londoño et al., 2010; Mauricio-Castillo et al., 2014). El modelado *in silico* de la estructura terciaria de las proteínas Rep de las cepas de RhMSinV mostró que el cambio en el residuo 71 modifica características relevantes de la región que es crítica para las interacciones específicas con el DNA. Los resultados de experimentos de replicación transitoria en protoplastos de células NT1 (*Nicotiana tabacum*), y de inoculación por biobalística de plantas susceptibles también fueron consistentes

con las expectativas teóricas de que RhMSinV-CO difiere del aislado original en su especificidad replicativa.

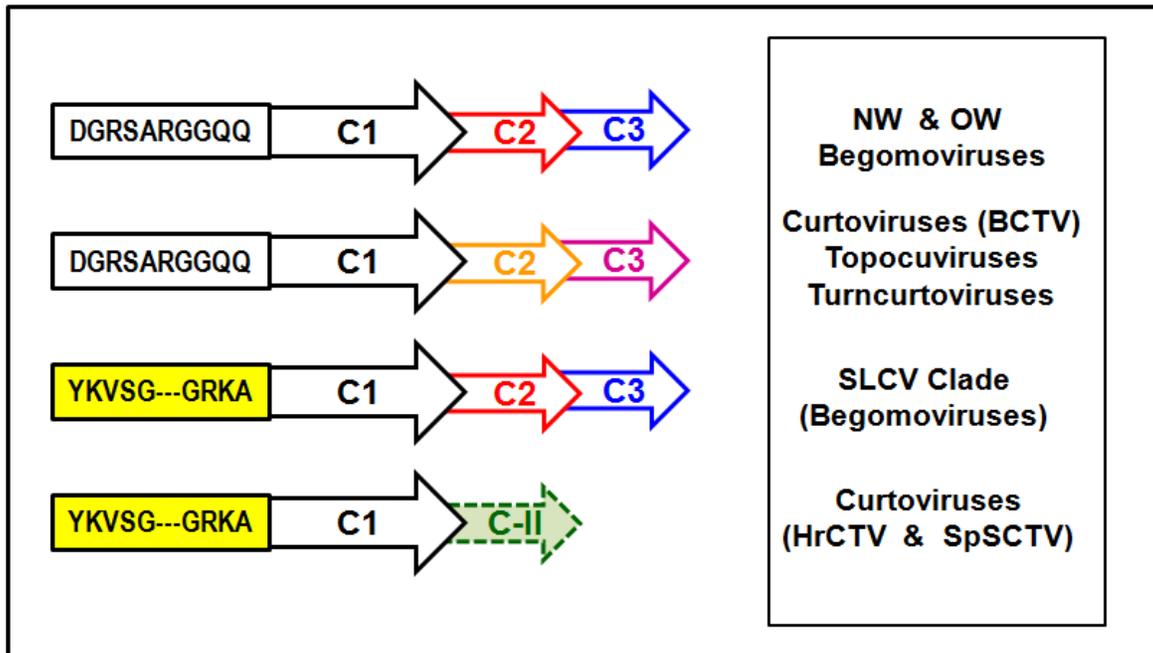


Figura 2.- Representación esquemática de los motivos conservados en las especies virales pertenecientes al clado SLCV. En el recuadro amarillo se ilustra uno de los motivos característicos del clado SLCV, los cuales también están presentes en las especies HrCTV y SpSCTV del género *Curtovirus*. C1 corresponde al gen que codifica la proteína asociada a la replicación (Rep), C2 y C3 representan los genes codificantes de las proteínas TrAP y REn respectivamente (Torres-Sepúlveda et al., manuscrito en preparación).

Métodos

Obtención de las muestras vegetales y extracción de DNA.

En enero del año 2012 se realizó un muestreo en un terreno suburbano ubicado en la ciudad de Manzanillo, Colima. Durante esta colecta se obtuvieron 10 muestras de plantas silvestres (malezas) que presentaban distorsión foliar, marchitez y mosaicos amarillos. El proceso de detección de begomovirus se inició extrayendo el DNA genómico por un método modificado de Dellaporta (Dellaporta *et al.* 1983). El DNA obtenido se utilizó para realizar una amplificación por círculo rodante (RCR) con el kit TempliPhi 100 (GE Healthcare, USA) de acuerdo a las instrucciones del fabricante. Una alícuota del DNA amplificado, se utilizó para la detección de los begomovirus mediante PCR.

Detección de los begomovirus mediante PCR

La detección de DNA begomoviral en la muestra se realizó mediante PCR. En las reacciones se utilizaron oligonucleótidos universales específicos para cada componente viral, previamente diseñados y probados por nuestro grupo de trabajo (Mauricio-Castillo, 2006; Gregorio-Jorge *et al.* 2010). La designación y secuencia de los oligonucleótidos utilizados se describen en la Tabla S1. Las condiciones de las reacciones de PCR fueron las siguientes: desnaturalización inicial a 94°C por 3 minutos, 35 ciclos de 1 minuto a 94°C, 1 minuto a 55°C y 1 minuto a 72°C y un periodo de extensión final de 5 minutos a 72°C. Los productos de PCR fueron analizados por movilidad electroforética en geles de agarosa al 1%.

Posteriormente, tanto para el componente A como para el componente B se utilizaron los siguientes iniciadores específicos para la detección de begomovirus pertenecientes al clado SLCV y virus emparentados con la especie *Rhynchosia mosaic Sinaloa virus*: SL-CP-225 For/ SL2310 Rev, Dif RhMSV-550 For/Dif BV RhMSV-56 Rev, DifRhMSV-940 For/Dif BV RhMSV- 275 Rev, Rh SinIR-For 2285/RhSin BV1 Rev 630 (Tabla S1).

Secuenciación de los genomas virales

Los fragmentos generados se clonaron en el vector pGEM T-easy (Promega). Las reacciones de ligación se utilizaron para transformar células de la cepa TOP-10 de *Escherichia coli*. La detección de clones recombinantes se realizó en medio LB-agar suplementado con X-gal/IPTG, en el que las colonias bacterianas con plásmidos recombinantes no se tiñen de azul. Las colonias seleccionadas fueron inoculadas en medio LB suplementado con carbenicilina (100µg/ml) e incubadas a una temperatura de 37 °C por un periodo de 16 hrs.

Posteriormente, se realizó extracción de DNA plasmídico por el método modificado de Birnboim (Sambrook y Russell, 2001). El DNA plasmídico se sometió a análisis por RFLP's (Restriction Fragment Length Polymorphism) con las enzimas de restricción *EcoRI* y *HinfI*. Los productos de las digestiones fueron resueltos por desplazamiento electroforético en geles de agarosa al 2.5%. Los fragmentos clonados fueron secuenciados mediante el equipo 3130 Genetic Analyzer (Applied Biosystems) en el Laboratorio Nacional de Biotecnología Ambiental y Médica (LANBAMA).

Análisis Filogenético

El análisis de los datos obtenidos por secuenciación, se realizó mediante BLAST (*Basic Local Alignment Search Tool*), con la finalidad de identificar los virus presentes en la muestra. El BLAST se realizó en la base de datos de GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). Las secuencias nucleotídicas completas de los componentes DNA-A y DNA-B de los virus aislados (que serán depositadas en breve en una base de datos pública) fueron comparadas con otras secuencias de begomovirus disponibles en la base de datos del GenBank. Los alineamientos fueron obtenidos por el método Clustal W en la aplicación MegAlign del paquete Lasergene (DNASTAR), utilizando los parámetros por default.

El análisis filogenético se realizó con secuencias completas de los DNA-A de los virus aislados y 29 genomas de virus pertenecientes al Viejo y Nuevo Mundo obtenidas del GenBank (Tabla S2). La secuencia del genoma de *Beet curly top virus* (BCTV) se utilizó como grupo externo. El árbol filogenético se construyó con el software Mega 6.06 con el método Neighbour-joining y 1,000 replicas bootstrap (Tamura *et al.*, 2011).

Modelado *in silico* de las proteínas Rep de RhMSinV-GU y RhMSinV-CO

El modelo estructural correspondiente a las proteínas Rep de RhMSinV-GU y RhMSinV-CO, se obtuvo mediante homología de secuencias en el servidor I-Tasser (Zhang, 2008; Ambrish *et al.*, 2010). Los modelos fueron construidos usando como molde la estructura tridimensional del dominio catalítico de la proteína Rep de *Tomato yellow leaf curl Sardinia virus* (TYLCSV), con número de

acceso PDB 1L2M (Campos-Olivas, 2010). La calidad de ambos modelos se evaluó mediante los servidores ProSa y Molprobitly (Wiederstein y Sippl, 2007; Sippl, 1993; Davis *et al.*, 2007). El software PyMOL se usó para la visualización y el análisis de los modelos estructurales.

Generación de clonas infectivas de RhMSinV-CO

Para la obtención de la secuencia completa del DNA-A de RhMSinV-CO se sintetizaron los iniciadores divergentes RhMCOA-For y RhMCOA-Rev diseñados sobre el sitio único para la enzima *KpnI* en el genoma viral (Tabla S1). Las condiciones de la PCR usadas fueron: desnaturalización inicial a 94 °C por 2 minutos, 30 ciclos de 94 °C 1 minuto, 60 °C por 1 minuto y 3 minutos a 72 °C con una extensión final a 72 °C por 5 minutos. Los amplicones generados de aproximadamente 2.7 Kb se clonaron en el vector pGEM-Teasy (Promega). Los componentes y cantidades de la reacción se adicionaron de acuerdo al protocolo del Kit de la Platinum Taq DNA Polymerase de Invitrogen by Life technologies.

Las clonas recombinantes fueron identificadas mediante complementación alfa, utilizando medio de cultivo LB suplementado con X-gal/ IPTG. La secuenciación se realizó en el LANBAMA.

Ensayos de replicación transitoria en protoplastos de células NT1

Los ensayos de replicación transitoria se llevaron a cabo en células del mesófilo de *Nicotiana tabacum* (NT1) crecidas en suspensión (Figura S2). El protocolo de preparación y aislamiento de protoplastos se realizó con el procedimiento modificado de Sankara *et al.*, 1995 (Material Suplementario).

Los protoplastos de NT1 fueron transfectados con los componentes virales mediante electroporación. Las condiciones utilizadas fueron: pulso de 500 μ F y 250 volts; 14 μ g de ADN total, correspondiente al DNA-A de RhMSinV-CO o RhMSinV-GU; y combinaciones con el DNA-B de RhMSinV-GU o PepGMV.

Las células electroporadas fueron recuperadas en medio MS con manitol en cajas de Petri (7 ml de volumen total) e incubadas a 25 °C por 48 horas, sin agitación y protegidas de la luz. Posteriormente los cultivos fueron cosechados y procesados para la extracción del DNA total (Dellaporta *et al.*, 1983).

Análisis tipo Southern Blot

El DNA obtenido fue digerido con la enzima *DpnI* y con enzimas de corte único en el componente viral (*KpnI*, *HindIII* o *BamHI*). Los productos se cargaron en un gel de Agarosa al 1%. Posteriormente el DNA se transfirió a una membrana de Nylon Amersham Hybond-N+ (GE Healthcare) y se fijó por irradiación con luz UV en el equipo UV Crosslinker (120 mJ/cm). Las sondas utilizadas consistieron de regiones específicas para los componentes A y B. La sonda para el DNA-A fue un fragmento de 175 pb del gen que codifica para Rep de RhMSinV obtenido por PCR con los oligonucleótidos SL 2060-For/SL 2200-Rev; y para el componente DNA-B una región de 265 pb del gen *BC1* amplificado con los oligonucleótidos BC1-290-For/BC1-80-Rev (Tabla S1). Cada fragmento se etiquetó en el extremo 3' OH con dUTP- Biotina con Biotin 3' End DNA Labeling Kit de acuerdo a las especificaciones del fabricante (Thermo Scientific). La etapa de hibridación se realizó a 55 °C en el horno (TECHNE Hybridiser HB-ID) durante toda la noche y procesada de acuerdo al protocolo descrito por Sambrook y Russell, 2001

(Material suplementario). La detección se realizó de acuerdo a las instrucciones de Chemiluminescent Nucleic Acid Detection Module de Thermo Scientific.

Experimentos de pseudorecombinación en plantas

Los ensayos de pseudorecombinación en *Nicotiana benthamiana* se realizaron utilizando el DNA-A y DNA-B de RhMSinV-GU, y el DNA-A de RhMSinV-CO, en sus formas monoméricas. Antes de la inoculación estos componentes genómicos se liberaron por digestión con las enzimas *Scal*, *BamHI* o *KpnI* respectivamente. Los DNA A y B de PepGMV se utilizaron en forma de hemidímeros, y no fue necesaria digestión alguna previa a la inoculación. Las combinaciones inoculadas por biobalística fueron las siguientes:

- 1)** RhMSinV-GU, DNA-A y DNA-B
- 2)** PepGMV, DNA-A y DNA-B
- 3)** RhMSinV-CO, DNA-A y RhMSinV-GU, DNA-B
- 4)** RhMSinV-CO, DNA-A y PepGMV, DNA-B

Las plantas de *Nicotiana benthamiana* utilizadas correspondían a una etapa de crecimiento de cuatro hojas verdaderas. La inoculación se realizó por biobalística utilizando el sistema de entrega de partículas Biolistic® PDS-100/He de BIO-RAD (Carrillo *et al.*, 2007). La cantidad de DNA viral fue de 6 µg de cada componente genómico (DNA-A y DNA-B) depositados en la superficie de micro-partículas de tungsteno (0.7 mm, BioRad, Hercules, CA). La presión de Helio usada para la dispersión de las partículas fue de 650 psi.

Las plantas inoculadas se mantuvieron en una cámara de crecimiento libre de insectos a una temperatura de 25 °C y con un ciclo de 16 horas luz/ 8 horas

oscuridad. El estado de infección de las plantas se monitoreó por inspección visual de los síntomas y fotografías desde el inicio hasta el fin del experimento. Se extrajo el DNA total (Dellaporta *et al.* 1983) de las hojas emergentes y la detección del DNA viral se realizó por PCR con los iniciadores SL2150- For/YMAC-Rev y BC1-290-For/BC1-80-Rev (Tabla S1).

Resultados

En el estudio de Gregorio-Jorge *et al.* (2010) descrito previamente, se demostró la incompatibilidad replicativa de dos aislados de la misma especie begomoviral, *Euphorbia mosaic virus*. La incompatibilidad observada entre esos begomovirus parece estar relacionada con sus diferencias en los presuntos determinantes de especificidad replicativa en *cis* (los iterones) y en *trans* (los SPDs de Rep).

Considerando la relevancia de los resultados obtenidos con los aislados de EuMV, decidimos emprender una búsqueda extensa de begomovirus que, al igual que estos últimos, presenten una alta identidad en secuencia genómica global, pero exhiban diferentes determinantes de especificidad potenciales, a fin de responder varias cuestiones todavía abiertas en el campo de las interacciones de Rep con el origen de replicación cognado.. En consecuencia, se realizaron colectas de plantas silvestres en diversas partes de la república mexicana, en busca de begomovirus con las características mencionadas. Una de ellas se realizó en un terreno agreste en los límites de la ciudad de Manzanillo, Colima (coordenadas 19°06'43.91" N; 104°20'21.84" W) en donde se obtuvieron muestras de 10 plantas que manifestaban síntomas característicos de infección por

geminivirus (i.e. mosaicos amarillos, enchinamiento de las hojas, clorosis, enanismo, etc.). Las muestras fueron transportadas al laboratorio de biología molecular de plantas del Instituto Potosino de Investigación Científica y Tecnológica (IPICYT) para realizar su análisis por técnicas moleculares.

Las muestras fueron etiquetadas de la siguiente manera: Col 1-*Malva* spp.; Col 2-planta No Identificada (NId); Col 3-*Rhynchosia mínima*; Col 4- *Sida* spp.; Col 5- *Euphorbia heterophylla*; Col 6-*Sida* spp.; Col 7-*Yellow* (NId); Col 8- *R. minima*; Col 9- *R. minima*; Col 10-*Malva* spp.

Análisis por métodos moleculares de las muestras vegetales

Para el análisis de las muestras se extrajo el DNA total de cada una de ellas y con el uso de diversos iniciadores (“primers”) degenerados y específicos (Tabla S1), se llevó a cabo la detección de DNA begomoviral por medio de la Reacción en Cadena de la Polimerasa (PCR) (ver Métodos). Los productos de PCR se clonaron en el vector pGEM T-easy, y las clonas recombinantes fueron analizadas por la técnica de RFLP’s utilizando las enzimas *EcoRI* y *HinFI*...Las clonas con distintos patrones de restricción fueron seleccionadas para su secuenciación.

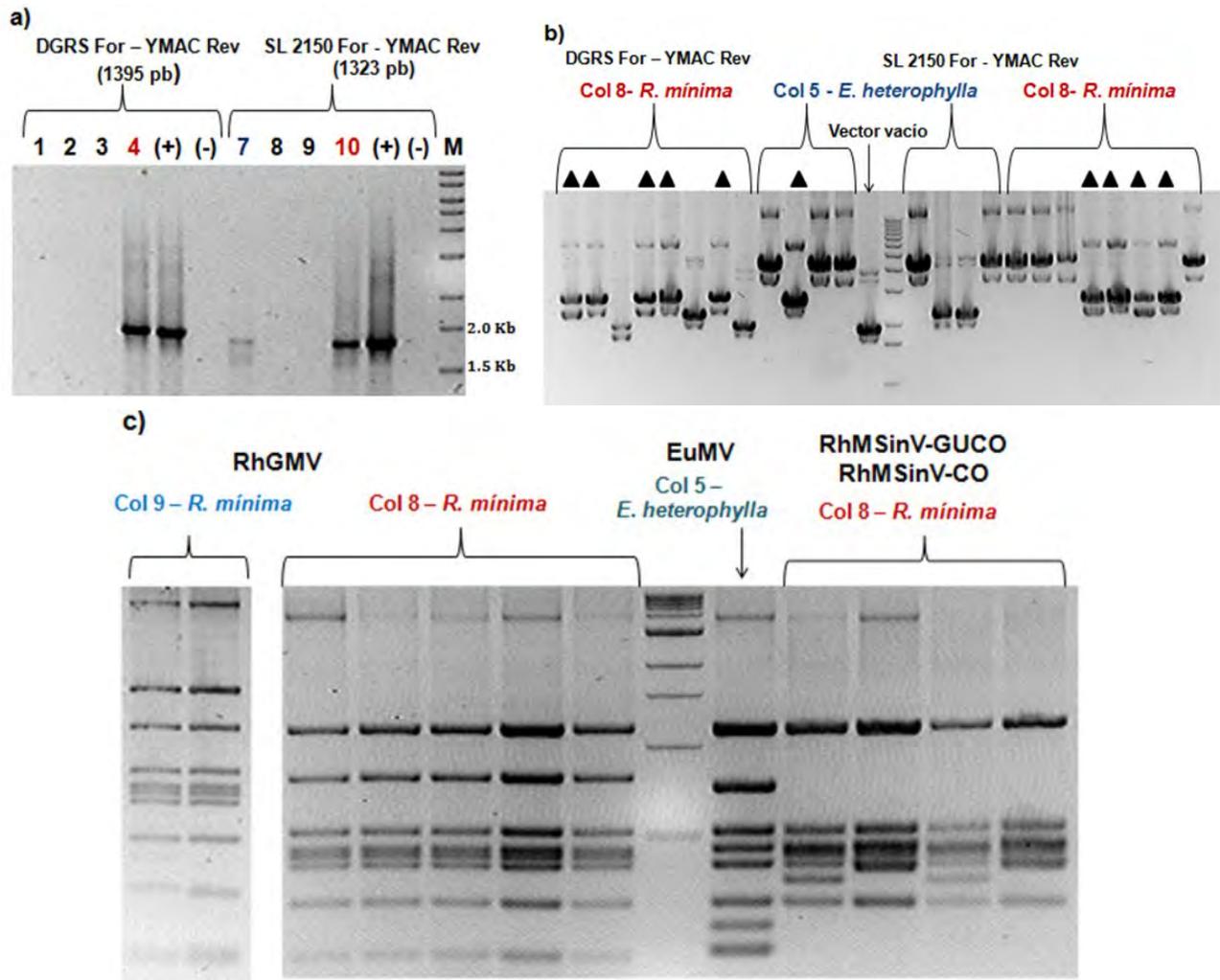


Figura 3.- Procesamiento de 4 muestras vegetales por el método LISOP. a)

Amplificación de DNA con las combinaciones de iniciadores DGRS For – YMAC Rev (1395 pb) y SL 2150 For - YMAC Rev (1323 pb); los carriles 1 y 7 muestran los productos de PCR de la muestra Col 5 - *E. heterophylla*; 2 y 8 de Col 6 - *Sida* spp.; 3 y 9 de Col 7- *Yellow Nid*; 4 y 10 de Col 8- *R. mínima*; (+) control positivo; (-) control negativo; (M) Marcador de peso molecular 1Kb; se observa amplificación en las muestras de los carriles 4, 7 y 10. **b)** Extracción de DNA plasmídico de las clonas transformadas con el vector pGEM-T easy más los amplicones generados; los triángulos negros indican a las clonas seleccionadas para la digestión con las enzimas *EcoRI* y *HinFI*. **c)** Análisis de RFLP de los virus RhGMV, EuMV, RhMSinV-GUco y RhMSinV-CO presentes en las muestras Col 9, Col 8 y Col 5.

En la Figura 3 se muestran algunos resultados que ilustran el procedimiento utilizado, denominado “método LISOP-RFLP”, en el que se usan pares de oligonucleótidos específicos de linaje (“DGRS” y “SL”) que hacen posible la amplificación diferencial de begomovirus “típicos” (ej., los que contienen el motivo conservado DGRSARGG en Rep; ver Fig.2) y los del clado SLCV, (que carecen de ese motivo y presentan varios otros). Los distintos patrones de restricción que corresponden a algunos de los virus aislados de las muestras como, RhGMV, RhMSinV-GUco, RhMSinV-CO y EuMV-Jal, se muestran en la figura 3c, Adicionalmente, en la figura S1 (Material suplementario) se ilustran los patrones de restricción correspondientes a otros begomovirus encontrados en las muestras Col 1-*Malva* spp., Col 2-*Nid*, Col 3-*R. mínima* y Col 8-*R. mínima*.

Los resultados del análisis LISOP-RFLP de las 10 muestras colectadas se resumen en la Tabla 1, que incluye una imagen de la planta sintomática de la que se tomó la muestra correspondiente. Entre los hallazgos del análisis se pueden destacar los siguientes: 1) Se identificó una nueva cepa de *Okra yellow mosaic Mexico virus* (OkYMMV-MZ), que aislamos de 3 muestras diferentes, Col1, Col2 y Col3; 2) Se identificó una nueva variante de RhMSinV-GU, que denominamos RhMSinV-GUco, la cual se encontró en tres muestras, Col3, Col8 y Col9; 3) Se descubrió una nueva cepa de RhMSinV, que denominamos RhMSinV-CO, la cual se aisló de la muestra Col-8, infectada por tres begomovirus; 4) Se aisló una variante de *Rhynchosia golden mosaic virus* (RhGMV) que presenta alta identidad en secuencia con varios aislados descritos previamente; este virus se encontró en las muestras Col8 y Col9; 5) Se aisló una variante de EuMV-Jal de las muestras Col5 y Col9; en el primer caso también se encontró una variante de EuMV-YP, y

en la muestra Col9 se amplificó (parcialmente) lo que parece un híbrido EuMV-RhMSinV (ver más adelante); 6) No se logró detectar DNA de begomovirus en cuatro muestras: Col4, Col6, Col7 y Col10. La descripción detallada de la caracterización molecular de los begomovirus mencionados se presenta más adelante.

Aislamiento de dos nuevos virus emparentados con *Rhynchosia mosaic Sinaloa virus*

El análisis preliminar de la muestra Col-8 proporcionó evidencia de que esa planta de *Rhynchosia minima* estaba infectada por 3 begomovirus diferentes, dos del clado SLCV y uno de un linaje de begomovirus “típicos”. Esto condujo a un esfuerzo para lograr la caracterización genómica completa de los tres virus co-infectantes. El begomovirus “típico” resultó ser una variante de *Rhynchosia golden mosaic virus* cuyos componentes genómicos A (2604 nt) y B (2587 nt) presentan una identidad de 97% y 80%, respectivamente, con los homólogos de RhGMV-Sb (GenBank No.:EU339936 y EU339937). De los dos miembros del clado SLCV que se encontraron en la muestra, solo se logró recuperar los dos componentes genómicos de uno de ellos, en tanto que del segundo solo se aisló el DNA-A. El primero posee un DNA-A de 2585 pb, y un DNA-B de 2553 pb, que presentan la mayor identidad en secuencia con los componentes homólogos de RhMSinV-GU, de 93.6% y 88%, respectivamente. Estos valores de identidad corresponden taxonómicamente a una “variante” de esa especie, que designamos RhMSinV-GUco; El segundo virus del clado SLCV también está relacionado con RhMSinV-

GU, pero su identidad en secuencia global es menor (91%) que en el caso anterior. De acuerdo a los criterios taxonómicos vigentes, los aislados de begomovirus que presentan una identidad nucleotídica del 88 al 92% en su DNA-A se consideran cepas (“strains”) diferentes de la misma especie (Fauquet *et al.*, 2008), por lo que este último virus de la muestra Col-8 representa una nueva cepa de RhMSinV, a la cual denominamos *Rhynchosia mosaic Sinaloa virus- Colima* (RhMSinV-CO). De este virus no encontramos evidencia de un DNA-B en la muestra vegetal, pese a repetidos y variados intentos que incluyeron el diseño y prueba de iniciadores especiales.

Caracterización genómica de otros begomovirus en las muestras colectadas.

A partir de los extractos de la muestra Col 1-*Malva* spp. se logró reconstruir el DNA-A completo (2615 pb) de un begomovirus que comparte un 91% de identidad en secuencia con el componente homólogo de *Okra yellow mosaic Mexico virus-Iguala* (OkYMMV). De acuerdo a los criterios taxonómicos actuales, (Fauquet *et al.*, 2008) este aislado constituye una nueva cepa de OkYMMV, a la que denominamos provisionalmente *Okra yellow mosaic Mexico virus- Manzanillo* (OkYMMV-MZ). Secuencias prácticamente idénticas a las de éste último fueron aisladas también de las muestras Col 2-Nld, y Col 3- *R.mínima*. (Tabla 1)

La muestra Col 3-*R.mínima* contenía, además de OkYMMV-MZ, un virus del clado SLCV, cuya secuencia fue muy similar (99%) a RhMSinV-GUco, aislado también de otras muestras de *R. mínima* (Col-8 y Col-9).

En el caso de la muestra Col-5 *Euphorbia heterophylla* se amplificaron secuencias parciales (1330 pb) de dos variantes de EuMV-Jal, una de las cuales presenta iterones y Rep-SPDs diferentes, análogas a los de EuMV-YP. Se intentará obtener los genomas completos de estas dos variantes, para posteriores estudios.

De la muestra Col9–*R. mínima* se obtuvieron secuencias parciales de los componentes A de cuatro virus distintos, amplificadas con dos pares de iniciadores (SL2150 Forward/YMAC-Reverse y DGRS-Forward/ YMAC-Reverse). Un amplicón de 1302 pb correspondió al DNA-A de RhMSinV-GUco, aislado de distintas muestras de esta colecta; un segundo amplicón, de 1442 pb, fue prácticamente idéntico a la variante de RhGMV encontrada en la muestra Col-8; la tercer secuencia amplificada (1331 pb), pertenece a EuMV-Jal (99% de identidad). El análisis de la cuarta secuencia parcial amplificada (1330 pb) reveló un DNA-A que proviene, aparentemente, de una recombinación entre EuMV-Jal y RhMSinV-GU, ya que la secuencia es similar a EuMV-Jal (~90% de identidad) y presenta baja identidad con RhMSinV-GU en los primeros 675 nt, en claro contraste con el fragmento que comprende los nucleótidos 670-1330, el cual exhibe alta identidad con RhMSinV-GU (94%). La secuencia parcial de este DNA-A híbrido EuMV/RhMSinV, y las secuencias completas de las nuevas cepas y variantes aisladas en este estudio se encuentran disponibles en el material suplementario. .

Tabla 1. - Begomovirus identificados, linaje al que pertenecen e imágenes de las plantas de las que se obtuvieron las muestras.

Muestra	Planta hospedera	Virus aislados	Linaje	Fotografía
Col 1-Malva spp.	Malva spp.	OkYMMV-MZ Nueva cepa DNA-A (2615 pb)	DGRS	
Col 2 - Nid	No identificada	OkYMMV-MZ Nueva cepa DNA-A (2615 pb)	DGRS	
Col 3 - R. mínima	Rhynchosia mínima	OkYMMV-MZ DNA-A (2615 pb) RhMSinV-Guco Nueva variante DNA-A (2585 pb)	DGRS SLCV	
Col 4 – Sida spp.	Sida spp.	Negativa	-----	
Col 5 - E. heterophylla	Euphorbia heterophylla	EuMV-YP DNA-A Parcial (1342 pb) EuMV-Jal DNA-A Parcial	SLCV	

Col 6 – <i>Sida</i> spp.	<i>Sida</i> spp.	Negativa	-----	
Col 7- Yellow	No identificada	Negativa	-----	
Col 8 – <i>R. mínima</i>	<i>Rhynchosia mínima</i>	RhGMV RhMSinV-GUco Nueva variante (DNA-A 2585 pb) RhMSinV-CO Nueva cepa (DNA-A 2575 pb)	DGRS SLCV	
Col 9 – <i>R. mínima</i>	<i>Rhynchosia mínima</i>	RhGMV Parcial(1442 pb) RhMSinV-GUco Parcial (1302 pb) EuMV-Jal Parcial (1331 pb) DNA-A Híbrido RhMSinV/EuMV Parcial	DGRS SLCV	
Col 10 - <i>Malva</i> spp.	<i>Malva</i> spp.	Negativa	-----	

Relaciones Filogenéticas

Utilizando el método neighbor-joining se generó un árbol filogenético con las secuencias completas del DNA-A de cuatro de los virus aislados de las muestras de Colima (RhMSinV-GUco, RhMSinV-CO, RhGMV y OkYMMV-MZ), 29 begomovirus de varios continentes y un curtovirus utilizado como grupo externo (Figura 4). El análisis mostró que RhMSinV-CO pertenece al clado SLCV, en tanto que RhGMV y OkYMMV se ubican en linajes diferentes.

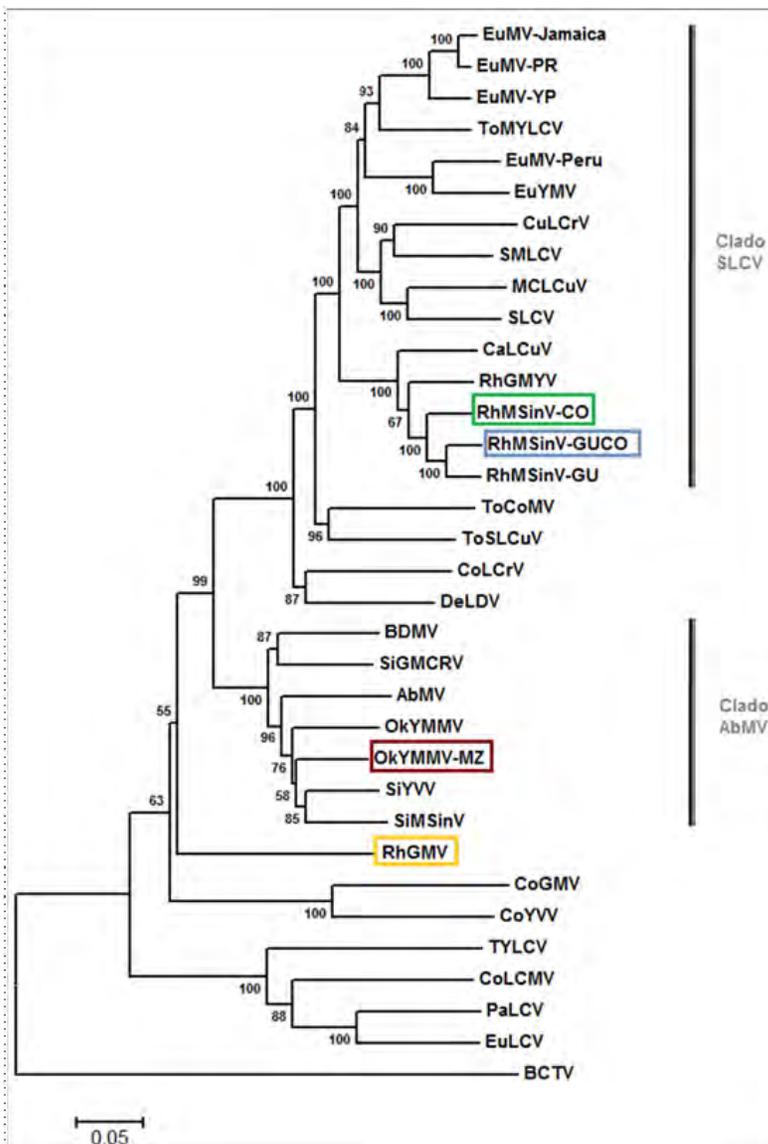


Figura 4.- Relaciones filogenéticas de los tres virus aislados. Árbol filogenético basado en la comparación de la secuencia completa del DNA-A. Los virus aislados en este trabajo se muestran en recuadros de colores: RhMSinV-Col/verde, RhMSinV-GU-Col/azul, RhGMV/amarillo y OkYMMV-MZ/Café. Los clados SLCV y AbMV se delimitan por líneas verticales. El árbol se construyó usando el algoritmo Neighbor-joining en el software MEGA 6.06 con 1000 replicas bootstraps y utilizando al curtovirus BCTV como grupo externo (Tamura *et al.*, 2011). Los nombres de los virus, acrónimos y números de acceso en GenBank se encuentran enlistados en la tabla suplementaria S2.

Diferencias en los presuntos determinantes de especificidad replicativa de RhMSinV-GU y RhMSinV-CO.

El análisis de la región del origen de replicación de los aislados RhMSinV-GUco y RhMSinV-CO mostró que éstos poseen iterones diferentes. El primero exhibe iterones similares a los del aislado original de Sinaloa, que tienen el consenso TGGAGGA, mientras que la segunda cepa posee iterones con la secuencia TGGAGTC (Figura 5a). Tanto RhMSinV-GU como RhMSinV-GUco y RhMSinV-CO muestran el arreglo y número de iterones típicos de los virus pertenecientes al clado SLCV, es decir, dos repetidos directos adyacentes a la caja TATA de la proteína Rep, un repetido invertido distal a los dos repetidos directos y dos iterones flanqueando la estructura tallo-asa (Figura 5a) (Argüello-Astorga *et al.*, 1994). Por otra parte, la comparación de las proteínas Rep codificadas por las dos cepas de RhMSinV reveló que difieren en el residuo en posición 71, uno de los residuos identificados por Londoño *et al.* (2010) como un potencial determinante de especificidad de esa proteína viral. La cepa de Colima presenta en esa posición una asparagina (N), como la mayoría de los begomovirus del Clado SLCV, en

tanto que la cepa de Guasave y su variante de Colima (GUco) exhiben un residuo de ácido aspártico (D), que no se presenta en la proteína Rep de ningún otro begomovirus (Fig. 5b).

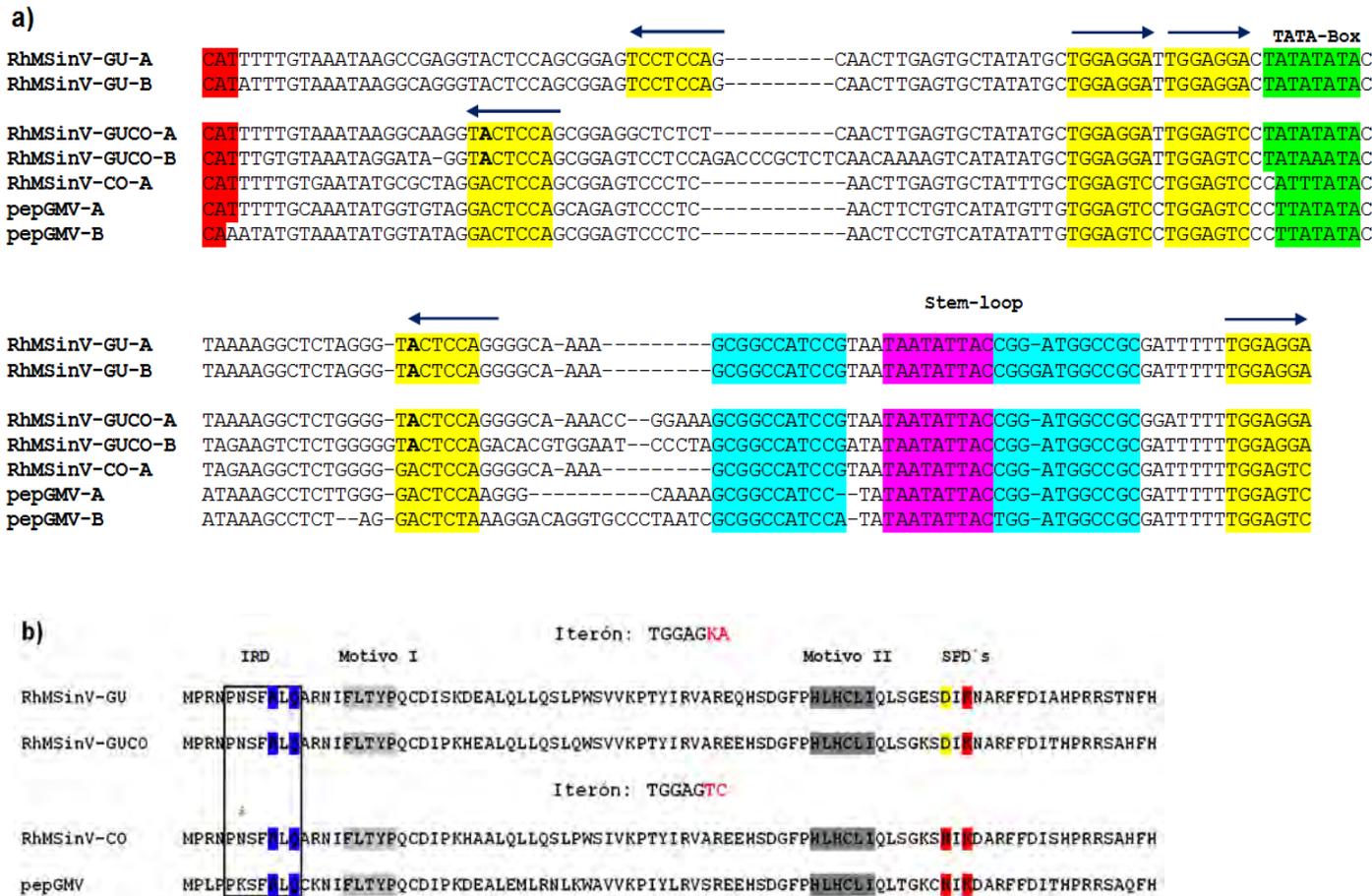


Figura 5.- Delimitación de los presuntos determinantes de especificidad replicativa de los aislados de RhMSinV. **a)** Ubicación y orientación de los iterones de los componentes A y B de RhMSinV-GU, RhMSinV-GUco y PepGMV y del DNA-A A de RhMSinV-CO. La caja TATA se encuentra sombreada de verde, el codón de inicio de Rep y BC1 de color rojo, la secuencia que forma una estructura tallo-asa se marca con colores azul y morado, los iterones están coloreados de amarillo y su orientación se muestra con flechas. **b)** Alineamiento de las proteínas Rep. El recuadro negro corresponde al Domino Relacionado al Iteron (IRD), las diferencias entre aminoácidos se encuentran marcadas con asteriscos. Sombreados de color gris están los Motivos I y II. Los SPD's están coloreados de rojo y amarillo.

Estructura tridimensional de las proteínas Rep de RhMSinV-GU y RhMSinV-CO

Los modelos estructurales de las proteínas Rep de RhMSinV-GU y RhMSinV-CO obtenidos con el servidor I-Tasser (Ambrish *et al.*, 2010), mostraron un C-score de 1.69 y 1.56, respetivamente. Estos valores son un parámetro para determinar la calidad de los modelos predichos. Un valor cercano a 2 corresponde a la mejor predicción de ubicación en el espacio de los residuos respecto a su estructura molde. Mientras que un valor inferior a 1 representa un modelo con baja similitud estructural al modelo de referencia.

Ambos modelos también fueron evaluados mediante el servidor ProSa (Wiederstein y Sippl. *et al.*, 2007). Los valores Z-score de ProSa fueron -6 para RhMSinV-GU y -5 para RhMSinV-CO, los dos valores se encuentran dentro del intervalo típico para estructuras de proteínas de longitud similar. Un análisis de estequiometría de los modelos realizado por el servidor Molprobit (Davis *et al.*, 2007) mostró que el 95.4% para RhMSinV-GU y 96.7% para RhMSinV-CO de los enlaces peptídicos se encontraban en las regiones favorecidas del gráfico de Ramachandran. El software PyMOL se usó para la visualización y el análisis de los modelos estructurales.

Los residuos de la proteína Rep que determinan su afinidad por su iterón cognado se localizan en los primeros 80 aminoácidos (Figura 6c). Los modelos estructurales y la comparación de los aminoácidos con la proteína Rep de TYLCSV muestra diferencias en los residuos que son considerados como determinantes de la especificidad (Figura 6) (Campos-Olivas, 2002). Los residuos 9, 11, 71 y 73 son importantes en el reconocimiento del iterón, de acuerdo a un

análisis comparado previo (Londoño *et al.*, 2010), y presumiblemente funcionan como determinantes de especificidad de la unión al DNA de Rep. Estos residuos, en la estructura de la proteína Rep, se encuentran en una pequeña hoja beta compuesta por las hebras $\beta 1$ y $\beta 5$. Las cadenas laterales de los residuos mencionados se encuentran expuestas hacia la superficie del dominio de unión al DNA (Figura 6a y 6b) Esto proporciona diferentes patrones de distribución de cargas eléctricas y puentes de hidrógeno potenciales para interacciones específicas con secuencias en el DNA de doble cadena (Fig.6c), que podrían explicar la mayor afinidad de Rep por iterones específicos (Londoño *et al.*, 2010).

Ensayos de replicación transitoria de RhMSinV-GU y RhMSinV-CO en protoplastos de tabaco.

Las diferencias observadas en los potenciales determinantes de especificidad de las dos cepas de RhMSinV sugieren la posibilidad de que las mismas sean incompatibles en replicación, de modo análogo a lo demostrado en el caso de las cepas EuMV-Jal y EuMV-YP (Gregorio-Jorge *et al.*, 2010). Para examinar la compatibilidad replicativa entre la cepa caracterizada en este estudio y la aislada originalmente en Guasave, Sin., procedimos a generar una clona infectiva del DNA-A de RhMSinV-CO. Este componente genómico fue amplificado por PCR utilizando oligonucleótidos divergentes parcialmente sobrelapados en el único sitio Kpn1 presente en el DNA-A de esa cepa, y clonado en pGEM-TEasy (ver Métodos). El componente genómico A de RhMSinV-GU fue previamente obtenido por medio de una metodología análoga por Mauricio-Castillo (Tesis Doctoral, 2012).

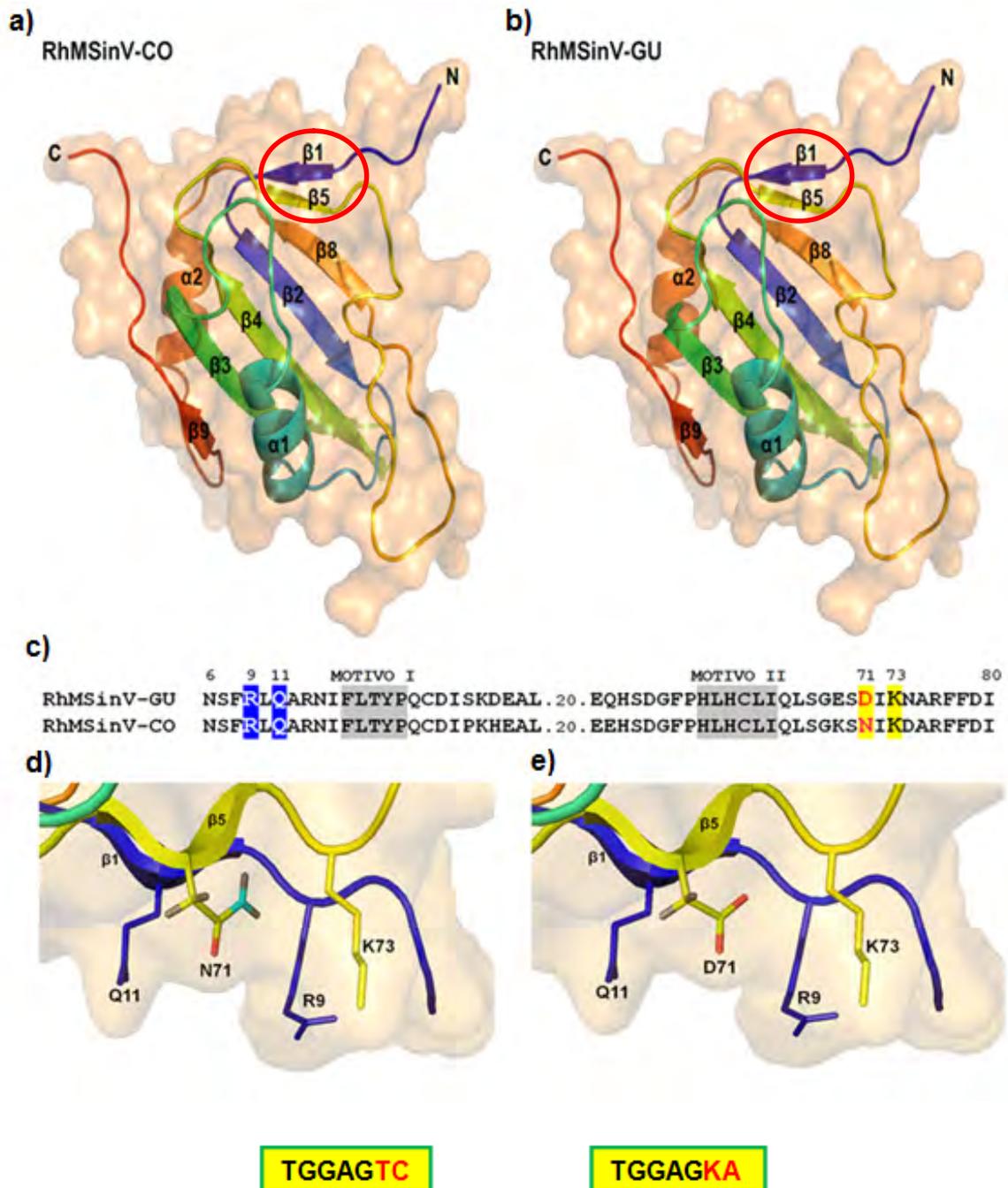


Figura 6.- Modelado por homología de las proteínas Rep de dos de los begomovirus aislados. Dominio de unión al DNA de a) RhMSinV-CO y b) RhMSinV-GU, en un círculo rojo se marca la formación de una mini hoja beta compuesta por las hebras $\beta 1$ y $\beta 5$. c) El alineamiento de los primeros 80 aminoácidos muestra diferencia en el residuo de la posición 71. d) y e) Los residuos 9, 11, 71 y 73 son críticos para el reconocimiento y la alta afinidad de la unión de Rep al DNA, estos se encuentran expuestos hacia la superficie del dominio de unión al DNA.

A partir del DNA obtenido de los protoplastos transfectados con los componentes virales se llevó a cabo un análisis de Southern blot con la sonda específica para el componente A (ver Métodos).

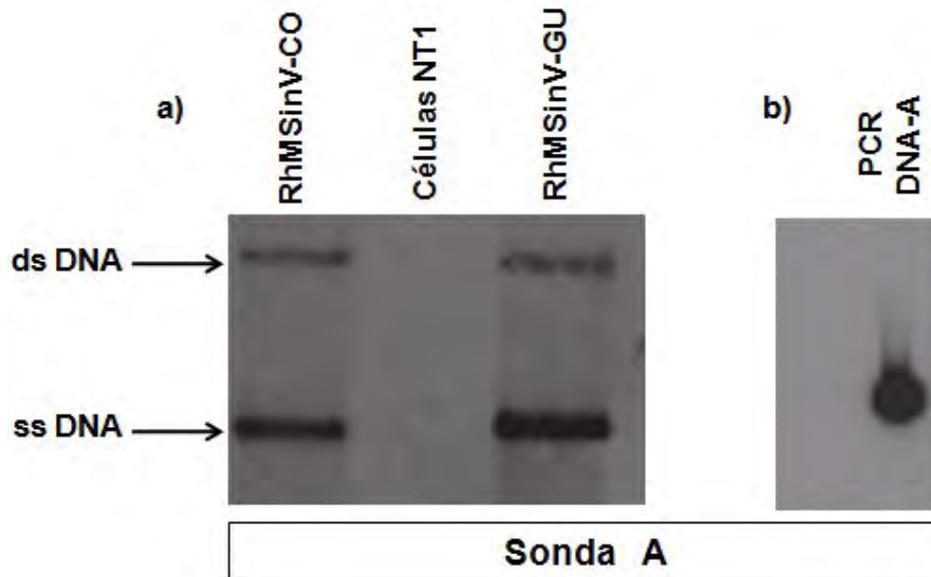


Figura 7.- Ensayo de replicación transitoria de los componentes A de RhMSinV-CO y RhMSinV-GU en protoplastos de *Nicotiana tabacum*. La hibridación se realizó con una sonda derivada del gen Rep obtenida por PCR con los primers SL2060-For/SL2200-Rev **a)** DNA viral en las diferentes formas replicativas del componente A de los virus RhMSinV-CO y RhMSinV-GU, en el carril de las células NT1 no transfectadas con DNA viral no se observa hibridación con la sonda. Cada línea contiene 10 µg de DNA total extraído del cultivo de protoplastos electroporados. **b)** Control positivo. PCR del DNA-A de los virus usados. Se indican las formas replicativas; dsDNA: DNA de cadena doble, ssDNA: DNA de cadena sencilla.

Mediante este análisis se detectó DNA viral generado *de novo* tanto en el caso del aislado original de Sinaloa como en el de la nueva cepa aislada en Colima, confirmando así que los DNA-As de ambos virus se pueden autorreplicar en las células vegetales (Figura 7a, líneas 1 y 3). Como control negativo se utilizaron células NT1 sin electroporar (Figura 7a, línea 2) y como control de hibridación se utilizó el producto de PCR obtenido con los primers SL2060-For/SL2200-Rev, usado como sonda (Figura 7b).

Incompatibilidad replicativa entre RhMSinV-GU y RhMSinV-CO

Para responder la cuestión relativa a la compatibilidad replicativa de las cepas de RhMSinV realizamos experimentos de replicación transitoria utilizando los dos componentes genómicos de RhMSinV-GU yPepGMV, y el componente A de RhMSinV-CO. Diferentes combinaciones de DNA-A+DNA-B fueron electroporadas en protoplastos de *N. tabacum*, y se evaluó la replicación del DNA viral mediante Southern blot, utilizando sondas específicas para los componentes A y B, respectivamente (Figura 8).

El análisis por Southern blot detectó replicación del componente B en los protoplastos electroporados con las mezclas homólogas (i.e., RhMSinV-GU (A+B), y PepGMV(A+B) (Figura 8a); sin embargo en las combinaciones heterólogas (i.e., RhMSinV-GU B + RhMSinV-CO A y PepGMV B + RhMSinV-CO A) se observaron diferencias. Por una parte, no se detectó replicación del DNA-B en la mezcla RhMSinV-CO A+ RhMSinV-GU B, que tienen diferentes iterones y Rep-SPD's. En contraste, la replicación del DNA –B sí fue detectada en la combinación RhMSinV-CO A+PepGMV-B, lo que es consistente con el hecho de que estos virus poseen

iterones similares, e SPD's en sus proteínas Rep (Figura 5 y 8a). Las células NT1 sin electroporar fueron utilizadas como control negativo (Figura 8a, carril 5) y como control positivo de la hibridación, el producto de PCR obtenido con los primers (BC1-290-For/BC1-80-Rev) (Figura 8b).

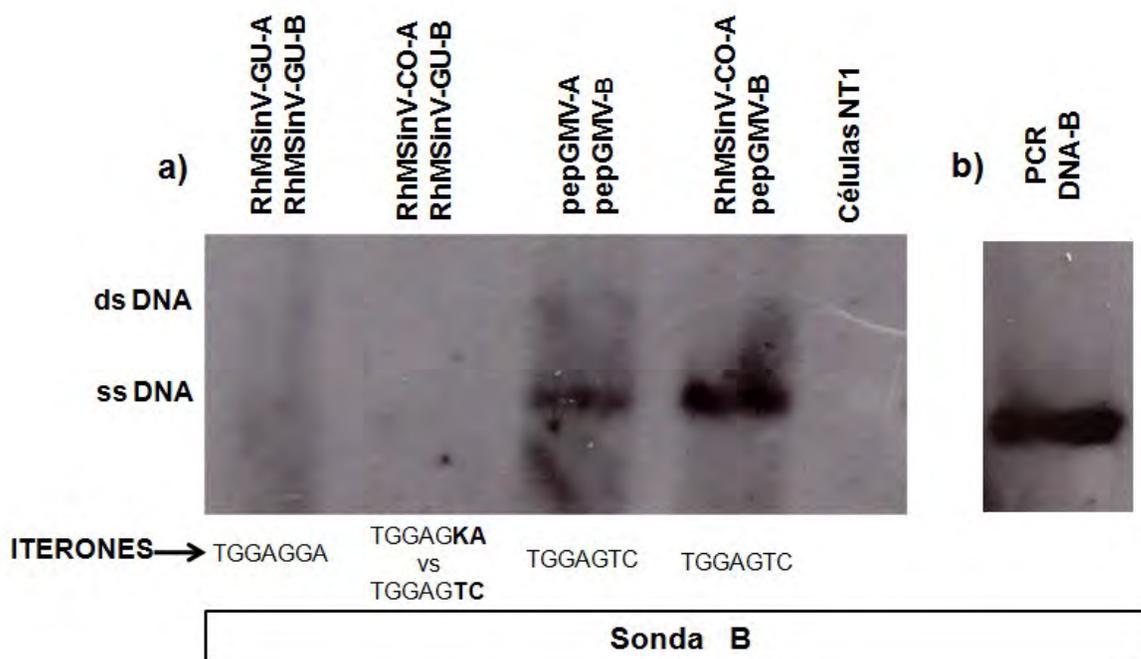


Figura 8. - Southern blot para detectar replicación del componente B de RhMSinV-GU y PepGMV en protoplastos de *Nicotiana tabacum*. La hibridación se efectuó con una sonda B obtenida por PCR con los iniciadores BC1-290-For/BC1-80-Rev. **a)** Detección de DNA-B en las cuatro combinaciones de componentes genómicos de RhMSinV-GU, RhMSinV-CO y PepGMV. Cada carril contiene 10 µg de DNA total extraído del cultivo de protoplastos electroporados y sin electroporar (NT1). La secuencia consenso de los iterones correspondientes a cada mezcla de componentes se muestra en la parte inferior de cada carril. **b)** Control positivo. PCR del DNA-B de los virus usados en éste Southern blot. Formas replicativas, dsDNA: DNA de cadena doble, ssDNA: DNA de cadena sencilla.

Experimentos de seudorecombinación en plantas

Los resultados de los ensayos de replicación en protoplastos mostraron que virus relacionados que difieren en sus potenciales determinantes de especificidad no se complementan en replicación, en tanto que virus lejanamente relacionados, como RhMSinV-CO y PepGMV, que poseen iterones con una secuencia similar y una proteína Rep con los mismos SPD's potenciales, sí se complementan replicativamente en células aisladas. Una cuestión que surge es si estos virus podrían ser capaces de formar pseudo-recombinantes viables en un huésped común e inducir una infección sistémica en la planta.

Para examinar esta cuestión y confirmar el resultado de la incompatibilidad replicativa de RhMSinV-GU y RhMSinV-CO, y la compatibilidad replicativa de PepGMV y RhMSinV-CO, llevamos a cabo experimentos de intercambio de componentes genómicos inoculando por métodos biobalísticos a plantas de *N. benthamiana*. Las combinaciones homólogas inoculadas fueron: RhMSinV-GU (A+B), PepGMV (A+B) y las heterólogas: RhMSinV-CO (DNA-A) + RhMSinV-GU (DNA-B) y RhMSinV-CO (DNA-A) + PepGMV (DNA-B).

El desarrollo de síntomas se observó a los 7 días post-inoculación (dpi) (Figura 9a) tanto en las plantas que fueron bombardeadas con las combinaciones homólogas como en las plantas inoculadas con la combinación RhMSinV-CO (DNA-A) y PepGMV (DNA-B). Los síntomas de amarillamiento, mosaicos y el rizamiento de las hojas se hicieron más evidentes a los 14 dpi (Figura 9b).

En contraste, las plantas bombardeadas con la combinación heteróloga del DNA-A de RhMSinV-CO + el DNA-B de RhMSinV-GU, y los controles negativos (plantas

sin inocular), no mostraron síntoma alguno incluso hasta los 21 dpi (Figuras 9a y 9b). La detección por PCR del DNA viral fue positiva para las plantas con sintomatología pero no así para las que se mantuvieron asintomáticas.

Discusión

En el presente trabajo reportamos la caracterización genómica de dos nuevas cepas y dos variantes de begomovirus previamente descritos, y la obtención de secuencias parciales correspondientes a otros dos begomovirus, todos ellos aislados de malezas colectadas en un terreno agreste de Manzanillo, Colima. Al concluir la caracterización de los aislados begomovirales, el estudio se enfocó principalmente en las cepas de RhMSinV aisladas en Colima y en Sinaloa, que perteneciendo a la misma especie viral presentan, sin embargo, pequeñas pero significativas diferencias en el origen de replicación viral y en la secuencia de la proteína que inicia la multiplicación del virus. Esas diferencias llevaron a la presunción de que las cepas RhMSinV-CO y RhMSinV-GU podrían ser incompatibles en replicación, lo cual decidimos examinar en dos tipos de experimentos diferentes: 1) ensayos de replicación transitoria en protoplastos de una línea celular de tabaco (NT1), y 2) experimentos de intercambio de componentes genómicos de virus diferentes (pseudo-recombinación) inoculando por biobalística a plantas de *Nicotiana benthamiana*.

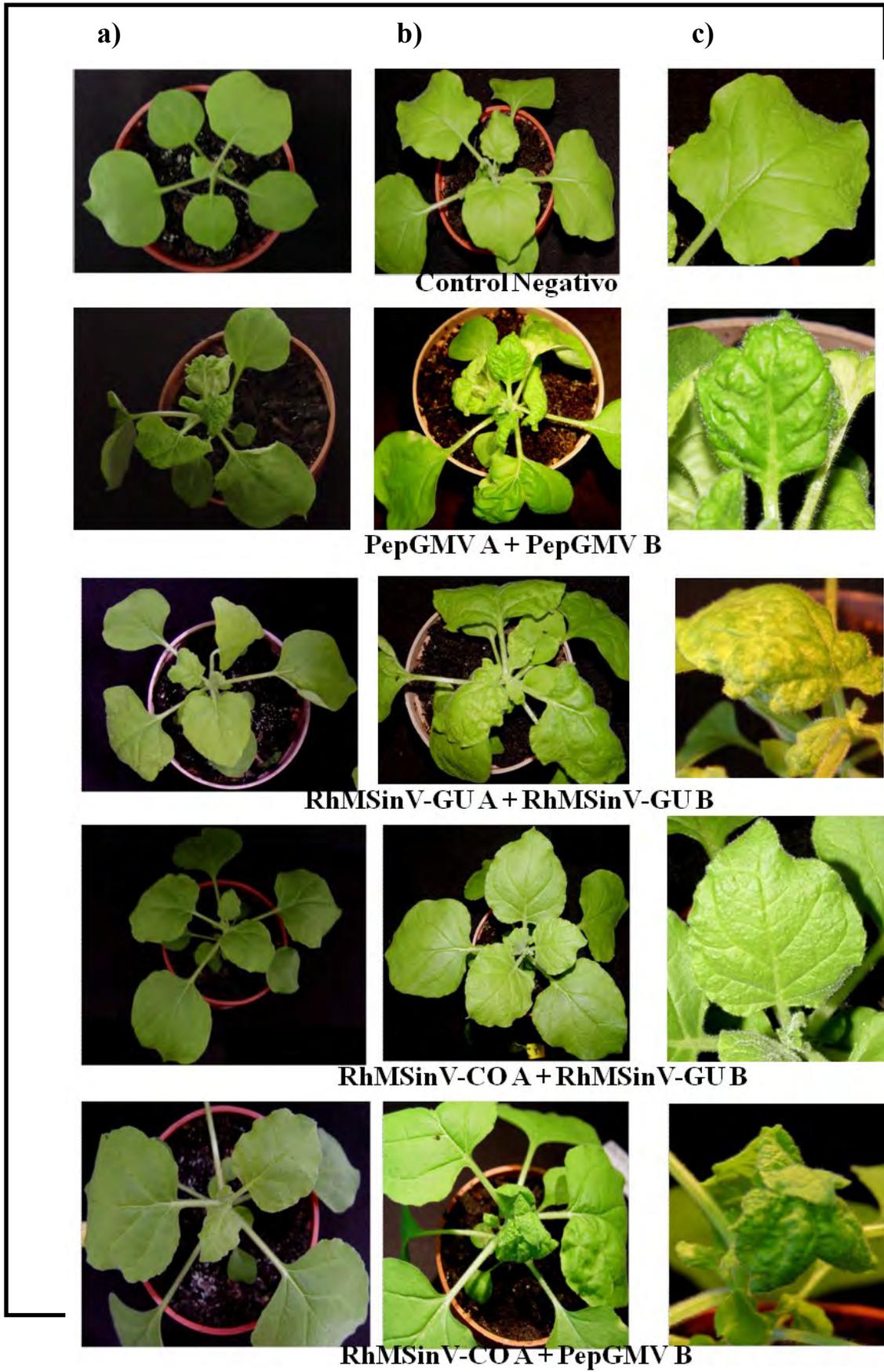


Figura 9. – Inducción de síntomas en plantas de *Nicotiana benthamiana* infectadas con los virus PepGMV, RhMSinV-GU y RhMSinV-CO. Las combinaciones homólogas y heterólogas de los componentes de PepGMV, RhMSinV-GU y RhMSinV-CO están indicadas en la parte inferior de cada imagen. Se muestran las plantas a los **a)** 7 dpi, **b)** 14 dpi y **c)** el acercamiento a la hoja de una planta a los 14 dpi. Las plantas que presentan síntomas de infección son las inoculadas con los DNAs homólogos de PepGMV y RhMSinV-GU (i.e., A +B), así como las plantas bombardeadas con la combinación RhMSinV-CO A + PepGMV B. Mientras que las asintomáticas son las inoculadas con la combinación RhMSinV-CO A + RhMSinV-CO B y las utilizadas como control negativo (sin bombardear).

Los resultados obtenidos en los experimentos de replicación transitoria en protoplastos (Figura 8) y de infectividad de pseudo-recombinantes en plantas conducen a la conclusión de que la nueva cepa aislada en este estudio (RhMSinV-CO) es incompatible en replicación con la cepa prototipo de la especie, RhMSinV-GU.

La diferencia en el residuo 71 de la proteína Rep de estos dos begomovirus parece ser un factor determinante en el reconocimiento específico de sus iterones cognados, ya que los presuntos SPDs localizados en el extremo N.terminal de las proteínas Rep de RhMSinV-GU y RhMSinV-CO son idénticos .

Se han publicado diversos experimentos que muestran la capacidad de formar pseudorecombinantes viables entre virus que contienen la misma secuencia de iterones y similares SPD's en sus proteínas Rep, como en el caso de *Tomato mottle Taino virus* (ToMoTV) y *Potato yellow mosaic virus* (PYMV), que formaron pseudorecombinantes infecciosos); estos virus comparten los mismos sitios de unión a la proteína Rep (i.e. AYTGGGGG); sin embargo no mostró la misma

habilidad con *Tomato mottle virus* (ToMoV), incluso cuando sus proteínas Rep y REn mostraban el más alto porcentaje de similitud comparándolo con otros virus, pero los iterones entre estos virus son diferentes (AYTGGGGG vs ATTGGAGT) (Ramos *et al.*, 2003).

Por otra parte Gregorio-Jorge *et al.*, 2010, investigaron la compatibilidad en replicación de dos aislados de *Euphorbia mosaic virus* (EuMV) que exhiben diferentes iterones y Rep-SPDs. Experimentos de pseudorecombinación con los dos virus (*Euphorbia mosaic virus- Peninsule Yucatan* y *Euphorbia mosaic virus- Jalisco*) demostraron que ambos virus no producían infección en plantas de *N. benthamiana*. Más recientemente, Silva *et al.*, 2014 demostraron la formación de pseudorecombinantes en todas las posibles combinaciones de genomas virales de *Tomato severe rugose virus* (ToSRV) y *Tomato rugose mosaic virus* (ToRMV). Ambos virus comparten el mismo DNA-B. Las secuencias completas de cada DNA-B comparten un 98.2% de identidad en secuencia además de la alta identidad de sus regiones comunes (96.2%) las cuales albergan las mismas secuencias iterativas (i.e. GGTAG) y sus proteínas Rep contienen idénticos SPD's en las cuatro posiciones predichas por Londoño *et al.*, (2010).

Los trabajos mencionados anteriormente y los análisis comparativos del origen de replicación de geminivirus y sus proteínas Rep concuerdan con los resultados presentados aquí y proveen una pista más acerca de la co-existencia de diferentes entidades virales en un mismo hospedero.

Referencias

- Adams, M.J.**, King, A.M.Q. & Carstens, E.B. (2013). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Arch Virol* **158**, 2023-2030.
- Ambrish, R.**, Kucukural, A. & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725-738.
- Argüello-Astorga, G.R.**, Guevara-González, R.G., Herrera-Estrella, L.R. & Rivera Bustamante, R.F. (1994). Geminivirus replication have a group-specific organization of iterative elements: A model for replication. *Virology* **203**, 90-100.
- Argüello-Astorga, G.R. & Ruiz-Medrano, R.** (2001). An iteron-related domain is associated to motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. *Arch Virol* **146**, 1465-1485.
- Argüello-Astorga, G.R.**, Ochoa-López, L., Kong, Ling-Jie., Orozco, M., Settlage, S.B., Hanley-Bowdoin, L. (2004). A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. *J Virol* **78**, 4817-4826.
- Bagewadi, B.**, Chen, S., Lal, S.K., Choudhury, N.R. & Mukherjee, S.K. (2004). PCNA interacts with Indian mung beans yellow mosaic virus rep and downregulates Rep activity. *J Virol* **21**, 11890-903.
- Böttcher, B.**, Unseld, S., Ceulemans, H., B-Russell, R., Jeske, H. (2004). Geminiate Structures of African Cassava Mosaic Virus Geminiate Structures of African Cassava Mosaic Virus. *J Virol* **78**, 6709–6714.
- Bridson, R.W.**, Patil, B.L., Bagewadi, B., Nawaz-ul-Rehman, M.S., Fauquet, C.M. (2010). Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evol Biol* **10**, 97.
- Campos-Olivas, R.**, Louis, J.M., Clerot, D., Gronenborn, B. & Gronenborn, A.M. (2002). The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. *Proc. Natl. Acad. Sci. USA* **99**, 10310-10315.

- Carrillo-Tripp, J.** Lozoya-Gloria, E., Rivera-Bustamante, R.F. (2007) Symptom remission and specific resistance of pepper plants after infection by Pepper golden mosaic virus. *Phytopathology*. **97**, 51-57.
- Davis, I.W.,** Feaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snowyink, J., Richardson, J.S. & Richardson, D.C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, 375-83.
- Dellaporta, S.L.,** Wood, J. & Hicks, J.B. (1983) A plant DNA miniprep: version II. *Plant Mol Biol Rep* **1**, 19-21.
- Fauquet, C.M.** & Stanley, J. (2005). Revising the way we conceive and name viruses below the species level: a review of geminivirus taxonomy calls for standardized isolate descriptors. *Arch Virol* **150**, 2151-2179.
- Fauquet, C.M.,** Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., & Zhou, X. (2008). Geminivirus strain demarcation and nomenclature. *Arch Virol* **153**, 783–821.
- Fontes, E.P.,** Eagle, P.A., Sipe, P.S., Luckow, V.A. & Hanley-Bowdoin, L. (1994). Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J Biol Chem* **269**, 8459-8465.
- Gregorio-Jorge, J.,** Bernal- Alcocer, A., Bañuelos-Hernández, B., Alpuche-Solís, A.G., Hernández-Zepeda, C., Moreno-Valenzuela, O., Frías-Treviño, G., Argüello-Astorga, G.R. (2010). Analysis of a new strain of *Euphorbia mosaic virus* with distinct replication specificity unveils a lineage of begomoviruses with short Rep sequences in the DNA-B intergenic region. *Virology* **7**, 275.
- Gutierrez, C.** (2000). DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.* **19**, 792–799.
- Ha, C.,** Coombs, S., Revill, P., Harding, R., Vu, M., Dale, J. (2006). *Corchorus* yellow vein virus, a New World geminivirus from Old World. *J Gen Virol* **87**, 997-1003.
- Ha, C.,** Coombs, S., Revill, P., Harding, R., Vu, M., Dale, J. (2008). Molecular characterization of begomoviruses and DNA satellites from Vietnam; additional evidence that the New World geminivirus were present in the Old World prior to continental separation. *J Gen Virol* **89**, 312-326.
- Hanley-Bowdoin, L.,** Settlege, S.B. & Robertson, D. (2004). Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. *Mol. Plant Pathol.* **5**, 149-156.

- Hanley-Bowdoin, L., R. Bejarano, E., Robertson, D. & Mansoor, S.** (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature* **11**, 777-788.
- Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J. & Gronenborn, B.** (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Res* **23**, 910–916.
- Jeske, H.** (2009). Geminiviruses. *Curr Top Microbiol-Immunol* **331**, 185-226.
- Kong, Ling-Jie., Orozco, B.M. Roe, J.L., Nagar, S., Ou, S., Feiler, H.S., Durfee, T., Miller, A.B., Grusse, W., Robertson, D., Hanley-Bowdoin, L.** (2000). A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. *EMBO J.* **19**, 3485-3495.
- Koonin, V.E. & Llyina, V.T.** (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *J Gen Virol* **73**, 2763-2766.
- Lefevre, P., Harkins, G.W., Lett, Jean-Michel, Briddon, R.W., Chase, M.W., Moury, B., Martin, D.P.** (2011). Evolutionary Time-Scale of the Begomoviruses: evidence from integrated sequences in the *Nicotiana* genome. *PLoS ONE*, 6(5): e19193.
- Llyina, V.T. & Koonin, V.E.** (1992). Conserved sequence motif in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archeobacteria. *Nucleic Acids Res* **20**, 3279-3285.
- Londoño, A., Riego-Ruiz, L. & Argüello-Astorga, G.R.** (2010). DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. *Arch Virol* **155**, 1033–1046.
- Mauricio-Castillo, J.R.** (2006). Métodos moleculares que potencian el descubrimiento de nuevas especies de begomovirus y la detección de infecciones mixtas. Tesis de Maestría. Instituto Potosino de Investigación Científica y Tecnológica, San Luis Potosí, SLP.
- Moffat, A.** (1999). Geminiviruses emerge as serious crop threat. *Science* **286**, 1835.
- Nash, T.E., Dallas, M.B., Reyes, M.I., Buhrman, G.K., Ascencio-Ibañez, J.T. & Hanley-Bowdoin, L.** (2011). Functional analysis of a novel motif conserved across geminivirus Rep proteins. *J. Virol.* **85**, 1182–1192.

- Orozco, B.M.**, Gladfelter, H.J., Settlege, S.B., Eagle, P.A., Gentry, R.N. & Hanley-Bowdoin, L. (1998). Multiple cis elements contribute to geminivirus origin function. *Virology* **242**, 346-356.
- Orozco, B.M.**, Miller, A.B., Settlege, S.B. & Hanley-Bowdoin, L. (1997). Functional domains of a geminivirus replication protein. *J. Biol. Chem.* **272**, 9840– 9846.
- Padidam, M.**, Maxwell, D.P. & Fauquet, C.M. (1997). A proposal for naming geminivirus. *Arch virol* **142**, 2553-62.
- Ramos, P.L.**, Guevara-González, R.G., Peral, R., Ascencio-Ibañez, J.T., Polston, J.E., Argüello-Astorga, G.R., Vega-Arreguín, J.C. & Rivera-Bustamante, R.F. (2003). *Tomato mottle Taino virus* pseudore-combines with PYMV but not with ToMoV: implications for the delimitation of cis- and trans-acting replication specificity determinants. *Arch Virol* **148**, 1697–1712.
- Rybicki, E.P.** (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch Virol* **139**, 49–77.
- Sambrook, J. & Russell.** (2001). *Molecular cloning: a laboratory manual*. 3rd ed. Cold Spring Harbor Laboratory press, Cold Spring Harbor, NY.
- Sanderfoot, A.A. & Lazarowitz, S.G.** (1996). Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. *Trends Cell Biol* **9**, 353-8.
- Sankara, K.R. & Prakash, H.** (1995). A simple method for the isolation of plant protoplasts. *J. Biosci* **20**, 645-655.
- Silva, F.N.**, Lima, A.T.M., Rocha C.S., Castillo-Urquiza, G.P., Alves-Júnior, & Zerbini, F.M. (2014). Recombination and pseudorecombination driving the evolution of the begomoviruses *Tomato severe rugose virus* (ToSRV) and *Tomato rugose mosaic virus* (ToRMV): two recombinant DNA-A components sharing the same DNA-B. *Virology Journal* **11**, 66.
- Sippl, M.J.**, (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins* **4**, 355-62.
- Stanley, J.**, Bisaro, D.M., Briddon, R.W., Brown, J.K., Fauquet, C.M., Harrison, B.D., Rybicki, E.P. & Stenger, D.C. (2005). *Geminiviridae*. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), *Virus Taxonomy*, VIIIth Report of the ICTV, 301-326. Elsevier/Academic Press, London.

- Tamura, K.**, Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.
- Varsani, A.**, Martín, D.P., Castillo-Navas, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Zervini, F.M. & Brown, J.K. (2014). Revisiting the classification of curtovirus based on genome-wide pairwise identity. *Arch Virol* **159**, 1873-1882.
- Wiederstein & Sippl.** (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research* **35**.
- Zhang, W.**, Olson, N.H., Baker, T.S., Faulkner, L., Agbandje-McKenna, M., Boulton, M.I., Davies, J.W. & McKenna, R. (2001). Structure of the Maize streak virus geminate particle. *Virology* **279**, 471–477.
- Zhang, Y.** (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40.

Material Suplementario

Tabla S1.- Oligonucleótidos universales utilizados para amplificar DNA-A y DNA-B de begomovirus y oligonucleótidos diferenciales diseñados como iniciadores para amplificar genomas virales pertenecientes a la especie RhMSinV

<i>Oligonucleótido</i>	<i>Secuencia</i>
SL2150-Forward	5'GACGGCRTTGGY GTCTTTGGC 3'
YMAC-N-Reverse	5' ACAAGCTTWGASGCATGNGTACATGCCA 3'
DGRS-Forward	5' CAGATCTCCTCCTCTAGCAGWTCTNCCGTC 3'
BV1310-Forward	5'AGGWACRGTNAARATYGARCGTGT 3'
BC1290- Reverse	5'CCSATMAGRTGYAACATAGATCTCC 3'
BC1-1840- Forward	5' GATCTAGAGYRCTYCTKGAAGCCCA 3'
BV1-310-Reverse	5' GAGAATTCGTTCAATYTTNACNGTWCCYTTRAAACG 3'
SL-CP-225-Forward	5' ATG SYTAAGMNGATGCCCCATGGCG 3'
SL-2310-Reverse	5' ATCMGAGTCGMMGAGAGGAACA 3'
Dif RhMSV-550-Forward	5' AATTTAAAGACTTAGGCGGTGCT 3'
Dif BVRhMSV-56-Reverse	5' TGTGAATACTCCTTCTATTTTAGAGCT 3'
Dif RhMSV-940-Forward	5' GATTTTGACCGTGGGCGCCCTAAATGGT 3'
Dif BVRhMSV-275-Reverse	5' GAAGCGCTTATTAGAAAGGGTTGTC 3'
Dif BV-RhMSV180 Reverse	5' -CGATGAGCTTATGCATAAATTGCCG - 3'
RhSinIR-Forward-2285	5'GAGAAGTGAAGAAGATGATAGGTTG 3'
RhSinBV1-Reverse-630	5' GGTTCACTCTTACCAAGAAGTGA 3'
RhMCOA Forward	5'ACCCCTATGGATTTTGGCC 3'
RhMCOA Reverse	5'ACCATAGGGTCTCCTATCC 3'
SL2060-Forward	5' CGGCGTTRRRTARACGTCGTC 3'
SL2200-Reverse	5' CAGAATTCGACMTYACTCACCC 3'
BC1-290-Forward	5'GAGAARTAGTGGAGATCTATGTRCA 3'
BC1-80-Reverse	5'GGTCTAGAGAYGARTATCARYTDTNCATGA 3'
CPc70 For-Forward	5'-GGT TGT GAA GGN CCN TGT AAG GTY CA -3'
CP- MYR-Forward	5' -GGCCYATGTAYAGRAAGCCAGGAT - 3'
SL-1890 Forward	5' -CCAAAATAMTCATCMGCCCA - 3'
CP-550 Reverse	5' -CTTKARCTTGATRTTYTCRTCCATCC - 3'
YMAC-Forward	5' -CTAAGCTTGTAYATGGCATGTACNCATGC- 3'
SL-2100 Forward	5' -CTAAGCTTGYCCGGAYTACAATAATCACC - 3'
BCI-470 Forward	5' -GATCTAGACNGTYGGTGCCCKGAANGG - 3'

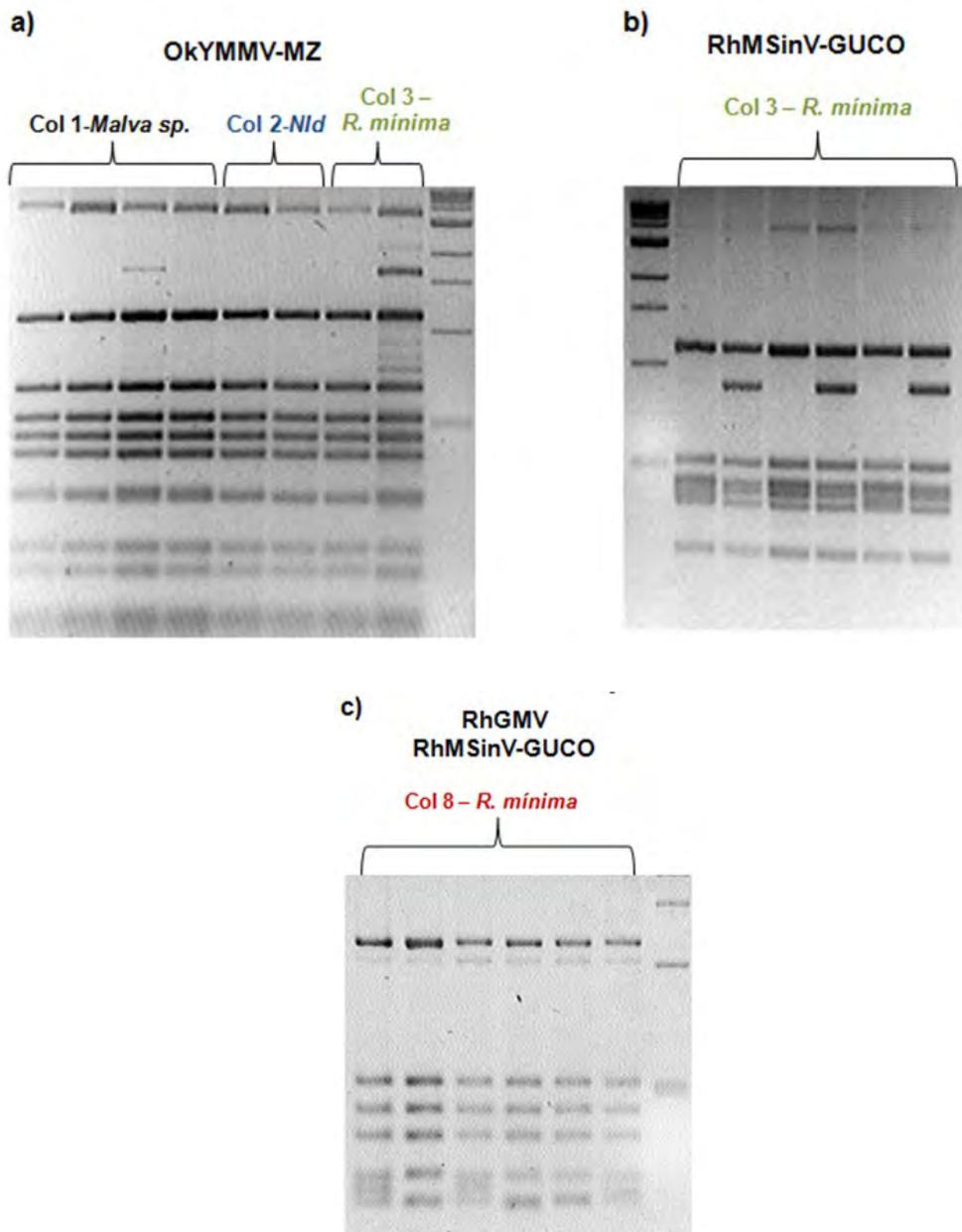


Figura S1.- Análisis de los polimorfismos en la longitud de fragmentos de restricción (RFLP). a) Patrón de restricción de clonas con amplicones YMAC-Forward/DGRS reverse; los ocho carriles corresponden al patrón de la nueva cepa de OkYMMV, la cual se encuentra en las muestras Col 1-*Malva sp.*, Col 2- *Nid* y Col 3- *R. minima*. **b)** RFLP de la nueva variante RhMSinV-GUco, encontrada en el extracto Col 3- *R. minima*, Col 8- *R. minima* y Col 9- *R. minima*. **c)** Análisis RFLP de las clonas conteniendo el fragmento amplificado con los iniciadores BV310 forward/BC1 290 reverse correspondiente al componente B de dos virus encontrados en la muestra Col 8- *R.*

Preparación de protoplastos de células NT1

- 1.-** En campana de flujo laminar, transferir 30 ml del cultivo de células NT1 (3-4 días de cultivo) a un tubo cónico y centrifugar a 2000 rpm por 2 minutos a temperatura ambiente.
- 2.-** Retirar el sobrenadante cuidadosamente.
- 3.-** Agregar 15-20 ml de Manitol 0.4 M y resuspender la pastilla suavemente.
- 4.-** Centrifugar a 2000 rpm por 2 minutos y retirar el sobrenadante.
- 5.-** Agregar 2 volúmenes de solución enzimática por cada volumen de células.
- 6.-** Mezclar suavemente hasta que toda la pastilla esté disuelta.
- 7.-** Incubar a 25°C con agitación constante a 65 rpm por unos 35 minutos.
- 8.-** Al término de este tiempo, tomar una alícuota (20 μ l) de la suspensión y observar al microscopio. En este tiempo aproximadamente el 90-95% de las células adquieren una forma redondeada. Nota: No se recomienda dejarlos en la solución enzimática más de una hora.
- 9.-** Centrifugar a 1500 rpm por 2 minutos. Retirar la mayor cantidad posible de sobrenadante (decantar). Tener cuidado ya que la pastilla es frágil y las células se pueden resuspender fácilmente.
- 10.-** Agregar 15 ml de Manitol 0.4 M. Mezclar suavemente y centrifugar a 1500 rpm por 2 minutos. Retirar la mayor cantidad posible de sobrenadante (decantar).
- 11.-** Repetir el lavado con Manitol 0.4 M.
- 12.-** Lavar las células con Buffer de electroporación. Agregar 15 ml de Buffer de electroporación y resuspender las células nuevamente.
- 13.-** Centrifugar a 1500 rpm por 2 minutos. Retirar lentamente el sobrenadante.
- 14.-** Repetir el lavado con Buffer de electroporación.
- 15.-** Después del segundo lavado con Buffer de electroporación. Resuspender suavemente las células con 1 a 2 volúmenes de buffer de electroporación.

Soluciones utilizadas:

- a) Manitol 0.4 M
- b) Solución Enzimática

Reactivo	Cantidad (g)/50 ml	Conc. final
Manitol	3.64	0.4 M
MES	0.213	20 mM
Celulasa	0.5	1%
Pectinasa	0.05	0.1%

- c) Buffer de electroporación

Reactivo	Cantidad (g)/500 ml	Conc. final
NaCl	4	0.8%
KCl	0.1	0.02%
KH ₂ PO ₄	0.1	0.02%
Na ₂ HPO ₄	5.5	0.11%
Manitol	36.43	0.4M

Ajustar el pH a 6.5 con HCL y esterilizar

Lavados

Después de terminar el tiempo de hibridación, realizar lavados a la membrana en un tubo de hibridación de la siguiente manera:

- 1.- Con 50 ml de solución #1 durante 15 min. Realizar este lavado 2 veces.

Solución #1 para 100 ml
SSC 2x 10 ml de SSC 20X
SDS 0.1% 500 µl de SDS 20%

- 2.- Hacer 1 lavado con 50 ml de solución #2 durante 30 min.

Solución #2 para 50 ml
SSC 1X 2.5 ml de SSC 20X
SDS 0.1% 250 µl de SDS 20%

- 3.- Por último, hacer 1 lavado con 50 ml de solución #3 durante 10 min.

Solución #3 para 50 ml
SSC 0.1% 250 µl de SSC 20x
SDS 0.1% 250 µl de SDS 20%



Figura S2.- Generación de protoplastos. a) Células del mesófilo de *Nicotiana tabacum* crecidas en suspensión (NT1), antes del tratamiento para la generación de protoplastos. **b)** Protoplastos generados después de 45 minutos con el tratamiento enzimático. Se observa su forma redondeada característica.

Tabla S2.- Nombres, acrónimos y números de acceso en GenBank de las secuencias de los geminivirus utilizados en el análisis filogenético

Nombre del virus	Acrónimo	Número de Acceso ADN-A
<i>Abutilon mosaic virus</i>	AbMV	NC_001928.2
<i>Beet curly top virus</i>	BCTV	NC_001412.1
<i>Bean dwarf mosaic virus</i>	BDMV	NC_001931.1
<i>Cabbage leaf curl virus</i>	CaLCuV	NC_003866.1
<i>Corchorus golden mosaic virus</i>	CoGMV	NC_009644.1
<i>Corchorus yellow vein virus</i>	CoYVV	NC_006358.1
<i>Cotton leaf crumple virus</i>	CLCrV	NC_004580.1
<i>Cotton leaf curl Multan virus</i>	CoLCMV	NC_004607.1
<i>Cucurbit leaf crumple virus</i>	CuLCrV	NC_002984.1
<i>Desmodium leaf distortion virus</i>	DesLDV	NC_008494.1
<i>Euphorbia leaf curl virus</i>	EuLCV	NC_005319.1
<i>Euphorbia mosaic virus-Jamaica</i>	EuMV-Jamaica	FJ407052.1
<i>Euphorbia mosaic virus-Perú</i>	EuMV-Perú	AM886131.1
<i>Euphorbia mosaic virus-Puerto Rico</i>	EuMV-PR	AF068642.3
<i>Euphorbia mosaic virus-Yucatan Peninsula</i>	EuMV-YP	NC_008304.1
<i>Euphorbia yellow mosaic virus</i>	EuYMV	NC_012553.1
<i>Melon chlorotic leaf curl virus</i>	MCLCuV	NC_003865.1
Okra yellow mosaic Mexico virus	OkYMMV	NC_014066.1
Papaya leaf curl virus	PaLCV	AJ436992.1
Rhynchosia golden mosaic virus	RhGMV	NC_010294.1
Rhynchosia golden mosaic Yucatan virus	RhGMYV	NC_012481.1
Rhynchosia mosaic Sinaloa virus-CO	RhMSinV-CO	Pendiente
Rhynchosia mosaic Sinaloa virus-Guasave	RhMSinV-GU	DQ406672.1
Rhynchosia mosaic Sinaloa virus-Guasave-Colima	RhMSinV-GUCO	Pendiente
<i>Sida golden mosaic Costa Rica virus</i>	SiGMCRV	NC_004657.1
<i>Sida mosaic Sinaloa virus</i>	SiMSinV	DQ520944.1
<i>Sida yellow vein virus</i>	SYVV	NC_004661.1
<i>Squash leaf curl virus</i>	SLCV	NC_001936
<i>Squash mild leaf curl virus</i>	SMLCV	NC_004645.1

Tomato common mosaic virus	ToCoMV	NC_010835.1
Tomato mild yellow leaf curl Aragua virus	ToMYLCV	NC_009490.1
Tomato yellow leaf curl virus	TYLCV	X15656.1
Tomato severe leaf curl virus isolate Rioverde-SLP2	ToSLCuV-SLP	DQ347947.1

Componentes virales aislados de las muestras de Manzanillo 1-10

OkYMMV- Nueva cepa (MZ) <91% con otros aislados (2615 pb)

ACCGGATGGCCGCGGATTTTTTCCCCCCTTCGCGCATGTGCTCCTCTTTAATTTGAATTAAGCGCTCCGC
 TTTCGTTTTCGTCCAATCATGTAACGTCTGACGAGCTTAGATATTTTCAACGACTTCGGGCCTAAGTAGTTGTG
 TCGGTTATAAATTAAGAGGCTTTGGCCACTACATTTAATTCAAAATGCCTAAGCGCGATCTCCCATGGCGC
 TCGATCGCGGGAAC TACAAAGGTTAGTCGTAATGCTAACTATTTCGCCACGTGCAGGTATTGGGCCAAGAGTTA
 ACAAGGCGGCTGAATGGGTAAATAGGCCCATGTACAGGAAGCCCAGAATCTACCGGACGCTGCGAACGCCCGA
 CGTTCCAAGAGGATGTGAAGGCCCTTGCAAGGTCCAGTCCATGAACAGCGACACGATATCTCACATGTGGGT
 AAGGTGATGTGTATATCTGATGTCACTCGTGGCAATGGTATCACCCACCGTGTGGGTAAGCGTTTTTGTGTCA
 AGTCTGTGTATATTCTAGGTAATACTGGATGGACGAGAATATCAAGCTCAAGAACCACACGAACAGTGTGAT
 GTTCTGGTTGGTCAGAGACCGTAGACCGTATGGCACGCCTATGGATTTTGGCCAAGTGTTC AACATGTTTCGAC
 AATGAGCCTAGCACTGCTACGGTGAAGAACGATCTACGCGATCGTTACCAGGTCATGCACAAGTTCTACGGCA
 AGGTGACAGGTGGACAGTATGCCAGCAATGAACAGGCGATAGTCAAGCGTTTTCTGGAAGGTCAACAATTATGT
 GGTGTACAATCATCAGGAGGCTGGCAAGTATGAGAATCACACGGAGAACGCATTATTACTGTATATGGCATGT
 ACCCATGCCTCAACCCTGTATATGCAACGCTTAAGATCCGGATCTATTTTTACGATTTCGATCATGAATTAATA
 AAGTTTGAATTTTATTGAATGATTTTCCAGTACATGATTTACATACGATCTGTCTGTTGCGAAACGAACAGCT
 CTGATTACATTGTTAATTGAAATGACGCCTAATTGATCTAAATACAAATTA ACTAAGTGTCTCAACCTAGCTA
 AATAAGTCGACCCAGAAGCTGTGTCGATGTGTCGTCAGACTTGGAAATTCAGGTAGGCTTTGTGGAGATGCAA
 CGCTCTCCTCAGGTTGTGGTTGAACCGTATCTGAACGTGGTAAACTCTGGTTCTCGTGTACAACGGTTCCCTCC
 ACGTTGTACATCTTGAAATAGAGGGGATTTGGAACCTCCCAAATAAAAACGCCATTCTCTGCCTGAGGTACAG
 TGATGAGTTCCCCTGTGCGTAAATCCTTTACCATTTTTGGCGCAGTTGATATGTATGTATATGGAGCACCCGC
 ACTCTAGATCAATCCGCTTCCGCTGATGGCCCTCCTCTTGGCTTCCCTGTGTTCTATCTTGATAGAGGGCGG
 ATGTGAGGGTGATGAAGATCGCATTCTTGATGGTCCAGTTTTTGGAGTATGCGTTTTCTTTTGTGCGAGGAA
 AGCTTTATAACTGGCACCCCTCACCTGGATTGCAAAGCACGATTGATGGGATTCCCTCCTTTAATTTGAACTGGC
 TTGCCGTACTTACAATTTGACTGCCAGTCCTTTTGGGCCCTATCAATTCTTTCCAGTGCTTTAGCTTTAGAT
 AGTGCGGTGCGACGTCATCAATGACGTTATACTCCACTTCGTTTGAATAGACTCGTGGGTTGAAGCCCAGATG
 TCCACTTAGATAATTATGTGGGCCTAGCGCACGAGCCCACATTGTCTTCCCTGTCTCGAATCACCTTCTACG
 ATGATACTAACTGGTCTCTCAGGCCGCGCAGCGGAACCCCTTCCAAAATAATCATCCGCCACTCTTGATCT
 CGTCCGGAACGTTAGTGAAAGAGGAGAGTTGAAACGGAGGAACCCAGCGTTCCGGAGCCTTTGCGAATATTCT
 TTCTAGGTTGGTGCAGGATGTTATGATTCTGCAGGATAAAAATCTTTTTGGTTGTTCCCTCCTTCAAAAATCGTCATG
 GCAGCTTGAACAGACCCAGCATTCAACGCCTTGGCATATGAATCGTTAGCTGACTGCTGACCTCCTCTAGCAG
 ATCTGCCGTGATCTGGAATTGCCCCATTCAATCGTATCTCCGTCTTGTGATGTAGGACTTGACGTCGGA
 GCTGGATTTATCTCCCTGTATGTTCCGGATGGAAATGTGCTGACCTGGTTGGGGACACCAGATCGAAGAATCTG
 TTATTTCGTGCACTGGTACTTGCCTTCGAATTGGATGAGAACGTGGAGATGAGGCTCCCCATTCTCATGAAGCT

CTCTGCAGATCTTGATAAATTTCTTGTTAACAGGAGTTTCTAGGGTTTGTAAATTGGGAAAGTGCTTCTTCTTT
GCTGAGAGAGCACTGGGGATATGTGAGGAAATAGTTTTTTTACTGAACCTCTAAATTTTCTAGGCTGTGGCATT
TTTTGTAATAAGATGGGGTACTCCAGATGAGCTCCTCTCAAACCTTGCTGTATGTTTTGGAGTATTAGAGTACTA
TATATACTAAAACCTCATTAACGGATTTGCAACACGTGGCGCCATCCGATATAATATT

DNA-A RhMSinV-Col 2575 pb

ACCGGATGGCCGCGATTTTTTTGGAGTCTGCGTGGGACCGCGCATAAAATGAAAATTAACCTGACCAATCATA
TACCGCCTGACGCGCTAGATATTTGTGACTTGGTGACCAAGTTGTGGCCTATATAACGACTAGAATTTTCA
TTTTATGCTTTAATTCAAAATGGTTAAGCGGGATGCCCATGGCGCCTATCTGCTGGGACCTCTAAGGTATCCC
GCTCTACCAATTATACACCTCGTGCAGGTATGGGCCCCAAATTTGATAAGGCCGCTGCTTGGGTGAACAGGCC
CATGTATAGGAAGCCCAGGATATATCGGACTTTGAGAAGCCCCGATGTCCCTAGAGGTTGTGAAGGGCCTTGT
AAAGTCCAGTCTTATGAGCAGCGGCATGATATCTCTCATGTTGGGAAGGTGATGTGTATTTCTGATATTACAC
GTGGTAATGGTATTACTCATCGTGTAGGAAAGCGTTTCTGCGTCAAGTCCGTGTACATTTTGGGCAAGGTATG
GATGGACGAGAATATCAAGCTCAAGAACCACACCAACAGCTGCATGTTTTGGTTGGTGAGGGATAGGAGACCC
TATGGTACCCCTATGGATTTTGGCCAGGTGTTCAACATGTTTCGATAACGAGCCCAGTACTGCTACCGTTAAGA
ACGATCTTCGTGATCGTTATCAAGTCATGCACAAGTTCTATGCTAAGGTTACGGGTGGACAGTATGCGAGCAA
CGAGCAGACGTTGGTCAAGCGTTTCTGGAAGGTCAACAACCTACGTGGTGTACAACCATCAAGAAGCTGGGAAA
TACGAGAATCATAACGAGAACGCTTTGTTATTGTATATGGCATGTACTCATGCCTCTAATCCTGTGTATGCGA
CATTGAAAATTCGGATCTATTTTTATGATTGATAACAAATTAATAAATTTTGAATTTTATTTTCATGATTCTC
GAGTACATCACTTACATAATGTTTATTTCGTCGCGAAAACGAACAGCTCTAATTACATTGTTTATGCTAATAACA
CCTAAGTTATCTAAATAAGACATTACAAGTAATTTGAATCTAGCTAAATATCTCTGCCAGAAGCTGTCATCG
AAGTCGTCCAGACTTGAAAATTGAAGTAGGCTTTGTGGAGAGCCAACGCTTTCCGTAGGTTGTGGTTGAACCG
GACTTCGATGTGGTATATCCTGGTTCTCGTGAATAACGGGTTCTCTACGTGGTGTATCCTGAAAAAGAGGGGA
TTTTGGAACCTCCCAAATAAAAACGGAACCTCTCTGCTTGAGCCGCAGTGATGCTCTCCCCTGTGCGTGAATCCA
TAGTCTGCGCAGTTAATGTGGAGGAATATTGAACAGCCACAGTTCAAGTCTATGCGTGTGACGAACAGCTC
TTTTTTTTAGCAATCCTGTGCTGTGCTTTGATAGAGGGGGGCTTCAAGGGTGACGAATTTAGCATTTTTGAGTG
TCCAAGCTCTTAGTGATGCATTTTCTCTTTGTTGAGGAAAGATATATAGCTGCTCCCCTCTCCTGGATTGCA
CAGCACGATTGATGGTATGCCACCTTTAATTTGAACTGGCTTTCCGTATTTGCAGTTTGTATTGCCAGTCCCTTT
TGGGCCCCTAGCAGCTCTTTCCAGTGCTTTAGCTTTAAATATTGCGGAGCTATGTCATCAATGACGTTATACT
CCACATCATTTGAATAGACCCTGGAATTGAAGTCGAGGTGACCACTCAAGTAATTATGTGGGCCTAGAGCACG
TGCCACATTGTTTTGCCTGTTGAGAATCACCTCAATTATGACACTAATAGGTCTTTCTGGCCGCGCAGCG
GAACTCCGACCAAAAATAATCATGAACCCAAGCTTTAATCTCGTCTGGAACGTTAGTAAAGGAGGAGAGTTGAA
ACGGAGGAGCCCATGGTTCCGGAGCCTTTTGGAAAGAGGCGCTCGATGTTTGACCTAACGTTATGATAACTAAC
AATAAATGTTTTGGGATCTCCAGCTTTTATTATGGCGAGAGCCTCTTCCGCACTTGCTGCATTAACCTGCGTTA
TGGTAGACGTCGTCCTTATTTGCCTTTGAACCCCCAGACACTTTGTATTCCCCGGATTCACAATAATCACCAT
CTTTGGTGATGTAATCCTTGACGGCATTGGTGTCTTTGGCTGCCTGAATGTTTGGGTGAAAATGGGCAGACCT
TCTGGGGTGACTGATGTGCAAAAATCTAGCATCCTTGATGTTAGACTTTTCTGAAAGTTGGATGAGACAGTGT
AAGTGTGGGAACCCGTCTGAATGTTTCTCTTTGCGACTCGTATGTACGTGGGTTTTGACGATTGACCATGGAA
GCGACTGAAGCAACTGAAGAGCTTCATGTTTGGGAATATCGCACTGGGGATATGTTAAGAAAATGTTTTCTGGC
TTGTAAGCGAAAAGAATTAGGGTTGCGTGGCATTTTTTGTGAATATGCGCTAGGACTCCAGCGGAGTCCCTCAA
CTTGAGTGCTATTTGCTGGAGTCTGGAGTCCCATTTATACTAGAAGGCTCTGGGGGACTCCAGGGGCAAAAG
CGCCATCCGTAATAATATT

RhMSinV- GUco (DNA-A 2585 pb)

ACCGGATGGCCGCGGATTTTTGGAGGATATGTGGGACCGCGCAATAAATGAAATTTGAACTGACCAATCACGT
TGCGCCTGACTAGCCTAGATATTTGTGACTTGGTGACCAGGTTGTGGGGCTATAAAAACGACGTGAAATTTTCAT
TTATTGCTTTAATTCAAAATGGTTAAGCGGGATGCCCGTGGCGCCTATCTGCTGGGACCTCTAAGGTATCCC
GCTCTACCAATTATACGCCTCGTTCAGGCATGGGCCCTAAATTTGATAAGGCCTCTGCCTGGGTTAACAGGCC
CATGTACAGGAAGCCCAGGATATATCGGACGTTGAGAAGCCCCGATGTTCCAAGAGGGTGTGAAGGGCCTTGC
AAGGTCCAGTCATTTGAATCAAGACATGACGTTACTCATGTTGGCAAGGTGTTGTGCATATCTGATGTGACAC
GTGGTAATGGTATTACCCATCGTGTAGGTAAGCGTTTTTTCGTTAAGTCCGTGTACATTTTGGGCAAGATATG
GATGGACGAGAATATCAAGCTCAAGAACCACACCAACAGCGTCATGTTTTGGTTGGTGCAGACAGGAGACCG
TATGGAACACCTATGGATTTTGGCCAGGTTTTCAACATGTTTGTATAACGAGCCAAGCACAGCTACTGTTAAGA
ACGATCTTCGTGATCGTTTTTCAAGTCATGCACAAGTTCTACGCTAAGGTGACGGGTGGACAGTATGCGAGCAA
CGAGCAGGCGTTGGTTAAGCGCTTTTGGAAAGGTCAACAACCTACGTGGTGTACAACCATCAGGAAGCAGCTAAA
TACGAGAATCATACTGAGAACGCTTTGTTATTGTATATGGCATGTACGCATGCGTCTAATCCTGTGTATGCGA
CATTGAAAATTCGGATCTATTTTATGATTCGATATCAAATTAATAAATGTTGAATTTTATTTTCATGATTCTC
GAGTACATCAGTTACATAATGTTTGTTCGTCGCGAAACGAACAGCTCTAATTACATTGTTTATTCCAATAACG
CCTAAGCTATCTAAATAAGAAATTACAAGTAATTTAAATCTAGCTAAATATCTCTGCCCAGAAGCTCTCATCG
AAGTCGTCCAAACTTGGAAATTGAAGAAGGCTTTGTGGAGACCCAACGCTTTCCGTAGGTTGTGGTTGAACCG
GACTTCGATGTGGTATATCCTGGTTCTCGTGAATAACGGGTTCTCTACGTGGTGTATCTTGAAATAAAGGGGA
TTTTGGAACCTCCCAAATAAAAACGGAACCTCTCTGCCTGAGCCGAGTGATGCTCTCCCCTGTGCGTGAATCCA
TTGTCTGCGCAGTTAATGTGGAGGAATATTGAACAGCCACAGTTCAAGTCTATGCGTCTACGAACAGCTC
TTCGTTTTAGCAATCCTGTGCTGTGCTTTGATAGAGGGGGGCTTCAAGGGTGACGAATTTAGCATTTTTGATTG
TCCAGGCTCTTAATGATGCGTTTTCTCTTTGTTGAGGAAAGATATATAGCTGCTCCCCTCTCCTGGATTGCA
CAGCACGATTGACGGTATGCCACCTTTAATTTGAACTGGCTTTCCGTATTTGCAGTTTTGATTGCCAGTCCCTTT
TGGGCCCCATTAGCTCTTTCCAATGCTTTAGCTTTAAGTATTGCGGAGCTATGTCATCAATGACGTTGTACT
CCACATTATTTGAAAACACCTTGAATTGAAGTCGAGGTGACCACTCAAGTAATTATGTGGGCCTAGAGCACG
TGCCACATTGTTTTGCCTGTTTCGAGAATCACCTTCAATTATGATACTAATAGGTCTTTCTGGCCGCGCAGCG
GGACTCCGACCAAAATAATCATTGCCCCAATCTTGAATCTCGTCTGGTACGTTAGTAAAGGAGGAGAGTTGAA
ACGGAGGAGCCCATGGTTCCGGAGCCTTATGAAAGAGGCGCTCGATATTTGACTTAACGTTATGATAACTAAC
AATAAATGTTTTGGGATCACCTGCCCTAATTATTGCGAGAGCCTCTGCCACACTTGCTGCATTAACCTGCGTTA
TGGTAGACATCGTCTTTATTTGCCTTGAACCCCCAGACACTTTGTATTGCCCGGATTACAATAATCACCAT
CTTTGGTGATGTAATTCTGGACGGCATTGGTGTCTTTGGCTGCCTGAATGTTTGGGTGAAAATTGGCAGACCT
TCTGGGGTGAGTGATGTCGAAAAATCTAGCATTCTTGATGTCAGACTTGCCCTGAGAGTTGGATGAGACAGTGT
AAGTGTGGGAACCCGTCGGAATGTTCCCTCTCTGGCGACACGTATGTATGTGGGTTTGCAGACTGACCATTGTA
ACGAGTGAAGCAACTGAAGAGCTTCATCTTTTTGGTATATCGCACTGGGGATATGTTAAGAAAATGTTTTCTGGC
TTGTAAACGAAATGAATTAGGGTTTCGTGGCATTTTTGTAAATAAGCGTAGGTACTCCAGCGGAGGCTCTCTC
AACTTGAGTGCTATATGCAGGAGGATTGGAGTACTATATACTAAAAGGCTCTGGGGTACTCCAGGGGCAAA
ACCGGAAAGCGGCCATCCGTAATAATATT

RhMSinV- GUco (DNA-B 2554 pb)

ACCGGATGGCCGCGGATTTTTGGAGGATAGCCCACTTGGCGTACCCCTCCCCTTTGGCGTACCCACCTCTT
TGGCGTACCCACCCCTTTGGCGTACCCCCCTTTTGGCGTACCCCTTTTGGCGTACCCCTTTTGGCGTACCCCTTTT
GGATAAAGCTAATCGCGCAATTGCGTTTTGAATTTTGAATATTTTTCGCGGATATGTGTGTATGGACCATTTTA
AAACGATAAGGTATTCTGACACATTGGTGGACGTGTACCAATTAACCTCCGCTGCTGAGTTAAGTTATTTGT
AATTTGACTGGTTTTGGTGTATATAATGGGGCGGAGTTTTATTTTTTAAAATCGTCTTTTGCATTAATAATATAC

ATACTGTGTTATTTAATATGTATACGTCTAAGTATCGACGCGGGTCATCATATACTCAACGACGATTTTATCC
ACGTAATCAATTGTTTTAAACGAGCAACCGCGGTGAAACGTATCGATGGCAAACGTGATCAAGTCATGTAAAT
AAGGTTTCATGAAGAGCCCAAGTTGACGTCGCAACGCATACATGAGAACCAATATGGCCCTGAATTTGTTATGA
GTCATAATGGAGCTATTTTCGACGTTTATTAATTTTCCAGTTCTTGGTAAGAGTGAACCCAATCGTAGTCGGTC
TTATATAAAAATTGAAGCGTTATCTTTTTAAAGGAACCGTCAAGATTGAGCGTGTTTCATGCTGATGTGAACATG
GACGGATTAAGCTCTAAAATAGAAGGAGTATTCACACTTGTGGTTGTGATTGATCGTAAACCACATGTAAGTT
CTTCTGGTGGTTTACATACATTTGATGAATTATTTGGTGCAAGAATCCACAGTCACGGCAATTTATGCATAAG
CTCATCGTTGAAAGATCGTTATTACGTGCGTCACGTGTTGAAGCGCGTTGTGTCCGTTGAGAAAGATAGTGTT
ATGCTAGACCTTGAAGGATCGACAACCCCTTTCTAATAAGCGCTTCAATTGTTGGTCTAGTTTTAAGGATCTTG
AGCATGATTCATGTAACGGTGTTTATGCGAATATAAGCAAGAACGCCATTTTAGTTTTATTATTGTTGGATGTC
GGATGTGATGTCTAAGGCATCGACATTTGTATCGTATGATCTTGACTATGTTGGGTGATCAATAATAAAAATGA
TGAACCTGTTTCATATTATGACTTGGCAACAAAAACTAATTTATTTAATTTAAAGACTTAGGCGGTGCTGGA
TTACAGTTATTGTTAATACACTCCTGCACCGTTGTTCTAACAAGTTCGTTTAATTGGGCCATTGACATGGTAA
TGTTGGATTGGGCCCTCTGTAACCCCACTTGTGACGCTGAATCCCCCGGTCTAAGACGCTCCCTCCTAGACG
ATTTAGTTCTCTATACGGATGTAATTCCGAGTCCGTCCTGTGTGAGTCCGCAATTGTGCTTCTTGAAGCC
CACGACTCACCTGGTTTTAATTCAATTGGGCCTGTCAGCCCAAGACGTGACATGGATGTGGACCTTATTGTCT
TCCTTTCCCATGGGCCGTAGTCTACATGCGAGAAGTCCACATCCTTATGGGAGAAGTGTGTTGAAAGGATTTT
GACCGTGGGCCCCCTAAATGGTATATCTACAGAGTGTGTTGCTGTGGATAGCTTTAATTTGCCTTTGAATTTG
GCAAAATGAGTCCTTTGATGGACATTTGTATCGGACACTCTGTAGTATAATTTCCATGGAATTGGATCCTTTA
GGGAAAAGAATGACGATGAGAAATAATGGAGGTCTATGTTGCAGCGTATCGGAAATGTCCAGGACGCTTGTAG
CGATTGCTTGTCCGTCATTCTTTTGTGATGAATTTCTACAATGACCGACCCTGTTGCGTTAATCGGTACTTGT
TGTCGGTATTCAATGACGCAATGGTCAATTTTCATACAGCTACGATTGAGTCTTGCTGTTAATTGTGACGCTG
TGGAAGGGAATTGAAGGACTATCTCAGTTAGGTCATGAGACAGTTGATATTTCATCCCTGTGAGACTCTATATA
ATTGAAAGCATTGCGGTGCGTTGGCTAACTGAGATTCCATATATGAATAATTGGCCGCGCAGCGGAAGTGTGTTG
CAGAAGAACAGGACTTTGTTATGAAGGTTGTTGCGTATAAGAACAATTGGAACAAGAAGATATCGTTATGAGG
TTGTTGCGGTTGTGTTGAGCAACAAATTAATAAGTAATACAACAACAATAGTTTTAAAGAATGTAAAAGTGATG
AACAAGTGAAGACTGAAGAAGATGATAGGTTGGTTAATTGAAGATGATGTTGTGATGAACTGCAACGAGAGTG
TTTCTTATATAGGCATACCATGTGTTGACTTATTAAGTTAAAGGAAAACCCATGATAGTGGCATTGTTGTGTA
TAGGATAGGTAAGTCCAGCGGAGTCCAGACCCGCTCTCAACAAAAGTCATATATGCTGGAGGATTGGAGTC
CTATAAATACTAGAAGTCTCTGGGGTACTCCAGACACGTGGAATCCCTAGCGGCCATCCGATATAATATT

RhGMV- Col (DNA-B 2587 pb)

ACCGGATGGCCGACCGCTGTACGACCCCCCCCCCTTTGGCGTACCCCAACCCCTTTGGCGTACTCACTCCCCTT
TGGCGTACCCCACTCTTGTGTTGGCGTACTCACTCCCCTTTGGCGTACTTTCCTTTAGCCGTGGCGCATTCTCT
ATCGTTGGATAACGCTTATCTATCGTTATGTATTTTGAATTTTGAATTTGAATTTGTTTGCAGCGGATATCGGTA
ATTGGGACCATGTGCATAACGATAATGGTTTCTGACACATTGGGCGTTGTGTTGAACGTGGCCCAATAAAAAC
TCCGCTGGCGAGTTAAGTTATTTTTTAATTTGAATCTTGTGCTCTATATATTGGAGCGGAGTTTTTTTATTCAA
ATCGTCTTTTACATTAACCTTCAATATTTCAAAAATTTGATTATGTATTTGTCTAGGTTTCGACGTGGGGTTT
CATATCCGCAAAGACGATTTTATCCTCGTAATTTTGCCTTTAAACGATCAAGCGTTGTTAAACGTAACGATGG
TAAACGTGACATGTTGTTGTAATAAGGTTAATGACGAGCCGAAGATGTCGCTGCAGCGCATTACAGAGAAC
CAATATGGGCCTGAATTTGTTATGGCCATAATGCAGCCATATCGACGTTTCAATTAATTTCCCCCATTTGGGTA
AATGTGAACCCAATCGTACCAGGTCATTCATTAAGTTAAAACGGCTTCGTTTTAAGGGTACGGTAAAGATTGA
GCGTGTTCGGTGGATCTGAATATGGACGGTTCACTTCAGAAAGTTGAAGGCGTTTTCTCCCTTGTGTTGTA
ATTGATCGCAAACCTCATTTGAGTTCATCAGGATGTTTGCACACATTTGACGAGTTATTCCGTGCAAGAATCC
ACAGCCATGGAAATTTAGCTATTAACCTTCATTGAAGGATCGTTTTTATATACGTCACGTCTTCAAGCGTGT
TCTGTCCGTGGAGAAAGACAGTTTGTATGATAGACGTGGAAGGATCTACTTGGTTATCTAGTCGGCGCTTTAAT
TGTTGGTCTATTTTTAAGGATTTGGATCACGATTCATGTAACAGTGTGTTATGACAATATTAGCAAAAACGCCA

TTTTAGTTTATTATTGTTGGATGTCGGATGTTGTGTCTAAGGCATCCACATTTGTATCGTTTGACCTTGACTA
TATTGGTTGAATATGAATAAACTAATGAACTTGTTCATATTATGACATGCTAAAAAATAATTTTATTTTC
AACGATTTTCGGTTGTGTTCGGATTACAATTATTTTTGATACACTCTTGGACCGTTGTACGTACAAGCTCGTTTA
ATTGGGCCATTGACATGGTTATGTTGGATTGGGCCCTCTCTAACCTACTTGTGATGCCGAATCACCTGGGTC
TAATACGCTTCCCCCTAAACGATTAAGTTCTCTATACGGGTGTATCTCTGAGTCCGTCTCTGTGTGTGTCCGT
CCAATTGTGCTTCTGGAAGCCCATGACTCACCCGGTTTTTAATTCAATTGGGCCTGTTAGCCCCAAGTCGATTCA
TGGATGTGGACCTTATTGTCTTTCGGTTCCACGTGCCGTAATCTACGTGTGAGAAATCCACATCCTTGTGCGA
GAATTGTTTGGAAAGGATCTTCACCGTCGGTGCACGGAAAGGGATATCTACAGAGTGTTCGCGGTGGAGAGC
TTTAATTTGCCTTTGAACTTCGCGAAATGTGTCTGCTGATGTACGTTGGTATCTGACACTCTGTAGTATAGTT
TCCATGGAATAGGGTCTTTAAGCGAGAAGAATGATGACGAAAAATAGTGGAGATCTATGTTACACCTGATGGG
AAATGTCCATGATGCTTGTAAACGATTTCGTTGTCCGTCATTCTTCTGTTCATGAATCTCCACTATGACGGAGCCT
GTTGCGTTAATTGGAACCTTGTGTCTGTATTCTATGACGCAATGGTCAATTTTCATACAGCTACGATTAAGTC
TTGCGGTGAATTGCGACGCTGTGGAAGGAAATTGGAGGATAATCTCAGTTAGGTCATGAGATAAATTGATATTC
ATCTCTGTGCGACTCTATGTAATTAAGGCATTAGGTGCATTTGCTAACTGAGATTCCATATATGAAAACCTTG
GCCGCGCAGCGGAAGTGTTCGAGCAAGATGAAGCTGTGAGGGAAGTTGTTCTGTTTGAAGTTGTACTTAACCC
AGAATGGTTGAACAAGACAAAGATGTTAAGATTGTGGTGTTCGTTGTTGACAAGGAACAGAAATGGTGTGGAA
ACTTGTGTAGAAGAGTTGTTTGCCTGGGTTAACTATGATATGAACAACCTATACGCTAGGCTTGTGTGTTTAT
GAAAGTGATGAGTGTTCGTTGCTGTGTATGCGAGACTCGAAATGATATGTTAAACCCAGTTTCGTGTGAGTTAA
ATAGAAATGTCTTGTGATAGTGAGCGAGAGCGTCTGAGTGGCATTTCGTTAGTTAGAAGGGTGTACCCCGATT
GAGGCTCTCTCAAACCTGGCTGATTATATCCGGTGTATTGGTATCCTATATATACTATCTTCCTTTCTATCACT
TTTGACACGTGGAGGCCATCCGTTATAATATT

Secuencia parcial del virus híbrido EuMV-Jal/ RhMSinV-GU DNA-A (1468 pb)

CGGCATTGGCGTCTTTGGCAGCCTGGATGTTTCGGGTGAAAACAGGCTGACCGTCTGGGGTGAGATATGTCGAA
AAATCCAGCATCCTTGATGTTGGACTTGCCAGATAAATTGTATGAGGCAGTGTAAGTGCGGGAACCCATCGCGA
TGTTCTCTCTGGCGACACGGATGTACGTTCGGTTTACGATCGACCAAGGAAGATGTTGAAGCATCTGAAGAA
CTTCATCTTTGGGAACATCGCACTGCGGATATGTTAAGAAGATATTTCTGGCTTGTATACGAAATGAATTAGG
GTTTCGTGGCATCTTTGTAATTATGACCCAGGACTCCAAGGGAGCTCTCCCATAAAACCTATTGTTTGTGGAG
TCCTGGAGTCCCATTTATACATAAAGTCCTCTAGGACTCCAAGGGGCACGTGGCGGCCATCCGCAATAATATT
ACCGGATGGCCGCGATTTTTTCGGAGTCCCTCGGCCAGGCCCAAACGCAAAGCTATTCTCGTCCTTTTATGG
GACCACTACAACATTTACTGACCAATCATTTTTCGCTCTGACGCGTCTAGATAATCATGACTTACCCACCAAG
CTGTGGTCCCTATAAATAAAGATGATTAGTCCCATACCTCTTTAATTCAAATGGTTAAGAGGGATGCCCCAT
GGCGACTATCGGCGGGGACCTCTAAGGTATCCCGCTCTACCAATTATACGCTCGTGCAGGCATGGGCCAAAA
ATTTGATAAAGCCTCTGCCTGGGTTAACAGGCCCATGTACAGGAAGCCCAGGATCTATCGGACGTTGAGAAGC
CCCGATGTTCCAAGAGGTTGTGAAGGGCCTTGTAAAGGTCCAACCTTTTGTAGTCCAGACATGACGTCACCTCATG
TTGGCAAAGTGTGTGTATATCTGACGTGACACGTGGTAATGGTATTACCCATCGTGTAGGTAAGCGTTTTTTG
TGTGAAGTCTGTGTACATTTTAGGCAAAATATGGATGGACGAGAATATCAAGCTCAAGAACCACACCAACAGC
GTCATGTTTTGGTTGGTGCAGAGACAGGAGACCGTATGGAACACCTATGGATTTTGGCCAAGTGTTCACATGT
TCGATAACGAGCCTAGCACTGCTACTGTTAAGAACGATCTTCGTGATCGTTTTTCAAGTCATGCACAAGTTCTA
CGCTAAGGTCACCTGGTGGACAGTATGCGAGCAACGAGCAGGCGTTGGTTAAACGCTTTTGGAAAGGTCAACAAC
TACGTGGTGTACAACCATCAGGAAGCAGCTAAATACGAGAATCATACTGAGAACGCTTTGTTATTGTATATGG
CATGTACGCATGCGTCTAAGCTTGTAAATCACTAGTGAATTCGCGGCCGCTGCAGGTCGACCATATGGGAGAG
CTCCAACGCGTTGGATGCATAGCTTGTAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAATCATGTCATA
GCGTTCCA