

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

**POSGRADO EN CIENCIAS EN BIOLOGIA MOLECULAR**

**"Delimitación experimental del dominio de unión  
específica al DNA de la proteína de replicación de  
los geminivirus"**

Tesis que presenta

**Bernardo Bañuelos Hernández**

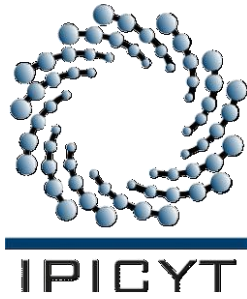
Para obtener el grado de

**Doctor en Ciencias en Biología Molecular**

**Director de la Tesis:**

**Dr. Gerardo Rafael Arguello Astorga**

San Luis Potosí, S.L.P., Junio del 2012



## Constancia de aprobación de la tesis

La tesis “**Delimitación experimental del dominio de unión específica a DNA de la proteína de replicación de los geminivirus**” presentada para obtener el Grado de de Doctor en Ciencias en Biología Molecular fue elaborada por **Bernardo Bañuelos Hernández** y aprobada el **23 de Mayo de 2012** 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.

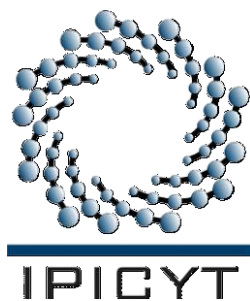
\_\_\_\_\_  
Dr. Gerardo Rafael Arguello Astorga  
Director de la tesis

\_\_\_\_\_  
Dr. Rafael Francisco Rivera Bustamante  
Asesor de la tesis

\_\_\_\_\_  
Dr. Angel Gabriel Alpuche Solis  
Asesor de la tesis

\_\_\_\_\_  
Dr. Sergio Rosales Mendoza

Asesor de la tesis

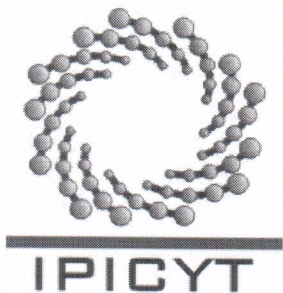


## Créditos Institucionales

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

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (**No. 188816**) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.

Los proyectos de investigación descritos en esta tesis fueron financiados con recursos otorgados al Dr. Gerardo Rafael Argüello Astorga por el CONACYT (Proyectos SEP/CONACYT [Ciencias Básicas] **42639-Q** y **84004**).



# 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 043 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Ciencias en Biología Molecular está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 11 días del mes de junio del año 2012, se reunió a las 12:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

<b>Dr. Sergio Rosales Mendoza</b>	<b>Presidente</b>	<b>UASLP</b>
<b>Dr. Rafael Francisco Rivera Bustamante</b>	<b>Secretario</b>	<b>CINVESTAV</b>
<b>Dr. Ángel Gabriel Alpuche Solís</b>	<b>Sinodal</b>	<b>IPICYT</b>
<b>Dr. Gerardo Rafael Argüello Astorga</b>	<b>Sinodal</b>	<b>IPICYT</b>
<b>Dr. J. Sergio Casas Flores</b>	<b>Sinodal</b>	<b>IPICYT</b>

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

**DOCTOR EN CIENCIAS EN BIOLOGÍA MOLECULAR**

sustentó el C.

**Bernardo Bañuelos Hernandez**

sobre la Tesis intitulada:

*Delimitación experimental del dominio de unión específica al DNA de la proteína de replicación de los geminivirus*

que se desarrolló bajo la dirección de

**Dr. Gerardo Rafael Argüello Astorga**

El Jurado, después de deliberar, determinó

**APROBARLO**

Dándose por terminado el acto a las 13:50 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición del interesado y para los fines que al mismo convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 11 días del mes de junio de 2012.

  
**Mtra. Ivonne Lizette Cuevas Vélez**  
Jefa del Departamento del Posgrado

  
**Dr. Marcial Bonilla Marín**  
Secretario Académico



## **Dedicatorias**

A mis padres y hermanos

A mi esposa e hijo

A la Dra. Natividad Turrubiartes y su familia

## **Agradecimientos**

A mis padres Ricardo y Felicitas que han dedicado parte de su vida en mi; que impulsaron en todo momento mis deseos de estudiar y alcanzar mis metas planteadas

A Isabel por su comprensión, aceptación y apoyo

A la doctora Naty por su apoyo incondicional para la realización de mis estudios

A Don Héctor y a Doña Julia por compartir conmigo su hogar durante 5 años

Al Dr. Gerardo por sus enseñanzas y su dedicación al presente trabajo

Al Biólogo Salvador Ambríz por su apoyo técnico brindado en el laboratorio

A la señora Rosy y a Rosy por su gran ayuda en el laboratorio

A los compañeros del laboratorio (Yair, Mariana, Mayra, Armando, Omar, Benita, Elvira, Aurorita, Josefát, Pablo, Ileana, Marlene, Verito, Jorge, Lula).

A Harumi por su paciencia y enseñanzas

Al Dr. Francisco Rafael por las facilidades brindadas en su laboratorio para el desarrollo de los experimentos de replicación viral e infección de plantas

A mis asesores de tesis por sus acertados comentarios y correcciones

# Contenido

Constancia de aprobación de la tesis	ii
Créditos institucionales	iii
Acta de examen	iv
Dedicatorias	v
Agradecimientos	vi
Contenido	vii
Lista de tablas	viii
Lista de figuras	x
Anexos	xi
Abreviaturas	xii
Resumen	xiv
Abstract	xv

## **Delimitación experimental del dominio de unión específica a DNA de la proteína de replicación de los geminivirus**

<b>I.- Introducción</b>	<b>1</b>
I.1.- Generalidades	1
I.2.- Taxonomía de la familia <i>Geminiviridae</i>	2
I.2.1.- Begomovirus	2
I.2.2.- Curtovirus	3
I.2.3.- Mastrevirus	4
I.2.4.- Topocovirus	4
I.3.- Replicación de los geminivirus	6
I.3.1.- Síntesis de la cadena negativa	6
I.3.2.- Síntesis de la cadena positiva	6
I.3.2.1.- Mecanismo de replicación dependiente de recombinación	7
I.3.2.2.- Mecanismo de replicación por círculo rodante	9
I.4.- Interacción específica de la proteína Rep y el origen de replicación	10
<b>II.- Materiales y Métodos</b>	<b>16</b>
II.1.- Modificación del plásmido pBluescript SK II	16
II.2.- Clonación del DNA-A de <i>Tomato mottle virus</i> (ToMoV) en el vector pBS $\Delta$ Xmnl	16
II.3.- Diseño de construcciones virales	18
II.4.- Análisis de restricción de las construcciones	19
II.5.- Secuenciación de las construcciones realizadas	21
II.6.- Bombardeo de células NT1 con las construcciones virales	23
II.7.- Análisis por <i>Southern blot</i> en DNA total de células NT1 transfectadas	25
II.8.- Inoculación de plantas de <i>Nicotiana benthamiana</i>	26
II.9.- Análisis por PCR de los extractos de las plantas inoculadas	27
II.10.- Secuenciación de los productos de PCR de los extractos de las	

plantas inoculadas	28
<b>III.- Resultados</b>	29
III.1.- Obtención del plásmido pBluescript mutante pBS $\Delta$ Xmnl	29
III.2.- Clonación del genoma A de ToMoV en el vector pBS $\Delta$ Xmnl	29
III.3.- DNA-A quiméricos derivados de ToMoV	29
III.4.- Autorreplicación de los virus híbridos en células NT1	32
III.5.- Transreplicación de componentes B heterologos por los componentes A híbridos	34
III.6.- Infección de plantas <i>Nicotiana benthamiana</i> con los virus híbridos	37
III.7.- PCR de las plantas inoculadas y secuenciación de los productos obtenidos	42
<b>IV.- Discusión</b>	43
<b>V.- Conclusiones</b>	45
<b>VI.- Referencias</b>	47
<b>VII.- Anexos</b>	55
<b>VII.- Apéndices</b>	77



## Lista de tablas

<b>Tabla 1.</b> Enzimas diferenciales en los fragmento de PCR clonados en ToMoV A-pBS $\Delta$ Xmnl	21
<b>Tabla 2.</b> Plásmidos y enzimas usadas para el manejo de los genomas virales	24
<b>Tabla 3.</b> Virus híbridos inoculados en células NT1 para evaluar su autorreplicación	25
<b>Tabla 4.</b> Combinaciones de componentes A y B inoculados en células NT1 para evaluar la trans-replicación del componente B	25
<b>Tabla 5.</b> Combinaciones de componentes A y B inoculados en plantas <i>Nicotiana benthamiana</i>	27
<b>Tabla 7.</b> Descripción de los síntomas presentados por las plantas <i>Nicotiana benthamiana</i> inoculadas con cada una de las combinaciones de componentes A y B	41

## Lista de figuras

<b>Figura 1.</b> Organización genómica de los géneros de la familia <i>Geminiviridae</i>	5
<b>Figura 2.</b> Complejidad de las formas replicativas de los geminivirus	8
<b>Figura 3.</b> Principio de replicación por el mecanismo RDR	9
<b>Figura 4.</b> Mecanismo de replicación por círculo rodante	10
<b>Figura 5.</b> Clonación del genoma viral de ToMoV A en el plásmido pBSΔXmnI	18
<b>Figura 6.</b> Digestiones del DNA-A de ToMoV silvestre y de los virus híbridos con enzimas específicas que los diferencian	22
<b>Figura 7.</b> Análisis de RFLP para buscar las clonas con las construcciones de los híbridos TohibTaino y TohibPYMV	23
<b>Figura 8.</b> Vector ToMoVA-pBSΔXmnI digerido con las endonucleasas <i>XmnI/NcoI</i> y productos de PCR derivados de los 9 begomovirus utilizados en este estudio	31
<b>Figura 9.</b> Esquema gráfico de los virus híbridos derivados de ToMoV-A	32
<b>Figura 10.</b> Southern blot de extractos de DNA total de células NT1 inoculadas con ToMoV A o con los componentes A híbridos	34
<b>Figura 11.</b> Organización de los iterones y otros elementos diferenciales entre los begomovirus del clado del SLCV y los begomovirus “típicos”	34
<b>Figura 12.</b> Southern blot en DNA total de células NT1 inoculadas con el componente A de ToMoV o los virus híbridos y su respectivo componente B	36
<b>Figura 13.</b> Southern blot en DNA total de células NT1 inoculadas con el virus ToMoV A y los componentes B heterólogos	37
<b>Figura 14.</b> Plantas inoculadas con combinaciones diferentes del componente viral A y B	39
<b>Figure 15.</b> Detección por PCR de los componentes virales A y B en hojas apicales de las plantas de <i>Nicotiana benthamiana</i> inoculadas	40

## Anexos

1. Artículo publicado : 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 Journal* 55
2. Artículo en prensa : A new strain of Tomato severe leaf curl virus and a unique variant of Tomato yellow leaf curl virus from Mexico. *Archives of Virology* 70

## Abreviaturas

<b>ACMV</b>	Virus Africano del mosaico de la cassava
<b>Asp</b>	Acido aspartico
<b>Asn</b>	Asparagina
<b>BCTV</b>	Virus del enchinamiento de la punta del betabel
<b>Bla</b>	Betalactamasa
<b>CP</b>	Proteína de la cápside
<b>CR</b>	Circulo rodante
<b><i>E. coli</i></b>	<i>Escherichia coli</i>
<b>IRD</b>	Dominio relacionado al iteron
<b>LIR</b>	Región intergénica larga
<b>mL</b>	Mililitros
<b>MP</b>	Proteína de movimiento
<b>ng</b>	Nanogramos
<b>OkYMMV</b>	Virus mexicano del mosaico amarillo de la okra
<b><i>Ori</i></b>	Origen de replicación
<b>pBS</b>	pBluescript
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>PepGMV</b>	Virus del mosaico dorado del chile
<b>PHYVV</b>	Virus huasteco del amarillamiento de la vena del chile
<b>psi</b>	libras por pulgada cuadrada
<b>PYMV</b>	Virus del mosaico amarillo de la papa
<b>RCR</b>	Replicación por circulo rodante

<b>RDR</b>	Replicación dependiente de recombinación
<b>REn</b>	Proteína potenciadora de la replicación
<b>Rep</b>	Proteína asociada a la replicación
<b>RFLP</b>	Polimorfismo de tamaño de fragmentos de restricción
<b>RNM</b>	Resonancia magnética nuclear
<b>SIR</b>	Región intergénica corta
<b>SLCV</b>	Virus del enchinamiento de la hoja de la calabacita
<b>SMC</b>	Sitio múltiple de clonación
<b>TGMV</b>	Virus del mosaico dorado del tomate
<b>TLCV</b>	Virus del enchinamiento de la hoja de tomate
<b>TrAP</b>	Proteína transactivadora
<b>ToLCV-Nde</b>	Virus del enchinamiento de la hoja de tomate de Nueva Delhi
<b>ToMoTV</b>	Virus taino del moteado de tomate
<b>ToMoV</b>	Virus del moteado de tomate
<b>ToYSV</b>	Virus del punteado amarillo del tomate
<b>TPCTV</b>	Virus del pseudo-enchinamiento de la punta del tomate
<b>TYLCV</b>	Virus del enchinamiento y amarillamiento de la hoja del tomate
<b>ug</b>	microgramos
<b>ul</b>	microlitros

## Resumen

### **Delimitación experimental del dominio de unión específica a DNA de la proteína de replicación de los geminivirus**

Los geminivirus se replican por un mecanismo de círculo rodante utilizando las proteínas del huésped. La única proteína viral esencial para este mecanismo de replicación es la proteína relacionada a la replicación, denominada Rep. La proteína Rep reconoce de manera específica secuencias cortas repetidas (iterones) presentes en el origen de replicación y después da inicio a la replicación por círculo rodante al introducir un corte endonucleolítico en la secuencia TAATATTAC, universalmente presente en los geminivirus. Estudios realizados mediante comparación de secuencias de las proteínas Rep de más de 240 especies de geminivirus, y mediante el análisis estructural de la proteína Rep de un begomovirus, *Tomato yellow leaf curl virus* (TYLCV) han permitido delimitar un dominio que parece estar implicado en el reconocimiento específico de los iterones. Este segmento proteico denominado Dominio Relacionado al iteron (IRD) comprende los primeros 11-12 aminoácidos de la proteína Rep y se ha propuesto como el principal determinante de especificidad de la replicación geminiviral. En el presente trabajo el IRD del Virus de Moteado de Tomate (ToMoV) se reemplazó con el IRD de otras especies de geminivirus, generando así proteínas Rep quiméricas. Mediante ensayos de replicación transitoria en protoplastos de tabaco, y en experimentos de infectividad *in planta* se pudo demostrar que las proteínas quiméricas poseen la capacidad de reconocer el origen de replicación de los virus donadores del IRD de Rep, produciendo infecciones sistémicas cuando los DNAs genéticamente modificados se combinaron con los componentes genómicos B de las especies de las que se derivó el IRD. Estos resultados confirman experimentalmente que éste pequeño dominio de Rep es la región que determina principalmente la unión específica de la proteína viral a su origen de replicación cognado. Sin embargo, en experimentos de infectividad en plantas de *Nicotiana benthamiana* en los que los componentes genómicos A derivados de ToMoV se combinaron con el componente B de ese mismo begomovirus, se observó que las proteínas Rep quiméricas no perdieron del todo la capacidad de reconocer el origen de replicación de ToMoV, sugiriendo de ese modo la participación de otra región de la proteína Rep en el reconocimiento de los iterones, o alternativamente, que el origen de replicación de ToMoV es reconocido de modo relajado por las proteínas Rep quiméricas cuando éstas alcanzan una concentración elevada en las células vegetales.

## **ABSTRACT**

### **Experimental evaluation of the DNA specific-union domain in the replication associated protein of geminivirus**

Geminiviruses replicate by a rolling circle replication using host proteins. The only viral protein essential for this mechanism is the protein associated to replication (Rep). Rep recognizes iterative sequences in a specific form located in the origin of replication and gives way to the beginning of rolling circle replication. Using a theoretical approach and based on structural analysis, a domain has been localized in Rep that could be involved in the specificity of iteron recognition. The domain (iteron-related domain) is located in the first 12 amino acids and is presumed to be the major specificity determinant. In this work, the IRD of Tomato Mottle Virus (ToMoV) has been replaced with the IRD of heterologous geminivirus resulting in chimaeric Rep proteins able to recognize the heterologous origin of replication. This result demonstrates that the IRD is the major region involved in specific iteron recognition during geminivirus replication. However, despite the quimeric proteins recognize heterologous origin of replication they maintain a weakly recognition of ToMoV's replication origin. This observation suggests that there is another region in Rep protein that participates in the recognition of Replication origin or that the ToMoV, s replication Origin is binding without high specificity.

## I.- INTRODUCCIÓN

### I.1.- Generalidades

Los geminivirus son patógenos de plantas superiores que poseen un genoma de DNA de cadena sencilla que replican mediante un mecanismo de círculo rodante. La familia *Geminiviridae* debe su nombre a la morfología geminada de su cápside (Mumford, 1974; Zhang. *et al.*, 2001; Bottcher, *et al.*, 2004.), la cual es única entre el gran número de virus existentes. Los geminivirus han emergido como uno de los principales grupos virales que afectan la producción agrícola alrededor del mundo. Los cultivos de tomate, chile, calabaza, frijol, maíz, betabel, yuca (cassava), entre otros, son afectados por miembros de alguno de los cuatro géneros de esta familia viral (Akhter *et al.*, 2009; Patil y Fauquet, 2009; Legg y Thresh, 2000; Salati, *et al.*, 2002). Durante una infección las moléculas de DNA geminiviral interactúan con proteínas de la planta y las codificadas por el propio virus para replicarse en el núcleo de la célula y moverse luego a través del floema (Lazarowitz *et al.*, 1989, Fontes *et al.*, 1994; Lewiz y Lazarowitz, 2009; Briddon y Markham, 2001). En el núcleo de las células vegetales el DNA viral de cadena sencilla y sentido positivo interactúa con los factores de replicación del huésped (DNA polimerasas, primasas, helicasas, etc) y con ello permite la replicación de la cadena negativa al parecer de igual forma en la que se replica la cadena retrasada del DNA cromosómico (Saunders *et al.*, 1992). Una vez que se ha sintetizado la cadena negativa, el DNA de doble cadena se replica por un mecanismo de círculo rodante para multiplicar la cadena positiva; este proceso lo realiza con la participación de la proteína viral asociada a la replicación, denominada AC1, C1 o Rep. Aunque estas características, de manera general, son compartidas por los cuatro géneros de la familia, la organización de los genes, las interacciones de cada tipo viral con su huésped, su distribución geográfica así como el insecto vector que lo dispersa es particular para cada género. El mecanismo de replicación por círculo rodante (RCR) es un mecanismo primitivo de replicación usado por plásmidos de bacterias y arqueas (Ilyina y Koonin, 1992) y por otras familias virales como los nanovirus en plantas (Vega-Rocha *et al.*, 2007 b) y los circovirus en aves y mamíferos (Vega-Rocha *et al.*, 2007 a). La función de la proteína Rep es interactuar y unirse al origen de replicación viral para producir un corte endonucleolítico en la secuencia conservada TAATATTAC entre la séptima y la octava base generando un extremo 3-OH libre el cual usan las polimerasas de la planta como “primer” o cebador para sintetizar de *novo* una cadena positiva de DNA viral (Fontes *et al.*, 1994; Laufs *et al.*, 1995). El mecanismo RCR ha sido el más estudiado en geminivirus, sin embargo un proceso de replicación por recombinación se ha propuesto como otra forma en la que los virus pueden generar nuevas moléculas de DNA (Jeske, *et al.*, 2001). Estos dos mecanismos de replicación, y la interacción específica de la



proteína Rep con el origen de replicación, que es primordial en el proceso replicativo por CR, serán abordados de manera detallada en esta introducción para fundamentar el desarrollo de los experimentos que se realizaron en el presente trabajo de investigación, orientados a delimitar la región mínima de la proteína Rep que es capaz de reconocer de manera específica el origen de replicación de una especie viral determinada.

## **I.2.- Taxonomía de la familia *Geminiviridae***

En la familia *Geminiviridae* se distinguen 4 géneros: *Begomovirus*, *Mastrevirus*, *Curtovirus*, y *Topocuvirus*. Estos subgrupos comparten características como la cápside geminada, el mecanismo de replicación y una secuencia conservada de 9 nucleótidos “TAATATTAC” que se encuentra en el “asa” de un elemento con el potencial de formar una estructura de “horquilla” (tallo-asa). Sin embargo cada género presenta características particulares como los hospederos que infecta, el insecto vector que los transmite, y la organización de su genoma (Figura 1). Las características propias de cada género son expuestas en los siguientes apartados.

### **I.2.1-Begomovirus**

Los begomovirus constituyen el grupo más diverso de la familia *Geminiviridae*, con con más de 200 especies descritas (Fauquet *et al.*, 2008; Briddon *et al.*, 2010), que infectan plantas dicotiledóneas y son transmitidos exclusivamente por la mosquita blanca (*Bemisia tabaci*), un homóptero perteneciente a la familia Aleyrodidae. Según el número de moléculas de DNA que constituyen su genoma, existen dos tipos de begomovirus; los monopartitas y los bipartitas (Figura 1). Los begomovirus monopartitas poseen un genoma compuesto por una sola molécula de DNA. El componente único de los virus monopartitas contiene 6 genes, 4 que se transcriben en el sentido complementario y codifican las proteínas C1, C2, C3 y C4, y dos genes que se transcriben en el sentido del virión y codifican las proteínas V1 (CP) y V2 (Hanley-Bowdoin *et al.*, 1999). La proteína C1 (Rep) es el iniciador de la replicación del DNA viral por un mecanismo de CR; C2 (TrAP) es un activador de la transcripción de genes tardíos; C3 (REn) es una proteína potenciadora de la replicación, y C4 es un supresor de la respuesta de silenciamiento de la planta, función que también realiza C2. Las proteínas codificadas por los genes en el sentido del virión V1 (CP) y V2 son responsables del movimiento local y sistémico del virus. El virus *Tomato yellow leaf curl virus* (TYLCV) es un virus perteneciente al grupo de los begomovirus monopartitas en el que se han estudiado la mayoría de las propiedades de cada una de las proteínas

mencionadas. Los virus monopartitas han evolucionado en el Viejo Mundo, aunque en años recientes algunos de estos han sido introducidos en el continente americano (Lefeuvre *et al.*, 2010). Una característica común pero no universal de los virus monopartitas es su asociación con moléculas satélites, denominadas satélites alfa o beta, con las cuales se favorece su capacidad de infección (Dry *et al.*, 1997; Saunders *et al.*, 2000). Los begomovirus bipartitas, por otro lado, poseen un genoma compuesto por dos moléculas de DNA denominados componentes genómicos A y B, o simplemente “DNA-A” y “DNA-B”. Estos componentes son encapsulados de manera individual por la proteína de la cápside, pero la infección de una planta con un virus bipartita requiere de la presencia de ambos componentes virales en las células del huésped (Liu *et al.*, 1997). Los componentes A y B codifican para genes involucrados en diferentes procesos de la infección. El DNA-A incluye al gen *V1 (CP)* en la cadena en el sentido del virión, y a los genes *Rep*, *TrAP*, *REn* y *AC4* en la cadena complementaria. La proteína Rep, que se describe en detalle más adelante, es la encargada de la replicación de ambos componentes genómicos, lo cual es posible porque comparten un segmento de DNA no-codificante denominada “región común”, en la cual se encuentran los elementos necesarios para la replicación viral (Revington *et al.*, 1989). La proteína TrAP es una proteína multifuncional que transactiva a los genes tardíos *CP* y *BV1*, y está involucrada además en la supresión del silenciamiento mediado por RNAs pequeños (Rodríguez-Negrete *et al.*, 2007; Bisaro, 2006). La proteína REn por su parte potencia la replicación de los genomas virales interactuando con la proteína Rep (Settlage *et al.*, 2005). La proteína AC4 ha sido estudiada experimentalmente en algunos virus bipartitas del Viejo Mundo y se le ha encontrado una participación en la supresión del silenciamiento (Vanitharani *et al.*, 2004). En el DNA-B, por otra parte, solo se codifican 2 proteínas: una denominada BV1 o NSP (Nuclear-shuttle protein) que posee señales de localización nuclear y es responsable del movimiento del DNA viral del núcleo al citoplasma, y la segunda denominada BC1, la cual está implicada en el movimiento sistémico de las moléculas virales (Sanderfoot *et al.*, 1996; Florentino *et al.*, 2006; Zhou *et al.*, 2011).

### **I.2.2- Curtovirus.**

El género curtovirus se caracteriza por presentar una sola molécula de DNA con un tamaño aproximado de 3000 nt en la cual se codifican 7 proteínas. En la cadena del virión se codifica a la proteína de movimiento V3, la proteína V2 reguladora de la proporción de cadena sencilla y doble del DNA viral, y la proteína de la cápside V1 (Hormuzdi y Bisaro, 1993). En la cadena complementaria se codifican a las proteínas C1, C2, C3 y C4 (Hormuzdi y Bisaro, 1995). Las

funciones de estas proteínas son un tanto parecidas a las funciones de las proteínas homólogas en los begomovirus. C1 es la proteína encargada de la replicación; C2 es una proteína supresora del silenciamiento que, en comparación con TrAP, no tiene actividad de transactivador; C3 es potenciadora de la replicación y C4 es una proteína determinante de patogenicidad (Stanley y Latham, 1992) y parece estar involucrada en el movimiento de las moléculas virales del núcleo al citoplasma (Teng *et al.*, 2010). El virus más estudiado en este género es el virus del enchinamiento de la punta del betabel (BCTV, beet curly top virus). Los curtovirus son transmitidos por el insecto *Circulifer tenellus* de la familia *Cicadellidae* (chicharritas) e infectan plantas dicotiledóneas. Este género se encuentra distribuido en los Estados Unidos y norte de México y recientemente se han encontrado especies que divergen de los curtovirus americanos en Irán (Yazdi *et al.*, 2008).

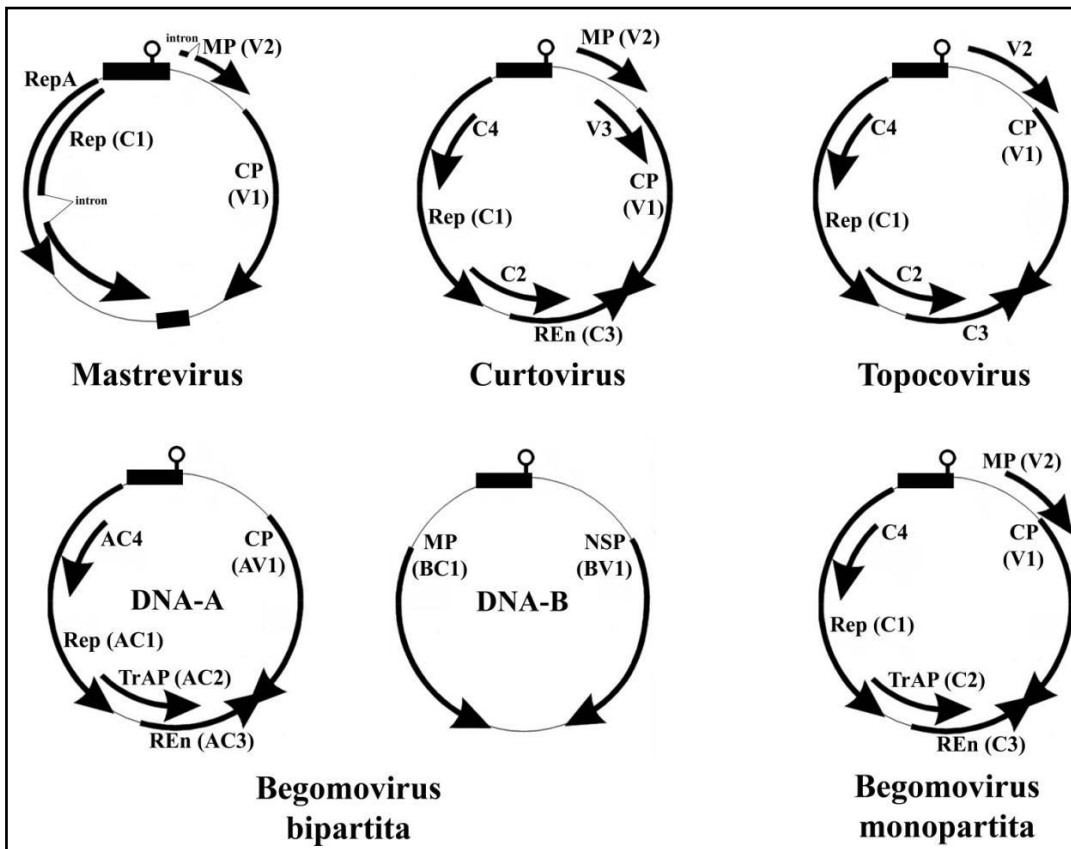
### **I.2.3- Mastrevirus.**

Los Mastrevirus son transmitidos por insectos de la familia *Cicadellidae*, infectan plantas monocotiledóneas y dicotiledóneas y poseen una única molécula de DNA de aproximadamente 2600nt. En esta molécula se encuentran codificados los genes *Rep A* y *Rep* en la cadena complementaria y los genes *MP* (V1) y *CP* (V2) en la cadena del virión. La molécula presenta dos regiones intergénica denominadas región intergénica corta (SIR) y región intergénica larga (LIR). En la región intergénica corta se inicia la replicación de la cadena negativa la cual es iniciada por una molécula pequeña de RNA (primer) que es también contenida en la cápside que contiene el genoma viral (Donson *et al.*, 1984 ; Hayes *et al.*, 1988). En la región intergénica larga se encuentran dos promotores en dirección opuesta que dirigen la expresión de los genes tanto en el sentido del virión como en el sentido complementario. En esta región se encuentran las secuencias necesarias para la replicación de la cadena positiva por el mecanismo de RCR. Los mastrevirus se encuentran distribuidos en varios continentes del Viejo Mundo (África, Europa, Australia) y causan severos daños a cultivos de maíz y trigo (Shepherd *et al.*, 2007; Van der Walt *et al.*, 2008). La especie mejor estudiada de este género es el virus del estriado del maíz (MSV; Maiz Streak Virus).

### **I.2.5- Topocuvirus.**

En el género *Topocuvirus* solo se clasifica actualmente a una especie denominada virus del pseudo-enchinamiento de la punta del tomate (TPCTV; *Tomato pseudo-curly top virus*). Este virus presenta 4 genes en la cadena complementaria y 2 en

el sentido del virion. En la cadena complementaria se codifican los genes C1, C2, C3 y C4. En la cadena del virion se codifican las proteínas V1 y V2. Este virus infecta plantas dicotiledóneas y es transmitido por el insecto *Micrutallia malleifera* de la familia *Membracidae* (Mcdaniel y Tsai, 1990). El análisis de la secuencia de este virus indica que surgió como una recombinación entre un begomovirus y un curtovirus (Bridson *et al.*, 1996). Es interesante en este virus conocer su interacción con el vector que lo transmite ya que es el único conocido dentro de la familia *Geminiviridae* que se transmite por homópteros comúnmente llamado “periquitos” (“treehoppers”). Como puede notarse poco son los estudios que se han realizado en este virus, sin embargo los procesos generales de replicación y transcripción se asume que son semejantes a los virus ampliamente estudiados en los curtovirus y begomovirus.



**Figura 1. Organización genómica de los géneros de la familia *Geminiviridae*.** Los marcos de lectura abierta de cada gen son indicados en flechas. El nombre de la proteína codificada por cada gen se indica al lado de cada marco de lectura, en cada caso fuera del paréntesis se indica el nombre común y dentro de paréntesis el nombre oficial de los genes. En los recuadros en negro se señala la región del origen de replicación de la cadena positiva que incluye también el elemento tallo-asa.

### **I.3.- Replicación de los geminivirus**

#### **I.3.1.- Síntesis de la cadena negativa**

Durante el ciclo infectivo de los geminivirus, una vez que el virión ha sido introducido por la mosquita blanca al floema, el DNA de cadena sencilla contenido en el mismo, llega al núcleo de las células mediante un mecanismo no bien entendido aun, en donde sirve como molde para generar el DNA intermediario de doble cadena. Una de las hipótesis propuestas para la llegada del DNA al núcleo es que la proteína de la cápside se desensambla en el citoplasma y forma un complejo con la molécula de DNA para transportarla al núcleo. Esto se ha sugerido ya que la proteína CP de los geminivirus contiene secuencias señal para la entrada y salida del núcleo (Sharma y Ikegami, 2009). Las moléculas de DNA de cadena sencilla que han alcanzado el núcleo sirven como molde para generar el DNA intermediario de doble cadena. La síntesis de esta segunda cadena (cadena complementaria o cadena negativa) es llevada a cabo por la maquinaria de la planta y requiere de las proteínas propias para la síntesis de la cadena retardada de la replicación del DNA celular (Figura 3). Esta síntesis es iniciada usando como cebador para la replicación secuencias cortas de RNA, a semejanza de lo que ocurre con los fragmentos de Okasaki (Saunders *et al.*, 1992). En el caso de los mastrevirus el “primer” para la síntesis de la cadena negativa es también encapsidado en las partículas virales (Donson *et al.* 1984; Hayes *et al.*, 1988) y esta característica los hace diferente en este aspecto a los otros géneros de geminivirus que no llevan el cebador en las partículas encapsidadas sino que es sintetizado una vez que está en el núcleo de la célula infectada. La molécula de DNA de doble cadena generada es entonces unida por histonas y empaquetada en un minicromosoma con la característica de tener libre de nucleosomas la región del origen de replicación viral (Pilartz y Jeske 2003). En esta región del minicromosoma es donde se regulan los procesos de transcripción y replicación de la cadena positiva del virus.

#### **I.3.2- Síntesis de la cadena positiva**

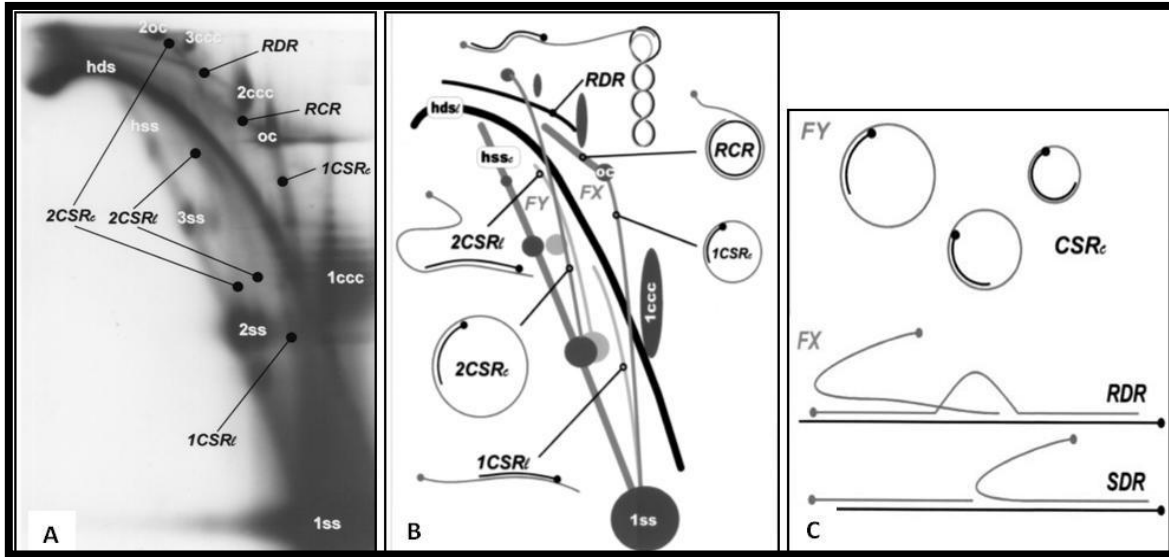
Una vez que el DNA viral ha alcanzado el núcleo y se ha sintetizado la cadena negativa, la molécula circular de doble cadena empaquetada en forma de minicromosoma es ahora transcripcionalmente activa, con la capacidad de expresar y regular cada uno de los genes codificados. En los begomovirus bipartitas los mecanismos de regulación de la transcripción de cada gen es regulada de manera temporal como lo han mostrado estudios en virus específicos (Shimada-Beltran y Rivera-Bustamante, 2007; Hung y Petty 2001). Esta regulación

se da en atención a las necesidades específicas del ciclo viral y de infección. En esta expresión temporal la primera proteína en expresarse es la proteína de replicación. La explicación detallada de la expresión de cada gen se deja para otro apartado para adentrarnos aquí en el proceso de replicación de la cadena positiva del DNA viral, que es el que genera nuevas moléculas de DNA circular de cadena sencilla.

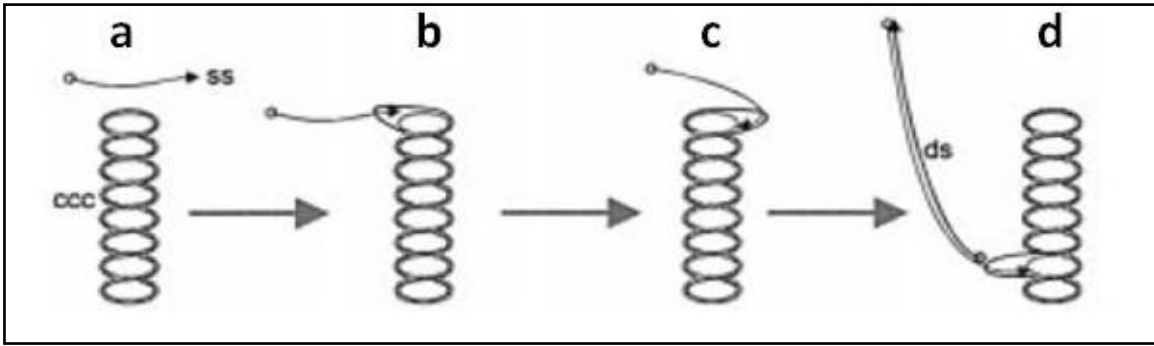
Para la replicación de los geminivirus en un inicio se considero solamente el mecanismo de RCR (Saunders *et al.*, 1991; Lazarowitz, *et al.* 1992; Hanley-Bowdoing, *et al.* 1999) en donde la proteína Rep es la proteína principal encargada de este mecanismo. Sin embargo el análisis más detallado del DNA en sus diferentes formas replicativas (Jeske *et al.*, 2001; Preiss y Jeske, 2002), muestra moléculas de DNA que no parecen encajar en los productos esperados para el mecanismo de circulo rodante. De estas observaciones se ha propuesto un mecanismo alternativo para la replicación de los geminivirus; el mecanismo de replicación-dependiente de recombinación (RDR). En los dos subtemas siguientes se explican estas dos formas de replicación usadas por los geminivirus.

### **I.3.2.1.- Mecanismo de replicación dependiente de recombinación**

La replicación por RCR genera monómeros de las moléculas de DNA viral, sin embargo el análisis de extractos de DNA de platas infectadas con geminivirus ha mostrado que se generan moléculas virales con un tamaño mayor al peso del monómero. La generación de multimeros de las moléculas de DNA viral se ha atribuido a la replicación dependiente de la recombinación (Erdmann *et al.*, 2010; Preiss y Jeske, 2003; Jeske *et al.*, 2001). En la figura 2 se presentan las posibles formas de DNA generadas por RDR. El paso básico de la recombinación dependiente de la recombinación es la invasión de una molécula de cadena positiva sencilla (completa o un fragmento de esta), en una molécula de doble cadena lineal o circular cerrada. Una vez que ha invadido al DNA doble cadena la molécula "intrusa" utiliza a la cadena negativa de este último como templado para elongarse (Figura 3). La recombinación de estas moléculas se cree que es mediada por proteínas de recombinación propias de la planta puesto que los geminivirus no codifican proteínas de recombinación (Jeske *et al.*, 2001). La figura 3 muestra el principio de esta replicación. Es importante resaltar que este mecanismo es solo complementario al mecanismo de RCR, ya que éste último es el único que tiene la capacidad de generar DNA de cadena sencilla.



**Figura 2. Complejidad de las formas replicativas de los geminivirus.** A).- Análisis por Southern blot de extractos de DNA de planta infectada con el virus AbMV. B).- interpretación de los spots generados en el Southern blot. C).- posibles formas de intermediarios de la replicación. **oc (open circular)**: DNA circular abierto, **ccc (covalently-closed circular)**: DNA circular covalentemente cerrado, **ss (single strand)**: DNA circular de cadena sencilla, **hds (heterogeneous double strand)**: DNA lineal de doble cadena de tamaño heterogeneo, **hss (heterogeneous single strand)**: DNA circular de cadena sencilla de tamaño heterogéneo. **CSR (complementary strand replication)**: intermediarios de replicación en la síntesis de la cadena negativa, **RDR (recombination dependent replication)**: intermediarios de replicación por recombinación, **SDR (splained dependent recombination)**: intermediario de replicación por desplazamiento de cadena. Los prefijos **1 2 y 3** indican el número de concatámeros y los sufijos **c y l (en cursiva)**, indican forma circular o lineal. Figuras tomadas y modificadas de Preiss y Jeske, 2002.



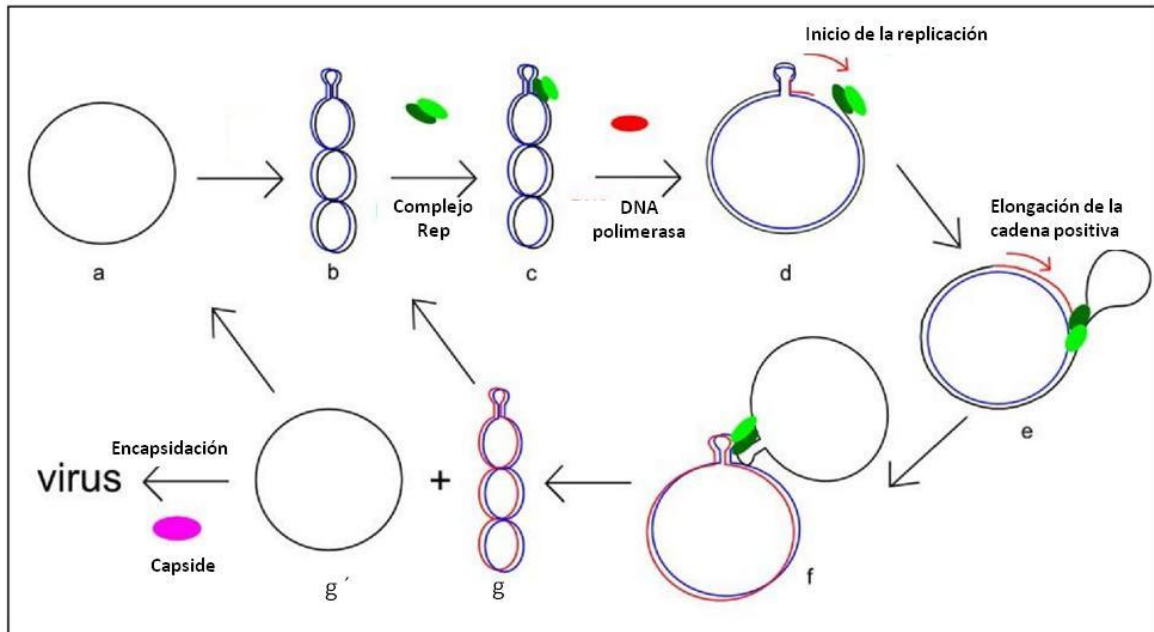
**Figura 3. Principio de replicación por el mecanismo RDR.** En la replicación dependiente de la recombinação una molécula de cadena sencilla(a), correspondiente al genoma viral se entrelaza por recombinação homologa con una molécula viral de doble cadena (b); a partir del sitio de entrecruzamiento la molécula de cadena sencilla se elonga usando como templado una de las cadenas del DNA de doble cadena(c y d). Tomada y modificada de Jeske, 2001.

### I.3.2.2- Mecanismo de replicación por círculo rodante

El mecanismo de replicación por círculo rodante es el más estudiado en los geminivirus y en este la proteína viral Rep tiene el papel central. La replicación de la cadena positiva en este mecanismo se genera cuando la proteína Rep se une al origen de replicación del virus. Esta unión de la proteína Rep es específica de cada especie viral, es decir, la proteína Rep de un virus hipotético X se une solo al origen de replicación del virus X y no al origen de replicación del virus Y. Una vez establecida la unión de la proteína Rep la capacidad de formar multímeros con otras unidades de Rep le permite posicionarse en la secuencia TAATATTAC que también se localiza en el origen de replicación y está conservada en todos los geminivirus, a excepción de uno recientemente descubierto (Yazdi *et al.*, 2008). En esta secuencia de 9 nucleótidos la proteína Rep ejerce su acción catalítica rompiendo el enlace fosfodiéster entre los nucleótidos 7 y 8 (TAATATT/AC, Laufs *et al.*, 1995). Una vez dada la ruptura la proteína Rep se mantiene unida al extremo 5' de la cadena (5'AC), y el extremo 3' (TAATATT 3') que se genera sirve de iniciador para que la maquinaria de replicación de la planta (DNA polimerasas y proteínas del complejo de replicación) comience la síntesis de la cadena positiva a partir de este punto (Luque *et al.*, 2002). La elongación de la cadena continúa hasta cerrar el círculo y completar la replicación de un genoma completo del virus. La proteína Rep que permanece unida en el extremo 5' de la cadena sencilla saliente al replicarse toda la molécula del virus corta la cadena sencilla y liga los extremos 5' y 3' (Laufs *et al.*, 1995). Este mecanismo genera así una molécula de cadena doble con una cadena positiva sintetizada *de novo* y una molécula de DNA



de cadena sencilla libre. Estas dos moléculas pueden entrar de nuevo al ciclo de replicación o, para el caso de la cadena sencilla, puede ser encapsulada para generar nuevos viriones. La figura 4 describe los pasos principales de la replicación por círculo rodante.



**Figura 4. Mecanismo de replicación por círculo rodante.** **a:** DNA de cadena sencilla que llega al núcleo después de la inoculación por la mosca blanca. **b:** Intermediario de doble cadena capaz de expresar la proteína Rep. **c:** Reconocimiento del origen de replicación por la proteína Rep y ruptura de una de las moléculas en la secuencia TAATATT/AC. **d** y **e:** unión del complejo de replicación de la planta y elongación de la cadena positiva. **f:** unión de los extremos 5' y 3' de la cadena positiva saliente (genoma completo). **g** y **g'**: generación de una molécula de cadena sencilla y restablecimiento de la molécula de doble cadena original con una de las cadenas sintetizadas *de novo*.

La importancia que tiene la unión de la proteína Rep al origen de replicación para iniciar la replicación viral me lleva a la redacción del siguiente apartado dedicado a exponer las propiedades de la proteína y las características de su unión con el origen de replicación.

#### I.4.- Interacción específica de la proteína Rep y el origen de replicación

Rep es una proteína multifuncional con actividades de helicasa (Choudhury et al., 2006; Clerot y Bernardi, 2006), ATPasa (Desbiez et al., 1995), Ligasa (Laufs et

*al.*, 1995), que puede interactuar con ella misma y con la proteína viral AC3 para formar multímeros ( Orozco, *et al.*, 2000) y además interactúa con proteínas del huésped como la proteína relacionada al retinoblastoma (Argüello-Astorga, *et al.*, 2004), el antígeno de proliferación nuclear (Bagewadi *et al.*, 2004) y la enzima de sumoilación (SUMO) de la planta (Sanchez-Duran *et al.*, 2011; Castillo *et al.*, 2004) entre otras. La proteína Rep ha sido muy ampliamente estudiada por ser la única proteína viral esencial para la replicación de los geminivirus. Este descubrimiento se hizo cuando se hicieron mutaciones de los genes individuales que forman el componente A del begomovirus bipartita TGMV y vieron que al mutar el gen que codifica para la proteína Rep (AL1) la capacidad de replicación del virus cuando era agro-infiltrado en *Nicotiana benthamiana* se abolía (Elmer *et al.*, 1988). A la postre, experimentos donde se crearon plantas transgénicas que expresaban la proteína Rep, cuando se agroinoculaban estas plantas con el componente B del virus TGMV este era replicado en ausencia del componente A (Hanley-Bowdoin *et al.*, 1990) indicando la actividad de replicación de la proteína Rep expresada por la planta. Estos experimentos se hicieron conociendo las secuencias del propio TGMV a si como también de otro begomovirus bipartita, el entonces denominado Cassava Latent Virus CLV y que ahora conocemos como Virus Africano del mosaico de la cassava (ACMV). En estos dos virus se había observado que en las secuencias de los componentes A y B de cada virus existe una región conservada en secuencia de entre 230 y 250 pares de bases la cual se sugirió que podría ser el origen de replicación de ambos componentes, dado que Rep podía reconocer ambas moléculas (Elmer *et al.*, 1988). Los experimentos que demostraron la funcionalidad de la región común como origen de replicación se realizaron años más tarde empleando los virus TGMV (*Tomato Golden Mosaic Virus*) y SLCV (*Squash Leaf Curl Virus*). Estos experimentos observaron que cuando la región común era puesta en plásmidos bacterianos estas moléculas al agro-inocularse con los virus en plantas de *N. benthamiana* también eran replicados (Lazarowitz *et al.*, 1992). En estos experimentos un hallazgo de gran relevancia para lo que concierne en nuestro trabajo es que los orígenes de replicación de estos dos virus no podían ser intercambiados llevando a la deducción de que el reconocimiento del origen de replicación era específico de las secuencias que estos contenían en cada especie. Para analizar de manera más fina la interacción de la proteína con el origen de replicación se hicieron ensayos de inmunoprecipitación en los que se ponía a interactuar la proteína Rep con diferentes fragmentos pertenecientes al origen de replicación hipotético. Mediante los ensayos de inmunoprecipitación y ensayos de competencia de unión a DNA se logró identificar en el *Ori* (origen de replicación) de TGMV una secuencia de 52 nucleótidos a la cual era altamente afín la proteína Rep (Fontes *et al.*, 1992). Posteriormente estudiando este mismo virus mediante mutaciones dirigidas la secuencia de reconocimiento se delimitó a trece pares de bases (5'-

GGTAGTAAGGTAG) y dentro de esas 13 pares de bases la secuencias críticas para la interacción fueron los repetidos directos de 5 pares de bases, GGTAG, de la secuencia. De los dos repetidos, el repetido localizado hacia el extremo 3' mostro tener más importancia para la unión de la proteína Rep (Fontes *et al.*, 1994). Experimentos bajo el mismo fundamento se realizaron con virus del género *Curtovirus* en donde dos variantes del virus del enchinamiento de la punta del betabel (BCTV-Logan y BCTV-CFH) presentan repetidos en el origen de replicación que son los importantes para la interacción con la proteína Rep (Choi y Stenger 1996; Choi y Stenger 1995; Stenger 1998). En la búsqueda de secuencias de este tipo en los diferentes virus se realizó un estudio de análisis de secuencias de los virus conocidos hasta ese momento y en el cual se encontró que los 4 géneros de los geminivirus (mastrevirus, begomovirus, curtovirus y topocovirus), presentan elementos repetidos en su origen de replicación, se encontró también que estos repetidos son específicos de las especies y que el arreglo, el número y la orientación que presentan varía según el género y el origen geográfico del linaje (Argüello-Astorga *et al.*, 1994). Los elementos repetidos presentes en el origen de replicación de los geminivirus fueron denominados “iterones”.

Una vez que se había dilucidado los elementos del origen de replicación que interactuaban con la proteína Rep, la pregunta lógica que surgió fue: ¿Cuál es la región de la proteína Rep que reconoce de manera específica las secuencias repetidas presentes en el origen de replicación de los virus? En adelante los experimentos realizados por diferentes grupos estudiosos de los geminivirus se condujeron para encontrar la región de la proteína Rep que pudiera reconocer e interactuar de manera específica con los iterones del *Ori*.

Uno de los primeros trabajos experimentales fue realizado con dos cepas del virus del enchinamiento y amarillamiento de la hoja de tomate (TYLCV), la cepa de Cerdeña y la cepa de Israel (TYLCV-Sar y TYLCV-Is). Estas dos cepas altamente similares en secuencia general tienen iterones diferentes y las proteínas Rep no pueden complementarse para replicar. En un primer ensayo intercambiaron las regiones intergénica de los virus y observaron que los virus híbridos obtenidos no se replicaron cuando fueron introducidos en protoplastos. Posteriormente en un segundo experimento realizaron proteínas Rep quiméricas en donde cambiaron la región carboxilo terminal de la proteína de TYLCV-Sar por la de TYLCV-Is dejando 116 aminoácidos de la región amino terminal de la proteína de TYLCV-Sar. Este mismo cambio lo hicieron en el virus híbrido de TYLCV-Sar que llevaba la región intergénica de la cepa TYLCV-Is. Cuando transformaron protoplastos con estas construcciones observaron que la proteína quimérica solo podía replicar el virus original que lleva los 116 aa de la cepa de Cerdeña y su región intergénica respectiva, en tanto que el virus que llevaba la región intergénica de la cepa de

Israel y la proteína quimérica no replicaba. De esto dedujeron que dentro de estos 116 aminoácidos amino terminales de la proteína Rep residían aminoácidos implicados en reconocer los iterones. (Jupin *et al.*, 1996). En un experimento semejante al realizado por Jupin y colaboradores se trabajo con el virus BCTV que pertenece al género de los curtovirus. Las cepas CFH y logan del virus BCTV no son compatibles en replicación pero al hacer quimeras de la proteínas Rep observan que intercambiando el dominio que abarca del aminoácido 3 al 89 se podía cambiar la especificidad de replicación (Choi y Stenger, 1995). En este experimento, cuando se trasladan los 89 aminoácidos amino terminal más la región intergénica de una cepa a otra, el virus hibrido resultante tiene la capacidad de replicar. En el análisis de la secuencia de esta región de 89 aminoácidos se observa que existen 12 posiciones distribuidas que difieren entre las dos proteínas y sobre las cuales se pensó que podría el reconocimiento específico. En este interés científico en el área de geminivirus de definir que aminoácidos le daban la alta especificidad de Rep por los iterones en el *Ori* y que permitía la replicación del DNA viral, Chatterji y colaboradores realizaron un experimento que analizó de manera puntual el residuo aminoácido 10 del virus del enchinamiento de la hoja de tomate de Nueva Delhi (ToLCV-Nde) el cual propuso que interactuaba con la base 3(T) del iteron GGTGT presente en el repetido imperfecto GGTGTCGGAGTC del origen de replicación (Chatterji *et al.*, 1999). En este experimento se trabajo con dos variantes del virus de nueva Delhi, la cepa severa que en su proteína Rep tiene en la posición 10 el residuo Asn y los iterones que reconoce tienen la secuencia GGTGTCTGGAGTC y la cepa moderada que en la posición 10 tiene el residuo Asp y reconoce los iterones GGCGTCTGGCGTC. En estos virus cuando a la cepa moderada se le cambia la Asp por Asn deja de replicarse y al hacer una segunda mutación en la que se muta la tercera base del iteron C por T se restable la replicación del virus.

Otra forma en la que se analizó la correlación en variabilidad entre la proteína Rep y los iterones fue mediante un análisis de alineamiento de secuencias de los virus conocidos para el año 2001 en este trabajo por comparación de secuencias de la proteína Rep y de las iterones se encontró que los primeros 11 aminoácidos de las proteínas Rep son altamente variables entre virus que poseen iterones diferentes y por el contrario son conservados en virus que poseen iterones iguales. Esta región variable entre las proteínas que se relaciona con la variabilidad en la secuencia de los iterones se denomino Dominio Relacionado al Iteron, IRD, por sus siglas en ingles de Iteron Related Domain (Argüello-Astorga y Ruiz-Medrano, 2001). Otra característica de estos aminoácidos es que están adyacentes al motivo conservado I en las proteínas de replicación por círculo rodante. Este motivo I había sido expuesto por Ilyina y Koonin en 1992 como un motivo conservado dentro de las proteínas iniciadoras de la RCR en plásmidos de

bacterias y en los geminivirus. Además de este motivo I las proteínas Rep presentan en su extremo amino terminal los dominios conservados II y III (Ilyina y Koonin, 1992). Para los motivos II y III se ha encontrado experimentalmente que funcionan para unir iones divalentes ( $Mg^{+2}$  o  $Mn^{+2}$ ) y para realizar el corte endonucleolítico, respectivamente (Orozco y Hanley-Bowdoin 1998; Laufs *et al.*, 1995), sin embargo para el motivo I no se había tenido inferencias sobre su funcionalidad. Del análisis de secuencia realizado por Argüello-Astorga que ubica al dominio relacionado con el iteron adyacente al motivo I se propuso que estos aminoácidos conforman un dominio en conjunto que le permite reconocer de manera específica los iterones en el origen de replicación y establecer las interacciones iniciales para unirse a este. Un trabajo más en aportación al estudio de la proteína Rep fue la obtención de la imagen tridimensional de la región amino- terminal (primeros 120aa) de la proteína Rep del virus TYLCV, la cual conserva las propiedades de unión y catálisis del origen de replicación. En la imagen obtenida por resonancia magnética nuclear (RMN) se revela que los aminoácidos 13 a 19, en los cuales está presente el motivo I “15-FLTY-20”, forman una estructura de lámina beta. Esta lamina beta tiene una extensión estructural que forma una lamina beta pequeña conformada por los residuos 9, 10, 11 y 12 (SIKA). Esta hoja de lamina beta pequeña en la configuración tridimensional esta a un lado de la hoja de lamina beta 5 conformada por los residuos 68, 69, 70, 71 y 72(KYNCT). Juntas, estas dos hojas de lamina beta (beta 1 y beta5) forman un dominio de lamina beta plegada que se asemeja a el dominio de la proteína del antígeno SV40 que interactúa con DNA (Campos-Olivas *et al.*, 2002).

En resumen todos estos experimentos han aportado el conocimiento para delimitar la región de interacción específica de la proteína Rep con los iterones, en los estudios experimentales en planta lo han hecho analizando bloques de la proteína Rep que conservan el reconocimiento específico, y mediante el análisis de un aminoácido puntual para el caso de las cepas de TYLCV nueva Delhi, el análisis teórico de Argüello-Astorga y el análisis por RNM de Campos-Olivas *et al.*, proponen un dominio de 4 aminoácidos adyacente al motivo I como el encargado del reconocimiento específico de los iterones. Sin embargo, aunque la mayoría de los estudios de interacción de la proteína Rep con los iterones son consistentes en encontrar que existe una alta especificidad de unión entre la proteína y los iterones, existen algunos reportes en donde la proteína de algunos virus puede reconocer iterones con secuencias diferentes. Uno de los principales casos de reconocimiento con baja especificidad se da en los begomovirus del Viejo Mundo en donde los DNA satélites de estos virus pueden ser replicados por más de una especie viral (Lin *et al.*, 2003; Dry *et al.*, 1997). Otro de los casos en donde se ha visto in-especificidad de replicación es en algunos virus bipartitas. Por ejemplo, Andrade y colaboradores reportaron en 2006 que el componente A del virus

TGMV puede establecer una infección en planta si es inoculado con el componente B del virus del punteado amarillo del tomate (ToYSV), aun cuando en estos dos virus el IRD y los iterones son un tanto diferentes (Andrade *et al.*, 2006). Este proceso en el que el componente A de un begomovirus puede establecer una infección con el componente B de otra especie se conoce como pseudorecombinación y es comúnmente dado entre virus que poseen IRD e iterones iguales. Otro caso interesante es el estudio realizado en el virus TLCV en donde mutaciones en los iterones impiden la unión de la proteína Rep, sin embargo cuando los mismos virus mutantes son inoculados en plantas estos conservan la capacidad de replicarse (Lin *et al.*, 2003, Chakraborty *et al.*, 2008).

Como se puede ver a través de esta revisión de la literatura existente, es claro que la proteína Rep reconoce de manera específica el origen de replicación a través de su extremo amino terminal, y en los estudios relacionados de Campos-Olivas *et al.* y Arguello-Astorga y Ruiz-Medrano, se sugiere que los aminoácidos adyacentes al motivo son los responsables directos de las interacciones específicas con los iterones del origen de replicación de los geminivirus. En el trabajo que hemos realizado para el desarrollo de este proyecto de tesis se ha adoptado la hipótesis de trabajo de que los 11 aminoácidos iniciales de la proteína Rep son suficientes para conferir la especificidad de reconocimiento hacia los iterones y se han diseñado experimentos *in planta* para explorar esta conjetura.

En el siguiente apartado de este escrito se describirá el diseño de los experimentos y las metodologías aplicadas. Los resultados obtenidos implican al IRD como principal dominio en el reconocimiento específico, aunque parecen existir otros factores que contribuyen a la interacción de Rep con el origen de replicación viral. Los datos obtenidos conducen a nuevas interrogantes en las que se habrá de trabajar para dilucidar de manera más completa las interacciones moleculares entre la proteína Rep y el DNA.

## II.- MATERIALES Y MÉTODOS

### II.1.- Modificación del plásmido pBluescript SK II

La secuencia **GAAAACGTTC** que es reconocida por la enzima *XmnI* fue mutada a **GAAAgCGcTC** utilizando el método de “QuickChange<sup>®</sup> site-directed Mutagenesis Kit” de Stratagene (Apendice A), usando los oligonucleótidos Bla-mutag2 directo (5'-GTGCT CATCATTGGAAA**gCGc**TCTTCGGGGCG-3') y BLA-Mut 1 reverso (5'-CGC CCCGAAGAg**CGc**TTTCCAATGATGAGCAC-3'), que se alinean de la posición 2608 a 2659 en el plásmido pBlueScript SK II+ (**5'-2607-gtgcctcatcattggaaaacgttcttcggggcg- 2660-3'**). En el producto de PCR obtenido los tripletes ACG y TTC (subrayado en la secuencia de pBS), cambia a GCG y CTC. En la cadena complementaria del plásmido que es la que codifica el gen *Bla* los cambios serian CGT a CGC y GAA a GAG. En el primer caso ambos tripletes CGT y CGC codifican para el aminoácido arginina y en el segundo GAA y GAG codifican para el aminoácido glutamato. De esta manera los cambios de secuencia conservan el marco de lectura para la expresión del gen *Bla* que da la resistencia a la ampicilina. Este vector modificado se identifico como pBS $\Delta$ XmnI.

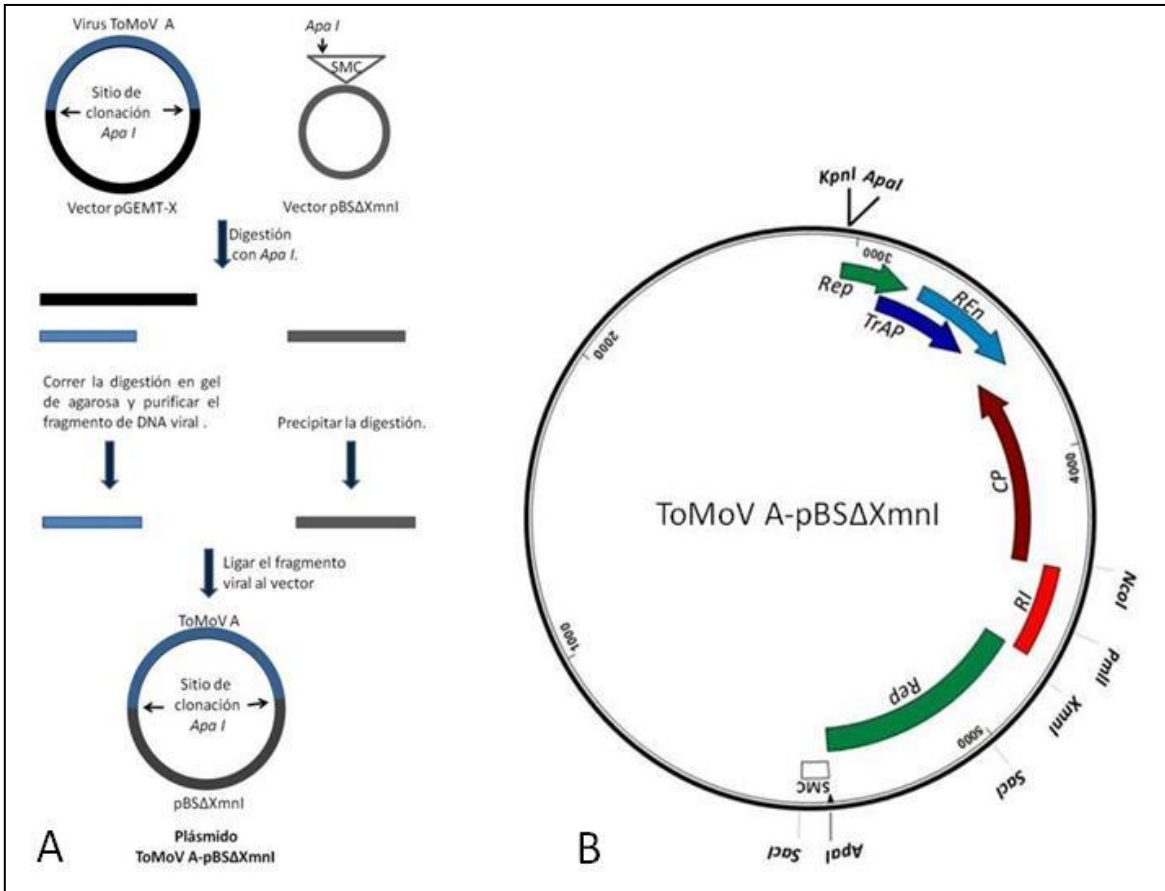
### II.2.- Clonación del DNA-A de *Tomato mottle virus* (ToMoV) en el vector pBS $\Delta$ XmnI

El componente genómico A de ToMoV A, que consta de 2601 nt, se escindió del vector pGEMTX mediante digestión completa con la enzima *Apal* (1 ug DNA plasmídico, 1uL *Apal*, 1uL buffer *Apal*, 7ul agua MQ estéril, 25°C por 12 horas). La reacción de digestión se corrió en gel de agarosa al 1% del cual se aisló la banda correspondiente al DNA viral y se purifico por columna acorde a las indicaciones del kit (Wizard<sup>®</sup> SV Gel and PCR Clean-Up System, Promega). El DNA purificado se eluyó con 30ul de agua MQ estéril, de los cuales se corrieron 3ul en gel de agarosa 1% para verificar la purificación y estimar la concentración del DNA purificado.

El plásmido pBS $\Delta$ XmnI se digirió con la enzima *Apal* hasta que todo estuviera completamente linearizado. Para ello de la mezcla de digestión se corrió una alícuota en gel al 1% para revisar la progresión de la digestión. Normalmente a las 2 horas la digestión es completa usando 1ul de la enzima y 1ug de DNA. Una vez linearizado el DNA se precipito agregando a la reacción de digestión 2 volúmenes de etanol absoluto y un décimo de acetato de sodio 3M pH 5.2 (Apendice B).

El plásmido pBS $\Delta$ Xmnl digerido se ligo con el genoma viral de ToMoV previamente aislado (mezcla de ligación; 1uL T4 ligasa, 1uL buffer T4 ligasa 10X, 200ng vector en 1ul y 600ng de inserto en 1ul, 6ul de agua MQ estéril, temperatura 25 °C, 2 horas). De la ligación se tomaron 5ul para transformar por choque térmico la cepa TOP10F' de *E. coli*. De las clonas obtenidas se eligieron 15 de ellas y se extrajo DNA plasmídico mediante el protocolo de Birboin (Sambrook and Russell, 3er ed, 2001). El DNA de cada clona se analizó por restricción con la enzima *ApaI* para seleccionar aquellas que tuvieran el inserto y posteriormente mediante digestión doble con las enzima *ApaI* / *SacI* para determinar la orientación del genoma viral en el plásmido. Con este proceso obtuvimos el plásmido ToMoV A- pBS $\Delta$ Xmnl cuyo mapa general se muestra en la figura 5.





**Figura 5. Clonación del genoma viral de ToMoV A en el plásmido pBSΔXmnl.** A) Esquema de clonación del DNA-A de ToMoV en el plásmido modificado. B) Mapa general del plásmido obtenido. Entre el sitio *SacI* y *ApaI* se encuentra el sitio múltiple de clonación del vector pBS. El sitio *ApaI* es el sitio de clonación del genoma viral. Los sitios *SacI*, *XmnI*, *PmlI* y *NcoI* son únicos en el virus. El recuadro rojo entre el gen *Rep* y *CP*, indicado como RI, señala la región intergénica viral. La región SMC corresponde al sitio multiple de clonación de pBluescript SK.

### II.3.- Diseño de construcciones virales

El plásmido ToMoV-A-pBSΔXmn (ToMoV A, número de acceso en el GenBank NC\_001938) se digirió con las enzimas *XmnI* y *NcoI* (*Biolabs, New England company*) durante toda la noche a 37°C. 1ul de la digestión se corrió en un en gel de agarosa al 1% para verificar que el plásmido se hubiera digerido completamente. La mezcla de digestión restante se precipito con etanol absoluto y acetato de sodio 3M pH. 5.2. El DNA precipitado se disolvió en 30ul de agua de

los cuales se corrieron 3 para verificar la presencia de DNA y estimar su concentración.

Usando los oligonucleótidos degenerados Rep-Mot-Gus forward (5'GAGTCT AGATGGATANGTDAGGAAATARTTCTTRGC3') y CP-Xba-Gus reverso (5'GAGTCT AGAACGCCATGGRGCATCNCGCTTAGSCAT3') se amplificó por PCR la región equivalente a la comprendida entre el sitio *XmnI* y *NcoI* de ToMoV usando como templado DNA de los virus *Tomato mottle taino virus* (ToMoTV), *Potato yellow mosaic virus* (PYMV), *Pepper huasteco yellow vein virus* (PHYVV), *Tomato golden mosaic virus* (TGMV) y *Pepper golden mosaic virus* (PepGMV). También se utilizaron como templado clonas de pGEM-Teasy con fragmentos de virus desconocidos denominados provisionalmente como *Okra yellow mosaic Mexico virus* (OkYMMV), un begomovirus aislado de *Datura stramonium*, un begomovirus de *Ipomea* spp. y un virus aislado de frijol colectado en Sinaloa. El oligonucleótido Rep-Mot-Gus forward lleva el sitio de reconocimiento de la enzima *XmnI* y el primer CP-Xba-Gus reverso el sitio de reconocimiento de la enzima *NcoI* los cuales se denotan en subrayado en la secuencia de los oligonucleótidos. Los productos de PCR obtenidos fueron digeridos mediante doble digestión con las enzimas *XmnI* –*NcoI* (*New England Biolabs*) y luego purificados en columnas (Wizard® SV Gel and PCR Clean-Up System, Promega) para ligarlos al vector pBSΔ*XmnI*-ToMoVA digerido previamente con las mismas enzimas. De la mezcla de ligación se tomaron 5ul para transformar *E. coli* TOP10F' (Apendice C). De las transformantes obtenidas se extrajo DNA plasmídico de 10 de ellas y se analizaron por restricción. Primeramente se digirió con las enzimas *XmnI* y *NcoI* (0.3ul de cada enzima, 600ng de DNA, buffer 4 de *New England Biolabs*), para identificar las clonas con el inserto, y luego con una enzima que cortara de manera específica en el fragmento de PCR clonado (Tabla 1). Mediante este proceso obtuvimos nueve diferentes virus híbridos derivados de ToMoV-A, los cuales de acuerdo al virus usado como donador del fragmento *XmnI*/*NcoI* etiquetamos del siguiente modo: **Tohib-ToMoTV (el cual para una mejor distinción nombrare en adelante Tohib-Taino), TohibPYMV, TohibPHYVV, TohibTGMV, TohibPepGMV, TohibOkYMMV, TohibDatura, TohibIpomea y TohibSinaloa.** Estas quimeras virales tienen un genoma base de ToMoV, en tanto que toda la región intergénica y los 12 codones iniciales del gen *Rep* provienen del virus heterólogo usado para la PCR.

#### II.4.- Análisis de restricción de las construcciones

Las diferentes construcciones generadas se cortaron con enzimas de restricción que tienen un sitio de reconocimiento en la región intergénica del fragmento del

virus insertado pero no en la región intergénica de ToMoV. De esta manera se revisaron los virus híbridos de PHYVV (*Hind III*), TGMV (*AvaI*), y PepGMV (*XbaI*), OkYMMV(*SacI*), Datura(*ApaI*) e Ipomea(*ApaI*). En la **Tabla 1 y la Figura 6** se indican las bandas diferenciales a obtener en cada caso. Para los virus híbridos con PYMV y ToMoTV no existe una enzima adecuada que diferencie con un solo corte la región intergénica de estos respecto a la de ToMoV, por lo que la estrategia seguida fue la siguiente. El DNA de 5 clonas obtenidas se sometió a PCR con los oligonucleótidos AI1-Mot-Gus y CP-XbaI-Gus. De igual manera se hizo PCR para el Virus ToMoV, ToMoTV y PYMV. Los productos de PCR obtenidos fueron digeridos con la enzima *HhaI* y se corrieron en gel de agarosa el 1.5%. Las clonas con el patrón de restricción semejante al generado por el producto de PCR del virus que deseamos unir al ToMoVA-pBS $\Delta$ XmnI fueron seleccionadas (**Figura 7**). Para descartar que estas clonas fueran fondo de los virus silvestres del que se hizo la PCR se les realizó una digestión con la enzima *XmnI* y *NcoI* sobre el DNA plasmídico.

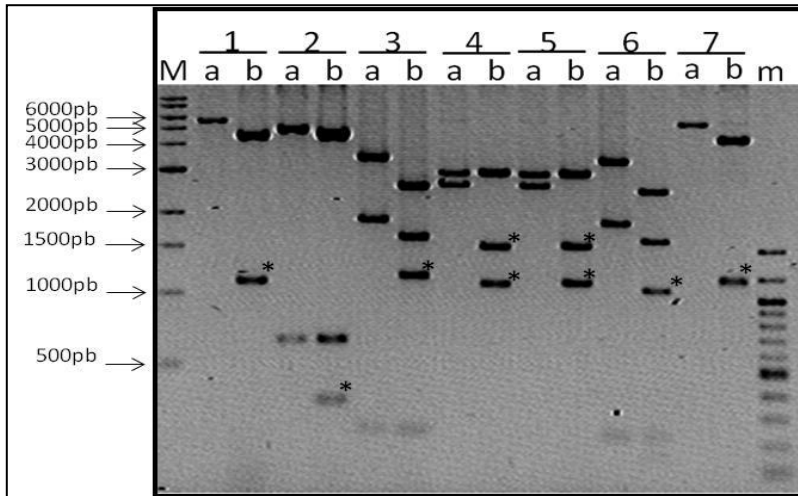
**Tabla 1.** Enzimas diferenciales en los fragmento de PCR clonados en ToMoV A-pBSΔXmnI.

Begomovirus donador de la RI y el IRD	Enzima con sitio único en el producto de PCR a insertar	Tamaño aproximado de banda diferencial en la construcción	Observaciones
ToMoTV	Inexistente	NA	Clonas seleccionadas por RFLP.
PHYVV	<i>Hind III</i>	1200pb	
TGMV	<i>Ava II</i>	1250pb	
PYMV	Inexistente	NA	Clonas seleccionadas por RFLP.
PepGMV	<i>Xba I</i>	1200	
OkYMMV	<i>Sac I</i>	350pb	
Datura	<i>Apa I</i>	1100 y 1500pb	<i>ApaI</i> es el sitio de clonación de ToMoV A en el plásmido pBS.
Ipomea	<i>Apa I</i>	1100 y 1500pb	<i>ApaI</i> es el sitio de clonación de ToMoV A en el plásmido pBS.
Sinaloa	<i>Ava II</i>	1100	

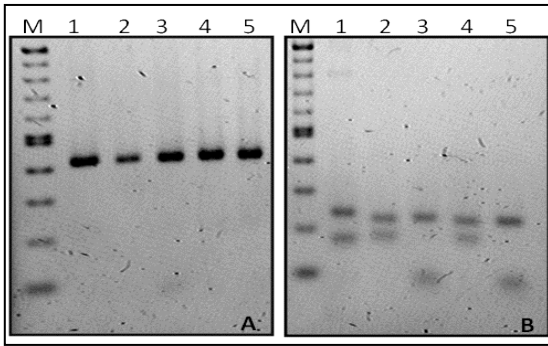
## II.5.- Secuenciación de las construcciones realizadas

De cada una las construcciones realizadas se realizo PCR con los oligos DGRS-forward (5´GAGTCTAGATGCTGACCTCCTCTAGCWGATCTGCCGTC3´) y CP-XbaI-Gus (ya citado). El Oligo DGRS está situado aproximadamente 400nt arriba del sitio *XmnI* en ToMoV-A. Los productos de PCR da cada construcción fueron ligados en el vector PGEM-Teasy (Promega inc.), y enviados a secuenciar con el

“primer” universal M13. Las secuencias obtenidas se analizaron en la base de datos de NCBI y mediante alineamiento local con Clustal W.



**Figura 6. Digestiones del DNA-A de ToMoV silvestre y de los virus híbridos con enzimas específicas que los diferencian.** Los DNA virales están clonados en el plásmido pBSΔXmnl. **1.-** Digestión de ToMoV A(a) y del híbrido TohibPHYVV(b) con la enzima **HindIII**. **2.-** Digestión de ToMoV A(a) y del híbrido TohibOkYMMV (b) con la enzima **SacI**. **3.-** Digestión de ToMoV A(a) y del híbrido TohibTGMV(b) con la enzima **Avall**. **4.-** Digestión de ToMoV A(a) y del híbrido TohibDatura(b) con la enzima **Apal**. **5.-** Digestión de ToMoV A(a) y del híbrido TohibIpomea(b) con la enzima **Apal**. **6.-** Digestión de ToMoV A(a) y del híbrido TohibSinaloa(b) con la enzima **Avall**. **7.-** Digestión de ToMoV A(a) y del híbrido TohibPepGMV(b) con la enzima **Xbal**. **M.-** marcador de Peso molecular de 1kb (*New England Biolabs*). **m.-** Marcador de peso molecular de 100pb (*Invitrogen*). A la izquierda se indican los pesos de las bandas del marcador de peso molecular de 1Kb. Con asterisco (\*), se señalan las bandas diferenciales que se obtienen después de la inserción del producto de PCR de cada virus.



**Figura 7.** Análisis de RFLP para buscar las clonas con las construcciones de los híbridos TohibTaino y TohibPYMV. **A).**- Amplificación por PCR con los oligos A11-Xba-Gus forward y CP-Xba1-Gus reverso. **B).**- Digestión de los productos de PCR con la enzima *HhaI*. **M.**- Marcador de peso molecular de 100pb, **1.**- ToMoV A **2.**- ToMoTV A, **3.**- PYMV A, **4.**- TohibTaino, **5.**- TohibPYMV.

## II.6.- Bombardeo de células NT1 con las construcciones virales

Las células NT1 cultivadas en medio líquido durante tres días se colectaron por filtración al vacío en papel filtro millipore de 2 micras. Las células colectadas en el papel se colocaron en cajas de Petri chicas (120 x 7mm) con medio NT1 osmótico durante tres horas previas a la inoculación por bombardeo. Cada uno de los DNA virales a bombardear se escindió del plásmido mediante digestión con la enzima apropiada (Tabla 2). De cada plásmido con DNA viral se digirieron 6.5ug. Terminada la digestión, las enzimas usadas en cada caso se inactivaron por temperatura (65°C por 20min) y los DNA digeridos se usaron para recubrir partículas de tungsteno M10 (Apéndice D), con las cuales se inocularon por biobalística las células NT1 mantenidas en medio osmótico (Apéndice E). Para los ensayos de autorreplicación las partículas se recubrieron con el componente A de ToMoV o con los virus híbridos (Tabla3), mientras que para los ensayos de trans-replicación de componentes B las partículas se recubrieron con la mezcla de los componentes A y B digeridos (Tabla 4).

Después de ser inoculadas, las células NT1 se transfirieron de las cajas con medio NT1 osmótico a cajas con medio NT1 normal en el cual se mantuvieron por 72 horas en obscuridad (cubiertas con aluminio), a una temperatura de 25°C. Transcurridas las 72 horas se colectaron las células en tubos Eppendorf de 1.5ml que contenían 250ml de NT1 líquido. De las células colectadas se extrajo DNA total mediante el protocolo de extracción de DNA de protoplastos (Shimada-Beltran *et al.*, 2007.; apéndice F). En caso de no proceder consecutivamente con

la extracción las células se guardaron en un ultracongelador a menos 70°C hasta su extracción.

**Tabla 2.** Plásmidos y enzimas usadas para el manejo de los genomas virales.

Plásmido/peso molecular	Resistencia	Virus clonado	Sitio de clonación	Observaciones
PGEMX-1 / 4000pb	Carbenicilina	ToMoV B	<i>Apa I</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	ToMoV A	<i>Apa I</i>	Monómero
pBluescript SK II / 2964pb	Carbenicilina	ToMoTV A	<i>Xba I</i>	Monómero
pZero / 3300pb	Kanamicina	ToMoTV B	<i>Xba I</i>	Monómero
pBluescript SK II / 2964pb	Carbenicilina	PHYVV A	<i>Hind III</i>	Hemidímero
pBluescript SK II/ 2964pb	Carbenicilina	PHYVV B	<i>BamHI</i>	Hemidimero
PBluescript SK II/ 2964pb	Carbenicilina	PepGMV A	<i>EcoRI</i>	Dímero
pBluescript SK II/ 2964pb	Carbenicilina	PepGMV B	<i>Hind III</i>	Dímero
Desconocido/3700pb	Carbenicilina	TGMV A	<i>EcoRI</i>	Monómero
pBSΔXmnl / 2964pb	Carbenicilina	TGMV B	<i>Cla I</i>	Monómero
PGEM/3000pb	Carbenicilina	PYMV A	<i>Apa I</i>	Monómero
PGEM/3000pb	Carbenicilina	PYMV B	<i>EcoRI</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenecilina	TohibTaino	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	TohibPYMV	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenecilina	TohibPHYVV	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	TihibTGMV	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	TohibPepGMV	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenecilina	TohibOkYMMV	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenecilina	TohibSinaloa	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	TohibDatura	<i>Apal</i>	Monómero
pBSΔXmnl/ 2964pb	Carbenicilina	TohibIpomea	<i>Apal</i>	Monómero

**Tabla 3.** Virus híbridos inoculados en células NT1 para evaluar su autorreplicación.

1	ToMoVA(wt)
2	TohibTaino
3	TohibPYMV
4	TohibPHYVV
5	ToHibTGMV
6	ToHibTGMV
7	ToHibOkYMMV
8	ToHibSinaloa
9	ToHibPepGMV

**Tabla 4.** Combinaciones de componentes A y B inoculados en células NT1 para evaluar la trans-replicación del componente B.

Combinación	Componente A	Componente B
1	ToMoV A	ToMoV B
2	ToMoV A	ToMoTV B
3	ToMoV A	PHYVV B
4	ToMoV A	TGMV B
5	ToMoV A	PYMV B
6	TohibTaino	ToMoTV B
7	TohibPHYVV	PHYVV B
8	TohibTGMV	TGMV B
9	TohibPYMV	PYMB

## II.7.- Análisis por Southern blot en DNA total de células NT1 transfectadas

Del DNA total extraído de las células NT1 se tomaron 40ug y se digirieron con 0.5ul de la enzima *Dpn1* a 37°C durante toda la noche, posteriormente a esta digestión se le agrego 1ul de la enzima *ApaI* para linearizar el genoma viral y se dejó actuar al menos durante 2 horas a 25°C. El DNA digerido se fraccionó por electroforesis en gel al 1% y se transfirió a una membrana de nylon hybond N+ (Amersham) con SSC 6X (Sambrook and Russell, 3era ed, 2001). Una vez transferido el DNA, la membrana se fijo mediante radiación UV en el entrecruzador de DNA usando la opción de "optimal crosslinker" (SPRECTOLINKER™, XL-1500 UV ESPECTROLINKER, MODEL XL-1500). La membrana con el DNA fijado se híbrido a 50°C durante 12 hrs usando como sonda radiactiva el fragmento



*ApaI/SacI* del gen *Rep* de ToMoV A marcado con el isótopo radiactivo fósforo 32 ( $^{32}\text{P}$ ), siguiendo las indicaciones del kit (Amersham inc.). Transcurridas las 12 horas de hibridación la membrana se lavó tres veces por 15 min con la solución I (SSC2X/SDS0.1%). La membrana lavada se colocó en un casete de exposición y en un ambiente oscuro se expuso en ella una película Kodac por 3 horas al cabo de las cuales esta fue revelada en el equipo Kodac para revelado.

En los ensayos de trans-replicación, para la detección del componente B, se siguió el mismo procedimiento para el Southern blot usando como sonda el fragmento *Apa I/ HincII* del gen BC1 de ToMoV B.

## **II.8- Inoculación de plantas de *Nicotiana benthamiana***

Se cultivaron semillas de *N. benthamiana* en un preparado comercial de tierra para germinación (Sunshine #3, SUN GRO Horticulture). Las plántulas obtenidas se trasplantaron a macetas con preparado de tierra comercial para crecimiento (Sunshine #6, SUN GRO Horticulture). Cuando las plantas presentaron su 3era o 4ta hoja verdadera se inoculó el DNA en sus hojas apicales mediante biobalística a baja presión, 120psi, (Carrillo-Tripp *et al.*, 2006). Las partículas usadas para la inoculación fueron recubiertas con la mezcla de 1ug del componente A más 1ug del componente B para cada planta. El DNA fue previamente digerido para escindir el DNA viral del plásmido. La preparación de las partículas de tungsteno con el DNA se hizo para un lote de 6 disparos como se describe en el apéndice D. Las plantas fueron inoculadas con diferentes combinaciones de componente A y componente B (Tabla 5). De manera general lo que se hizo fue co-inocular el componente A de ToMoV con los componentes B de ToMoTV, PYMV, PHYVV, TGMV y el propio ToMoV B.

Los virus híbridos por su parte fueron bombardeados en combinación con el componente B heterólogo o con el componente B de ToMoV. En el caso del híbrido TohibPYMV, este se co-bombardeó además con el componente B de ToMoTV con el cual se ha reportado que es capaz de infectar, ya que en un primer experimento no fue capaz de infectar con el componente B de PYMV.

Para cada una de las combinaciones se inocularon 6 plantas por experimento y se observaron diariamente para ver la aparición de síntomas y registrar los datos (Tabla 6). Transcurridas tres semanas post-inoculación, tiempo en el que ya se había observado sintomatología en la mayoría de las plantas, se extrajo DNA tanto

de las plantas sintomáticas como de las asintomáticas (apéndice G), y se analizaron por PCR.

**Tabla 5.** Combinaciones de componentes A y B inoculados en plantas *Nicotiana benthamiana*.

Combinación	Componente A	Componente B
1	ToMoV A	ToMoV B
2	ToMoV A	ToMoTV B
3	ToMoV A	PHYVV B
4	ToMoV A	TGMV B
5	ToMoV A	PYMV B
6	TohibTaino	ToMoTV B
7	TohibTaino	ToMoV B
8	TohibPHYVV	PHYVV B
9	TohibPHYVV	ToMoV B
10	TohibTGMV	TGMV B
11	TohibTGMV	ToMoV B
12	TohibPYMV	PYMV B
13	TohibPYMV	ToMoV B
14	TohibPYMV	ToMoTV B

## II.9.- Análisis por PCR de los extractos de las plantas inoculadas

De cada uno de los extractos de las plantas sintomáticas o asintomáticas se tomo un microlitro de DNA total y se usó como templado en una reacción de PCR con los oligonucleótidos degenerados DGRS forward/CP-Xba-Gus en las siguientes condiciones: desnaturalización inicial, 5min 94°C; desnaturalización 94°C 1min; alineamiento 55°C 1min; elongación 72 °C 1min, con 35 ciclos y una elongación final de 5 min a 72°C. Estos oligonucleótidos dirigen la amplificación de un producto de 900pb del componente A que abarca del codón 134 del gen *Rep* al codón 9 del gen *CP*.

Para detección del componente B la PCR se hizo con los oligonucleótidos BC1290 forward y BV1310 reverso los cuales producen un fragmento de amplificación de 1480pb que van del nucleótido 290 del gen *BC1* al nucleótido 310 del gen *BV1*. Las temperaturas utilizadas fueron iguales a las usadas para la amplificación del componente A.

Los productos de PCR del componente A y del componente B se corrieron en gel de agarosa al 1% para visualizar la obtención del producto de amplificación.

## **II.10.- Secuenciación de los productos de PCR de los extractos de las plantas inoculadas**

Los productos de PCR obtenidos tanto para el componente A como para el componente B fueron clonados en el vector pGEMT-easy de acuerdo a las indicaciones del fabricante (Promega). Para el PCR de cada planta se secuenciaron tres clonas. Las secuencias obtenidas de las muestras de cada una de las plantas se alinearon con la secuencia del virus ToMoV A y con las secuencias de las quimeras previamente determinadas. El alineamiento se hizo utilizando el método Clustal W del programa DNA Star. En el alineamiento se analizaron posibles modificaciones en la región sustituida; se buscaron cambios puntuales en los iterones, cambios en la región codificante del IRD y posibles arreglos por recombinación (esto en el componente B).

### III.- RESULTADOS

#### III.1.- Obtención del plásmido pBluescript mutante pBS $\Delta$ XmnI

Mediante el Kit “QuickChange® Site-directed Mutagenesis” se eliminó del plásmido pBlueScript SK II el sitio de reconocimiento para la enzima *Xmn* I, localizado en el gen de resistencia a Ampicilina. Los cambios realizados en la secuencia fueron mutaciones silenciosas en los codones CGT a CGC y GAA a GAG que no alteran la secuencia de aminoácidos de la proteína betalactamasa. En el primer caso ambos tripletes CGT y CGC codifican para arginina y en el segundo GAA y GAG codifican para glutamato. El plásmido obtenido lo denominamos pBS $\Delta$ XmnI.

#### III.2.- Clonación del genoma A de ToMoV en el vector pBS $\Delta$ XmnI

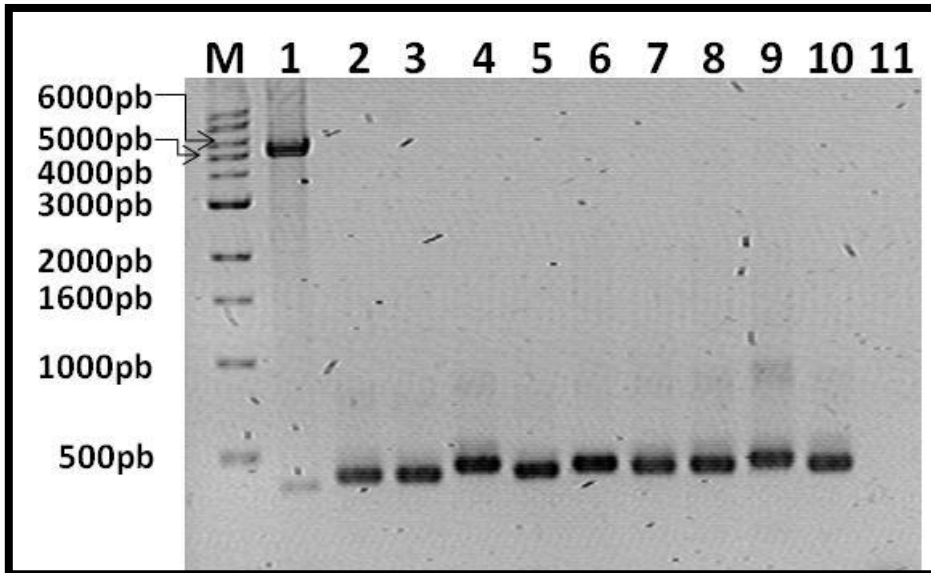
El DNA-A de ToMoV fue escindido con la enzima *Apal* del vector pBluescript y clonado en el plásmido pBS $\Delta$ XmnI. La eliminación previa del sitio *XmnI* en el vector pBS $\Delta$ XmnI permitió que al clonar en este plásmido el componente A de ToMoV se pudiera usar el sitio único *Xmn* I presente en ese DNA viral. El plásmido con el DNA-A de ToMoV se denominó ToMoVA-pBS $\Delta$ XmnI y su mapa se muestra en la figura 5.

#### III.3.- DNA-A quiméricos derivados de ToMoV

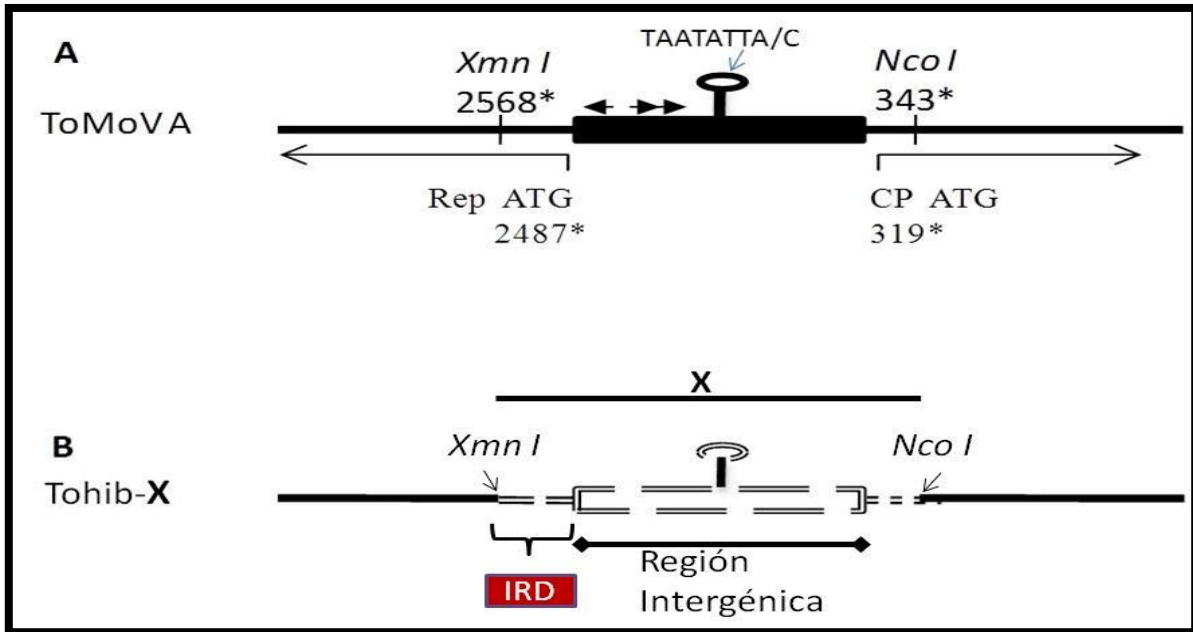
El plásmido ToMoV A-pBS $\Delta$ XmnI posee los sitios *XmnI* y *NcoI* como únicos, como se muestra en la figura 5. Estos sitios flanquean lo que teóricamente es el módulo de replicación específica del virus. Esto es, el sitio de restricción *XmnI* ubicado al lado izquierdo del origen de replicación delimita los primeros 13 codones del gen *Rep*, de los cuales los primeros 11 son altamente variables entre proteínas *Rep* con diferente especificidad de replicación y son considerados como el Dominio Relacionado al Iteron o IRD, por sus siglas en inglés (Argüello-Astorga y Ruiz-Medranol 2001). Por otro lado, el sitio de restricción *NcoI* se ubica en el 8vo codón del gen *CP*. *NcoI* fue usado en nuestro caso para delimitar, la región intergénica del virus por el extremo 3' de la secuencia viral. La región intergénica comprende el segmento del virus que va del codón de inicio del gen *Rep* al codón de inicio del gen *CP*. La ubicación de estos sitios de restricción permiten escindir el módulo IRD-Región intergénica del virus ToMoV mediante digestión doble (Figura 8). Por otro lado estos sitios de restricción se ubican en zonas semi-

conservadas del genoma viral, lo que permitió el diseño de los oligonucleótidos degenerados Rep-Mot-for y CP-Xba-rev, los cuales flanquean esta región e integran ambos sitios de restricción en el producto de PCR de los virus que no los poseen en forma natural, sin alterar la región de los genes *Rep* y *CP* en la que se alinean. Con estos oligonucleótidos pudieron amplificarse las regiones homólogas de otras especies virales (Figura 8). Las especies utilizadas para estos PCR fueron los siguientes: ToMoTV, PHYVV, TGMV, PYMV, PepGMV, OkYMMV, y begomovirus aislados de una planta de frijol colectada en Sinaloa, y de dos malezas, *Ipomoea sp.* y *Datura stramonium*, respectivamente, los cuales fueron caracterizados solo parcialmente. Los productos de PCR comprenden el módulo de replicación específica de los virus heterólogos, los cuales fueron insertados individualmente en el DNA-A de ToMoV luego de ser digeridos con las endonucleasas *XmnI* y *NcoI* y ligados en los sitios equivalentes de ToMoV-A.

Esta metodología nos permitió obtener 9 híbridos derivados de ToMoV diferentes, los cuales denotamos como **TohibTaino**, **TohibPHYVV**, **TohibTGMV**, **TohibPYMV**, **TohibPepGMV**, **TohibOkYMMV**, **TohibDatura**, **TohibIpomoea** y **TohibSinaloa** (Figura 9). Lo que se espera, de acuerdo al conocimiento existente sobre la replicación de geminivirus, es que los DNA-As híbridos serán capaces de autorreplicarse sólo si el IRD de la proteína Rep quimérica generada en cada caso presenta elevada afinidad por los iterones presentes en el origen de replicación heterólogo, el cual reemplazó completamente al *Ori* de ToMoV. Es importante señalar aquí que las proteínas quiméricas generadas no presentan ninguna otra diferencia con la proteína Rep de ToMoV más que la substitución de los once aminoácidos del extremo N-terminal de la proteína, que incluyen al IRD.



**Figura 8. Vector ToMoVA-pBS $\Delta$ Xmnl digerido con las endonucleasas *Xmnl/NcoI* y productos de PCR derivados de los 9 begomovirus utilizados en este estudio. M).- marcador de peso molecular de 1Kb (New England) 1).- ToMoVA-pBSaXmnl digerido con *Xmnl* y *NcoI*. 2-11).- Productos de PCR con los oligos RepMotGus y CPXbaGus usando diferentes virus como templados; 2).- Amplicon de ToMoTV, 3).- PYMV, 4).- PHYVV, 5).- TGMV, 6).- PepGMV, 7).- OkYMMV, 8).- Datura 9).- Ipomea, 10).- Sinaloa 11).- Control negativo. Estos productos de PCR fueron digeridos con *Xmnl* y *NcoI* y ligados al fragmento superior de ToMoVA-pBS $\Delta$ Xmnl purificado.**



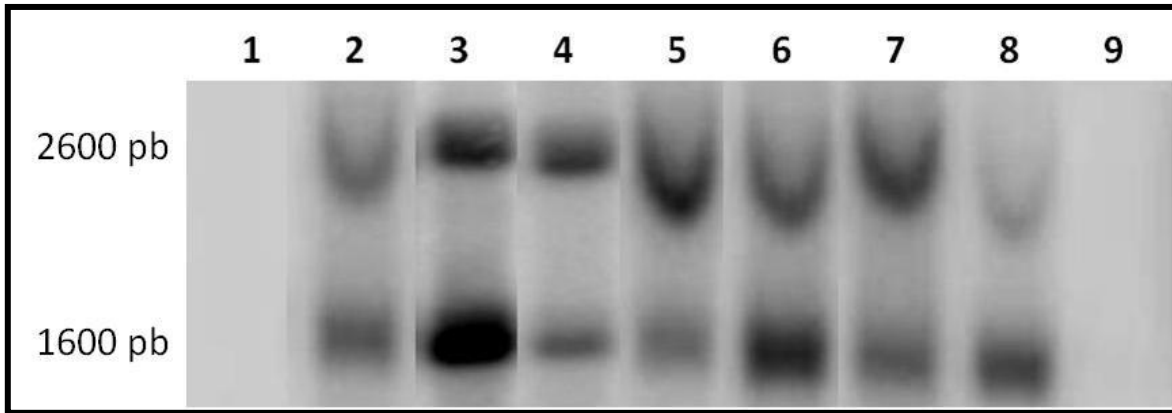
**Figura 9. Esquema grafico de los virus híbridos derivados de ToMoV-A.** A).- ToMoV A original. Con asterisco se indican las coordenadas de los sitios de restricción *Xmn I* (2568) y *Nco I* (343) y de los codones de inicio de los genes *Rep* (2487) y *CP*(319). Las flechas negras indican la posición de los iterones de ToMoV A y la secuencia TAATATTAC presente en todos los geminivirus es señalada en el “asa” de la estructura de horquilla ilustrada en la figura. B).- Esquema general de la organización de los DNA-As híbridos, los cuales contienen en la región comprendida entre los sitios *Xmn I* y *Nco I* el producto de PCR obtenido a partir de los diferentes virus usados en el presente trabajo. **X** representa la región de ToMoV que fue substituida por el segmento homólogo de cada uno de los virus heterólogos. Los productos de PCR contienen la región codificante del IRD y la región intergénica completa del virus heterólogo.

#### III.4.- Autorreplicación de los virus híbridos en células NT1

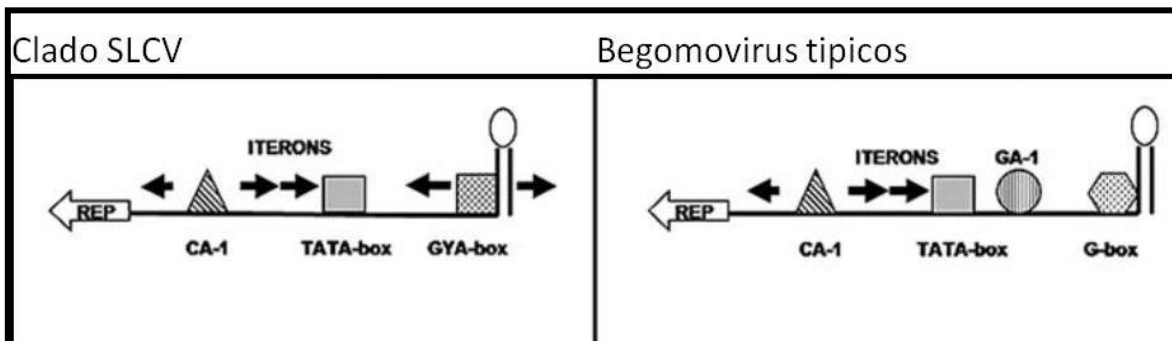
Los componentes A híbridos y el DNA-A del ToMoV silvestre se inocularon por biobalística en células NT1 para explorar sus propiedades de replicación. En el caso de los virus híbridos TohibDatura y Tohibpomea no se realizaron las inoculaciones debido a que en el proceso de digestión para separar el virus del vector con la enzima *Apa I*, el DNA viral se corta en dos por la presencia del sitio de restricción en la región intergénica de estos dos virus (Figura 6). Las células NT1 inoculadas con los DNA virales se mantuvieron a 25°C por 72hrs, tiempo en el cual ToMoV-A realiza varios ciclos de replicación. De las células inoculadas se extrajo el DNA total, el cual se trato con las enzimas *Dpn I* y *Apa I*. La enzima *Dpn I*

reconoce DNA metilado y permite eliminar el DNA plasmídico que se introdujo a las células, mientras que la enzima *ApaI* lineariza el DNA que se genera en la replicación viral. El DNA digerido se corrió en un gel de Agarosa al 1%, se transfirió a una membrana de Nylon y se procedió a realizar un Southern blot usando como sonda el fragmento *ApaI/ SacI* de ToMoV-A. En la figura 10 se muestra el resultado obtenido en el Southern blot, en el que observamos dos bandas que hibridan con la sonda utilizada, una de ~2600pb y otra de ~1600nt. La banda de 2600pb corresponde al DNA viral de cadena doble y la banda de 1600nt corresponde a DNA circular de cadena sencilla generada en el proceso de replicación viral por CR. La capacidad de autorreplicación en células NT1 de 6 de los 7 DNA-As híbridos indica que la proteína Rep quimérica es capaz de reconocer los iterones del virus donador del IRD. El único caso en el que no se dio la autorreplicación fue en el virus híbrido TohibPepGMV. La incapacidad para autorreplicarse de este DNA-A híbrido puede estar relacionada al hecho de que el virus donador del módulo replicativo, (PepGMV) es el único entre los begomovirus examinados que pertenece a un linaje diferente al de ToMoV, el llamado clado del *Squash leaf curl virus* (SLCV). Los begomovirus de este linaje poseen características distintivas en el arreglo, posicionamiento y número de iterones que presentan, además de tener en su origen de replicación elementos como la caja GYA adyacente a la estructura de tallo-asa, la cual podría interactuar con factores de la planta. Por otro lado algunos otros elementos como la caja G y GA-1 presentes en los begomovirus típicos como ToMoV no se encuentran en los virus del clado SLCV (Arguello-Astorga *et al* 1994, Hernández-Zepeda *et al.*, 2009; Figura 11). Las mencionadas diferencias en el origen de replicación de ToMoV y PepGMV sugieren que la proteína Rep de PepGMV en su forma silvestre probablemente establece interacciones con ciertos factores de la planta para formar un complejo proteínico necesario para la replicación; este complejo podría no ensamblarse adecuadamente en presencia de la proteína quimérica de TohibPepGMV.





**Figura 10. Southern blot de extractos de DNA total de células NT1 inoculadas con ToMoV A o con los componentes A híbridos.** 1) Vector vacío pBluescriptSK-II; 2).- ToMoV-A; 3).- TohibTaino; 4).- TohibPYMV; 5).- TohibPHYVV; 6).-TohibOkYMMV; 7).- TohibSinaloa; 8).-TohibTGMV; 9).- TohibPepGMV.

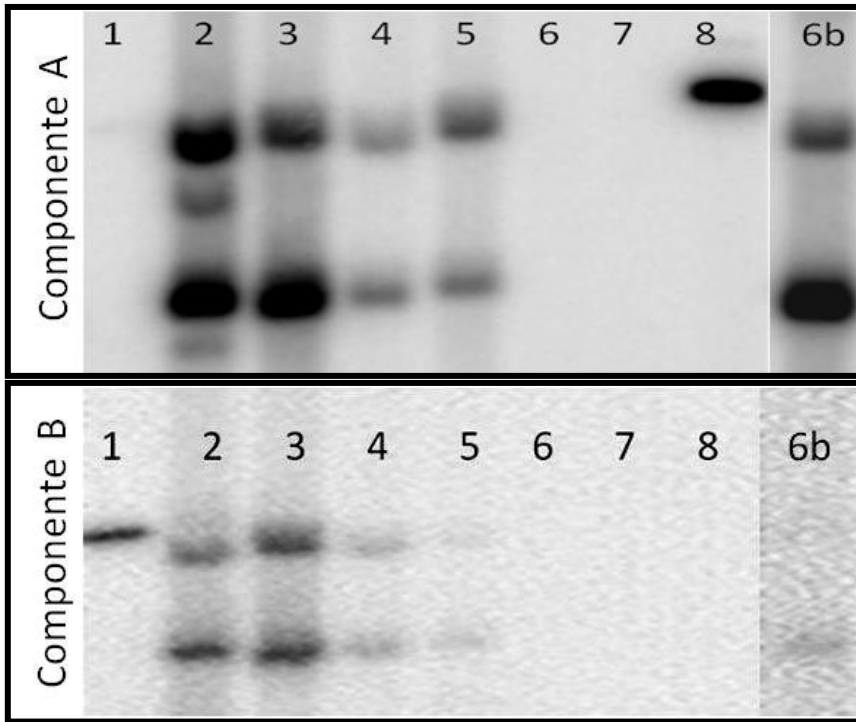


**Figura 11. Organización de los iterones y otros elementos diferenciales entre los begomovirus del clado del SLCV y los begomovirus “típicos”.** El número de iterones, representados con flechas, es mayor en los begomovirus del clado SLCV que en los begomovirus típicos: además dos de los iterones en el clado SLCV existen en la proximidad del elemento tallo-asa. Otros elementos diferenciales como las cajas GA-1, GYA y G son mostrados. (Modificada de Hernández-Zepeda *et al.*, 2009)

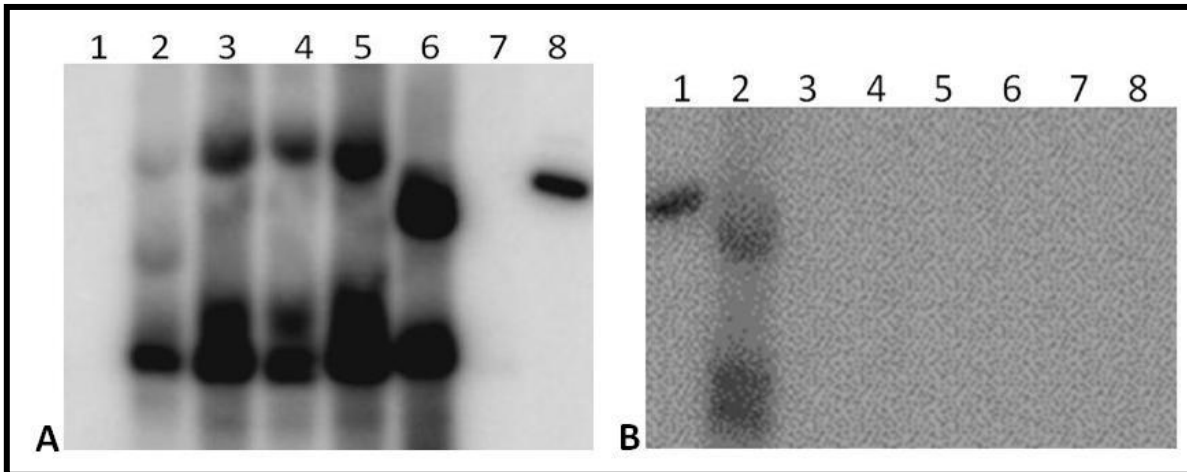
### III.5.- Transreplicación de componentes B heterólogos por los componentes A híbridos

Habiendo evaluado la capacidad de autorreplicación de los virus híbridos el siguiente paso fue evaluar la habilidad de replicar los componentes B heterólogos

con el origen de replicación en común. Para llevar a cabo este experimento se cobombardaron los híbridos de TohibTaino, TohibPYMV, TohibPHYVV y TohibTGMV con su respectivo componente B heterólogo (Tabla 4). Los híbridos de OkYMMV y BGV-Sinaloa no se usaron en este experimento ya que no se contaba con los componentes B de estos virus. Como control negativo en el experimento de trans-replicación también se hicieron las coinoculaciones de ToMoV-A tipo silvestre con cada uno de los componentes B heterólogos (Tabla 4). El Southern blot realizado al DNA total de las células coinoculadas (Figura 12) mostró que los cuatro virus híbridos son capaces de trans-replicar el componente B heterólogo respectivo; lo que contrasta con el resultado obtenido con el virus ToMoV-A silvestre el cual no trans-replicó ninguno de estos componentes heterólogos (Figura 13). Es de recalcar en este punto que la proteína Rep quimérica producida por cada uno de los virus híbridos solo difiere de la producida por ToMoV-A en los 11 primeros aminoácidos. En la figura 12 se muestra en la parte superior la auto-replicación del componente A. La variación en la intensidad de la señal en cada caso puede estar relacionada con la eficiencia de inoculación de las células durante el bombardeo, para esta réplica en particular, más que con la eficiencia de replicación *per se*, ya que en otras réplicas la intensidad de señal para TohibPYMV y TohibPHYVV fue igual o superior a la intensidad observada con el virus silvestre ToMoV A (Figura 10). La parte inferior de la figura 12 nos muestra la trans-replicación del componente B en las mismas muestras en que se evaluó la autorreplicación. Se observa una correlación entre la señal que se obtiene para el componente A y la obtenida para el componente B en los casos de TohibTaino, TohibPHYVV y TohibPYMV, sin embargo en el caso de TohibTGMV aun cuando tenemos una señal muy alta en el componente A la señal en el componente B es muy baja. Otra característica encontrada en el virus híbrido de TGMV fue que de manera consistente mostramos una mayor intensidad de banda para la cadena sencilla que para la cadena doble. Ambas observaciones podrían estar reflejando algún defecto en la proteína quimérica para reconocer el origen de replicación o para interactuar con proteínas del complejo de replicación. La figura 12 muestra la réplica del carril 6 dado que el gel corrido especialmente para obtener la imagen, no mostramos señal para la muestra de TohibTGMV, hecho atribuido a error humano ya que en todas las otras réplicas realizadas siempre mostramos señal.



**Figura 12. Southern blot en DNA total de células NT1 inoculadas con el componente A de ToMoV o los virus híbridos y su respectivo componente B.** 1).- ToMoV B banda purificada de Gel de agarosa; 2).- ToMoV A / ToMoV B; 3).- TohibTaino / ToMoTV B; 4).- TohibPHYVV/ PHYVV B; 5).- TohibPYMV /PYMV B; 6.- TohibTGMV/TGMV B, 7).- pBS, 8).- ToMoV A banda purificada de gel de agarosa, 6b).- Experimento replica de TohibTGMV/TGMVB. La membrana se hibridó con una sonda específica para el componente A (arriba) y después de obtenida la imagen se regeneró e hibridó con una sonda específica para el componente B (abajo).

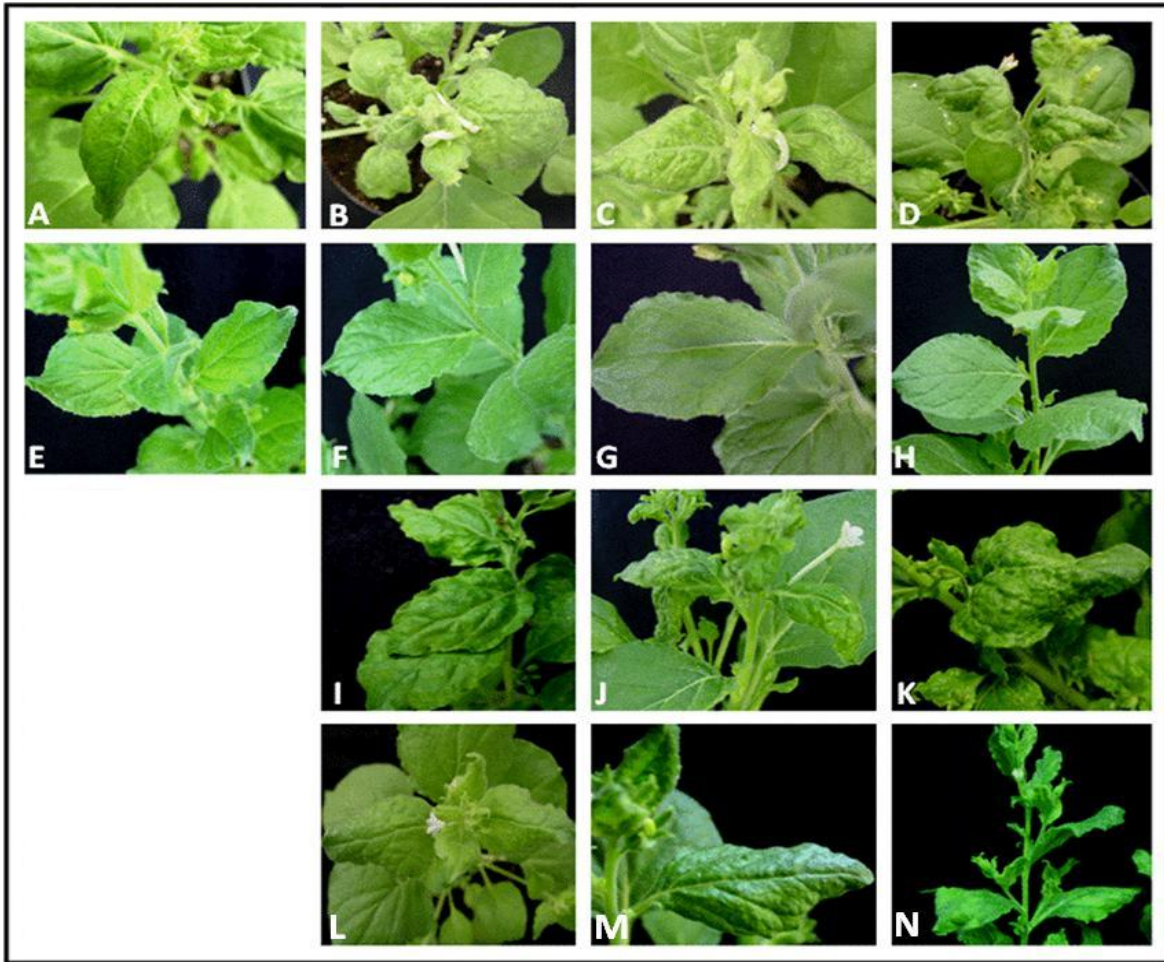


**Figura 13. Southern blot en DNA total de células NT1 inoculadas con el virus ToMoV A y los componentes B heterologos.** 1).- ToMoV B purificado de gel de agarosa. 2).- ToMoVA/ToMoVB, 3).-ToMoVA/ToMoTVB, 4).- ToMoVA/PYMVB, 5).- ToMoVA/PHYVVB, 6).- ToMoVA/TGMVB, 7).- pBS, 8).- ToMoV A purificado de gel de agarosa. En A la membrana fue hibridada con una sonda radiactiva para el componente A y después de tomarse la imagen se regenero y se hibrido con la sonda para el componente B.

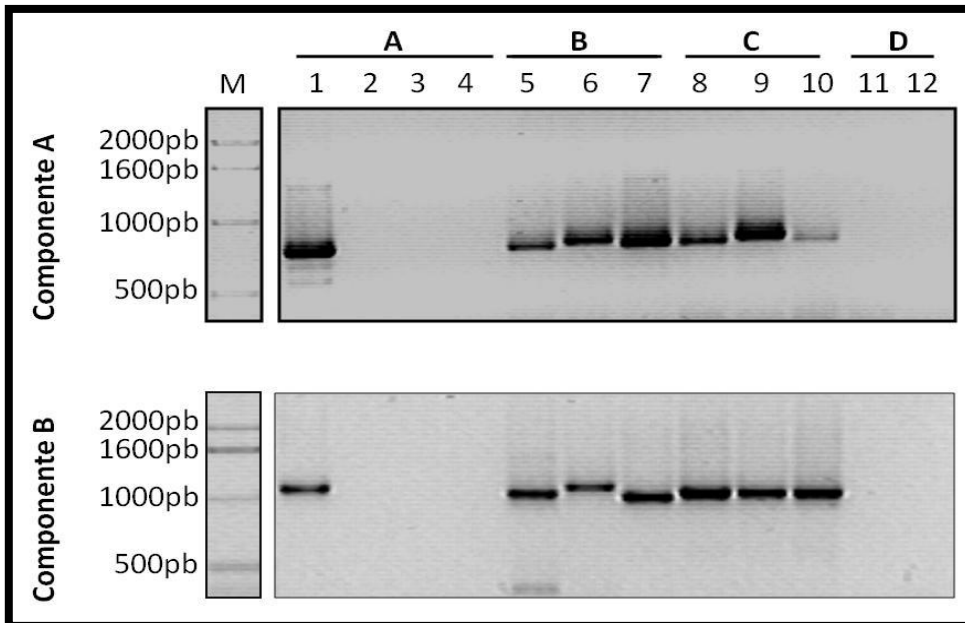
### III.6.- Infección de plantas *Nicotiana benthamiana* con los virus híbridos

La habilidad de los virus híbridos para autorreplicarse y trans-replicar los componentes B heterologos en el sistema de células NT1, y a la vez la variabilidad que encontramos de un resultado a otro en el sistema de células y la cual no es posible eliminar, nos llevó a plantear la posibilidad de evaluar la capacidad de los híbridos para infectar plantas. Para llevar a cabo el experimento se inocularon plantas de *N. benthamiana* mediante biobalística y se mantuvieron en observación para detectar el desarrollo de síntomas. A continuación se explican las combinaciones con las que se infectaron las plantas (Tabla 6). Como grupo control se infectaron las plantas con las combinaciones siguientes: ToMoV A con ToMoV B (6 plantas), ToMoV A con Taino B (6 plantas), ToMoV A con PHYVV B (6 plantas) y ToMoV A con TGMV B (6 plantas), de estas plantas solo las inoculadas con la combinación de ToMoV A con ToMoV B dio síntomas mientras que las otras combinaciones no produjeron síntoma alguno (Figura 14, incisos A, F, G y H; Tabla 6). Otro grupo de plantas fue inoculado con la combinación de cada uno de los virus híbridos y el componente B de ToMoV por un lado y por otro con el componente B heterólogo correspondiente al híbrido (Tabla 6). En este caso, las plantas inoculadas con cada virus híbrido y su componente B heterólogo correspondiente mostraron una sintomatología muy severa (Figura 14, incisos I, J

y K) mientras que las plantas inoculadas con cada virus híbrido y el componente B de ToMoV, de manera no esperada, también mostraron síntomas aunque estos fueron de manera atenuada y con un retraso de al menos 5 días con respecto a la infección tipo silvestre (Figura 14, incisos L, M y N; Tabla 6). Los síntomas observados en las plantas inoculadas con cada híbrido y componente B heterólogo fueron más semejantes a los síntomas que produce la infección de los virus heterólogos silvestres (Figura 14, incisos B, C y D) que a la producida por ToMoV (Figura 14 inciso A). De esta manera la infección por TohibTaino con ToMoTV B genera un enchinamiento tenue y amarillamiento de las hojas de manera parecida (no idéntica) a la infección de ToMoTV-A con ToMoTV-B (Figura 14 incisos I y B). La infección de TohibPHYVV con PHYVV-B genera una deformación severa de la hoja, con un enrollamiento hacia el envés de la hoja y amarillamiento de las venas (Figura 14, incisos J y C), y el híbrido TGMV con TGMV-B genera enchinamiento y moteados amarillos de la hoja de manera similar a la infección de TGMV A con TGMV B (Figura 14, incisos D y K). En contraste con lo anterior los síntomas de las plantas inoculadas con TohibTaino y ToMoV B, TohibPHYVV y ToMoV B y TohibTGMV y ToMoV B se asemejan más a los síntomas generados por ToMoVA con ToMoV B mostrando un enrollamiento de la hoja hacia el envés sin amarillamientos moteados o en las venas. También es de notar que el enrollamiento en las infecciones con los híbridos se da de manera menos acentuada que en el caso de la infección con el virus silvestre (Figura 14, incisos L, M, N y A). En este experimento la presencia del DNA viral en las plantas sintomáticas, así como la ausencia en las plantas asintomáticas, se corroboró por PCR con los oligonucleótidos DGRS forward / CP-Xba-Gus para el componente A y con los oligonucleótidos BC1290 Forward/BV1 310 Reverso para el componente B (Figura 15). Estos resultados nos llevan a pensar que la interacción de las proteínas quiméricas con el origen de replicación del componente B heterólogo es fuerte y eficiente ya que se encuentran los sitios de reconocimiento (iterones) para el IRD específico mientras que para el componente B de ToMoV que no los posee el reconocimiento es débil y se refleja en una atenuación y retardo en la aparición de los síntomas (Tabla 6). Otra posible explicación es que el reconocimiento dado entre las proteínas Rep quiméricas y el componente B de ToMoV se deba a que este componente haya recombinado con el componente A híbrido y con ello haya obtenido elementos del origen de replicación suficientes para establecer la infección. Para dilucidar esta última hipótesis los PCR realizados para el componente A y componente B (Figura 15) se clonaron en el vector PGEM-Teasy y se secuenciaron para analizarla progenie viral en las plantas.



**Figura 14. Plantas inoculadas con combinaciones diferentes del componente viral A y B. A).**- ToMoV A:ToMoV B, **B).**-ToMoTV A:ToMoTV B, **C).**-PHYVVA: PHYVV B, **D).**-TGMV A:TGMV B, **E).**- pBluescript, **F).**- ToMoV A: ToMoTV B, **G).**- ToMoV A : PHYVV B, **H).**- ToMoV A : TGMV B, **I).**- TohibTaino : ToMoTV B, **J).**- TohibPHYVV : PHYVV B, **K).**- TohibTGMV : TGMV B, **L).**- TohibTaino : ToMoV B, **M).**- TohibPHYVV : ToMoV B, **N).**- TohibTGMV : ToMoV B.



**Figure 15. Detección por PCR de los componentes virales A y B en hojas apicales de las plantas de *Nicotiana benthamiana* inoculadas. A) Plantas coinoculadas con ToMoV A y los componentes B heterologos. A1) ToMoV A: ToMoV B, A2)ToMoV A:ToMoTV B, A3)ToMoV A:PHYVV B, A4) ToMoV A:TGMV B,. B) plantas inoculadas con los virus híbridos y su respectivo componente B heterólogo. B5) TohibTaino:ToMoTV B, B6) TohibPHYVV:PHYVV B, B7) TohibTGMV:TGMV B. C) Plantas coinoculadas con los virus híbridos y el componente B de ToMoV. C8) TohibTaino:ToMoV B, C9) TohibPHYVV:ToMoV B, C10) TohibTGMV:ToMoV B. D) planta inoculada con el vector vacío (11) y planta sin inocular (12).**

**Tabla 7.** Descripción de los síntomas presentados por las plantas *Nicotiana benthamiana* inoculadas con cada una de las combinaciones de componentes A y B.

Construcción	Componente B	Observaciones en las plantas	Plantas infectadas / plantas inoculadas	Días post inoculación de aparición de síntomas
ToMoV A (tipo silvestre)	ToMoV	Amarillamiento (moteado) y enrollamiento apical de hojas	11/11	15
	ToMoTV	Asintomáticas	0/6	NA
	PYMV	Asintomáticas	0/6	NA
	PHYVV	Asintomáticas	0/6	NA
	TGMV	Asintomáticas	0/6	NA
TohibToMoTV	ToMoTV	Amarillamiento de hoja	6/11	21
	ToMoV	Amarillamiento y enrollamiento apical de la hoja de manera atenuada	6/6	21
TohibPYMV	PYMV	Asintomáticas	0/6	NA
	ToMoTV	Asintomáticas	0/6	NA
	ToMoV	Asintomáticas	0/6	NA
TohibPHYVV	PHYVV	Severo enrollamiento apical de la hoja, enchinamiento y moteado amarillo	6/11	20
	ToMoV	Amarillamiento y enrollamiento apical de la hoja de manera atenuada	6/6	21
TohibTGMV	TGMV	Mosaico amarillo y enchinamiento de la hoja	7/12	21
	ToMoV	Amarillamiento y enrollamiento apical de la hoja de manera atenuada	2/6	23



### III.7.- PCR de las plantas inoculadas y secuenciación de los productos obtenidos

De cada una de las plantas de los experimentos previamente descritos se extrajo DNA total y de este se tomo 1ul para utilizarlo en pruebas de PCR usando oligonucleótidos para amplificar tanto el componente A como el componente B de la progenie viral. Del componente A se obtuvo una banda de 900 pb que comprende desde el codón 125 del gen *Rep* hasta el inicio del gen *CP*. Para el componente B se usaron el par de oligonucleótidos BC1290 forward y BV1310 reverso que amplifican 1100pb. Ambos productos de PCR abarcan la región intergénica y en el caso de del componente A abarca hasta el codón 134 de la proteína *Rep*. Estos productos de PCR se clonaron en PGEM-Teasy y se secuenciaron. Las secuencias obtenidas se alinearon con las secuencias de ToMoV y de los híbridos inoculados para detectar posibles re-arreglos genómicos. Los alineamientos mostraron que la progenie viral era idéntica a la usada para la infección por lo que se deduce que el reconocimiento es dado por la proteína quimérica sin modificaciones y que el reconocimiento es aparentemente más débil debido a la menor capacidad de interactuar con el origen de replicación. Estos resultados correlacionan con aquellos ejemplos en los que el componente B de un virus puede ser trans-replicado por el componente A de otro aún cuando parecería no haber compatibilidad entre el IRD y los iterones (Andrade *et al.*, 2007). También da pie este resultado a la discusión de cómo los satélites de DNA que acompañan a los virus monopartitas en el Viejo Mundo, pueden ser trans-replicados por más de una especie viral con diferentes IRD (Saunders *et al.*, 2008).

#### IV.- DISCUSIÓN

La unión de la proteína Rep al origen de replicación viral ha mostrado que es altamente específica y que esta especificidad está determinada por elementos en *cis* (iterones) presentes en el *Ori* (Lazarowitz *et al.*, 1992; Fontes *et al.*, 1992; Fontes *et al.*, 1994; Argüello-Astorga *et al.*, 1994; Choi y Stenger 1995; Choi y Stenger 1996; Stenger 1998;). Por otro lado los determinantes de la replicación específica en *Trans* que interactúan con los iterones han sido mapeados dentro de los primeros 120 aminoácidos de la región amino terminal de la proteína Rep (Jupin *et al.*, 1996; Choi y Stenger, 1995; Campos-Olivas *et al.*, 2002). En una aproximación más fina se mostro que para el virus ToLCV-New Delhi el aminoácido 10 de la proteína es responsable de interactuar específicamente con el tercer nucleótido de los iterones del virus (Chatterji *et al.*, 1999). En otro trabajo, mediante análisis comparativo de secuencias virales, se propuso la hipótesis de que dentro de los primeros 11 aminoácidos de la proteína se encontraban los residuos responsables de la replicación específica de los geminivirus, estos residuos de aminoácidos forman un dominio denominado Dominio relacionado al iteron (IRD) (Argüello-Astorga y Ruiz-Medrano, 2001). En nuestros experimentos observamos que el componente A de ToMoV es incapaz de manera silvestre de establecer una infección con los componentes B heterologos de ToMoTV, PHYVV, y TGMV. Pero, en cambio los virus híbridos TohibTaino, TohibPHYVV y TohibTGMV que producen una proteína Rep quimérica con el IRD heterólogo si pueden establecer la infección con el componente B de ToMoTV, PHYVV y TGMV respectivamente. Este resultado muestra que el IRD es suficiente para conferir a la proteína Rep de ToMoV la capacidad de unir de manera específica el origen de replicación del componente B heterólogo.

Una observación inesperada fue que los virus híbridos no pierden al cien por ciento su capacidad para establecer una infección con el componente B de ToMoV si no que mantuvieron su capacidad infectiva aunque los síntomas fueron tenues y tardíos. Esta observación nos hace pensar que, además del IRD, existe otro u otros dominios en la proteína Rep quimérica que al no ser alterados pueden interactuar con baja eficiencia con los elementos de replicación en *cis* en el componente B de ToMoV. El trabajo que determino la estructura tridimensional de la región amino-terminal de Rep y en el análisis teórico del las secuencias de Rep (Argüello-Astorga y Ruiz-Medrano, 2001; Campos-Olivas *et al.*, 2002), señalan que el domino beta 5 al estar formando la lamina beta con el dominio beta 1(IRD) también interactúa con el origen de replicación. En las proteínas quiméricas este dominio permanece intacto y podría ser el causal del reconocimiento débil de ToMoV B por los virus híbridos. Por otro lado podemos citar aquellos trabajos en los que el componente A de un virus puede trans-replicar

componentes B de otros que no poseen iterones iguales (Andrade *et al.*, 2006) o el caso de los begomovirus del viejo mundo que trans-repican DNAs beta satélites sin poseer iterones semejantes (Lin *et al.*, 2003; Dry *et al.*, 1997). Las proteínas Rep quimericas de ToMoV tienen un comportamiento muy semejante a las de los virus ToYSV y virus del viejo mundo analizados en estos trabajos. Los autores de estos trabajos sugieren que la proteína Rep es un tanto versátil en la replicación y que no es un requisito la conservación de la relación de los iterones y el IRD. Según nuestros resultados parece haber cabida a ambas ideas, por un lado el IRD es de gran importancia durante la replicación que el hecho de cambiarlos en la proteína de ToMoV le permitió reconocer de manera muy eficiente los iterones de los virus heterologos y por otro lado puede permitirse cambios en el IRD de la proteína Rep de ToMoV sin perder de todo su capacidad replicativa. Ahora la tarea es buscar en la proteína Rep esos factores además del IRD que permiten la replicación o buscar factores presentes en la planta que pudieran colaborar con la proteína Rep para formar el complejo inicial con el DNA.

Una indicación de la importancia de los factores de la planta para la acción de la proteína Rep nos la da el virus híbrido de TohibPepGMV en donde llevando aparentemente los elementos virales necesarios, IRD e iterones, este no puede replicarse.

## V.- CONCLUSIONES

En nuestro trabajo se ha evaluado experimentalmente la participación de los 11 aminoácidos iniciales como un factor que determinan la unión específica de la proteína Rep de geminivirus a los iterones en el origen de replicación. Los resultados nos indican que los 11 aminoácidos de las proteínas Rep de ToMoTV, PHYVV y TGMV son suficientes para conferir especificidad hacia su origen de replicación cognado logrando que las proteínas quiméricas de ToMoV pudieran autoreplicar el virus híbrido y pudieran transreplicar los componentes B heterologos tanto en la línea celular NT1 como en las plantas de *Nicotiana benthamiana*.

Para la proteína Rep de ToMoV, puesto que las quimeras de la proteína aun mantienen la capacidad de unir el componente B de ToMoV, queda abierta la pregunta de cuan específica es y de qué participación pudieran tener otras regiones de la proteína en reconocer los iterones cognado. Si analizamos la secuencia de los iterones de ToMoV, GGAGTATTAGAGTA, observamos que la primera base de los iterones directos no se conserva, llevándonos a pensar en una baja especificidad necesaria para ser reconocidos por la proteína Rep. Si este planteamiento no fuera acertado otra posibilidad que explique porque las tres diferentes quimeras de Rep de ToMoV mantienen el reconocimiento de ToMoV B puede encontrarse en los títulos de Rep necesarios para replicar el componente B. Así, el hecho de que la sintomatología en la plantas de *Nicotiana benthamiana* se dé con un retraso de 3 o más días nos indica la necesidad de una mayor concentración de la proteína Rep quiméricas para poder reconocer los iterones y con una baja eficiencia reflejada en la atenuación de los síntomas.

Una tercera conclusión podemos obtener del resultado con la proteína quimérica de TohibPepGMV, en este caso el IRD no fue capaz de reconocer los iterones cognado. Estos nos sugiere que para este virus, que pertenece a un clado diferente al de ToMoV, es necesaria alguna otra región, probablemente la región beta 5, para poder interactuar con los iterones o con factores de las células que permitan la replicación. Una característica encontrada por Arguello-Astorga y Ruiz-Medrano (2001) en los virus del clado al que pertenece PepGMV fue que, aunque existen virus con iterones similares a virus del clado de ToMoV, su IRD no se asemeja. Esto puede implicar que existe alguna particularidad en la proteína Rep de los virus de este clado (SLCV) para interactuar con los iterones.

De manera general los experimentos demuestran la importancia de los primeros 11 aminoácidos de la proteína Rep de geminivirus para la replicación y pensando en estrategias de control del patógeno estos aminoácidos son blanco para bloquear la interacción de la proteína con el DNA, inhibir la replicación y prevenir

la infección. Por otro lado pueden generarse proteínas Rep trucas sin capacidad de replicación pero con afinidad específica hacia el DNA de un geminivirus específico que compitan por el sitio de unión inhibiendo la replicación viral en una infección.

## VI.- REFERENCIAS

Akhter, A., Qazi, J., Saeed, M., Mansoor, S. (2009). A severe leaf curl disease on chilies in Pakistan is associated with multiple begomovirus components. *Plant disease*, 93: 962.

Andrade, E.C., Manhani, G.G., Alfenas, P.F., Calegario, R.F., Fontes, E.P. y Zerbini, F.M. (2006). Tomato yellow spot virus, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from *Sida* sp., forms pseudorecombinants with begomoviruses from tomato but not from sida. *J Gen Virol*, 87: 3687-3696.

Arguello-Astorga, G.R., Lopez-Ochoa, L., Kong, L.J., Orozco, B.M., Settlege, S.B. y Hanley-Bowdoin L. (2004). A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. *J Virol*, 78(9): 4817-4826.

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(8): 1465-1485.

Argüello-Astorga, G.R., Guevara-González, R.G., Herrera-Estrella, L.R. y Rivera-Bustamante, R.F. (1994). Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology*, 203(1): 90-100.

Bagewadi, B., Chen, S., Lal, S.K., Choudhury, N.R. y Mukherjee SK. (2004). PCNA interacts with Indian mung bean yellow mosaic virus rep and downregulates Rep activity. *J Virol*, 78(21): 11890-11903.

Bisaro, D.M. (2006). Silencing suppression by geminivirus proteins. *Virology*, 344(1): 158-168.

Bottcher, B., Unseld, S., Ceulemans, H., Russell, R.B., Jeske, H. (2004). Geminata structures of *African cassava mosaic virus*. *J Virol*, 78: 6758–6765.

Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X., Fauquet, C.M. (2008). Recommendations for the classification and nomenclature of the DNA-beta satellites of begomoviruses. *Arch Virol*, 153(4): 763-781.

Briddon, R.W. y Markham, P.G. (2001). Complementation of bipartite begomovirus movement functions by topocoviruses and curtoviruses. *Arch Virol*, 146: 1811–1819.

Briddon, R.W., Bedford, I.D., Tsai, J.H. y Markham, P.G. (1996). Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, Tomato pseudo-curly top virus, suggests a recombinant origin. *Virology* 219(2): 387-394.

Campos-Olivas, R., Louis, J.M., Clerot, D., Gronenborn, B., y Gronenborn, A.M. (2002). The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. *PNAS*, 99(16): 10310-10315.

Carrillo-Tripp, J., Shimada-Beltrán, H. y Rivera-Bustamante, R. (2006). Use of geminiviral vectors for functional genomics. *Curr Opin Plant Biol*, 9(2): 209-215.

Castillo, A.G., Kong, L.J., Hanley-Bowdoin, L. y Bejarano, E.R. (2004). Interaction between a geminivirus replication protein and the plant sumoylation system. *J Virol*, 78(6): 2758-2769.

Chakraborty, S., Vanitharani, R., Chattopadhyay, B. y Fauquet, C.M. (2008). Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India. *J Gen Virol*, 89: 818-828.

Chatterji, A., Padidam, M., Beachy, R.N. y Fauquet, C.M. (1999). Identification of replication specificity determinants in two strains of tomato leaf curl virus from New Delhi. *J Virol*, 73(7): 5481-5489.

Choi, I.R. y Stenger, D.C. (1996). The strain-specific cis-acting element of beet curly top geminivirus DNA replication maps to the directly repeated motif of the ori. *Virology*, 226(1): 122-126.

Choi, I.R. y Stenger DC. (1995). Strain-specific determinants of beet curly top geminivirus DNA replication. *Virology*, 206(2): 904-912.

Choudhury, N.R., Malik, P.S., Singh, D.K., Islam, M.N., Kaliappan, K. y Mukherjee SK. (2006). The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. *Nucleic Acids Res*, 34(21): 6362-6377.

Clérot, D. y Bernardi, F. (2006). DNA helicase activity is associated with the replication initiator protein rep of Tomato yellow leaf curl geminivirus. *J Virol*, 80(22): 11322-11330.

Desbiez, C., David, C., Mettouchi, A., Laufs, J. y Gronenborn, B. (1995). Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *PNAS*, 92(12): 5640-5644.

Donson J, Morris-Krsinich BA, Mullineaux PM, Boulton MI, Davies JW. (1984). A putative primer for second-strand DNA synthesis of Maize streak virus is virion-associated. *EMBO J*, 3(13): 3069-3073.

Dry, I.B., Krake, L.R., Rigden, J.E., Rezaian, M.A. (1997). A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. *PNAS*, 94(13): 7088-93.

Elmer, J.S., Brand, L., Sunter, G., Gardiner, W.E., Bisaro, D.M. y Rogers S.G. (1988). Genetic analysis of the Tomato golden mosaic virus. II. The product of the AL1 coding sequence is required for replication. *Nucleic Acids Res*, 16: 7043-7060.

Erdmann, J.B., Shepherd, D.N., Martin, D.P., Varsani, A., Rybicki, E.P. y Jeske, H. (2010). Replicative intermediates of Maize streak virus found during leaf development. *J Gen Virol*, 91: 1077-1081.

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(4): 783-821.

Florentino, L.H., Santos, A.A., Fontenelle, M.R., Pinheiro, G.L., Zerbini, F.M., Baracat-Pereira, M.C., y Fontes E.P.B. (2006). A PERK-like receptor kinase interacts with the geminivirus nuclear shuttle protein and potentiates viral infection. *J Virol*, 80(13): 6648-6656.

Fontes, E.P.B., Gladfelter, H.J., Schaffer, R.L., Petty, I.T.D., y Hanley-Bowdoin L. (1994). Geminivirus replication origins have a modular organization. *The plant cell*, 6: 405-416.

Fontes, E.P.B., Eagle, P.A., Sipe, P.S., Luckow, V.A. y Hanley-Bowdoin L. (1994). Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J Biol Chem*, 269(11): 8459-8465.

Fontes, E.P.B., Luckow, V.A., y Linda Hanley-Bowdoin, L. (1992). A geminivirus replication protein is a sequence-specific DNA binding protein. *The plant cell*, 4: 597-608.

Hanley-Bowdoin, L., Settlege, S.B., Orozco, B.M., Nagar, S. y Robertson, D. (1999). Geminiviruses: models for plant DNA replication, transcription and cell cycle regulation. *Crit Rev Plant Sci*, 18: 71–106.

Hanley-Bowdoin, L., Elmer, J.S. y Rogers, S.G. (1990). Expression of functional replication protein from Tomato golden mosaic virus in transgenic tobacco plants. *PNAS*, 87(4): 1446-1450.

Hayes, R.J., Mcdonal, H., Coutts, R.H.A., y Buck K.W. (1988). Priming of complementary DNA synthesis in vitro by small DNA molecules tightly bound to virion DNA of Wheat dwarf virus. *J. gen. Virol.*, 69: 1345-1350.



Hernández-Zepeda, C., Argüello-Astorga, G., Idris, A.M., Carnevali, G., Brown, J.K. y Moreno-Valenzuela, O.A. (2009). Molecular characterization and phylogenetic relationships of *Desmodium* leaf distortion virus (DeLDV): a new begomovirus infecting *Desmodium glabrum* in Yucatan, Mexico. *Virus Genes*, 39(3): 371-374.

Hormuzdi, S.G. y Bisaro, D.M. (1995). Genetic analysis of Beet curly top virus: examination of the roles of L2 and L3 genes in viral pathogenesis. *Virology*, 206(2): 1044-1054.

Hormuzdi, S.G. y Bisaro, D.M. (1993). Genetic analysis of Beet curly top virus: evidence for three virion sense genes involved in movement and regulation of single- and double-stranded DNA levels. *Virology*, 193(2): 900-909.

Hung, H.C. y Petty, I.T. (2001). Functional equivalence of late gene promoters in Bean golden mosaic virus with those in Tomato golden mosaic virus. *J Gen Virol*, 82: 667-672.

Ilyina, T.V., y Koonin, E.V. (1992). Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Research*, 20(13): 3279-3285.

Jeske, H., Lütgemeier, M., Preiss, W. (2001). DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon* mosaic virus. *EMBO J.* 20(21): 6158-6167.

Jupin, I., Hericourt, F., Benz, B. y Gronenborn B. (1995). DNA replication specificity of TYLCV geminivirus is mediated by the amino-terminal 116 amino acids of the Rep protein. *FEBS*, 362(2): 116-120.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S.G., Schell, J. y Gronenborn B. (1995). In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of Tomato yellow leaf curl virus. *PNAS*, 92(9): 3879-3883.

Lazarowitz, S.G., Wu, L.C., Rogers, S.G. y Elmer, J.S. (1992). Sequence-specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. *Plant Cell*, 4(7): 799-809.

Lazarowitz, S.G., Pinder, A.J., Damsteegt V.D. y Rogers S.G. (1989). Maize streak virus genes essential for systemic spread and symptom development. *The EMBO Journal*, 8(4): 1023-1032.

Lefeuvre, P., Martin, D.P., Harkins, G., Lemey, P., Gray, A.J., Meredith, S., Lakay, F., Monjane, A., Lett, J.M., Varsani, A., Heydarnejad, J. (2010). The spread of

Tomato yellow leaf curl virus from the Middle East to the world. *PLoS Pathog*, 6(10): e1001164.

Legg, J.P. y Thresh, J.M. (2000). Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus research*, 71(1-2): 135-149.

Lewis J.D. y Lazarowitz S.G. (2009). Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. *PNAS*, 107(6): 2491–2496.

Lin, B., Akbar Behjatnia, S.A., Dry, I.B., Randles, J.W. y Rezaian MA.(2003). High-affinity Rep-binding is not required for the replication of a geminivirus DNA and its satellite. *Virology*, 305(2): 353-363.

Liu, S., Bedford, I.D., Briddon, R.W. y Markham, P.G. (1997). Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *J Gen Virol*, 78: 1971-1974.

Luque, A., Sanz-Burgos, A.P., Ramirez-Parra, E., Castellano, M.M. y Gutierrez C.(2002). Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. *Virology*, 302(1): 83-94.

Mumford, D.L. (1974). Purification of Curly top virus. *Phytopathology*, 64: 136-139.

Orozco, B.M., Kong, L.J., Batts, L.A., Elledge, S., Hanley-Bowdoin, L. (2000). The multifunctional character of a geminivirus replication protein is reflected by its complex oligomerization properties. *J Biol Chem*, 275(9): 6114-6122.

Orozco, B.M. y Hanley-Bowdoin, L. (1998). Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. *J Biol Chem*, 273(38): 24448-24456.

Patil, B.L. y Fauquet, C.M. (2009). Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol Plant Pathol*, 10(5): 685-701.

Pilartz, M. y Jeske, H. (2003). Mapping of Abutilon mosaic geminivirus minichromosomes. *J Virol*, 77(20): 10808-10818.

Preiss, W. y Jeske, H. (2003). Multitasking in replication is common among Geminiviruses. *J Virol*, 77(5): 2972–2980.

Revington, G.N., Sunter, G. y Bisaro, D.M. (1989). DNA sequences essential for replication of the B genome component of Tomato golden mosaic virus. *Plant Cell*, 10: 985-992.

Rodríguez-Negrete, E.A., Carrillo-Tripp, J. y Rivera-Bustamante, R.F. (2007). RNA silencing against geminivirus: complementary action of posttranscriptional gene silencing and transcriptional gene silencing in host recovery. *J Virol*, 83(3): 1332-1340.

Salati, R., Nahkla, M.K., Rojas, M.R., Guzman, P, Jaquez, J, Maxwell, D.P., y Gilbertson, R.L. (2002). Tomato yellow leaf curl virus in the Dominican Republic: Characterization of an infectious clone, virus monitoring in whiteflies, and identification of reservoir hosts. *Phytopathology*, 92: 487-496.

Sambrook, L.; Russell, D.W. (2001). *Molecular cloning: a laboratory manual*. 3rd Ed. New York, USA: Cold spring Harbor Laboratory Press.

Sánchez-Durán, M.A., Dallas, M.B., Ascencio-Ibañez, J.T., Reyes, M.I., Arroyo-Mateos, M., Ruiz-Albert, J., Hanley-Bowdoin, L. y Bejarano, E.R. (2011). Interaction between geminivirus replication protein and the SUMO-conjugating enzyme is required for viral infection. *J Virol*, 85(19): 9789-9800.

Sanderfoot, A.A., Ingham, D.J. y Lazarowitz, S.G.(1996). A viral movement protein as a nuclear shuttle. The geminivirus BR1 movement protein contains domains essential for interaction with BL1 and nuclear localization. *Plant Physiol*, 110(1): 23-33.

Saunders, K., Briddon, R.W. y Stanley, J. (2008). Replication promiscuity of DNA-beta satellites associated with monopartite begomoviruses; deletion mutagenesis of the *Ageratum* yellow vein virus DNA-beta satellite localizes sequences involved in replication. *J Gen Virol*, 89: 3165-3172.

Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M. y Stanley, J. (2000). A unique virus complex causes *Ageratum* yellow vein disease. *PNAS*, 97(12): 6890-6895.

Saunders, K., Lucy, A. y Stanley, J. (1992). RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. *Nucleic Acids Research*, 20(23): 6311-6315.

Saunders, K., Lucy, A. y Stanley, J. (1991). DNA forms of the geminivirus African cassava mosaic virus consistent with a rolling circle mechanism of replication. *Nucleic Acids Res*, 19(9): 2325-2330.

Settlage, S.B., See, R.G. y Hanley-Bowdoin, L. (2005). Geminivirus C3 protein: replication enhancement and protein interactions. *J Virol*, 79(15): 9885-9895.

Sharma, P. y Ikegami, M.(2009). Characterization of signals that dictate nuclear/nucleolar and cytoplasmic shuttling of the capsid protein of Tomato leaf curl Java virus associated with DNA beta satellite. *Virus Res*, 144:145-153.

Shepherd, D.N., Martin, D.P., Van Der Walt, E., Dent, K., Varsani, A. y Rybicki EP. (2010). Maize streak virus: an old and complex 'emerging' pathogen. *Mol Plant Pathol*, 11(1): 1-12.

Shimada-Beltrán, H. y Rivera-Bustamante, R.F.(2007). Early and late gene expression in Pepper huasteco yellow vein virus. *J Gen Virol*, 88 : 3145-3153.

Stanley, J., y Latham, J.R. (1992). A symptom variant of Beet curly top geminivirus produced by mutation of open reading frame C4. *Virology*, 190(1): 506-9.

Stenger, D.C. (1998). Replication specificity elements of the Worland strain of Beet curly top virus are compatible with those of the CFH strain but not those of the Cal/Logan strain. *Phytopathology*, 1998 88(11): 1174-1178.

Teng, K., Chen, H., Lai, J., Zhang, Z., Fang, Y., Xia, R., Zhou, X., Guo, H. y Xie Q. (2010). Involvement of C4 protein of Beet severe curly top virus (family Geminiviridae) in virus movement. *PLoS One*, 5(6): e11280.

Van der Walt, E., Martin, D.P., Varsani, A., Polston, J.E. y Rybicki EP. Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. *Virol J*, 5: 104.

Vanitharani, R., Chellappan, P., Pita, J.S. y Fauquet, C.M. (2004). Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. *J Virol*, 78(17): 9487-9498.

Vega-Rocha, S.,Byeon, I.J., Gronenborn, B., Gronenborn, A.M., y Campos-Olivas, R.(2007a). Solution structure, divalent metal and DNA binding of the endonuclease domain from the replication initiation protein from Porcine circovirus 2. *J Mol Biol*, 367(2): 473-487.

Vega-Rocha, S., Gronenborn, B., Gronenborn, A.M., y Campos-Olivas, R.(2007). Solution structure of the endonuclease domain from the master replication initiator protein of the nanovirus Faba bean necrotic yellows virus and comparison with the corresponding geminivirus and circovirus structures. *Biochemistry*, 46(21): 6201-6212.

Yazdi, H.R., Heydarnejad, J. y Massumi, H. (2008). Genome characterization and genetic diversity of Beet curly top Iran virus: a geminivirus with a novel nonanucleotide. *Virus Genes*, 36(3): 539-545.

Zhang, W., Olson, N.H., Baker, T.S., Faulkner, L., Agbandje-McKenna, M., Boulton, M.I., Davies, J.W., McKenna, R. (2000). Structure of the Maize streak virus geminate particle. *Virology*, 279: 471-477.

Zhou, Y., Rojas, M.R., Park, M.R., Seo, Y.S., Lucas, W.J. y Gilbertson, R.L.(2011). Histone H3 interacts and colocalizes with the nuclear shuttle protein and the movement protein of a geminivirus. *J Virol*, 85(22): 11821-11832.

## VII.- ANEXOS

### Anexo A

Artículo publicado como co-autor, en el que estuve a cargo de los experimentos de inoculación en *Nicotiana benthamiana* con todas las combinaciones posibles entre los componentes genómicos A y B de dos cepas del begomovirus *Euphorbia mosaic virus*, a través de los cuales se demostró la incompatibilidad replicativa de estas cepas debida a las diferencias existentes entre sus determinantes de especificidad replicativa en *cis* (los iterones) y en *trans* (el residuo 11 de Rep)

Gregorio-Jorge et al. *Virology Journal* 2010, 7:275  
<http://www.virologyjournal.com/content/7/1/275>



#### RESEARCH

#### Open Access

## 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

Josefat Gregorio-Jorge<sup>1†</sup>, Artemiza Bernal-Alcocer<sup>2†</sup>, Bernardo Bañuelos-Hernández<sup>1</sup>, Ángel G Alpuche-Solis<sup>1</sup>, Cecilia Hernández-Zepeda<sup>3</sup>, Oscar Moreno-Valenzuela<sup>3</sup>, Gustavo Frías-Treviño<sup>2</sup>, Gerardo R Argüello-Astorga<sup>1\*</sup>

#### Abstract

**Background:** *Euphorbia mosaic virus* (EuMV) is a member of the SLCV clade, a lineage of New World begomoviruses that display distinctive features in their replication-associated protein (Rep) and virion-strand replication origin. The first entirely characterized EuMV isolate is native from Yucatan Peninsula, Mexico; subsequently, EuMV was detected in weeds and pepper plants from another region of Mexico, and partial DNA-A sequences revealed significant differences in their putative replication specificity determinants with respect to EuMV-YP. This study was aimed to investigate the replication compatibility between two EuMV isolates from the same country.

**Results:** A new isolate of EuMV was obtained from pepper plants collected at Jalisco, Mexico. Full-length clones of both genomic components of EuMV-Jal were biologically inoculated into plants of three different species, which developed symptoms indistinguishable from those induced by EuMV-YP. Pseudorecombination experiments with EuMV-Jal and EuMV-YP genomic components demonstrated that these viruses do not form infectious reassortants in *Nicotiana benthamiana*, presumably because of Rep-iteron incompatibility. Sequence analysis of the EuMV-Jal DNA-B intergenic region (IR) led to the unexpected discovery of a 35-nt-long sequence that is identical to a segment of the *rep* gene in the cognate viral DNA-A. Similar short *rep* sequences ranging from 35- to 51-nt in length were identified in all EuMV isolates and in three distinct viruses from South America related to EuMV. These short *rep* sequences in the DNA-B IR are positioned downstream to a ~160-nt non-coding domain highly similar to the CP promoter of begomoviruses belonging to the SLCV clade.

**Conclusions:** EuMV strains are not compatible in replication, indicating that this begomovirus species probably is not a replicating lineage in nature. The genomic analysis of EuMV-Jal led to the discovery of a subgroup of SLCV clade viruses that contain in the non-coding region of their DNA-B component, short *rep* gene sequences located downstream to a CP-promoter-like domain. This assemblage of DNA-A-related sequences within the DNA-B IR is reminiscent of polyomavirus microRNAs and could be involved in the posttranscriptional regulation of the cognate viral *rep* gene, an intriguing possibility that should be experimentally explored.

#### Background

The members of the family *Geminiviridae*, one of the two largest natural groups of plant viruses, are characterized

by a circular, single-stranded DNA (ssDNA) genome encapsidated within virions whose morphology is unique in the known virosphere, consisting of two joined, incomplete T = 1 icosahedra [1,2]. Geminiviruses are classified into four genera, based on their genome organization, plant host range, and insect vector. Members of the most diversified genus, *Begomovirus*, are transmitted by the whitefly *Bemisia tabaci* (Hemiptera; Aleyrodidae), infect

\* Correspondence: [grarguel@picytedu.mx](mailto:grarguel@picytedu.mx)

† Contributed equally

<sup>1</sup>Instituto Potosino de Investigación Científica y Tecnológica, A.C., Camino a la Presa San José, 78216 San Luis Potosí, SLP, México

Full list of author information is available at the end of the article



© 2010 Gregorio-Jorge et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

a wide range of dicotyledonous plant species, and have either monopartite or bipartite genomes [3]. In recent decades, these viruses have emerged as major threats to food and fiber crop production throughout the world, apparently as a result of a great increase in vector population densities, expansion of crop monocultures, transport of plant materials between geographically distant regions, and introduction of foreigner whitefly biotypes [4,5].

Approximately 200 species of begomoviruses are currently known, grouped into two major lineages based on their genomic sequences: the Old World (OW; Europe, Africa, the Indian subcontinent, Asia, and Australasia) and the New World (NW; the Americas) begomoviruses [6,7]. The OW begomoviruses have either monopartite or bipartite genomes, while all NW begomoviruses (for simplicity, NW-Beg) have two genomic components, known as DNA-A and DNA-B. The DNA-A component of NW-Beg has one open reading frame in the virion sense (*AV1* or *cp* gene) encoding the coat protein, and four overlapped ORFs in the complementary sense (*AC1* or *rep* gene, *AC2* or *trap* gene, *AC3* or *ren* gene, and *AC4*) that encode proteins involved in DNA replication, regulation of viral gene expression and suppression of host-defense responses [1,8]. The DNA-B component contains only two ORFs, one in the virion sense (*BV1* or *nsp* gene) and other in the complementary sense (*BC1* or *mp* gene), encoding proteins involved in intra- and intercellular movement of the virus [9,10]. The two genomic components are very different in overall nucleotide sequence, with the exception of a ~180-nt segment of the intergenic region (IR) displaying high sequence identity, termed the "common region" (CR). This region includes several repeated sequences (5 to 8-nt in length) called "iterons", which are closely associated to a ~30-nt conserved element that has the potential to form a hairpin structure that harbors in its apex the invariant nonanucleotide 5'-TAATATTAC-3' [1]. Both the iterons and the conserved nonanucleotide in the hairpin element are functional targets for Rep, the virus-encoded protein that initiates the DNA replication by a rolling-circle (RCR) mechanism. Rep recognizes and binds specifically to the iterons and subsequently introduces a nick into the invariant nonanucleotide to initiate the RCR process [11,12].

The NW-Beg have radiated to a great extent since its arrival to the American continent, and several secondary lineages or "clades" have been identified in phylogenetic studies [6,13,14]. The most atypical of the NW-Beg clades is the one named after the *Squash leaf curl virus* (SLCV) that encompasses more than 15 viral species distributed from Southern EUA to Brazil [7,13]. Members of the SLCV clade are differentiated from other NW-Beg by two main features: 1) the number and

arrangement of the iterons in their replication origin, that are distinctive, and 2) the N-terminal domain (i.e., residues 1 to 150) of their Rep proteins display low aa sequence identity (< 50%) with proteins encoded by typical NW-Beg, lacking several amino acid motifs which are conserved in both NW- and OW- begomovirus Rep proteins [[15-17]; unpublished data].

Among the earliest recorded members of the SLCV-clade is *Euphorbia mosaic virus* (EuMV), which was associated with symptomatic *Euphorbia heterophylla* plants throughout the Caribbean basin and the tropical Americas since the 1970's [18,19]. However, its molecular characterization was not carried out until 2007, when the complete genome sequence of EuMV-YP, the isolate associated with the former plant host in the Yucatan Peninsula of Mexico, was reported [20]. Complete DNA-A sequences from two additional EuMV isolates were available at GenBank at that time, one from Puerto Rico (EuMV-PR) and the isolate whose complete sequence is now reported here, from Jalisco, Mexico (EuMV-Jal). According to their full-length DNA-A sequence identity, the EuMV isolates were classified into two different strains, simply termed "A" and "B". The first strain was represented by EuMV-YP and EuMV-PR, while EuMV-Jal was the only member of the "B-strain" [7]. However, the recently described EuMV-JM, from Jamaica [21], displays a very similar sequence identity to both EuMV-PR (A-strain, 95% identity) and EuMV-Jal (B-strain, 95.4% identity). Therefore, the relationship between EuMV isolates belonging to supposedly distinct strains should be experimentally addressed.

In this work we report the complete molecular characterization of EuMV-Jal, which was found infecting peppers and weeds in Jalisco, Mexico, and was shown to be incompatible in replication with EuMV-YP in reassortment experiments. The genomic analysis of this novel EuMV strain led to the unforeseen discovery of an assemblage of DNA-A homologous sequences in the intergenic region of its DNA-B, whose position and arrangement is conserved in several begomovirus species, hence suggesting the intriguing possibility of a functional role of those atypical sequences in the infective cycle of EuMV and its relatives.

## Results

### Isolation of a new strain of *Euphorbia mosaic virus*

During Autumn 2005, a survey of farming fields infested with whiteflies in the state of Jalisco, Mexico, was undertaken. Pepper plants exhibiting a variety of symptoms (including leaf curling and crumpling, yellow veins, deformed fruits, and stunted growth) were observed in fields of three Jalisco localities. Leaf samples from 63 symptomatic weeds and pepper plants were



collected, and total DNA extracts were tested for the presence of begomoviruses using polymerase chain reaction (PCR) with several pairs of degenerated primers (see Methods). More than 80% of the examined samples were PCR-positive and sequence analyses of the amplicons revealed that the majority of the symptomatic plants were infected by begomoviruses belonging to two different species, *Pepper huasteco yellow vein virus* (PHYVV) and *Pepper golden mosaic virus* (PepGMV), which commonly infect pepper and tomato crops throughout the north and central areas of Mexico [22-24]. Partial DNA-A sequences of a third begomovirus were obtained from two pepper samples from the Castillo locality (close to the Pacific coast, coordinates 19°45'00" N; 104°23'30" W), one *Nicotiana glauca* plant ("tabaquillo") collected at Sayula (coordinates 19° 47'55" N; 103°46'05" W) and one *Euphorbia heterophylla* plant collected at Teocuitatlán (coordinates 20°12'30" N; 103°30'00" W). In the four cases the plants were co-infected with either PHYVV or PepGMV. The complete sequence of the DNA-A and DNA-B genomic components of the unidentified begomovirus was obtained from overlapped PCR products derived from one pepper plant co-infected with PHYVV (see Methods). Comparisons with sequences available at the GenBank database using BlastN showed that the third pepper-infecting virus was an isolate of *Euphorbia mosaic virus*, displaying a DNA-A overall sequence identity of 95.4%, 92.8% and 92.1% with EuMV isolates from Jamaica [GenBank: DQ395342], Puerto Rico [GenBank: AF068642] and the Yucatan Peninsula [GenBank: DQ318937], respectively.

#### Genome organization of EuMV-Jal

The EuMV-Jal genome exhibited a genetic organization typical of NW-Beg. The DNA-A molecule [GenBank: DQ520942] was 2609 nt in length, and encoded five genes (*cp*, *rep*, *trap*, *ren* and *AC2*). The DNA-B molecule [GenBank: HQ185235] was 2590 nt in size, and contained two major ORFs (*BV1* and *BC1*). The common region (CR) of EuMV-Jal DNA-A and DNA-B encompassed 169 and 170 nt, respectively, with 98% identity. The CR contained the origin of replication comprising the conserved hairpin element and five iterations (GGAGTCC) that displayed the characteristic arrangement of the viruses belonging to the SLCV-cluster [15,16]. Comparisons of EuMV-Jal CR with the homologous region of other EuMV isolates revealed that EuMV-Jal and EuMV-JM have a DNA-A replication origin with a composition of putative *cis*-acting elements different to the homologous *Ori* of EuMV-YP and EuMV-PR. Indeed, in addition to harbor iterative elements with a distinct nucleotide sequence, the EuMV isolates from Jalisco and Jamaica display a G-box motif in the immediate vicinity of the conserved hairpin

element, which is absent in the DNA-A of EuMV-PR and EuMV-YP (Figure 1A). The later viruses display instead a conserved motif (GGGGCAAAA) that is characteristic of most members of the SLCV-clade (our unpublished data). In contrast with the differences observed between the DNA-A components, comparisons of the DNA-B CR revealed a similar modular organization in all EuMV isolates, with a G-box motif adjacent to the hairpin element (Figure 1B). A similar organization of the DNA-B CR is observed in *Euphorbia yellow mosaic virus* (Fernandes et al, unpublished) [GenBank: FJ619507 and FJ619508], a recently described begomovirus from Brazil, that is a distant relative of EuMV (Figure 1B).

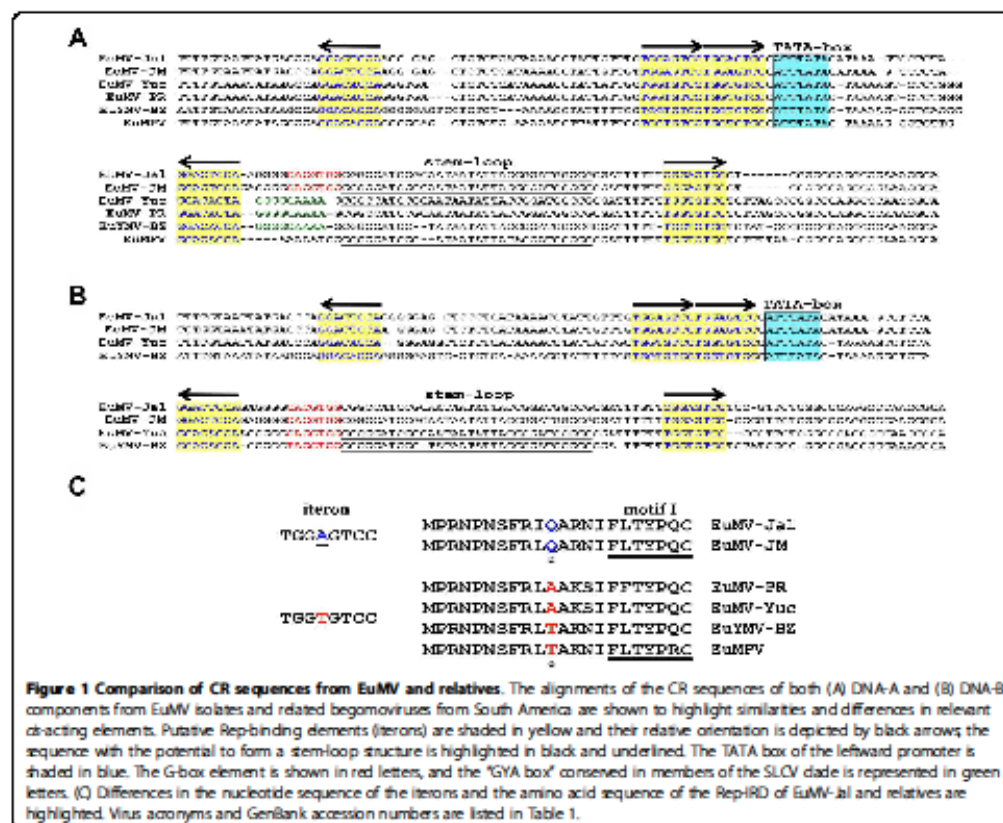
#### Phylogenetic relationships

A phylogenetic tree based on the full-length DNA-A of four EuMV isolates, 20 NW-Beg and several bipartite and monopartite OW-Beg (Table 1), was generated using the neighbor-joining method with 1,000 bootstraps replications (Figure 2). The analysis indicated a close relationship between the EuMV isolates from Mexico and the Caribbean basin with the following three begomoviruses from South America: *Tomato mild yellow leaf curl Aragua virus* (TMYLCAV) from Venezuela [GenBank: AY927277], *Euphorbia mosaic Peru virus* (EuMPV) [25], and *Euphorbia yellow mosaic virus* (EuYMV) from Brazil. This grouping was well-supported by both the phylogenetic analysis (bootstrap value 84) and the pairwise-identity analyses (Table 2), thus defining a sub-lineage within the SLCV clade that is broadly distributed in the American continent. A phylogenetic analysis based on the full-length DNA-B sequences produced similar results for the EuMV subclade and the group of cucurbit-infecting viruses (data not shown), but not for other members of the SLCV lineage that were placed into groups that are not congruent with the phylogeny derived from their DNA-A sequences. The incongruent phylogenies of DNA-A and DNA-B components of some begomoviruses is generally indicative of recombination and/or reassortment events [6,26].

#### Recombination analysis

The differences between the strains A and B of EuMV regarding nucleotide sequence and modular organization of the *Ori* region could be indicative of either divergent molecular evolution or intermolecular recombination between co-infecting begomoviruses [27,28]. To search for potential recombinant sequences in the genome of EuMV strains, we analyzed sequence alignments that included the DNA-A of the four EuMV isolates under exam, as well as diverse sets of begomoviruses of the SLCV clade, using the suite of programs for detection of recombinant breakpoints integrated within the RDP package [29]. The





analysis identified a ~210-nt long EuMV genomic region (recombinant breakpoints at positions 2432 and 33 of EuMV-Jal DNA-A) as a fragment of possible recombinant origin, which includes the entire common region (~170-nt) as well as the first 44 nucleotides of the *rep* gene, encompassing the IRD-coding sequence [17]. The plausible recombinant origin of this DNA fragment is underscored by direct comparisons of the DNA-A components from EuMV-JM and EuMV-PR, which are members from different strains exhibiting very high sequence identity (97.4%) along a segment encompassing ~2,400 out of the 2,609-nt of its DNA-A, a fact that is in clear contrast with the low sequence identity (77.5%) displayed in the 210-nt genomic region flanked by the recombinant breakpoints detected by our analysis.

The assembled data suggest that EuMV A-strain viruses are the product of an intermolecular recombination event involving an EuMV-JM-related virus (the major parent) and a virus closely related to *Calopogonium golden mosaic*

*virus* (CpGMV) [GenBank: AF439402] which might have donated the ~210-nt fragment with the viral replication module. This DNA segment, which is entirely identical in sequence between EuMV-PR and EuMV-YP, is shared with CpGMV at 90% of nucleotide identity. Two additional observations support the hypothesis of intermolecular recombination: (1) The absence of a G-box element within the CR of the DNA-A component of EuMV-YP, that is nevertheless present in their cognate DNA-B component (see Figure 1); and (2) The lower than expected sequence identity of the EuMV-YP common region (i.e., 86%) that is in contrast with the high identity of the CR of both EuMV-Jal and EuMV-JM (98% and 96%, respectively) [20,21].

#### Experimental infection of host plants

EuMV-Jal was identified in four field samples that contained an additional, distinct begomovirus, as mentioned above. In order to examine experimentally EuMV-Jal in

**Table 1 Names, acronyms, and GenBank accession numbers of the geminiviruses used in this study**

Name	Acronym	Accession number	
		DNA-A	DNA-B
Abutilon mosaic virus	AbMV	NC_001928	NC_001929
African cassava mosaic virus	ACMV	NC_001467	NC_001468
Ageratum yellow vein virus	AYV	NC_004626	
Bean calico mosaic virus	BGMV	NC_003504	NC_003505
Bean dwarf mosaic virus	BDMV	NC_001981	NC_001930
Bean golden yellow mosaic virus	BGYMV	NC_001489	NC_001438
Beet curly top virus	BCTV	NC_001412	
Beet mild curly top virus	BMCTV	NC_004753	
Cabbage leaf curl virus	GaLCV	NC_003866	NC_003887
Chino del tomate virus	GdTV	NC_003830	NC_003831
Corchorus golden mosaic virus	CoGMV	NC_009644	NC_009646
Corchorus yellow vein virus	CoYV	NC_006358	NC_006359
Cotton leaf crumple virus	CLGV	NC_004580	NC_00481
Cotton leaf curl multan virus	CLCuMV	NC_004607	
Cucurbit leaf crumple virus	CuLGV	NC_002984	NC_002985
Desmodium leaf distortion virus	DeLDV	NC_008494	NC_008495
Euphorbia leaf curl virus	EuLCV	NC_005319	
Euphorbia leaf curl India virus	EuLCM	EJ194914	
Euphorbia mosaic Peru virus	EuMPV	AM886131	
Euphorbia mosaic virus-Jalisco	EuMV-Jal	DQ520842	HQ185235
Euphorbia mosaic virus-Jamaica	EuMV-JM	FJ407052	EU740969
Euphorbia mosaic virus-Puerto Rico	EuMV-PR	AF086642	
Euphorbia mosaic virus- Yucatan	EuMV-YP	NC_008304	NC_008305
Euphorbia yellow mosaic virus	EuYMV	NC_012553	NC_012554
Papaya leaf curl virus	PaLCuV	AJ436992	
Pepper golden mosaic virus	PepGMV	NC_004101	NC_004096
Pepper huasteco yellow vein virus	PHYV	NC_001359	NC_001369
Rhynchosia golden mosaic Yucatan virus	RhGMVucV	NC_012481	NC_012482
Sida golden mosaic virus	SIGMV	NC_002046	NC_002047
Squash leaf curl virus	SLCV	NC_001986	NC_001937
Squash mild leaf curl virus	SMLCV	NC_004645	NC_004646
Squash yellow mild mottle virus	SYMMLV	NC_003865	NC_003860
Tomato common mosaic virus-Brazil	ToCoMV-BZ	NC_010835	NC_010836
Tomato golden mosaic virus	TGMV	NC_001507	NC_001508
Tomato mild yellow leaf curl Aragua virus	TMYLCAV	NC_009480	NC_009491
Tomato mottle virus	ToMoV	NC_001988	NC_001939
Tomato severe leaf curl virus	ToSLCV	DQ347947	
Tomato yellow leaf curl Thailand virus	TYLCTHV	X63015	X63016
Tomato yellow leaf curl virus	TYLCV	X15656	
Watermelon chlorotic stunt virus	WmCSV	NC_003708	NC_003709

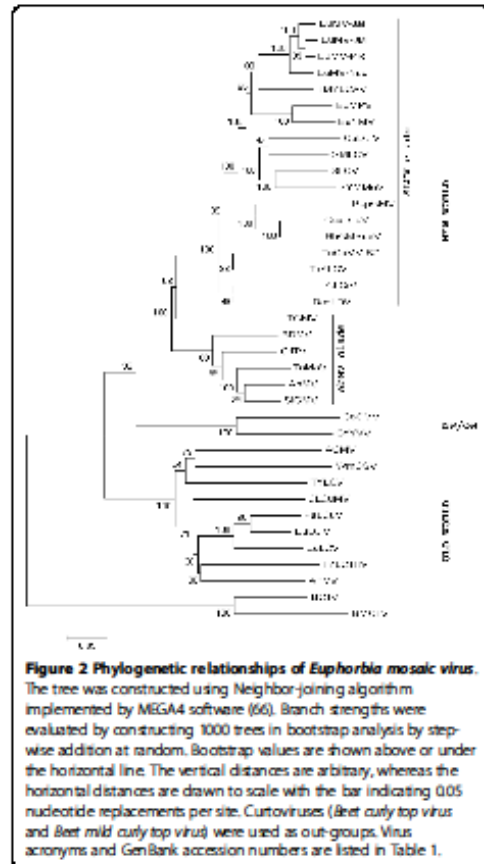
single plant infections, we generated infectious clones of both DNA-A and DNA-B components (see Methods), and carried out biolistic inoculation of these clones into

four plant species: *Datura stramonium*, *Nicotiana benthamiana*, pepper (*Capsicum annum*), and zucchini (*Cucurbita pepo*). All solanaceous species were susceptible and developed systemic symptoms at 10-12 dpi, while the zucchini plants did not show symptoms and no viral DNA was detected by PCR in their tissues at 14 dpi. Symptoms of EuMV-Jal infection varied between plant species. In *N. benthamiana* the symptoms included leaf crumpling, greenish mosaics and shortened internodes (Figure 3A). In pepper plants the first symptom was the appearance of small green spots that progressed into a pale green mosaic and moderate downward leaf curling; a few small necrotic spots were also observed in several plants (Figure 3B). The most severe symptoms were observed in *D. stramonium* plants, whose leaves showed deformation and extensive green and yellow mottle covering most of the foliar surface, progressing in time to necrotic lesions leading to the destruction of significant parts of the foliar lamina (Figure 3C). In all, the symptoms induced by EuMV-Jal in the examined three plant species were very similar to those generated by infection of EuMV-YP [20], hence suggesting that these viruses express equivalent pathogenesis factors, as expected from the high amino acid sequence identity of their predicted proteins (Table 2).

#### EuMV-Jal and EuMV-YP are incompatible in replication

The replication modules of the EuMV strains A and B exhibit two main differences: 1) their iterons display a different nucleotide N within the GGNGTCC core, and 2) the iteron-related domain of their Rep proteins have a different amino acid residue at position X<sub>3</sub> of the IRD core FX<sub>1</sub>LX<sub>3</sub> [17], that is either FRLA or FRLT in A-strain viruses, and FRLQ in B-strain members (Figure 1C). These observations suggest the intriguing possibility that EuMV strains A and B could be incompatible in replication. To answer this question we carried out re-assembly experiments with the EuMV-Jal and EuMV-YP genomic components. The four possible combinations A+B of the cloned viral DNAs were biolistically inoculated into *N. benthamiana* plants, that were subsequently scored for the appearance of disease signs.

Systemic symptoms developed at 10-12 dpi in most plants inoculated with the homologous mixtures (i.e., EuMV-Jal [A+B], and EuMV-YP [A+B]); in contrast, the plants bombarded with the heterologous combinations (i.e., EuMV-Jal [A]/-YP [B] and its reciprocal, EuMV-YP [A]/-Jal [B]) displayed no symptoms at 12 dpi, and remained symptomless until the end of the experiment, at 30 dpi (Figure 4A). These experiments were repeated three times, six plants for each combination, with similar results obtained (data in Figure 4B). All plants inoculated with cognate viral components scored positive for presence of both EuMV DNA-A and DNA-B, based on



PCR detection of a ~1300-bp fragment encompassing a part of the *rep* and *cp* genes and the entire DNA-A intergenic region, and a ~1400-bp segment comprising the DNA-B IR and a part of both *BV1* and *BC1* genes. In contrast, none of the newly emerged leaves of plants bombarded with the heterologous combinations of EuMV genomic components tested positive for presence of EuMV DNA-B, although a few plants (5 out 36) were PCR-positive for DNA-A at 14 dpi, but not at 28 dpi (data not shown). These results indicate that viral factors required for replication are not exchangeable between EuMV-Jal and EuMV-YP.

#### EuMV *BV1* promoter contains a short sequence homologous to *Rep* gene

In the course of a meticulous scrutiny of the DNA-B intergenic region of EuMV-Jal to identify potential *cis*-

regulatory elements involved in the transcriptional control of the *BC1* and *BV1* genes, we unexpectedly discovered a 35-bp DNA stretch displaying 100% sequence identity with a segment of the homologous *rep* gene. This sequence is located ~150-nt upstream to the *BV1* gene (nucleotides 337-372) and contains the coding information for aa residues 15 to 25 of EuMV-Jal Rep (i.e., FLTYPQCDVPK) that includes the conserved Motif I of the RCR initiators [30]. No additional sequences homologous to the *rep* gene were found in the *BV1* promoter region. The finding of a short sequence apparently derived from the cognate DNA-A within the noncoding region of EuMV-Jal DNA-B was intriguing and prompted further scrutiny of other EuMV DNA-B components. In all the examined cases a short *Rep* homologous sequence (*sRepHS*) was found within the *BV1* promoter region, which in EuMV-JM is similar to the EuMV-Jal element in both sequence and length (35-nt), but that is longer in EuMV-YP that displays a DNA stretch 51-nt in length identical to a segment of its cognate *rep* gene (Figure 5). A search for analogous elements in the DNA-B IR from all members of the SLCV clade revealed that *sRepHS* elements are not common, being identified only in two close relatives of EuMV, namely, TMYLCAV from Venezuela and EuYMV from Brazil. The TMYLCAV *sRepHS* element is similar but not identical in both length (36-nt) and nucleotide sequence (88% identity) to the equivalent sequence of EuMV-Jal (Figure 5). In contrast, the *sRepHS* identified in EuYMV DNA-B is different in both length (45-nt) and nucleotide sequence (< 30% identity) to the analogous elements of EuMV strains. Indeed, the EuYMV *sRepHS* element corresponds to a distinct segment of the cognate *rep* gene, encoding the Rep aa residues 40-53 (i.e., VVKPTYIRVARELH) instead of Rep residues 15-25 encoded by the *sRepHS* elements of TMYLCAV and EuMV. Notwithstanding its divergent nucleotide sequence, the EuYMV *sRepHS* element is 100% identical in nucleotide sequence to a segment of its cognate *rep* gene, like in EuMV and TMYLCAV (Figure 5) and is located at a position equivalent to the *sRepHS* in the latter viruses.

#### *sRepHS* upstream sequences are similar to CP promoters

The conservation of *sRepHS* elements in the DNA-B intergenic region of EuMV and their relatives suggests that those atypical sequences might play a defined role in the infective cycle of these viruses. Since the *sRepHS* elements do not contain a start codon and are not a part of a distinctive ORF, it seems plausible that its function, if any, involves an intermediary RNA molecule. This notion naturally led us to suggest the existence of a functional promoter next to the *sRepHS* element.



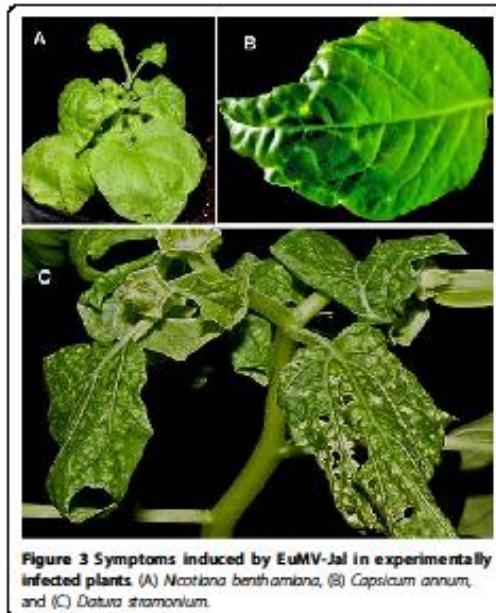
**Table 2 Percentages of sequence identities between EuMV-Jal and selected begomoviruses (DNA and predicted proteins\*)**

	DNA-A	IR-A	CP*	AC1*	AC2*	AC3*	AC4*	DNA-B	IR-B	BV1*	BC1*
Virus											
ACMV	45	25	66	49	43	42	19	27	22	24	41
BCaMV	76	50	92	86	78	77	64	55	28	73	83
BGYMV	64	37	91	63	70	78	11	48	22	67	80
CdTV	67	43	92	63	67	78	30	51	27	71	78
CoYVV	51	24	87	43	51	43	19	41	22	52	71
CuLCrV	77	46	91	83	71	71	72	51	27	66	76
DesLDV	72	44	91	80	66	73	58	50	23	64	77
EuMPV	77	52	93	86	81	76	58	-	-	-	-
EuYMV	77	51	90	85	80	76	62	52	35	73	82
EuMV-JM	<b>95</b>	<b>91</b>	<b>98</b>	<b>97</b>	<b>97</b>	<b>95</b>	<b>88</b>	<b>86</b>	<b>73</b>	<b>96</b>	<b>98</b>
EuMV-PR	92	82	<b>99</b>	96	93	91	<b>91</b>	-	-	-	-
EuMV-YP	92	80	99	93	93	91	87	85	63	94	<b>98</b>
PepGMV	72	50	90	80	71	75	14	48	25	64	74
PHYW	59	33	89	49	50	63	12	47	25	66	74
RhGMVY	76	54	94	86	70	70	66	51	31	69	78
SLCV	78	57	94	82	72	80	77	50	30	63	80
ToCoMV-BZ	73	43	90	85	64	72	57	52	31	63	77
TMYLCAV	84	66	95	88	87	80	82	56	43	75	83
TYLCTHV	48	28	68	48	43	39	22	25	19	21	39

In order to identify potential IR internal promoters, we analyzed the sequences upstream to *sRepHS* in all members of the EuMV lineage using a phylogenetic-structural approach. This methodology entails the identification of "phylogenetic footprints" (i.e., putative binding sites for transcription factors) and conserved arrays of them, named "Conserved Modular Arrangements" (CMAs), in non-coding regions of evolutionarily-related DNA sequences [31,32]. The new analysis exposed a DNA-B IR domain ~160-bp-long exhibiting a remarkable similarity both in overall nucleotide sequence and modular organization, to CP promoters of viruses that belong to the SLCV clade. The example showed in Figure 6 illustrates the remarkable similarity between the CP promoter-like (CPprom-L) domain of EuMV-Jal IR and a 156-nt segment of the CP promoter of *Rhynchosia golden mosaic Yucatan virus* (RhGMVYU), a recently described virus of the SLCV lineage [33]. The similarity between these DNA-B and DNA-A sequences, respectively, includes nine phylogenetic footprints in a definite order, and it is extended beyond the start codon of RhGMVYU *cp* gene including a block of 8-nt of coding sequence that is conserved in the non-coding sequence of EuMV-Jal DNA-B.

The demarcated CPprom-L domain of the DNA-B IR includes several putative *cis*-regulatory elements that were identified by consulting plant transcription factors databases like PlantCare [34] and PLACE [35].

Among the identified potential *cis*-acting motifs there were well-characterized regulatory elements such as the "Conserved Late Element" (CLE) [36], the CCAAT box, and several elements that confer responsiveness to a variety of plant hormones (see Figure 6 legend). Among these sequences there is a 12-bp long element (consensus: CTTTAATTCAAA) which is identical to a conserved sequence immediately adjacent to the *cp* gene in more than 75% of the known begomoviruses from America (Cardenas-Conejo et al., unpublished data). The AATTCAAA motif of the former element is both a putative ethylene-responsive element (ERE) and a binding-site for nuclear factors of carnation, tomato and *Solanum melongena* [37-39]. In addition, this motif constitutes the 8-nt long leader sequence of the CPmRNA of *Tomato golden mosaic virus* (TGMV) [40]. The ERE-like motif is located downstream to the actual TATA-box of NW-Beg CP promoters, at a similar distance (21-29 bp) to that observed between the ERE and a putative TATA box in the CPprom-L domain [Additional file 1: Supplemental Figure S1a]. Taken as a whole, these remarkable similarities between noncoding DNA regions from two different genome components of separate begomovirus species, can hardly be explained by random sequence convergence; rather, they strongly suggest that the DNA-B CPprom-L domain of EuMV and relatives is evolutionarily derived from a begomovirus CP promoter.

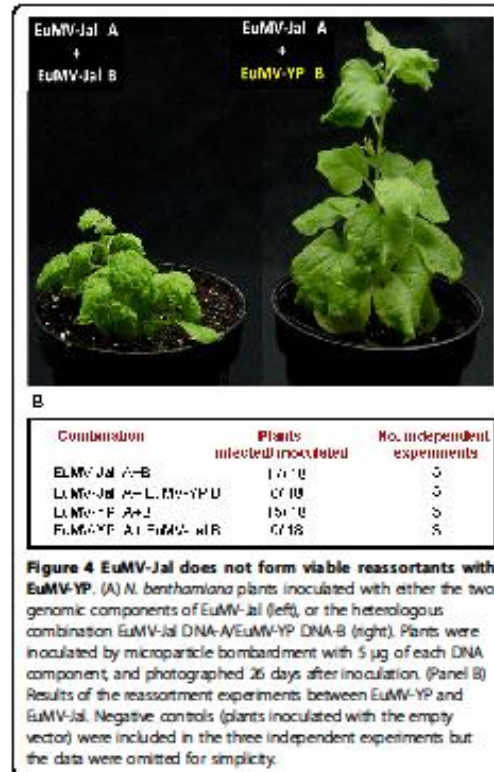


**Figure 3** Symptoms induced by EuMV-Jal in experimentally infected plants (A) *Nicotiana benthamiana*, (B) *Capsicum annuum*, and (C) *Datura stramonium*.

#### Distantly related begomoviruses contain sRepHS elements

The existence of *sRepHS* elements in the DNA-B IR of viruses belonging to a minor lineage of the SLCV clade is an interesting evolutionary enigma. To determine whether analogous elements actually exist in other viral lineages, we searched for *rep* homologous sequences in the DNA-B IR of begomoviruses belonging to 12 major and minor clades, distributed in several continents. The analysis of ~60 members of those lineages led us to the identification of only two additional begomoviruses displaying *sRepHS* in the *BV1* upstream region: TGMV and the recently described *Cleome leaf crumple virus* (CLCrV) [41]. These viruses are native from Brazil, like EuYMV, but do not belong to the SLCV clade. The *sRepHS* element of CLCrV is 100% identical to a 46-nt-long segment of its cognate *rep* gene, encoding the aa residues 97 to 110 (SSSDVKSYVDKDGD), that comprise the conserved RCR Motif 3 (underlined) [30]. On the other hand, the TGMV *sRepHS* element is only 88% identical to a 52-nt-long segment of its cognate *rep* gene, encoding the aa residues 255 -271 (NKVEYNVIDDVTPQYLK) of this replication initiator, that include the Walker B-motif (underlined), a critical aa sequence of the protein ATPase/helicase domain [42,43].

The upstream sequences of TGMV and CLCrV *sRepHS* elements were examined, but no significant similarity between them nor with the *BV1* promoter region of EuMV lineage viruses was found. However, a



**Figure 4** EuMV-Jal does not form viable reassortants with EuMV-YP. (A) *N. benthamiana* plants inoculated with either the two genomic components of EuMV-Jal (left), or the heterologous combination EuMV-Jal DNA-A/EuMV-YP DNA-B (right). Plants were inoculated by microparticle bombardment with 5 µg of each DNA component, and photographed 26 days after inoculation. (Panel B) Results of the reassortment experiments between EuMV-YP and EuMV-Jal. Negative controls (plants inoculated with the empty vector) were included in the three independent experiments but the data were omitted for simplicity.

careful re-examination of sequences nearby to the 5' end of CLCrV *sRepHS* revealed a 23-bp sequence with partial dyad symmetry that is well-conserved both in sequence and in position relative to the *sRepHS* element in all viruses of the EuMV cluster [Additional file 1: Suppl. Figure S1b]. The consensus of this conserved sequence includes a palindromic core with the repeated motif TTGTGGTCC, similar to the CLE, a functional target of plant transcriptional activators [44,45] that has been involved in TrAP-mediated activation of the CP promoter in some begomoviruses [36]. No sequence similar to the latter symmetric element was found in the *BV1* promoter region of TGMV. In fact, the *sRepHS* of the latter virus differs from the analogous elements in CLCrV and the EuMV subclade viruses in several other important features: (1) It is not 100% identical to the corresponding segment of its cognate *rep* gene; (2) It has opposite polarity compared to all other known *sRepHS* elements; (3) It is closely located downstream to a putative internal promoter that does not exhibit significant similitude with CP promoters of SLCV clade viruses (data not shown). It is relevant to point out here



that TGMV and CLCrV are grouped, on the basis of their full-length DNA-A sequences, within the Brazilian cluster of NW-Beg [41], but they have very divergent DNA-B components. Thus, our finding of the *sRepHS*-associated semi-palindromic sequence in CLCrV DNA-B suggests an actual relationship of the latter with the homologous genomic components of EuMV and relatives, a notion that is supported by a recent study that groups the CLCrV DNA-B with viruses of the EuMV lineage [41].

### Discussion

In this study, we described the molecular and biological characterization of a novel strain of *Euphorbia mosaic virus* that was isolated from pepper plants in the state of Jalisco, Mexico, near to the Pacific shoreline. This virus displays 92% sequence identity with EuMV-YP, that was isolated in the same country but in a distant region, close to the Atlantic coastline [20]. These viruses differ in two important features of their DNA-A replication origin region: the nucleotide sequence of their iterons, and the presence or absence of a G-box element, a *cis*-acting sequence which is critical for *Rep* promoter activity in some NW-Beg [46]. The differences observed in the predicted *Rep*-binding sites of EuMV-Jal and EuMV-YP prompted us to explore experimentally their ability to form viable reassortants in pseudorecombination tests. The results of these experiments confirmed the presumption of replication incompatibility between EuMV-YP and EuMV-Jal, thus demonstrating that the latter is a new, biologically-defined strain exhibiting different replication specificity.

The finding of begomovirus strains that are not able to form viable reassortants is somehow bewildering because the common definition of a virus species is "A... class of viruses that constitutes a replicating lineage and occupies a particular ecological niche." [47,48]. Accordingly, it is not expected that strains of a virus species would be incompatible in replication because that implies that they do not constitute an actual replicating lineage. Nonetheless, it is generally recognized that several strains of begomoviruses probably are not complementary in replication because they display different putative *cis*- and *trans*-acting replication specificity determinants [7,17]. There is at least one report of strains belonging to a bipartite begomovirus that are not equivalent in replication functions (the "severe" and "mild" strains of *Tomato leaf curl New Delhi virus*, ToLCNDV) [49]. However, that case is different from the one examined here because the "mild" phenotype of one ToLCNDV strain seems to be related to an inefficient *trans*-replication of the "cognate" DNA-B, which displays *Rep* binding-sites different to those of the associated DNA-A [49,50].

The case of the EuMV strains is significant because it is paradigmatic of an apparently common theme in begomovirus evolution, i.e., the sudden change of virus replication specificity determinants by intermolecular recombination between co-infecting viruses [27,51]. Indeed, the recombination analysis of EuMV isolates indicates that viruses of the EuMV A-strain probably evolved by an event of DNA intermolecular exchange involving a member of the EuMV B-strain and a virus related to CpGMV, which had donated a ~210-bp DNA segment encompassing the region of the virus replication origin and the first 44 nucleotides of the *rep* gene. If this hypothetical scenario is accurate, then the recombination event should have changed simultaneously both the iterons and the *Rep* aa residues interacting with them, thus maintaining the proper matching of *cis*- and *trans*-acting replication determinants in the recombinant DNA-A component.

Diverse studies have identified the sequences encompassing the viral strand replication origin and the *rep* gene segment encoding the *Rep* N-terminal domain, as the regions of geminivirus genomes most frequently exchanged during recombination [28,51-53]. This is consistent with the known genome localization of the *Rep*-binding sites and the coding sequence of the *Rep* domain that contains the putative DNA-binding specificity determinants of this protein, which have been theoretically mapped into the first 75 aa residues [17,54]. Consequently, a recombination event involving a genome portion as small as 200 to 360-bp might confer a completely different replication phenotype to begomoviruses involved in mixed infections, as presumably is the case for the EuMV strains.

Since that intermolecular recombination is/has been a major force in the evolution of geminiviruses, the concepts of both "species" and "strains" should be adapted to the peculiar nature of these entities, that are genetic mosaics in continual change, different in quality to cellular organisms. In fact, it is altogether possible that a significant part of the currently recognized begomovirus species would not constitute "replicating lineages" in a strict sense, as would be the case of EuMV, according to our experimental data. For instance, a thorough sequence analysis entailing the identification of the putative *cis*- and *trans*-acting Replication Specificity Determinants (RSDs) of the 182 recognized begomovirus species summarized by Fauquet et al. in 2008 [7] revealed the existence of 34 species that include at least two groups of viruses exhibiting distinct putative RSDs, analogous to the strains A and B of EuMV. Furthermore, some ICTV-accepted species as *Ageratum yellow vein Huallan virus*, *Honeysuckle yellow vein virus*, *Tomato leaf curl Bangalore virus*, *Tomato leaf curl Philippines virus*, *Tomato leaf curl Taiwan virus*, and





An intriguing observation is that the identified *sRepHS* elements reproduce sequences encoding conserved motifs which are critical for Rep functions. For example, the *sRepHS* of EuMV strains and TMYLCAV correspond to the coding sequence of RCR Motif 1; the equivalent element of CILCrV encodes the RCR Motif 3, and the analogous *sRepHS* of TGMV duplicate the *rep* sequence encoding the Walker B motif of ATPases/helicases. An apparent exception is the *sRepHS* of EuYMV, which displays the coding sequence of a conserved Rep motif of unknown function. The evolutionary conservation of *sRepHS* elements and the associated sequence motifs, suggests that those atypical elements play a definite but hitherto unknown function in the viral infective cycle. In absence of any factual data it is only feasible to speculate about the possible function(s) of the *sRepHS* on the basis of their common characteristics.

Certainly, the most remarkable feature of the *sRepHS* elements is its complete identity in nucleotide sequence with a specific segment of the *rep* gene in the cognate DNA-A component, because the evolutionary preservation of such an absolute matching between specific segments of distinct, physically separated DNA molecules, should involve very strong selective pressures against mutations diminishing the identity between the former DNA sequences. Therefore, the function of the *sRepHS* elements is most likely related to a process that requires a perfect or very high complementarity between DNA and/or RNA molecules, such as the gene regulation by microRNAs (miRNAs).

The miRNAs are ~22-nt-long noncoding RNAs that posttranscriptionally regulate gene expression by binding to specific mRNAs, thus repressing its translation and/or inducing its degradation [55]. Several DNA viruses (i.e., herpesviruses, adenoviruses, ascoviruses and polyomaviruses) encode miRNAs which participate in the regulation of some processes of the viral infection cycle [56,57]. For example, the simian virus 40 (SV40) encodes a single miRNA which lie antisense to the viral mRNA encoding the T-antigen, a multifunctional protein essential for virus replication. This miRNA is expressed late in infection, hence promoting the T-antigen mRNA degradation and downregulating the synthesis of this protein at late stages of the SV40 replication cycle [58]. In close analogy with SV40 miRNA, the *sRepHS* elements of begomoviruses are single, discrete noncoding DNA sequences highly similar to a specific segment of the gene encoding the viral replication protein. Further analogies between those heterologous viral sequences are the following: (1) The genomic location of the miRNA, but not its nucleotide sequence, is conserved among polyomaviruses (i.e., SV40, Merkel cell virus, human BK virus, JC virus, and mouse polyomavirus) [59-61]; similarly, the location of *sRepHS* elements

within the DNA-B intergenic region, but not its specific sequence, is conserved among begomoviruses (data from this study); (2) The temporal expression of the SV40 miRNA, that is restricted to the late stage of infection, is similar among all the examined polyomaviruses [57,59]; likewise, although the temporal expression of begomovirus transcripts including the *sRepHS* region (if any) is unknown, it is plausible than them would be late expressed, because the hypothetical promoter that lead its transcription is similar to begomovirus *CP* promoters, which are typically active at the late phase of the viral infection cycle [1,36]; (3) Like the polyomavirus pre-miRNAs, the DNA-B sequences encompassing *sRepHS* and the neighboring sequences, have the potential to form extensive hairpin structures susceptible to cleavage by RNase III enzymes (i.e., Droscha and Dicer) involved in the processing of pre-miRNAs (data not shown). Taken together, these lines of indirect evidence suggest a potential function of the *sRepHS* elements in the posttranscriptional regulation of Rep expression, a hypothesis that must be experimentally examined.

## Conclusions

The evidence gathered in this study indicates that EuMV-YP and EuMV-Jal, which are members from the strains A and B of *Euphorbia mosaic virus* respectively, are actually incompatible in replication, hence implying that these viruses probably represent distinct replicating lineages in natural ecosystems. The scenario we propose for the origin of the EuMV A-strain viruses involves a recombination event that substituted the DNA-A core replication module of an EuMV B-strain virus, with the analogous genomic region of a virus related to CpGMV. This intermolecular exchange suddenly changed the replication specificity of the recombinant DNA-A, thus triggering the process that led to the evolutionary differentiation of EuMV into two distinct strains. The fact that more than 30 recognized begomovirus species include two or more classes of viruses with distinct putative RSDs (i.e., analogous to the EuMV strains) suggests that intermolecular recombination events that involve the virion-strand origin of replication and the first part of the *rep* gene, are quite common in this group of ssDNA viruses, as has been previously pointed out [51, 52, 53]. Another relevant result from this study is the discovery of atypical sequences within the intergenic region of the DNA-B component from some NW-begomoviruses, mostly related to EuMV. These sequences include short fragments of the cognate *Rep* gene located downstream from a potential internal promoter very similar in modular organization to *CP* promoters of viruses of the SLCV clade. Even though we do not know the actual function of these *sRepHS* elements, several lines of indirect evidence suggest their





B the degenerate primers BC1- 290-rev (CCSAT-MAGRTGYAACATAGATCTCC) and BV1-310-for (AGGWACRGTNAARATYGCATGCTGT) were used. The PCR-products were cloned into the pGEM-T easy vector (Promega) and subjected to Restriction Fragment Length Polymorphism (RFLP)-analysis by using *EcoR* I in combination with either *Hinf*I or *Msp* I. The produced DNA molecules were fractionated in 2.5% agarose gels, and PCR clones with different restriction patterns were sequenced.

#### Generation of infectious EuMV clones

To clone the full-length genomic components A and B of EuMV-Jal, the DNA extract of one pepper plant infected with both EuMV-Jal and PHYVV collected at the Castillo locality (see Results) was subjected to rolling circle amplification (RCA) using the TempliPhi kit (GE Healthcare, USA) according to the manufacturer's instructions. The full-length EuMV DNA-A was obtained by cutting the RCA-amplified DNA with *Xho* I and subsequent cloning of the 2.6 Kb DNA molecule into a plasmid vector. The full-length EuMV DNA-B could not be obtained by a similar procedure after several attempts. Consequently, abutted divergent primers designed over the unique *Bam*HI site in the DNA-B of EuMV-Jal were used in a standard PCR procedure, and the generated 2.6 Kb amplicon was cloned into a pGEM-Teasy vector (Promega). The infectious clones of EuMV-Jal A and EuMV-Jal B were generated as follows: the 0.8-Kb *Bam*HI- *Xba*I fragment of EuMV-A containing the origin of replication was cloned into the *Bam*HI-*Xba*I sites of a modified pBlueScript plasmid to create pEu-oriA. Subsequently, a full-length DNA-A of EuMV digested with *Xba*I was inserted into the *Xba*I site of pEu-oriA to generate the viral replicon pEuMV1.33A. The infectious clone of EuMV-B was generated by an analogous procedure: the full-length DNA-B cloned into pGEM-T easy was digested with *Nco*I and re-ligated. This procedure deleted a portion of the viral genome, leaving intact all elements important for replication (1.3-kb), thus generating the pEu-oriB plasmid. Finally, a full-length DNA-B digested with *Bam*HI was cloned into the *Bam*HI site of pEu-oriB, yielding the infectious clone pEuMV1.5B.

#### Plant infection assays

*Nicotiana benthamiana*, *Capsicum annuum* and *Datura stramonium* plants were inoculated using a low-pressure biolistic method [64]. The target leaves (third- to four-leaf stage) were either directly shot at 100 to 120 psi helium pressure with tungsten particles (0.7 mm, BioRad, Hercules, CA) covered with the EuMV-A and EuMV-B viral DNAs (5 µg), or inoculated mechanically by using carborundum according to a procedure

recently described [65]. The inoculated plants were maintained in an insect-free growth chamber (27°C, daily cycle of 16 h light -8 h dark), and subsequently scored for the appearance of disease symptoms. The infection status of the inoculated plants was assessed by visual inspection of symptoms and by PCR analysis of all plants at the end of the experiment.

#### Reassortment experiments

Pseudorecombination experiments were carried out by biolistically inoculating seedlings of *N. benthamiana* plants with all possible pair combinations of A and B component clones of both EuMV-Jal and EuMV-YP. Infectious clones of EuMV-Jal were partial tandem repeats of either DNA-A or DNA-B components, as mentioned above, whereas in the case of EuMV-YP, cloned monomeric components were used as previously described [20]. A total of 18 seedlings (three independent experiments, six plants each replicate) were inoculated with each one of the four possible combinations of EuMV-Jal and EuMV-YP genomic components. Mock-inoculated negative controls were included for each replicate. The inoculated plants were maintained in a growth chamber (27°C, daily cycle of 16 h light -8 h dark) and scored for the appearance and development of disease symptoms during 4-5 weeks. All plants, both symptomless and symptomatic, were tested for the presence of viral DNA in newly emerged leaves at 14 dpi by PCR-based detection, using both DNA-A and DNA-B specific primers. Asymptomatic plants were re-examined by PCR at 28 dpi, to detect cases of delayed infection.

#### Phylogenetic analysis

Full DNA-A and DNA-B sequences from EuMV-Jal were compared with other New World and Old World begomoviruses available at the GenBank-NCBI database, using BLAST-N. The positions and sizes of EuMV-Jal open reading frames were predicted using EditSeq (DNASTAR Inc., Madison, WI). Paired alignments were obtained by the ClustalV and ClustalW methods in the MegAlign application of the Lasergene package (DNASTAR), using the default parameters. Neighbour-joining phylogenetic trees for EuMV DNA-A and DNA-B components were constructed using Mega 4.0 [66] with 1,000 bootstrap replicates and pairwise evolutionary distances calculated with a maximum likelihood nucleotide substitution model. Trees were drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

#### Recombination analysis

Detection of potential recombination breakpoints and recombinant sequences was carried out using the suite

of recombination detection methods implemented in RDP software (29). A sequence alignment containing four EuMV isolates and closely related species (sharing >76% of nucleotide identity with EuMV-Jal) was used as input data for RDP. The analysis was performed using the default settings in all detection methods, with a Bonferroni corrected P-value cut-off of 0.05.

### Additional material

**Additional file 1: Supplemental Figure S1: Conserved elements upstream to sRepHS elements.** (A) Comparisons of conserved modular arrangements (CMA4) composed by two cis-acting elements present in DNA-A of NW begomoviruses, and DNA-B of EuMV and relatives, respectively. The CPMRNA transcription start site of TGMV is indicated above the BRE-like motif. (B) Alignment of partially palindromic elements which are conserved in position relative to the sRepHS element of CLCrV and EuMV subclade members. The consensus of this symmetric element is indicated. Colors in boxes identify the distinct classes of sRepHS according to their nucleotide sequence, and numbers indicate the length (in base pairs) of those elements.

### Acknowledgements

We thank Salvador Ambriz-Granados for his technical assistance, and Dr. Braulio Gutierrez-Medina for critical reading of the manuscript. JGJ, ABA and BBH were supported by a PhD fellowship from the Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico). This research was partially supported by a CONACYT grant (SEP84004) to GAA.

### Author details

<sup>1</sup>Instituto Potosino de Investigación Científica y Tecnológica, AC., Camino a la Presa San José, 78216 San Luis Potosí, SLP, México. <sup>2</sup>Universidad Autónoma Agraria Antonio Naro, Departamento de Parasitología Agrícola, Bellavista, C.P. 25315, Saltillo, Coahuila, Mexico. <sup>3</sup>Centro de Investigación Científica de Yucatán, AC., Mérida, Yucatán, México.

### Authors' contributions

JGJ generated the infectious clones of EuMV-Jal, performed plant infection tests, carried out the phylogenetic analysis and helped to prepare the manuscript. ABA collected isolates, cloned and sequenced the viruses, analyzed the field data, and perform plant infection tests. BBH carried out the pseudorecombination experiments, and analyzed the experimentally infected plants. AAS helped in comparative sequence analyses, provided partial funding for the project's execution, and offered ideas and comments during manuscript preparation. CHZ carried out the recombination analysis, helped in plant infection tests. OMV provided the BuMV-YP clones and helped in plant infection tests with this virus. GFT collected isolates and helped to analyze the field data. GAA coordinated the project, carried out the comparative sequence analyses, secured funding for the project's execution, and prepared the manuscript. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 18 August 2010 Accepted: 19 October 2010

Published: 19 October 2010

### References

1. Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S, Robertson D: Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Crit Rev Plant Sci* 1999, 18:71-106.
2. Zhang W, Olson NH, Baker TS, Faulkner L, Agbandje-McKenna M, Boulton M, Davies JW, McKenna R: Structure of the Maize streak virus geminate particle. *Virology* 2001, 279:471-477.

3. Stanley J, Bizaro DM, Briddon RW, Brown JK, Fauquet CM, Harrison BD, Rybicki EP, Stenger DC: Geminiviridae. In *Virus Taxonomy With Report of the ICTV*. Edited by Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Bal LA. Elsevier/Academic Press, London; 2005:301-306.
4. Rojas MR, Hagen C, Lucas WJ, Gilbertson RL: Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. *Annu Rev Phytopathol* 2005, 43:361-394.
5. Seal SE, vanderBosch F, Jeger MJ: Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. *Crit Rev Plant Sci* 2006, 25:23-46.
6. Briddon RW, Patel BL, Bagewad B, Nawaz-ul-Rehman MS, Fauquet CM: Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evol Biol* 2010, 10:97.
7. Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X: Geminivirus strain demarcation and nomenclature. *Arch Virol* 2008, 153:783-821.
8. Bizaro DM: Silencing suppression by geminivirus proteins. *Virology* 2006, 344:158-68.
9. Hehnl S, Wege C, Jeske H: Interaction of DNA with the movement proteins of geminiviruses revisited. *J Virol* 2004, 78:698-706.
10. Sandriotto AA, Lazarowitz SG: Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. *Trends Cell Biol* 1996, 9:353-358.
11. Fontes BP, Gladfelter HJ, Schaffer RL, Petty IT, Hanley-Bowdoin L: Geminivirus replication origins have a modular organization. *Plant Cell* 1994, 6:405-416.
12. Singh DK, Malik PS, Choudhury NR, Mukherjee SK: MYMM replication initiator protein (Rep): roles at the initiation and elongation steps of MYMM DNA replication. *Virology* 2008, 380:75-83.
13. Rojas A, Nvamenhen A, Marcano D, Vallonen JP: Sequence characterization of tomato leaf curl Simola virus and tomato severe leaf curl virus: phylogeny of New World begomoviruses and detection of recombination. *Arch Virol* 2005, 150:1281-1299.
14. Rybicki EP: A phylogenetic and evolutionary justification for 3 genera of Geminiviridae. *Arch Virol* 1994, 139:49-77.
15. Argüello-Astorga G, Herrera-Estrella L, Rivera-Bustamante R: Experimental and theoretical definition of geminivirus origin of replication. *Plant Mol Biol* 1994, 26:553-556.
16. Argüello-Astorga GR, Guevara-González RG, Herrera-Estrella LR, Rivera-Bustamante RF: Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology* 1994, 203:90-100.
17. Argüello-Astorga GR, Ruiz-Medrano R: 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* 1994, 146:1465-1485.
18. Bird J, Sánchez J, Rodríguez RL, Juliá FJ: Rugaceous (whitefly-transmitted) viruses in Puerto Rico. In *Tropical Diseases of Legumes*. Edited by Bird J, Maramba K. New York, USA: Academic Press; 1975:3-25.
19. Costa AS: Whitefly-transmitted plant diseases. *Annu Rev Phytopathol* 1976, 14:429-49.
20. Hernández-Zepeda C, Idris AM, Carnevali G, Brown JK, Moreno-Valenzuela OR: Molecular characterization and experimental host range of Euphorbia mosaic virus Yucatan Peninsula, a begomovirus species in the Squash leaf curl virus clade. *Plant Pathol* 2007, 56:763-770.
21. Collins AM, Brown JK, Rehman MM, Roye ME: Complete nucleotide sequence of an isolate of Euphorbia mosaic virus that infects Euphorbia heterophylla and Wissadula amplissima in Jamaica. *Arch Virol* 2009, 154:1859-1860.
22. Méndez-Lozano J, Torres-Pacheco I, Fauquet CM, Rivera-Bustamante RF: Interactions between geminiviruses in a naturally occurring mixture: Pepper huasteco virus and Pepper golden mosaic virus. *Phytopathology* 2003, 93:270-277.
23. Torres-Pacheco I, Garzón-Tiznado JA, Brown JK, Becerra-Flora A, Rivera-Bustamante RF: Detection and distribution of Geminiviruses in Mexico and the Southern United States. *Phytopathology* 1996, 86:1186-1192.
24. Torres-Pacheco I, Garzón-Tiznado JA, Herrera-Estrella L, Rivera-Bustamante RF: Complete nucleotide sequence of pepper huasteco virus analysis and comparison with bipartite geminiviruses. *J Gen Virol* 1993, 74:2225-2231.



25. Shepherd DN, Martin DP, Lefeuve P, Morjane AL, Owor BE, Rybicki EP, Varsani A: A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. *J Virol Methods* 2008, 149:97-102.
26. Hernández-Zepeda C, Argüello-Astorga G, Idris AM, Carnevali G, Brown JK, Moreno-Villanueva OA: Molecular characterization and phylogenetic relationships of Desmodium leaf distortion virus (DeLDV): a new begomovirus infecting *Desmodium glabrum* in Yucatan, Mexico. *Virus Genes* 2009, 39:371-374.
27. García-Andrés S, Tomás DM, Sánchez Campos S, Navas-Castillo J, Moriones E: Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. *Virology* 2007, 365:210-219.
28. van der Walt E, Rybicki EP, Varsani A, Polston JE, Bilharz R, Donaldson L, Morjane AL, Martin DP: Rapid host adaptation by extensive recombination. *J Gen Virol* 2009, 90:734-746.
29. Martin DP, Williams C, Posada D: RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* 2005, 21:260-262.
30. Ilyina TV, Koonin EV: Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res* 1992, 20:3279-3285.
31. Argüello-Astorga G, Herrera-Estrella L: Ancestral multipartite units in light-responsive plant promoters have structural features correlating with specific phototransduction pathways. *Plant Physiol* 1996, 112:1151-1166.
32. Argüello-Astorga G, Herrera-Estrella L: Evolution of light-regulated plant promoters. *Annu Rev Plant Physiol Plant Mol Biol* 1998, 49:25-555.
33. Hernández-Zepeda C, Brown JK, Moreno-Villanueva OA, Argüello-Astorga G, Idris AM, Carnevali G, Rivera-Bustamante RF: Characterization of *Rhynchosia yellow mosaic Yucatan virus*; a new recombinant begomovirus associated with two fabaceous weeds in Yucatan, Mexico. *Arch Virol* 2010, 155:1571-1579.
34. Leisler M, Déhais P, Morsau Y, De Moor B, Rouzé P, Rombauts S: PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 2002, 30:325-327.
35. Higo K, Ugawa Y, Iwamoto M, Korenaga T: Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research* 1999, 27:297-300.
36. Rub-Medrano R, Guevara-Gonzalez RG, Argüello-Astorga GR, Monsalve-Fonseca Z, Herrera-Estrella LR, Rivera-Bustamante RF: Identification of a sequence element involved in AC2-mediated transactivation of the *Pepper huasteco virus* coat protein gene. *Virology* 1999, 253:162-169.
37. Itohaki H, Maxson JM, Woodson WR: An ethylene-responsive enhancer element is involved in the senescence-related expression of the camaron glutathione-S transferase (GST1) gene. *Proc Natl Acad Sci USA* 1994, 91:8925-8929.
38. Montgomery J, Goldman S, Dellman J, Margosian L, Fischer RL: Identification of an ethylene responsive region in the promoter of a fruit ripening gene. *Proc Natl Acad Sci USA* 1993, 90:5939-5943.
39. Rawat R, Xu Z-F, Yao K-M, Chye ML: Identification of cis-elements for ethylene and circadian regulation of the *Solanum melongena* gene encoding cysteine proteinase. *Plant Mol Biol* 2005, 57:629-643.
40. Petty ITD, Coutts RHA, Buck KW: Transcriptional mapping of the coat protein gene of tomato golden mosaic virus. *J Gen Virol* 1988, 69:1359-1365.
41. Paprota T, Metzler V, Jeske H: The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. *Virology* 2010, 404:148-157.
42. Hanson SF, Hoogstraaten RA, Ahlquist P, Gilbertson RL, Russell DR, Maxwell DP: Mutational analysis of a putative NTP-binding domain in the replication-associated protein (AC1) of Bean golden mosaic geminivirus. *Virology* 1995, 211:1-9.
43. Koonin EV: A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res* 1993, 21:2541-2547.
44. Cazzonelli CI, Burke J, Velten J: Functional characterization of the geminiviral conserved late element (CLE) in uninfected tobacco. *Plant Mol Biol* 2005, 58:465-481.
45. Velten J, Money KJ, Cazzonelli CI: Plant viral intergenic DNA sequence repeats with transcription enhancing activity. *Viol J* 2005, 2:16-24.
46. Eagle PA, Harley-Bowdoin L: Cis elements that contribute to geminivirus transcriptional regulation and the efficiency of DNA replication. *J Virol* 1997, 71:6947-6955.
47. Gibbs AJ, Gibbs MJ: A broader definition of the virus species. *Arch Virol* 2006, 151:1419-1422.
48. Van Regenmortel MH: Applying the species concept to plant viruses. *Arch Virol* 1989, 104:1-17.
49. Chatterji A, Padidam M, Beachy RN, Fauquet CM: Identification of replication specificity determinance in tomato leaf curl virus from New Delhi. *J Virol* 1999, 73:5481-5489.
50. Padidam M, Beachy RN, Fauquet CM: Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *J Gen Virol* 1995, 76:25-35.
51. Lefeuve P, Lett JM, Varsani A, Martin DP: Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol* 2009, 83:697-707.
52. Lefeuve P, Lett JM, Reynaud B, Martin DP: Avoidance of protein fold disruption in natural virus recombinants. *PLoS Pathog* 2007, 3:e181.
53. Varsani A, Shepherd DN, Morjane AL, Owor BE, Erdmann JB, Rybicki EP, Peterschmitt M, Briddon RW, Marilham PG, Oluwalafemi S, Windram OP, Lefeuve P, Lett JM, Martin DP: Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *J Gen Virol* 2008, 89:2069-2074.
54. Londoño A, Riego-Ruiz L, Argüello-Astorga GR: DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. *Arch Virol* 2010, 155:1033-1046.
55. Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004, 116:281-297.
56. Nair V, Zavolan M: Virus-encoded microRNAs: novel regulators of gene expression. *Trends Microbiol* 2006, 14:169-175.
57. Skolny RL, Cullen BR: Viruses, microRNAs, and host interactions. *Annu Rev Microbiol* 2010, 64:123-141.
58. Sullivan CS, Grundhoff AT, Teverthia S, Pipas JM, Ganem D: SV40 encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* 2005, 435:682-686.
59. Cullen BR: Five Questions about Viruses and MicroRNAs. *PLoS Pathog* 2010, 6:e1000787.
60. Seo GJ, Fink LH, O'Hara B, Atwood WJ, Sullivan CS: Evolutionarily conserved function of a viral microRNA. *J Virol* 2008, 82:9823-9828.
61. Sullivan CS, Sung CK, Park CD, Grundhoff A, Lukacher AE: Murine Polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection. *Virology* 2009, 387:157-167.
62. Delaporta SL, Wood J, Hicks JB: A plant DNA miniprep: version II. *Plant Mol Biol Rep* 1983, 1:19-21.
63. Mauricio-Castillo JA, Argüello-Astorga GR, Alpuche-Solis AG, Monreal-Vargas CT, Díaz-Gómez O, de la Torre-Almaraz R: First Report of Tomato severe leaf curl virus in México. *Plant Dis* 2006, 90:1116.
64. Camillo-Tripp J, Lozoya-Gloria E, Rivera-Bustamante RF: Symptom remission and specific resistance of pepper plants after infection by *Pepper golden mosaic virus*. *Phytopathology* 2007, 97:51-57.
65. Ascencio-Ibañez JT, Settlage SB: DNA abrasion onto plants is an effective method for geminivirus infection and virus induced gene silencing. *J Virol Methods* 2007, 142:198-203.
66. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24:1596-1599.
67. Sunter G, Bisaro DM: Identification of a minimal sequence required for activation of the Tomato golden mosaic virus coat protein promoter in protoplasts. *Virology* 2003, 305:452-462.

doi:10.1186/1748-4226-7-275

Cite this article as: Gregorio-Jorge et al.: Analysis of a new strain of *Baphorbia mosaic virus* with distinct replication specificity unravels a lineage of begomoviruses with short Rep sequences in the DNA-B intergenic region. *Virology Journal* 2010 7:275.

## Anexo B

### Artículo aceptado como primer autor

A new strain of Tomato severe leaf curl virus and a unique variant of Tomato yellow leaf curl virus from Mexico. *Archives of Virology (en prensa)*

The final publication is available at Springer via <http://dx.doi.org/10.1007/s00705-012-1358-z>

Arch. Virol.  
DOI 10.1007/s00705-012-1358-z

#### 1 ANNOTATED SEQUENCE RECORD

### 2 A new strain of tomato severe leaf curl virus and a unique variant 3 of tomato yellow leaf curl virus from Mexico

4 B. Bañuelos-Hernández · J. A. Mauricio-Castillo ·  
5 Y. Cardenas-Conejo · R. G. Guevara-González ·  
6 G. R. Arguello-Astorga

7 Received: 19 January 2012 / Accepted: 23 April 2012  
8 © Springer-Verlag 2012

9 **Abstract** The complete genome sequence of a distinct  
10 variant of tomato yellow leaf curl virus-Israel (TYLCV-IL)  
11 and the DNA-A sequence of a new strain of tomato severe  
12 leaf curl virus (ToSLCV) isolated in San Luis Potosí,  
13 Mexico, are described and analyzed. The TYLCV-  
14 IL [MX:SLP:11] variant differs from all TYLCV-IL isolates  
15 described so far by a unique 42-nt duplicated sequence  
16 comprising a part of the conserved stem-loop element of the  
17 virion-strand replication origin and adjacent regulatory  
18 sequences. TYLCV-IL [MX:SLP:11] was associated with  
19 tomato chino La Paz virus (ToChLPV-B [MX:SLP:11]) in a  
20 *Solanum pimpinellifolium* plant, and with pepper huasteco  
21 yellow vein virus (PHYVV [MX:SLP:11]) and ToSLCV-  
22 GT [MX:SLP:11] in a *Solanum lycopersicum* plant. In  
23 addition, a distinct ToSLCV exhibiting low sequence  
24 identity (<89 %) to other ToSLCV isolates from Mexico  
25 was found in a tomato plant collected in the same field.  
26 Sequence analysis of this new ToSLCV strain indicates that

it is a recombinant of close relatives of ToSLCV-  
GT [MX:SLP:11] and ToChLPV-B [MX:SLP:11] found in  
mixed infections with TYLCV-IL [MX:SLP:11].

The family *Geminiviridae* is a monophyletic, highly  
diversified group of plant viruses encompassing over 220  
species distributed worldwide [3]. Members of this family  
have genomes composed of one or two circular single-  
stranded DNA (ssDNA) molecules of about 2.6–3.0 kb,  
packed into twinned quasi-icosahedral capsids [17]. The  
geminiviruses transmitted by the whitefly *Bemisia tabaci*  
(Gennadius) are classified into the genus *Begomovirus*,  
which includes some of the most devastating viruses  
affecting dicotyledonous crop plants in the tropical and  
temperate regions of the world [25]. The genomes of  
begomoviruses ("BGVs") that are native to the New  
World are bipartite, consisting of two components desig-  
nated as DNA-A and DNA-B. In contrast, most of the  
known Old World BGVs have monopartite genomes  
consisting of a single DNA molecule homologous to the  
DNA-A component of bipartite BGVs [17, 25]. Tomato  
(*Solanum lycopersicum* L.) is an economically important  
vegetable that is severely affected by BGVs, of which  
tomato yellow leaf curl virus (TYLCV) is the most dam-  
aging, most widely distributed and most extensively  
studied [1]. TYLCV is a monopartite virus that is native to  
the Mediterranean basin and has considerably extended its  
geographical distribution in the last two decades [18]. In  
Mexico, TYLCV was first found in 1996 in the Yucatan  
Peninsula close to the Caribbean [2], and a decade later, it  
was also reported in Sinaloa [4, 10], ~1800 km away, on  
the Pacific coast. According to molecular phylogeny  
studies, these distant Mexican isolates descended from  
different TYLCV introductions [8, 18]. In addition to

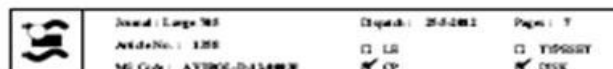
A1 **Electronic supplementary material** The online version of this  
A2 article (doi:10.1007/s00705-012-1358-z) contains supplementary  
A3 material, which is available to authorized users.

A4 B. Bañuelos-Hernández · J. A. Mauricio-Castillo ·  
A5 Y. Cardenas-Conejo · G. R. Arguello-Astorga (✉)  
A6 Instituto Potosino de Investigación Científica y Tecnológica  
A7 (IPICYT), A.C., Camino a la Presa de San José, 78216 San Luis  
A8 Potosí, S.P., Mexico  
A9 e-mail: gmguel@ipicyt.edu.mx

A10 J. A. Mauricio-Castillo  
A11 Unidad Académica de Agronomía, Universidad Autónoma  
A12 de Zacatecas, Apdo. Postal 336, 98000 Zacatecas, Zac, México

A13 R. G. Guevara-González  
A14 C.A. Ingeniería de Biosistemas, Facultad  
A15 de Ingeniería, Universidad Autónoma de Querétaro,  
A16 76010 Santiago de Querétaro, Qro, México

Springer



62 TYLCV, several BGVs that are native to the American  
63 continent have been described affecting tomato crops in  
64 Mexico [22, 28]. Two of those viruses, tomato chilo La  
65 Paz virus (ToChLPV) [12, 13] and tomato severe leaf curl  
66 virus (ToSLCV) [20, 24] are particularly interesting  
67 because both have been found exclusively in mixed  
68 infections with other BGVs, and their genomic component  
69 B has not been identified until now in spite of efforts by  
70 several laboratories to isolate them [12, 20, 23, 24]. Here,  
71 we describe a new strain of ToSLCV of recombinant  
72 origin and a TYLCV variant exhibiting an intergenic  
73 region (IR) with unique features that were found in mixed  
74 infections with ToSLCV and ToChLPV in tomato plants  
75 collected in a locality of the Huasteca region of San Luis  
76 Potosí (SLP), Mexico.

77 Leaf material from one wild tomato (*Solanum*  
78 *pimpinellifolium*; sample #1-Spm) and two cultivated  
79 tomato plants (*Solanum lycopersicum*; samples #2-Slyc and  
80 #3-Slyc) exhibiting mosaic and leaf curling symptoms  
81 were collected during July 2011 in a home garden located  
82 in Aquismón, a county of the SLP state (21°36' 08.10" N,  
83 96°06'21.43" W). Total DNA was extracted from these  
84 samples by a modified version of the Dellaporta method  
85 [7], and the presence of BGVs was assessed by PCR using  
86 the degenerate primer cpYMAC-rv in combination with  
87 two lineage-specific degenerate primers, namely, the  
88 primer repDGR-fw, specific for BGVs from all continents  
89 [21], and the primer SL2150-fw, specific for viruses of the  
90 squash leaf curl virus (SLCV) clade, a lineage of New  
91 World BGVs displaying distinctive features in their replica-  
92 tion-associated protein (Rep) and the region encom-  
93 passing the origin of virion-strand replication [11]. The  
94 amplicons obtained with the pair of primers cpYMAC-rv/  
95 SL2150-fw are ~1.35 kb in length, whereas the PCR  
96 products obtained with the pair of primers cpYMAC-rv/  
97 repDGR-fw are ~1.4 kb in length for New World BGVs,  
98 and ~1.55 kb for Old World BGVs. To obtain the com-  
99 plete DNA-A sequence of BGVs in the plant samples, two  
100 sets of primers, SL2150-rv/CP70-fw and repDGR-rv/  
101 CP70-fw [11, 21], respectively, were used to amplify a  
102 ~1.7-kb fragment overlapping along a ~500-bp segment  
103 with the previously generated amplicons. The partially  
104 overlapping PCR products derived from each sample were  
105 cloned and sequenced separately, and full-length contigs  
106 were then assembled. Two sets of degenerate primers,  
107 BC1-290fw/ BV1-470rv and BC1-290rv/ BV1-310fw [11],  
108 were utilized to amplify DNA-B sequences. Four primers  
109 (A to D) complementary to sequences within the replica-  
110 tion origin of ToSLCV were generated to assist the PCR  
111 amplification of its cognate DNA-B: primer A, 5'-TAAAT  
112 GGGACACCAAG-3'; primer B, 5'-ATTTTGC  
113 TGGTGTCTGTGTCC-3'; primer C, 5'-AATTTGCC  
114 TTGGTGTCCCA-3'; primer D, 5'-TAAAACCTCTCTGG

115 GGACACCAAG-3'. Viral amplicons were cloned into  
116 pGEM-T Easy Vector (Promega) and subjected to restric-  
117 tion fragment length polymorphism (RFLP) analysis using  
118 *EcoRI* and *HinfI* enzymes. The digestion products were  
119 separated by electrophoresis in 2.5 % agarose gels, and  
120 PCR clones exhibiting different restriction fragment pat-  
121 terns were identified. Three independent clones of each  
122 type of amplicon with a distinctive restriction pattern were  
123 sequenced. Automated sequencing was carried out at  
124 LANBAMA-IPICYT (San Luis Potosí, Mexico) using a  
125 3130 Genetic Analyzer (Applied Biosystems). Pairwise  
126 nucleotide identities were calculated using the Clustal-V  
127 algorithm of the MegAlign application of Lasergene  
128 (DNAstar Inc., Madison, WI, USA). Phylogenetic trees  
129 were constructed using the neighbour-joining algorithm  
130 and the number of differences nucleotide substitution  
131 model available in Mega 5.0 [27]. Recombination analysis  
132 was performed using the suite of recombination detection  
133 methods in the RDP3 software [19].

134 All plant samples examined produced two or three viral  
135 amplicons with different RFLP patterns (Supplementary  
136 Fig. 1a), hence indicating mixed infections. In sample #1-  
137 Spm, two distinct DNA A-like sequences were assembled,  
138 one of them consisting of 2,632 nt that displayed the  
139 highest level of nucleotide sequence identity (SI) to ToCh-  
140 LPV-B [MX:RioMM4:05] (accession no. DQ347949) at  
141 95.4 %, whereas the second one exhibited features typical  
142 of Old World BGVs, with two ORFs in the virion-sense  
143 and four ORFs in the complementary sense. BLAST  
144 analysis of this assembled genome sequence (2,794 nt in  
145 length) showed the highest SI to a TYLCV-IL isolate from  
146 Texas (EF110890), at 99.1 %. This new variant was termed  
147 TYLCV-IL [MX:SLP:11]. The two *Solanum lycopersicum*  
148 samples contained different combinations of BGVs. In  
149 sample #2-Slyc, three BGVs were identified: pepper  
150 huasteco yellow vein virus (PHYVV), ToSLCV and  
151 TYLCV-IL. The PHYVV DNA-A consisted of 2,630 nt  
152 with the highest identity (SI = 95.8 %) to PHYVV-  
153 [MX:Tam] (X70418), whereas the ToSLCV DNA-A  
154 (2,594 nt in length) showed the highest SI to ToSLCV-  
155 [MX:Rio2:05] (DQ347947), at 96.7 %. The monopartite  
156 BGV in this sample had a sequence that was nearly iden-  
157 tical (SI = 99.8%) to that of TYLCV-IL [MX:SLP:11]. The  
158 sample #3-Slyc, on the other hand, contained two different  
159 DNA-A molecules, one of them displaying 98.9 %  
160 sequence identity to the PHYVV isolate from sample #2-  
161 Slyc (termed PHYVV-[MX:SLP:11]). The second DNA-A,  
162 consisting of 2593 nt, displayed the highest SI to ToSLCV  
163 isolates from Nicaragua (AJ508785 and AJ508784), at  
164 93.4 % and 93.3 %, respectively. PCR amplification of  
165 DNA-B sequences from this sample produced amplicons of  
166 PHYVV but not of ToSLCV (data not shown). Interest-  
167 ingly, no DNA-B sequences could be amplified from the



168 two TYLCV-infected plants (samples #1-Spim and  
169 #2-Slyc) despite repeated attempts using PCR and rolling-  
170 circle amplification with Phi29 polymerase. Moreover, in a  
171 systematic endeavor to amplify DNA-B sequences of  
172 ToSLCV by PCR, several oligonucleotide primers that  
173 specifically target critical sequences within the replication  
174 origin (*Ori*) of this virus were used in combination with  
175 degenerate primers complementary to *BCI* and *BVI* con-  
176 served sequences. No PCR product was obtained from  
177 samples #2-Slic and #3-Slic after a number of attempts,  
178 although amplicons of the expected size were obtained from  
179 those samples when the ToSLCV *Ori*-specific primers were  
180 combined with degenerate primers complementary to *CP*  
181 and *Rep* conserved sequences (Supplemental Figure 2). An  
182 analogous absence of DNA-B was observed previously in  
183 pepper plants co-infected with TYLCV-IL and ToChLPV in  
184 Baja California Sur, Mexico [5]. These observations sug-  
185 gest that ToSLCV and ToChLPV are complemented in  
186 movement functions by TYLCV in mixed infections.

187 The variety of BGVs identified in the three field samples  
188 examined and their GenBank accession numbers are sum-  
189 marized in Table 1. Two of these viruses were particularly  
190 interesting, namely, the ToSLCV isolated from sample #3-  
191 Slyc and TYLCV-IL [MX:SLP:11]. The former, provision-  
192 ally termed ToSLCV-S#3, exhibited an overall sequence  
193 identity of 90.9 % with ToSLCV-GT [MX:SLP:11] and  
194 88.1–88.6 % with other isolates of the ToSLCV-GT strain,  
195 including two isolates from Rioverde, San Luis Potosí:  
196 ToSLCV-GT [MX:Rio:05] (DQ347946) and ToSLCV-  
197 GT [MX:Rio:2:05] (DQ347947). Recombination analysis  
198 using the RDP3 program and based on a Clustal W align-  
199 ment of all available ToSLCV (seven isolates) and ToCh-  
200 LPV (six isolates) sequences, including those obtained in  
201 this study, revealed that ToSLCV-S#3 is a recombinant,  
202 with the sequences spanning the IR 5' segment and all of  
203 complementary-sense genes (i.e., *Rep*, *TRAP*, *REN* and *AC4*)  
204 originating from ToSLCV-GT [MX:SLP:11], and the IR  
205 right segment and *CP* originating from ToChLPV-  
206 B [MX:SLP:11] (the "minor" parent). The putative recom-  
207 bination event between ToSLCV and ToChLPV was well  
208 supported by recombination detection methods in RDP3  
209 (RDP,  $P = 8.24 \times 10^{-67}$ ; Geneconv,  $P = 1.61 \times 10^{-63}$ ;  
210 Bootscan,  $P = 5.61 \times 10^{-66}$ ). The presumed recombinant  
211 fragment from the minor parent is 860 nt in size (coordi-  
212 nates 15–875) and displays high sequence identity (98.4 %)  
213 to ToChLPV-B [MX:SLP:11] but only 73.4 % identity to  
214 ToSLCV-GT [MX:SLP:11]. In contrast, the remainder of  
215 the DNA-A molecule (coordinates 876 to 2612, and 1 to 14)  
216 exhibits 99.1 % sequence identity to the last ToSLCV iso-  
217 late but only 76.1 % identity to ToChLPV-B [MX:SLP:11]  
218 (Fig. 1). Moreover, close inspection of alignments of all  
219 available ToSLCV and ToChLPV sequences strongly sup-  
220 ported the hypothesis that a variant of ToChLPV-

221 B [MX:SLP:11] is the minor parent of ToSLCV-S#3, 221  
222 because a 19-nt insertion exclusively present in this isolate 222  
223 is also observed in the former virus (Fig. 2). A phylogenetic 223  
224 tree based on an alignment of the complete DNA-A 224  
225 sequences of ToSLCV and ToChLPV isolates identified in 225  
226 this study and a number of selected New World BGVs, was 226  
227 generated (Fig. 3). The dendrogram reveals two interesting 227  
228 facts: (1) ToSLCV and ToChLPV form a "distinct" cluster 228  
229 that is basal to the SLCV clade, notwithstanding that ToS- 229  
230 LCV, but not ToChLPV, exhibits all the distinctive features 230  
231 of members of this lineage; (2) ToSLCV-S#3 seems to be 231  
232 more closely related to isolates from Nicaragua than the 232  
233 ones from Mexico and Guatemala. However, the data 233  
234 illustrated in Figs. 1 and 2 strongly suggest that the major 234  
235 parent of ToSLCV-S#3 is a close relative of ToSLCV- 235  
236 GT [MX:SLP:11]. Considering that ToSLCV-NI is also a 236  
237 recombinant of ToSLCV and ToChLPV, as demonstrated 237  
238 previously [24], it is reasonable to presume that the high 238  
239 sequence identity (93.4 %) between ToSLCV-S#3 and 239  
240 ToSLCV-NI is just a consequence of their independent 240  
241 origin as recombinants from the same pair of viruses, not a 241  
242 reflection of their actual phylogenetic relationships. Taking 242  
243 into account those observations and the current taxonomic 243  
244 criteria for BGVs establishing that a DNA-A sequence 244  
245 identity lower than 93 % with other isolates of a viral 245  
246 species is indicative of a separate strain [9], we concluded 246  
247 that ToSLCV-S#3 represents an isolate of a new ToSLCV 247  
248 strain for which the designation ToSLCV-Huasteca (ToS- 248  
249 LCV-HU) is proposed. 249

250 Sequence analysis of the monopartite BGV identified in 250  
251 this study, TYLCV-IL [MX:SLP:11], revealed two distinc- 251  
252 tive features of its IR: (1) A deletion that eliminated a 252  
253 29-nt long sequence located between the Rep-binding sites 253  
254 (iternons) and a set of four putative ethylene-responsive 254  
255 elements (EREs) adjacent to the conserved stem-loop ele- 255  
256 ment (Fig. 4a); and (2) A perfect duplicate of a 42-nt 256  
257 sequence adjacent 3' to the Rep nick-site in the stem-loop 257  
258 element, comprising the "stem" right arm and a DNA 258  
259 motif named Conserved Late Element (CLE) [26]. This 259  
260 duplication is not found in any other TYLCV isolate 260  
261 described to date. The duplicated segment was observed in 261  
262 all sequenced PCR clones, and its presence in TYLCV-IL 262  
263 [MX:SLP:11] DNA was confirmed by sequencing a full- 263  
264 length genome clone obtained by RCA from sample 264  
265 #2-Slic. A BlastN search revealed the existence of analog- 265  
266 ous repeated sequences in other Old World BGVs: (i) In 266  
267 the first described isolate of Indian cassava mosaic virus 267  
268 (ICMV) (but no other isolates of this viral species), which 268  
269 exhibits a 41-nt sequence containing the stem right arm, a 269  
270 CLE and a G-box element, which is reported three times in 270  
271 tandem in the IR of both genomic components [14]; (ii) in 271  
272 several BGVs related to East African cassava mosaic virus 272  
273 (EACMV), namely, EACMCV, EACMKV, EACMMV, 273

Table 1 Begomoviruses identified in this study

Plant host	Begomovirus	Acronym	GB accession number
<i>Solanum pimpinellifolium</i> (sample 41-Spin)	Tomato yellow leaf curl virus-strain Isreal [Mexico: San Luis Potosí-2:2011]	TYLCV-IL [MX:SLP:11]	JQ303121
	Tomato chilo La Paz virus-strain B [Mexico: San Luis Potosí:2011]	ToChLPV-B [MX:SLP:11]	JN676150
<i>Solanum lycopersicum</i> (sample 42-Slyc)	Tomato yellow leaf curl virus-strain Isreal [Mexico: San Luis Potosí-1:2011]	TYLCV-IL [MX:SLP:11]	JN680353
	Tomato severe leaf curl virus-strain Guatemala [Mexico: San Luis Potosí:2011]	TSLCV-GT [MX:SLP:11]	JN676151
	Pepper huasteco yellow vein virus-[Mexico: San Luis Potosí:2011]	PHYVV-[MX:SLP:11]	JQ303122
<i>Solanum lycopersicum</i> (sample 43-Slyc)	Tomato severe leaf curl virus-strain Huasteca* [Mexico: San Luis Potosí:2011]	TSLCV-HU [MX:SLP:11]	JN680352
	Pepper huasteco yellow vein virus-[Mexico: San Luis Potosí-2:2011]	PHYVV-[MX:SLP:2:11]	Not submitted

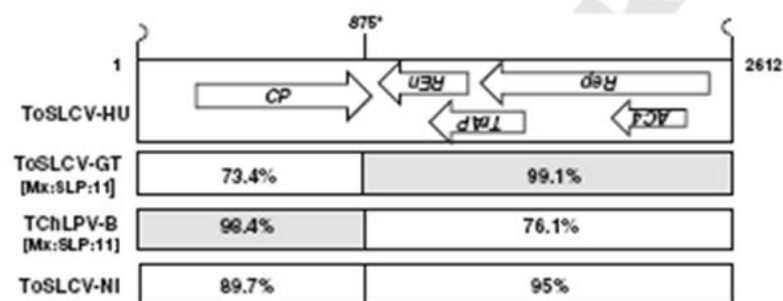


Fig. 1 Recombinant regions detected in ToSLCV-HU [MX:SLP:11]. The schematic representation of ToSLCV-HU DNA-A is at the top of the figure, and the equivalent genomic sequences of its putative "major" and "minor" parents, as well as that of ToSLCV-NI, are represented below. Segment coordinates are nucleotide positions of

detected recombination breakpoints. The nucleotide sequence identity to the homologous region of ToSLCV-HU is indicated at each genomic section. Shaded segments correspond to genomic fragments presumably contributed by the "major" and "minor" parents

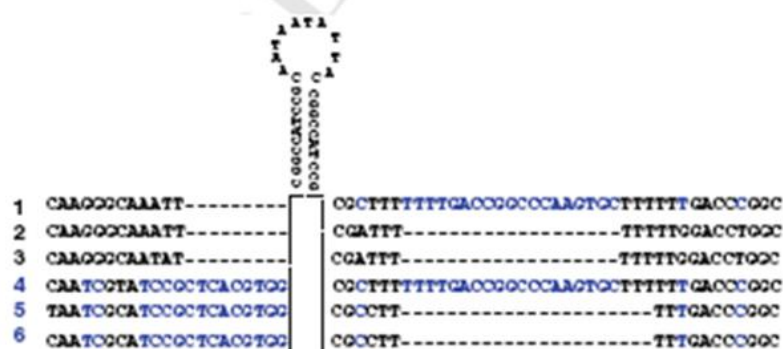
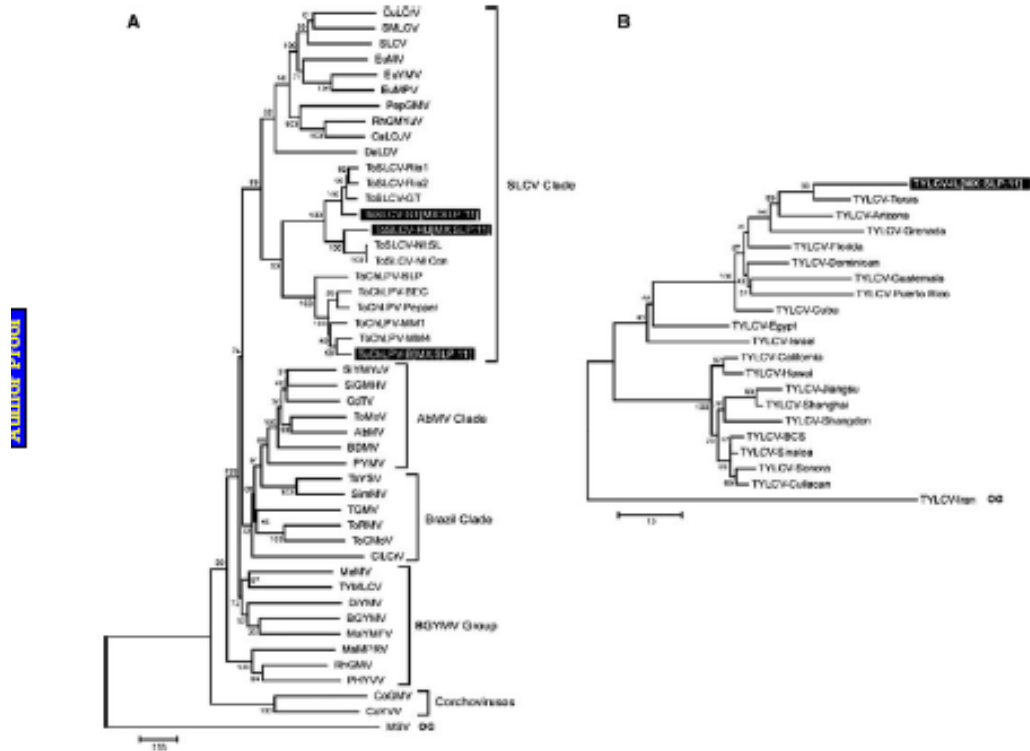


Fig. 2 Comparisons of nucleotide sequences adjacent to the conserved stem-loop of ToSLCV-HU with the homologous region of selected TSLCV and ToChLPV isolates. The compared sequences correspond to the following viruses: (1) ToSLCV-HU [MX:SLP:11];

(2) ToSLCV-GT [MX:SLP:11]; (3) ToSLCV-NI [NI:Con]; (4) ToChLPV-B [MX:SLP:11]; (5) ToChLPV-B [MX:RioM4:05]; (6) ToChLPV-B [MX:BCS]. Letters in blue denote nucleotides of ToChLPV that are not shared with ToSLCV





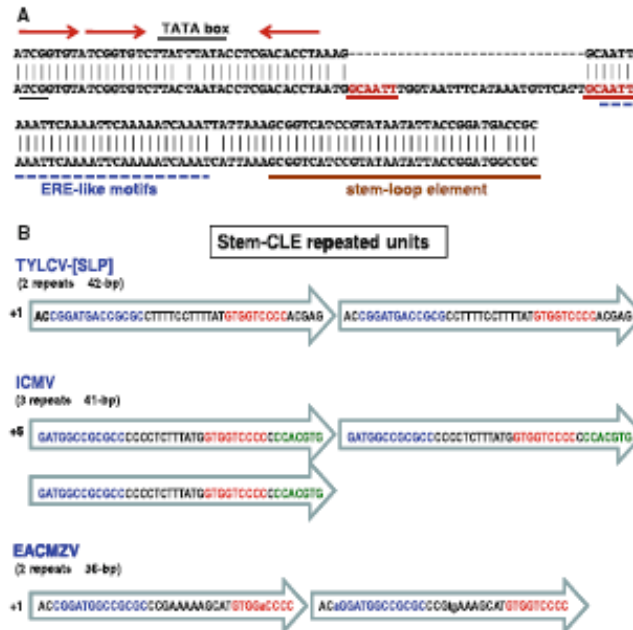
**Fig. 3** Phylogenetic trees based on alignments of the full-length DNA-A sequence of viruses isolated in this work (shaded boxes) and selected begomoviruses. The trees were generated using the neighbor-joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000

replicates) are shown next to the branches. The evolutionary distances were computed using the number of differences method and are in units of number of base differences per sequence. Evolutionary analysis was conducted using MEGA5. Virus names, acronyms and GenBank accession numbers are listed in Supplementary Table 1

274 and EACMV, all of which exhibit two imperfect repeats  
 275 of 36 nt in length containing the stem right arm and a CLE  
 276 motif (Fig. 4b). These sequence similarities between dis-  
 277 tantly related BGVs and the presence of duplicate or tri-  
 278 plicate IR segments only in certain isolates of TYLCV and  
 279 KMV are intriguing observations that are difficult to  
 280 explain at present. The function of the duplicated frag-  
 281 ments encompassing the stem right arm and a CLE is not  
 282 known, but they could be involved in the transcriptional  
 283 regulation of viral genes, because the CLE motif is a well-  
 284 characterized cis-regulatory element of CP gene expression  
 285 in both BGVs and curtoviruses [6, 15, 26], whereas the  
 286 stem right arm and 3' adjacent elements of TYLCV are  
 287 very similar in sequence (SI = 85 %) to a 34-nt fragment  
 288 (-146 to -112) of cotton leaf curl Multan virus  
 289 (CLCuMV) whose deletion abolished the activity of its Rep

290 promoter [29]. Therefore, it is reasonable to expect that the  
 291 duplication of the 42-nt IR segment in TYLCV-  
 292 ILJ[MX:SLP:11] would alter in some way the activity and/  
 293 or the regulatory properties of its Rep and CP promoters—  
 294 an interesting question that we are currently examining.

295 Phylogenetic analysis of New World TYLCV isolates  
 296 showed that TYLCV-ILJ[MX:SLP:11] is more closely  
 297 related to isolates from the USA than to other Mexican  
 298 isolates (Fig. 3b). Moreover, PCR clones of TYLCV-  
 299 ILJ[MX:SLP:11] displayed RFLP patterns that were clearly  
 300 different from those of TYLCV isolates from Sinaloa and  
 301 Baja California Sur (Supplementary Fig. 1b). Since the  
 302 SLP isolate exhibits a 29-nt deletion that is uniquely shared  
 303 with isolates from Texas and Arizona [16] but displays in  
 304 addition a distinctive 42-nt duplication, it may be inferred  
 305 that this molecular trait evolved after the deletion event,



**Fig. 4** Distinctive features of the intergenic region of TYLCV-II[MX:SLP-11]. (A) Alignment of the virion-strand origin region of this virus (upper line) with the homologous region of TYLCV-II[USA: Flo] (lower line) illustrating the 29-nt deletion shared with TYLCV-II isolates from Texas and Arizona. The Rep-binding sites (terons) are represented by arrows, and a set of four ERE-like motifs (AATTCAAA) in tandem is underlined. The 6-nt direct repeats of TYLCV-II[USA: Flo] matching the borders of the deletion (red letters) probably were involved in the excision event. (B) Repeated sequences encompassing part of the conserved stem-loop element and adjacent regulatory motifs in the IR of three distantly related

begomoviruses. The repeats of TYLCV-[SLP] and ICMV are perfect (SI = 100 %), whereas those of EACMZV are imperfect, and there are no gaps between individual repeats. The relevant elements in the repeated sequences are distinguished by colored letters: the "stem" right arm (CGGATGGCCCGC), the CLE motif (GTGGTCCCGC) and the G-box element (CCACGTG). The first AC dinucleotide in the TYLCV and EACMZV repeats is the last dinucleotide of the geminivirus invariant nonanucleotide (TAATATTAC). Accession numbers of ICMV and EACMZV isolates are Z34758 and AJ717567, respectively.

306 which probably took place within the last 15 years  
307 according to the historical record of TYLCV introductions  
308 to America [8, 18]. This illustrates the intrinsic potential of  
309 BGVs to evolve rapidly through spontaneous point muta-  
310 tions, deletions and duplications in addition to the pro-  
311 cesses of reassortment of genomic components and  
312 intermolecular recombination, which are the main mecha-  
313 nisms of evolution in this subgroup of geminiviruses.

314 **Acknowledgments** We thank Salvador Ambrós-González for his  
315 technical assistance and two anonymous reviewers by thoughtful  
316 comments and suggestions. BBH, JAMC and YCC were supported by  
317 a PhD fellowship from the Consejo Nacional de Ciencia y Tecnología

(CONACYT, Mexico). This research was supported by a FOMEX-San  
Luis Potosí grant (FMSLR-2008-C01-86994) to GIRA.

## References

1. Althay M, Pail BL, Fauquet CM (2007) Molecular biodiversity, taxonomy, and nomenclature of tomato yellow leaf curl-like viruses. In: Cooseck H (ed) Tomato yellow leaf curl virus disease: management, molecular biology, breeding for resistance. Springer, Berlin, pp 85–118
2. Ascencio-Ibanez JT, Diaz-Plaza R, Mendes-Luzano J, Monsalve-Romero ZI, Aguillo-Astorga GR, Rivas-Bustamante RF

- 329 (1999) First report of *Tomato yellow leaf curl geminivirus* in  
330 Yucatan, Mexico. *Plant Dis* 83:1173
- 331 3. Briddon RW, Patel BL, Bagewadi B, Nawaz-ul-Rehman MS,  
332 Fauquet CM (2010) Distinct evolutionary histories of the DNA-A  
333 and DNA-B components of bipartite begomoviruses. *BMC Evol*  
334 *Biol* 10:97
- 335 4. Brown JK, Idris AM (2006) Introduction of the exotic mono-  
336 partite *Tomato yellow leaf curl virus* into west coast Mexico.  
337 *Plant Dis* 90:1360
- 338 5. Cardenas-Conejo Y, Argüello-Astorga GR, Poghosyan A, Her-  
339 nández-González I, Letsky V, Holguín-Peña J, Medina-Her-  
340 nández D, Vega-Peña S (2010) First report of *Tomato yellow leaf*  
341 *curl virus* co-infecting pepper with *Tomato chilo La Paz virus* in  
342 Baja California Sur, México. *Plant Dis* 94:1266
- 343 6. Cassoelli CI, Burke J, Velten J (2005) Functional character-  
344 ization of the geminiviral conserved late element (CLE) in  
345 uninfected tobacco. *Plant Mol Biol* 58:465–481
- 346 7. Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA mini-  
347 preparation: version II. *Plant Mol Biol Rep* 1:19–21
- 348 8. Duffy S, Holmes EC (2007) Multiple introductions of the Old  
349 World begomovirus *Tomato yellow leaf curl virus* into the New  
350 World. *Appl Environ Microbiol* 73:7114–7117
- 351 9. Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J,  
352 Zerbini M, Zhou X (2008) Geminivirus strain demarcation and  
353 nomenclature. *Arch Virol* 153:783–821
- 354 10. Gómez-Jiménez C, Romero-Romero JL, Santos-Chrvantes MB,  
355 Leyva-López NE, Méndez-Lomazo J (2009) *Tomatillo (Physalis*  
356 *incorpae)* as a natural new host for *Tomato yellow le of curl virus*  
357 in Simolco, México. *Plant Dis* 93:545
- 358 11. Gregorio-Jorge J, Bernal-Alcocer A, Baduelos-Hernández B,  
359 Alpuche-Solis AG, Hernández-Zepeda C, Moreno-Valenzuela O,  
360 Frías-Treviño G, Argüello-Astorga GR (2010) Analysis of a new  
361 strain of *Euphorbia mosaic virus* with distinct replication speci-  
362 ficity unveils a lineage of begomoviruses with short Rep  
363 sequences in the DNA-B intergenic region. *Viral J* 7:275
- 364 12. Holguín-Peña RJ, Argüello-Astorga GR, Brown JK, Rivera-  
365 Bustamante RF (2006) A new strain of *Tomato chilo La Paz*  
366 *virus* associated with a leaf curl disease of tomato in Baja Cal-  
367 ifornia Sur, México. *Plant Dis* 90:973
- 368 13. Holguín-Peña RJ, Vázquez-Juárez R, Rivera-Bustamante RF  
369 (2005) A new begomovirus causes tomato leaf curl disease in  
370 Baja California Sur, Mexico. *Plant Dis* 89:341
- 371 14. Hong YG, Robinson DJ, Harrison BD (2003) Nucleotide  
372 sequence evidence for the occurrence of three distinct whitefly-  
373 transmitted geminiviruses in cassava. *J Gen Virol* 74:2437–2443
- 374 15. Hur J, Choi E, Buddiey KJ, Lee S, Davis KR (2008) Identification  
375 of a promoter motif involved in *Curtovirus* sense-gene expression  
376 in transgenic *Arabidopsis*. *Mol Cells* 26:131–139
- 377 16. Idris AM, Guerrero JC, Brown JK (2007) Two distinct isolates of  
378 *Tomato yellow leaf curl virus* threaten tomato production in  
379 Arizona and Sonora, México. *Plant Dis* 91:910
- 380 17. Jolke H (2009) Geminiviruses. *Curr Top Microbiol Immunol*  
381 331:185–226
- 382 18. Lefeuve P, Martin DP, Harkins G, Lemey P, Gray AJ, Meredith  
383 S, Lakay F, Monjan A, Lett JM, Vassani A, Heydarejad J  
384 (2010) The spread of tomato yellow leaf curl virus from the  
385 Middle East to the world. *PLoS Pathog* 6(10):e1001164
- 386 19. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuve P  
387 (2010) RDP3: a flexible and fast computer program for analyzing  
388 recombination. *Bioinformatics* 26:2462–2463
- 389 20. Mauricio-Castillo JA, Argüello-Astorga GR, Alpuche-Solis AG,  
390 Monreal-Vargas CT, Díaz-Gómez O, de La Torre-Almanaz R  
391 (2006) First Report of *Tomato severe leaf curl virus* in México.  
392 *Plant Dis* 90:1116
- 393 21. Mauricio-Castillo JA, Argüello-Astorga GR, Ambríz-González S,  
394 Alpuche-Solis AG, Monreal-Vargas C (2007) First Report of  
395 *Tomato golden mottle virus* on *Lycopersicon esculentum* and  
396 *Solanum rostratum* in México. *Plant Dis* 91:1513
- 397 22. Morales FJ (2006) History and current distribution of bego-  
398 moviruses in Latin America. *Adv Virus Res* 67:127–162
- 399 23. Nakhla MK, Sorenson A, Mejía L, Ramirez P, Karkashian JP,  
400 Maxwell DP (2005) Molecular characterization of tomato-  
401 infecting begomoviruses in Central America and development of  
402 DNA-based detection methods. *Acta Hort (ISHS)* 695:277–288
- 403 24. Rojas A, Kvamhøden A, Marcano D, Valkonen JP (2005)  
404 Sequence characterization of *Tomato leaf curl Simola virus* and  
405 *Tomato severe leaf curl virus*: phylogeny of New World be-  
406 gomoviruses and detection of recombination. *Arch Virol*  
407 150:1281–1299
- 408 25. Rojas MR, Hagen C, Lucas WJ, Gilbertson RL (2005) Exploiting  
409 clinks in the plant's armor: evolution and emergence of gemin-  
410 iviruses. *Annu Rev Phytopathol* 43:361–394
- 411 26. Ruiz-Medrano R, Guevara-González RG, Argüello-Astorga GR,  
412 Monsalvo-Fomiega Z, Herra-Estrella LR, Rivera-Bustamante  
413 RF (1999) Identification of a sequence element involved in AC2-  
414 mediated transactivation of the pepper huasteco virus coat protein  
415 gene. *Virology* 253:162–169
- 416 27. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S  
417 (2011) MEGA5: Molecular evolutionary genetics analysis using  
418 Maximum Likelihood, Evolutionary Distance, and Maximum  
419 Parsimony methods. *Mol Biol Evol* 28:2731–2739
- 420 28. Torres-Pacheco I, Garzon-Timado JA, Brown JK, Becerra-Floa  
421 A, Rivera-Bustamante RF (1996) Detection and distribution of  
422 geminiviruses in Mexico and the southern United States. *Phyto-*  
423 *pathology* 86:1186–1192
- 424 29. Xie Y, Liu Y, Meng M, Chen L, Zhu Z (2003) Isolation and  
425 identification of a super strong plant promoter from *Cotton leaf*  
426 *curl Malvar virus*. *Plant Mol Biol* 53:1–14
- 427

## VIII.- APÉNDICES

### A. Protocolo para la mutagénesis sitio-dirigida del plásmido pBluescript; Método de QuikChange site-directed mutagenesis kit

1.- Diseñar los oligonucleótidos con la o las mutaciones necesarias en la región de interés. Los oligonucleótidos deberán tener las siguientes consideraciones.

a.- longitud entre 25 y 45 nucleótidos con una  $T_m$  mayor o igual a 78°C.

b.- el sitio de la mutación debe ubicarse en la parte media del primer con 10 a 15 nucleótidos en ambos lados del oligonucleótido.

c.- tener un contenido de GC mínimo de 40% y terminar en una o más G o C.

2.- Preparar la mezcla de reacción como sigue:

- 5ul de buffer 10X para la *Pfu* DNA polimerasa.
- 2-50 ng de DNA templado.
- 125 ng de oligonucleótido directo.
- 125 ng de oligonucleótido reverso.
- 1ul de dNTPs.
- Agua desionizada estéril hasta un volumen de 50ul. Este volumen depende de la cantidad en microlitros que se hayan puesto de DNA templado y oligonucleótidos según la concentración a la que se encuentren.
- 1ul de DNA Polimerasa *Pfu* turbo (2.5U/ul).

3.- Someter la mezcla al programa de PCR. Los parámetros a utilizar son los siguientes:

- Desnaturalización inicial de 30seg a 95°C.
- 12 ciclos con tiempos de; 30seg a 95°C, 1min a 55°C y 3min 68°C. El tiempo para la temperatura de elongación (68°C) depende del tamaño del inserto, por cada kilobase de DNA dar un minuto de elongación.
- terminados los ciclos incubar la reacción en hielo por 2minutos.

4.- Agregar al producto de PCR 1ul (10U) de la enzima *Dpn I* e incubar a 37°C por al menos una hora (*Dpn I* digiere el DNA plasmídico original usado como templado).

5.- Usar el producto de PCR tratado para transformar células  $\text{CaCl}_2$ -competentes por choque térmico (apéndice C).

6.- Hacer el escrutinio de las colonias bacterianas obtenidas.

## **B. Precipitación de DNA con acetato de sodio y etanol 100%**

- 1.- Agregar al DNA 2 volúmenes de etanol absoluto más 1/10 de acetato de sodio 3M pH 5.2 y mezclar por inversión.
- 2.- Centrifugar 5min a 13000rpm y tirar el sobrenadante.
- 3.- Agregar 700ul de etanol al 70%. Lavar invirtiendo el tubo suavemente repetidas veces.
- 4.- Centrifugar 3min a 1300rpm y tirar el sobrenadante.
- 5.- Repetir los pasos 3 y 4.
- 6.- Dejar secar la pastilla de DNA y resuspender en 30ul de agua desionizada estéril.

## **C. Transformación por choque térmico de *E. coli* Top10F'**

- 1.- Tomar las células competentes de *E. coli* mantenidas a -70°C y colocarlas en hielo.
- 2.- Agregar el DNA a transformar a las células competentes y dejar en hielo por 15 minutos.
- 3.- Llevar las células competentes con el DNA a 42°C por 1min e inmediatamente pasarla a hielo.
- 4.- Mantener en el hielo por 5min y posteriormente en la campana (condiciones de esterilidad) agregar 1uL de medio LB.
- 5.- Llevar las células en medio LB a recuperación por 1hr a 37°C. (En caso de que no se observe nada turbio el medio, dejarlo hasta que adquiera turbidez, sobre todo si se trata de transformación de ligaciones).
- 6.- Centrifugar 2min a 8000rpm y eliminar 800ul del medio.
- 7.- De los 200ul restantes sembrar 100ul en cajas Petri con Medio LB solido más carbenicilina (ampicilina) a una concentración de 100ug/ml.
- 8.- Incubar a 37 ° C por 14 horas.
- 9.- Hacer el escrutinio de las colonias bacterianas obtenidas.
- 10.- En el paso 7 los 100ul restantes de células transformadas guardarlos a 4°C para su uso posterior, por ejemplo en caso de obtener demasiado crecimiento sembrar nuevamente para obtener colonias bacterianas mas separadas.

## **D. Preparación de partículas de tungsteno usadas para la transformación por biobalística**

### **I.- Limpieza de las partículas**

- 1.- Pesar 60mg de partículas de tungsteno (M10, BioRad) y ponerlos en un microtubo de 2ml.
- 2.- Resuspender en 2ml de ácido nítrico ( $\text{HNO}_3$ ) 0.1N y sonicar 20 minutos. Revisar continuamente que las partículas estén resuspendidas)
- 3.- Centrifugar a 10000rpm por 2 minutos.
- 4.- Eliminar el sobrenadante y agregar 2ml de etanol absoluto y sonicar brevemente.
- 5.- Centrifugar a 10000 durante 2 minutos.
- 6.- Eliminar el sobrenadante y resuspender en 1ml de agua desionizada estéril.
- 7.- Asegurándose de que las partículas estén bien resuspendidas dividir el mililitro contenido en el tubo en 4 tubos (250ul en cada tubo)
- 8.- Agregar a cada tubo 750ul de agua desionizada estéril.
- 9.- La concentración final de tungsteno en cada tubo es de 15ug/ul.
- 10.- Almacenar a  $-20^\circ\text{C}$ .

### **II.- Recubrimiento de las partículas de tungsteno con el DNA**

- 1.- Poner en un microtubo de 1.5ml (tubo eppendorf) el DNA a utilizar (en nuestro caso 6.5ug de DNA A y 6.5ug de DNA B)
- 2.- Agregar 50ul de partículas de tungsteno (equivalente a 50ug).
- 3.- Agregar 50ul de cloruro de calcio ( $\text{CaCl}_2$ ) 2.5M.
- 4.- Agregar 20ul de espermidina 0.1M.
- 5.- Sonicar brevemente (hasta que las partículas se vean resuspendidas).
- 6.- Centrifugar 10seg a 10000rpm.
- 7.- Tirar el sobrenadante y agregar 400ul de etanol absoluto.
- 8.- Centrifugar 10seg a 10000rpm.
- 9.- Tirar el sobrenadante y resuspender en 60ul de etanol absoluto.

10.- Colocar 10ul de las partículas en las membranas acarreadoras para biobalística. Dejar secar y proceder con el bombardeo.

### **E. Inoculación de células NT1**

1.- Filtrar las células NT1 de 3 días de crecimiento en papel filtro millipore de 200um.

2.- Colocar las células filtradas en cajas de Petri chicas con medio NT1 osmótico y dejar ahí por 3 horas.

3.- Montar el DNA a disparar en la pistola de bombardeo de microparticulas Dupont modelo PDS1000 de BioRad.

4.- Colocar las células a una distancia aproximada de 10cm.

5.- Hacer el disparo usando membranas de ruptura para una presión de 900psi.

6.- Pasar el papel filtro con las células inoculadas a cajas Petri chicas con medio NT1 (no osmótico).

7.- Incubar las células a 25°C en condiciones de obscuridad por 72 horas.

8.- Recolectar las células raspando suavemente con una espátula y colocarlas en tubos eppendorf de 1.5ml.

9.- Extraer DNA total de las células para su análisis o almacenar a -70°C.

### **F. Extracción de DNA de células NT1**

1.- Congelar las células NT1 contenidas en el microtubo en nitrógeno líquido.

2.- Triturar la muestra usando el homogeneizador o el pistilo.

3.- Agregar 250ul de buffer de extracción (Tris-HCl ph8 0.05M, EDTA ph 8 0.05M, NaCl 0.15M, N-lauril sarcosina 0.5% m/V, SDS 0.5% m/V, 2-mercaptoetanol 0.01M).

4.- Moler y agregar 500ul más de buffer de extracción.

5.- Incubar a temperatura ambiente por 30minutos.

6.- Agregar 750ul de fenol-cloroformo-alcohol isoamílico (25:24:1) y mezclar vigorosamente en un agitador (vortex).

7.- Centrifugar 10min a 13000 rpm y recuperar el sobrenadante.

- 8.- Agregar un volumen de fenol-cloroformo-alcohol isoamílico y mezclar en el vortex.
- 10.- Centrifugar 10min a 13000 rpm y recuperar el sobrenadante.
- 11.- Agregar 5ul de RNAasa (2ug/ul), e incubar 30min.
- 12.- Agregar un volumen de isopropanol frio e incubar al menos 20min a -20°C.
- 13.- Centrifugar 10min a 13000rpm.
- 14.- Eliminar el sobrenadante.
- 15.- Agregar 750ul de etanol al 70 % y mezclar por inversión suavemente.
- 16.- Centrifugar 3min a 13000rpm y eliminar el sobrenadante.
- 17.- Repetir los pasos 15 y 16.
- 18.- Dejar secar la pastilla y resuspender en 30 a 50ul de agua Milli-Q estéril.

## **G. Extracción de DNA de plantas**

- 1.- Tomar aproximadamente 50mg de tejido en un microtubo de 1.5ml (eppendorf) y colocarlos en nitrógeno líquido.
- 2.- Moler la muestra usando un pistilo estéril.
- 3.- Agregar 350ul de buffer de extracción y moler nuevamente.
- 4.- Agregar 350ul de buffer de extracción (Tris-HCl pH8 0.05M, EDTA pH 8 0.05M, NaCl 0.15M, N-lauril sarcosina 0.5% m/V, SDS 0.5% m/V, 2-mercaptoetanol 0.01M).
- 5.- Incubar a temperatura ambiente por 30minutos.
- 6.- Agregar 750ul de fenol-cloroformo-alcohol isoamílico (25:24:1) y mezclar vigorosamente en un agitador (Vortex).
- 7.- Centrifugar 10min a 13000 rpm y recuperar el sobrenadante.
- 8.- Agregar 5ul de RNAasa (2ug/ul), e incubar 30min.
- 9.- Agregar un volumen de isopropanol frio e incubar al menos 20min a -20°C.
- 10.- Centrifugar 10min a 13000rpm.
- 11.- Eliminar el sobrenadante.
- 12.- Agregar 750ul de etanol al 70 % y mezclar por inversión suavemente.



13.- Centrifugar 3min a 13000rpm y eliminar el sobrenadante.

14.- Dejar secar la pastilla y resuspender en 30 a 50ul de agua Milli-Q estéril.