



CENTRO DE INVESTIGACIONES BIOLÓGICAS
DEL NOROESTE, S.C.

Programa de Estudios de Posgrado

PREDICCIÓN *in silico* DE miRNAs DE *Solanum lycopersicum* COMO MECANISMO DE DEFENSA (RNAi) ANTE DIFERENTES BEGOMOVIRUS REPORTADOS EN BCS.

T E S I S

Que para obtener el grado de

Maestro en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales
(Orientación en Biotecnología)

P r e s e n t a

Omar Aranda López

La Paz, Baja California Sur, mayo de 2020.

ACTA DE LIBERACIÓN DE TESIS

En la Ciudad de La Paz, B. C. S., siendo las 10:00 horas del día 20 del Mes de mayo del 2020, se procedió por los abajo firmantes, miembros de la Comisión Revisora de Tesis avalada por la Dirección de Estudios de Posgrado y Formación de Recursos Humanos del Centro de Investigaciones Biológicas del Noroeste, S. C., a liberar la Tesis de Grado titulada:

"PREDICCIÓN *in silico* DE miRNAs DE *Solanum lycopersicum* COMO MECANISMO DE DEFENSA (RNAi) ANTE DIFERENTES BEGOMOVIRUS REPORTADOS EN BCS"

Presentada por el alumno:

OMAR ARANDA LÓPEZ

Aspirante al Grado de MAESTRO EN CIENCIAS EN EL USO, MANEJO Y PRESERVACIÓN DE LOS RECURSOS NATURALES CON ORIENTACIÓN EN BIOTECNOLOGÍA

Después de intercambiar opiniones los miembros de la Comisión manifestaron su **APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA



Dr. Ramón Jaime Holguín Peña
Director de Tesis



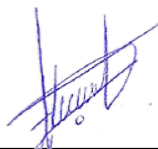
Dra. Diana Medina Hernández
Co-Tutora



Dr. Aaron Barraza Celis
Co-Tutor



Dr. Yair Cárdenas Conejo
Co-Tutor



Dra. Gracia Alicia Gómez Anduro,
Directora de Estudios de Posgrado y
Formación de Recursos Humanos.

Conformación de Comités

Comité Tutorial

Dr. Ramón Jaime Holguín Peña
Director de Tesis
Centro de Investigaciones Biológicas del Noroeste S. C.

Dra. Diana Medina Hernández
Co-Tutora de Tesis
Centro de Investigaciones Biológicas del Noroeste S. C.

Dr. Aarón Barraza Celis
Co-Tutor de Tesis
Centro de Investigaciones Biológicas del Noroeste S. C.

Dr. Yair Cárdenas Conejo
Co-Tutor de Tesis
Universidad de Colima

Comité Revisor de Tesis

Dr. Ramón Jaime Holguín Peña
Dra. Diana Medina Hernández
Dr. Aarón Barraza Celis
Dr. Yair Cárdenas Conejo

Jurado de Examen

Dr. Ramón Jaime Holguín Peña
Dra. Diana Medina Hernández
Dr. Aarón Barraza Celis
Dr. Yair Cárdenas Conejo

Suplente

Dr. Yair Cárdenas Conejo

Resumen

El RNAi o silenciamiento génico, es un sistema que se ha documentado para el control de infecciones begomovirales en cultivos de interés agrícola; pero su nivel de eficiencia es determinado por el uso de construcciones genéticas, las cuales contienen secuencias blanco al begomovirus, por lo que, el diseño de estas construcciones son de gran importancia; dentro del amplio número de investigaciones para este tema, el análisis in silico de los miRNAs tiene un rol importante, donde se agrupan criterios de selección para generar estas construcciones in silico. El objetivo de este estudio, es una predicción in silico de los miRNAs en *Solanum lycopersicum* que pueden unirse al genoma de los begomovirus ToChLPV, TYLCV, SLCuV y PepGMV reportados en BCS, con el potencial de utilizarse posteriormente en un sistema de silenciamiento génico, contra infecciones begomovirales mixtas. Con los resultados de este análisis in silico, se identificaron diferentes miRNAs que pueden unirse al genoma de los cuatro begomovirus mencionados, de los cuales algunos tienen como objetivo múltiples genes/región. El mayor número de sitios blanco para los miRNAs se encontró en los genes C1, C4 y la región intergénica (IR). El análisis de la similitud mostró los valores más bajos de identidad para las mismas regiones (C1, C4 e IR). Mientras que el análisis de diversidad y presión de selección mostraron el valor más alto de diversidad (π) para C4, C1 y V1. El gen C4 mostró estar bajo un mayor efecto de selección positiva, mientras C1 y V1 tienen efecto de presión negativa. Lo que se reflejó en una mayor acumulación de mutaciones sinónimas para V1/C1 y no sinónimas para C4. Estos datos refuerzan la idea de que los miRNA candidatos juegan un importante papel en el silenciamiento de estos begomovirus de manera natural y en infecciones mixtas. Con este análisis de predicción in silico se propone como candidatos a mir395 y mir9477 para el diseño de construcciones que activen el silenciamiento génico, porque que tienen sitios objetivos en TYLCV, SLCuV, PepGMV, ToChLPV, y sobretodo porque se ha documentado que estos miRNAs no tienen efectos secundarios para el hospedero.

Palabras clave: miRNAs, BGVs, infección mixta, RNAi, ToChLPV, SLCuV, PepGMV, TYLCV, in silico.

V.o.B.o.



Dr. Ramón Jaime Holguín Peña
Director de tesis

Summary

The RNAi or gene silencing, is a system that has been documented for the control of begomoviral infections in crops of agricultural interest; but its level of efficiency is determined by the use of genetic constructions, which contain target sequences to the begomovirus, so that, the design of these constructions is of great importance. Inside the wide number of research for this topic, the *in silico* analysis of miRNAs plays an important role, where are grouped selection criteria to generate these *in silico* constructions. The aims of this study, is an *in silico* prediction of the miRNAs in *Solanum lycopersicum* that can bind to the genome of the begomoviruses ToChLPV, TYLCV, SLCuV and PepGMV discribed for BCS, with the potential to be used later in a gene silencing system, against mixed infections of begomovirals. With the results of this *in silico* analysis, were identified different miRNAs that can bind to the genome of the four mentioned begomoviruses, some of which target multiple genes or region. The largest number of targeted sites for miRNAs was found in genes C1, C4 and the intergenic region (IR). The similarity analysis showed the lowest identity values for the same regions (C1, C4 and IR). While the analysis of diversity and selection pressure showed the highest value of diversity (π) for C4, C1 and V1. The C4 gene showed to be under a greater positive selection effect, while C1 and V1 have an effect of negative pressure. This was reflected in a greater accumulation of synonymous mutations for V1/C1 and not synonymous for C4. These data reinforce the idea that candidates miRNAs play an important role in the silencing these begomoviruses naturally and in mixed infections. With this *in silico* prediction analysis, mir395 and mir9477 are proposed as candidates to the design of constructions that activate gene silencing, because they have targeted sites in TYLCV, SLCuV, PepGMV, ToChLPV, and above all because been documented that these miRNAs have not secondary effects for the host.

Key words: miRNAs, BGVs, infection mixed, RNAi, ToChLPV, SLCuV, PepGMV, TYLCV, *in silico*.

Vo.Bo.



Dr. Ramón Jaime Holguín Peña
Director de tesis

Dedicatoria

A Dedico este documento, que refleja el esfuerzo de más de dos años de arduo trabajo y el cual no habría sido posible sin el amor incondicional de mi amada madre Honorina López Cruz y de mi abuela Alicia Cruz; gracias por su gran labor y esfuerzo en mi formación como ser humano y profesionista; también se lo dedico a mis hermanas Ana Laura; Elaid, Marisol Karla, a mi madrina Cristina López y a Yolanda Vallarino. A todas ellas quiero agradecer su apoyo en los momentos que he necesitado de su ayuda, han estado ahí a mi lado, confiando en mí cada día.

Agradecimientos

Al Centro de Investigaciones Biológicas del Noroeste, S. C. (CIBNOR), por haber sido la institución receptora de mis estudios de posgrado y por facilitarme el acceso y uso de sus instalaciones.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo brindado al otorgarme la beca No. 636346.

Quiero agradecer al Dr. Ramón Jaime Holguín Peña por los consejos y la paciencia; la Dra. Diana Medina Hernández por alentarme a no desistir para lograr mis objetivos; al Dr. Aarón Barraza Celis por su guía en la edición y redacción de mi documento y al Dr. Yair Cárdenas Conejo por compartir sus comentarios. Todos participaron con sus conocimientos, esfuerzo y tiempo, durante el proceso de mi formación y la culminación de este trabajo de investigación.

Al M. en C. Mario Arce Montoya, por brindarme sus conocimientos y sobre todo su invaluable amistad en los momentos más complicados.

Al técnico, M. en C. Julio Antonio Hernández González por las amenas charlas. A los técnicos del Laboratorio de Fitopatología M. en D. Martín Guadalupe de Jesús Aguilar García y la Hidrobióloga María Sofía Ramos Galván.

A mis amigos y compañeros del Laboratorio de Fitopatología, del CIBNOR: Jesús Ricardo Hernández Barrera (invaluable apoyo al final de esta larga carrera); Máyela Vargas Salinas, Diana Rojas Sánchez, Roció Solís, Jorge Alberto Alejandro Rosas; con quienes compartí momentos alegres y tristes, pero siempre al hombro de gigantes.

A la Biol. Patricia Hinojosa Baltazar, del Laboratorio de Fisiología Comparada por todo el apoyo en su laboratorio.

A la Directora de Posgrado La Dra. Gracia Alicia Gómez Anduro, por el apoyo incondicional; al staff de posgrado por la atención, paciencia, apoyo y facilidades brindadas: Lic. Osvelia Ibarra Morales, Tania Verónica Núñez Valdez, Lic. Leticia González Rubio Rivera, Lic. Claudia Elizabeth Olachea León, Adriana Teresa Franzoni García, Lic. Horacio Sandoval Gómez y a Lic. José Manuel Melero Astorga.

Al personal de la Biblioteca del CIBNOR, Lic. Ana María Talamantes Cota, Lic. María Esther Ojeda Castro, Susana Luna García, Elizabeth Guadalupe Sánchez Vázquez, por todo el apoyo y amabilidad.

A mis compañeros de generación de posgrado y de la red de apoyo, en los buenos y en los malos momentos de la vida, son un pilar incondicional de amistad y compromiso.

Contenido

Resumen	i
Summary	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	vi
Lista de figuras	viii
Lista de tablas	ix
1. INTRODUCCIÓN	1
1.1 Importancia del cultivo del tomate.....	1
1.2 Las enfermedades begomovirales como limitantes en la producción..	1
1.3 Las infecciones mixtas.....	2
1.4 Los begomovirus.....	3
1.5 El ciclo de infección de los BGVs.....	5
1.6 Silenciamiento génico.....	6
1.7 Clasificación de los sRNAs.....	7
1.8 Diseño de construcciones con tecnología RNAi.....	10
2. ANTECEDENTES	13
3. JUSTIFICACIÓN	19
4. HIPÓTESIS	20
5. OBJETIVOS	21
5.1 Objetivo general.....	21
5.2 Objetivos específicos.....	21
6.MATERIALES Y MÉTODOS	22
6.1 Selección de especies, aislamientos y variantes de los BGVs.....	22
6.2 Comparación nucleotídica para búsqueda de divergencia.....	24
6.3 Diversidad nucleotídica.....	25
6.4 Presión de selección.....	25
6.5 Análisis filogenético.....	25
6.6 Predicción de sitios semilla.....	26
6.7 Especificidad de los miRNAs candidatos con los genes objetivo en los BGVs y de los genes del hospedero.....	26
7.RESULTADOS	28
7.1 Comparación genómica de especies caracterizadas en BCS con los aislados reconocidos por el ICTV.....	28
7.2 Homología de las proteínas de los Begomovirus aislados de BCS ...	32
7.3 Diversidad nucleotídica	32
7.4 Presión de selección.....	33
7.5 Relación filogenética entre ToChLPV y BGVs presentes en BCS.....	40
7.6 Predicción de sitios semilla.....	43
7.7 Análisis por miRNA.....	45
7.8 Análisis por cada secuencia viral	48
7.9 Eficiencia in silico de los miRNAs con potencial de silenciamiento.....	50
8. DISCUSIÓN	53

9. CONCLUSIONES	63
10. LITERATURA CITADA	64

Lista de figuras

Figura 1. Estructura genómica de los BGVs, adaptado de ICTV.....	4
Figura 2. Un esquema simplificado del ciclo de replicación del DNA de los geminivirus... 5	5
Figura 3. Gráficas por especie de similitud de BGVs de referencia comparadas con aislamientos de BCS.....	31
Figura 4. Distribución de sustituciones sinónimas y no sinónimas en el gen C1.....	34
Figura 5. Distribución de sustituciones sinónimas y no sinónimas en el gen C2.....	35
Figura 6. Distribución de sustituciones sinónimas y no sinónimas en el gen C3.....	36
Figura 7. Distribución de sustituciones sinónimas y no sinónimas en el gen V1.....	37
Figura 8. Distribución de sustituciones sinónimas y no sinónimas en el gen C4.....	38
Figura 9. Distribución de sustituciones sinónimas y no sinónimas en el gen V2.....	39
Figura 10. Dendograma construido a partir de los aislamientos reportados en América de los virus que se encuentran en el estado en la historia evolutiva.....	42
Figura 11. Representación global del número de sitios reconocidos por los miRNAs en SLCuV, ToChLPV, PepGMV y TYLCV	46
Figura 12. Representación gráfica del número de aislamientos reconocidos por los miRNAs por cada una de los virus	48
Figura 13. Número sitios blanco reconocido por los miRNA en los diferentes componentes dentro de los diferentes aislamientos de los BGVs.....	49
Figura 14. Representación del número de sitios blanco por gen y por begomovirus	50

Lista de tablas

Tabla I. Secuencias seleccionadas de los begomovirus usados en este estudio, corresponde a aquellas secuencias reportadas como especies por el Comité Internacional de Taxonomía de Virus (ICTV).....	23
Tabla II. Porcentaje de identidad de los componentes de los Begomovirus ToChLV, TYLCV, SLCuV, PepGMV (identidad nucleotídica y similitud de Proteína	29
Tabla III. Tabla de diversidad nucleotídica del genoma y genes por especie.....	33
Tabla IV. Probabilidad de rechazar la hipótesis de presión de selección.....	40
Tabla V. Genes de BGVs reconocidos por miRNAs.....	43
Tabla VI. Comparación de genes blanco en el hospedero y BGVs para miRNA con los genes blanco en una sola especie	51
Tabla VII. Comparación de genes blanco en hospedero y virus para miRNA con múltiples especies blanco de BGVs	52

1. INTRODUCCIÓN

1.1 Importancia del cultivo del tomate

El tomate (*Solanum lycopersicum* L.) es uno de los cultivos más importantes a nivel mundial debido a la demanda de consumo, por su aportación alimenticia, principalmente vitaminas y antioxidantes (Cruz-Bojórquez *et al.*, 2013). En el 2018 se sembraron 4.7 millones de hectáreas (Ha) en el mundo con una producción de 182 millones de toneladas (Ton), que representan un valor comercial de aproximadamente 87 mil millones de dólares (FAOSTAT, 2018).

En México, el tomate se cultiva en dos zonas climáticas divididas por el Trópico de Cáncer: la zona templada que comprende al norte y noroeste del país y la zona tropical que comprende al centro y sur del país. En el 2018 la producción total reportada para México fue de 4.5 millones de Ton según la FAO (FAOSTAT, 2018). En el 2018 Baja California Sur (BCS) fue el sexto estado con la mayor producción (Ton) de este cultivo (SIAP, 2018) esto con una producción de 157,879 Ton de tomate, además este cultivo ha sido, económicamente, el más importante para el estado en los últimos años (SIAP, 2016, SIAP, 2017 y SIAP, 2018).

1.2 Las enfermedades begomovirales como limitantes en la producción

Además de los problemas inherentes a las condiciones ambientales en la producción de este cultivo y los relacionados a los sistemas de comercialización, uno de los problemas más recurrentes de este cultivo son la presencia de plagas y enfermedades (Camejo y Torres, 2000; Darré *et al.*, 2019). Dentro de la amplia diversidad de problemas fitopatológicos aquellos que ocasionados por virus tienen un nicho especial, pues las medidas tradicionales para la prevención y el manejo no se enfocan hacia los virus y en dado caso resultan en medidas indirectas de control. En los problemas relacionados con este tipo de enfermedades se pueden involucrar tres componentes epidemiológicos de suma importancia: a) los agentes causales, b) los insectos vectores y c) el amplio rango de hospederos (Brown *et al.*, 2015; Inoue-Nagata *et al.*, 2016). Uno de los modelos más estudiados en la

región noroeste son los virus de DNA de cadena sencilla del género *Begomovirus*, que son transmitidos por el complejo de la mosquita blanca (*Bemisia tabaci* Gennadius.) (Bañuelos-Hernández *et al* 2012; Holguín-Peña *et al.*, 2004; Holguín-Peña *et al.*, 2010; Medina-Hernández *et al.*, 2019). Los begomovirus (BGVs) afectan al cultivo principalmente por un daño generalizado a la planta, la cual presenta una diversidad de síntomas que terminan disminuyendo la producción y la calidad del fruto (Sharma *et al.*, 2019). Los síntomas asociados a este tipo de enfermedades van desde la presencia de clorosis (mosaicos), arrugamiento de las hojas hasta síntomas más severos como un enchinamiento drástico y enanismo que puede ocasionar pérdidas totales, especialmente cuando la infección es mixta (Inoue-Nagata *et al.*, 2016; Renteria-Canett *et al.*, 2011). En los últimos 20 años no se han podido establecer de manera apropiada sistemas de control para el insecto vector, a pesar de la importancia de los BGVs como patógenos a nivel mundial (Luan *et al.*, 2014). Al igual que en el resto del mundo en México estos patógenos han ocasionado pérdidas económicas de diferentes cultivos, entre los que se encuentra el tomate, más recientemente se han caracterizado parte de la diversidad de BGVs presentes en el Noroeste de México lo que nos da una pauta para su control (Holguín *et al.*, 2004; Holguín *et al.*, 2005; Holguín *et al.*, 2006, Medina *et al.*, 2019; Rodríguez-Negrete *et al.*, 2019).

1.3 Las infecciones mixtas

Los términos infección mixta o co-infección hacen referencia a la presencia de más de un patógeno en un organismo. Dado lo general de esta definición el término se usa para describir las posibles infecciones causadas por la combinación de dos o más patógenos los cuales pueden ser organismos de diferentes taxones o bien organismos del mismo taxón (Abrahamian *et al.*, 2015; Khalid *et al.*, 2017; Lebsky *et al.*, 2011; Reveles *et al.*, 2012). Las infecciones mixtas ocasionadas por más de un begomovirus son el caso de estudio de este trabajo, por lo tanto, es el sentido que se le dará ha dicho término a lo largo de este documento. Este interés nace en el hecho que las infecciones begomovirales frecuentemente se presentan como

infección mixta en la naturaleza, permitiendo que algunos BGVs interactúen de manera sinérgica, lo que se refleja en un aumento de la severidad de los síntomas durante una infección mixta (Méndez *et al.*, 2003; Rodríguez-Negrete *et al.*, 2019; Vaca-Vaca *et al.*, 2012). Del mismo modo se ha reportado que este fenómeno tiene un rol primordial en la evolución de estos virus, ya que durante este tipo de infecciones tienen lugar los fenómenos de recombinación y con ello el surgimiento de nuevas variantes, especies o quimeras (Hosseinzadeh *et al.*, 2014).

1.4 Los begomovirus

El género *Begomovirus* (BGVs) incluye a 596 especies, lo que lo hace uno de los géneros más diversos de la familia *Geminiviridae* (ICTV, 2020). Los BGVs son agentes virales de DNA de cadena sencilla (ssDNA) con un tamaño de 22x38 nm, con una cápside de doble icosaedro truncado que está formada por 22 capsómeros (110 subunidades) (Zhang *et al.*, 2001). Los virus de este género pueden ser monopartita cuando cuenta con un solo componente genómico de 2.5kb (conocido como componente A o DNA-A) o puede ser bipartita con un genoma completo de 5.2kb cuando presenta dos componentes (DNA-A y DNA-B) (Hanley-Bowdoin *et al.*, 2000).

Los BGVs monopartitas y bipartitas comparten una región intergénica (RI) la cual es una región con estructura tallo-asa donde se encuentra el origen de replicación viral (TAATATT*AC) así como los promotores y los elementos de regulación de la transcripción de sus genes, esta región también está presente en el genoma B de lo bipartitas (Borah *et al.*, 2016); en el caso de los BGVs monopartitas el genoma A contiene los genes del movimiento (V2 o MP), mientras que en las bipartitas los genes del movimiento (BC1 o MP) y los que codifican las proteínas del transporte nuclear (BV1 o NSP) están en el genoma B, para ambos tipos de BGVs, se localiza en el genoma A los genes (V1 o CP) que codifican para las proteínas de la cápside, los genes C1 que codifican para la proteína asociada a la replicación (C1 o Rep), la proteína activadora de la transcripción (C2 o TrAp), la proteína potenciadora de la

transcripción (C3 o REn) y la implicada en el control del ciclo celular (C4). (Ascencio-Ibáñez *et al.*, 1999; Borah *et al.*, 2016) (Fig. 1).

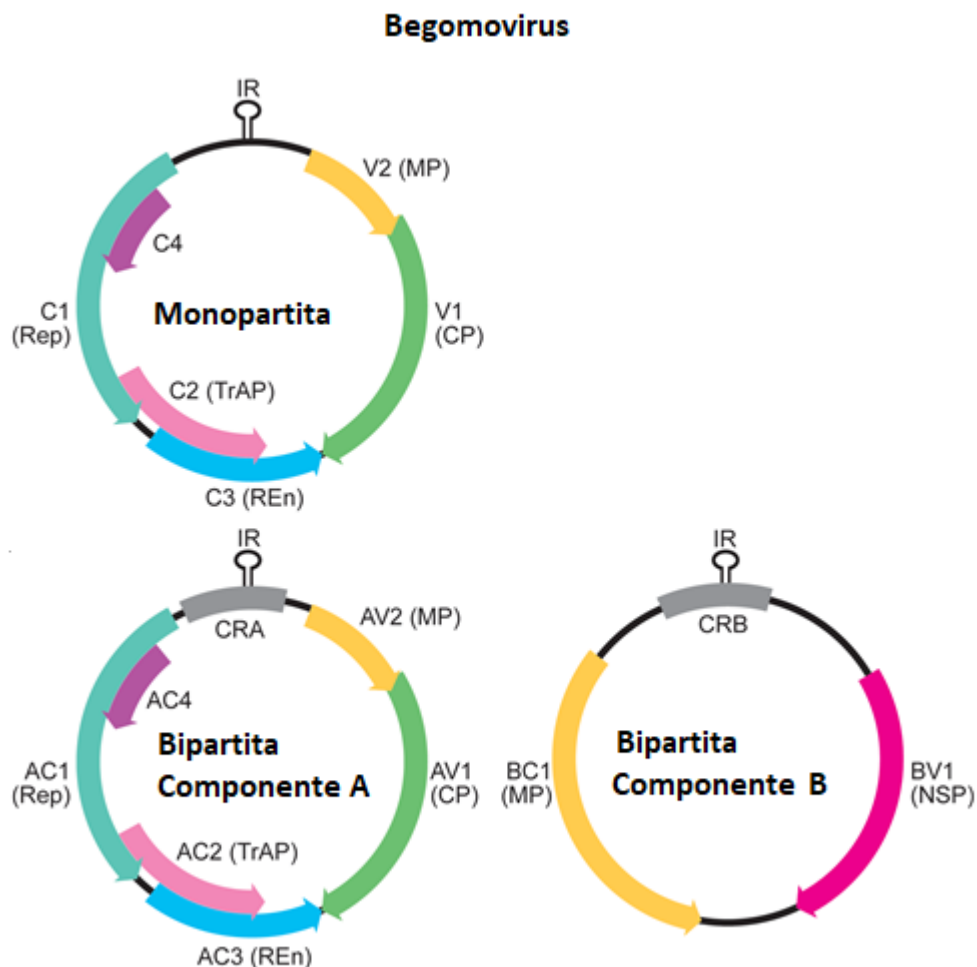


Figura 1. Estructura genómica de los BGVs, adaptado de ICTV (<https://talk.ictvonline.org/>). En la imagen superior, se muestra el genoma de los BGVs monopartita con los cinco ORFs comunes (C1, C2, C3, C4, V1), el componente V2 presente en los virus de nuevo mundo, así como la IR. En la imagen inferior, se muestran ambos componentes del genoma de un BGVs bipartita, el componente A muestra una estructura similar al de los BGVs monopartitas mientras que el componente B presenta una IR dentro de la región común y dos proteínas asociadas al movimiento intracelular (NSP) e intercelular (MP).

1.5 El ciclo de infección de los BGVs

El ciclo infectivo inicia cuando el vector inocula el virón (DNA viral sin cápsula proteica) a la planta hospedera, ingresa a la célula vegetal y se moviliza al núcleo de estas (Pradhan *et al.*, 2017). Una vez dentro del núcleo, a partir de DNA del viral se genera una doble hebra de DNA (dsDNA) a la cual se le unen los nucleosomas para darle estabilidad (Hanley-Bowdoin *et al.*, 2000). Posteriormente, los genes son transcritos por la RNA polimerasa II (RNAPolII) y finalmente traducidos a proteínas funcionales por los ribosomas. Una vez sintetizadas las proteínas RE_n y Rep ocurre la replicación del genoma mediante un mecanismo denominado replicación por círculo rodante (ACR) (Gutiérrez *et al.*, 2002). El ciclo de transmisión se completa luego de que el virus sale de la célula vegetal y entra al floema de la planta transportándose a otros tejidos por medio de las proteínas del movimiento (MP), para que este sea adquirido por otro insecto vector (Pradhan *et al.*, 2017) (Fig. 2).

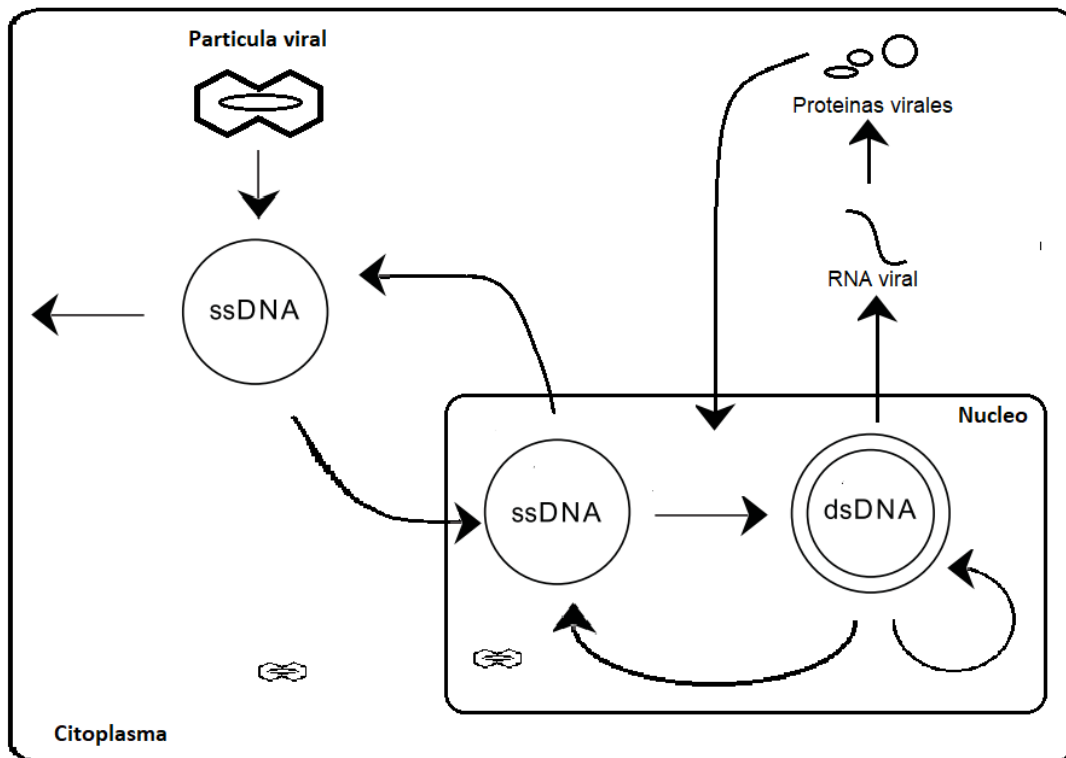


Figura 2. Un esquema simplificado del ciclo de replicación del DNA de los

geminivirus. Después de la liberación de la partícula del virus, el ssDNA genómico se convierte primero en un intermedio de replicación de dsDNA, que también es la plantilla transcripcionalmente activa. Este dsDNA entra en un proceso de replicación de círculo rodante que da lugar a la amplificación de las moléculas de ssDNA genómico, que se incorporan al grupo de replicación, se encapsulan y se transportan a las células vecinas.

1.6 Silenciamiento génico

Los RNA no codificantes (ncRNA) son un grupo de ácidos nucleicos en el que se engloban todas las moléculas de RNA que no se transcriben a proteínas pero que participan en diferentes procesos celulares (Ricaño-Rodríguez *et al.*, 2014). Los RNAs pequeños (sRNAs) son moléculas de RNA de menos de 200nt las cuales tienen diferentes funciones en la regulación génica en un sistema denominado silenciamiento génico por RNA de interferencia (RNAi) (Won *et al.*, 2014). Existen dos mecanismos de regulación/silenciamiento genético en los que participan los sRNAs, estos son conocidos como: silenciamiento génico transcripcional (TGS) y silenciamiento génico post-transcripcional (PTGS) dependiendo en la etapa que se afectan los procesos de transcripción o traducción (García *et al.*, 2016; Ricaño-Rodríguez *et al.*, 2014). El PTGS se lleva a cabo mediante la unión de un sRNAs con un RNA mensajero (mRNAs) para la inhibición de la traducción o degradación del mRNA objetivo, lo que depende de la complementariedad del sRNAs con el objetivo (Brodersen *et al.*, 2008). Por otro lado, en la regulación tipo TGS, los sRNAs se unen a los mRNA durante la transcripción para degradar el mRNA blanco o induciendo la metilación del DNA y derivando en la consecuente inhibición de la síntesis de un mRNA (Noriega *et al.*, 2016).

Estos mecanismos juegan un papel importante en células eucariotas, como en la regulación del desarrollo celular, como respuesta frente diferentes tipos de estrés abiótico y además se ha demostrado que en plantas este mecanismo juega un papel importante en la defensa contra infecciones virales (Raja *et al.*, 2010). Pese a que el silenciamiento génico y el silenciamiento contra virus comparten muchos mecanismos con la regulación de la expresión celular, la expresión de cada sRNAs

puede ser específica para tejido, etapas del desarrollo de la planta o como respuesta a ciertos estímulos (D' Ario *et al.*, 2017; Ricaño-Rodríguez *et al.*, 2014).

1.7 Clasificación de los sRNAs

Existen varias clases de sRNAs los cuales tienen diferentes precursores (orígenes biológicos) y biogénesis (mecanismo de síntesis). Sin embargo, en la biogénesis de todos ellos participan principalmente tres familias de proteínas. Por ejemplo; i) las enzimas endorribonucleasas tipo III o enzimas DICER (DCL, Dicer-Like Proteins), que son las encargadas de cortar los precursores de RNA bicatenarios o RNAs de doble hebra (dsRNAs, double-stranded RNAs) en fragmentos más pequeños 19 a 30nt o sRNAs funcionales, ii) las RNA polimerasas dependientes de RNA (RDR, RNAPol-Dependent of RNA), que son las encargadas de generar moléculas de dsRNAs a partir dos o más precursores de RNAs monocatenario o RNAs de una sola hebra (ssRNAs, single-stranded RNAs) unidos por complementariedad y iii) las proteínas argonautas (AGO) que son las encargadas de dirigir la unión del sRNA con su objetivo de RNA y degradación de este último (Borges *et al.*, 2015; D' Ario *et al.*, 2017).

De acuerdo a su estructura, tamaño y proceso de biogénesis los sRNAs endógenos se clasifican como microRNAs (miRNAs, micro RNAs) y RNAs pequeños de interferencia (siRNAs, small interfering RNA). Estas dos familias tienen muchas similitudes en su estructura y en su biogénesis, pues ambos son dsRNAs de pequeño tamaño y muchas enzimas participan en la biogénesis de ambas familias. Sin embargo, estas familias presentan sutiles diferencias, por ejemplo: hasta ahora se conoce que, de manera directa, los miRNAs solo participan en los mecanismos de PTGS, los miRNAs maduros tienen un rango de longitud de 20-24nt (Carthew *et al.*, 2009; Kamthan *et al.*, 2015) y mientras que los siRNAs tienen un tamaño de 19-25nt (Watanabe *et al.*, 2011).

Los miRNA solo tienen un tipo de precursor, que son codificados por genes de miRNAs, por lo que parte de su proceso de biogénesis tiene lugar en el núcleo. Es aquí donde se lleva a cabo la transcripción de genes que expresan para ssRNAs,

que posteriormente entran a un proceso de maduración para dar origen a los miRNA primarios (pri-miRNAs). Durante este proceso cada ssRNA adquiere una estructura secundaria de tallo asa, que consiste en un dsRNA que se une por una región de horquilla y un extremo con dos cadenas de RNA no complementarias. El extremo no complementario de los pri-miRNA es retirado por un complejo que contiene una enzima DCL1 y otras proteínas como: HYL1 (HYPOPLASTIC LEAVES1), SE (SERRATE), DDL (DAWDLE), CPL1 (C-TERMINAL DOMAIN PHOSPHATASE-LIKE), MOS2 y TGH (TOUGH), lo que da lugar a los miRNAs precursores (pre-miRNA) los cuales conservan en su estructura la región de horquilla. Posteriormente, los pre-miRNA es procesado la enzima DCL1 perdiendo esta horquilla lo que da entonces lugar al miRNA maduro (Won *et al.*, 2014). Una vez maduro este miRNA es metilado en el extremo 3' por la metiltransferasa (HEN1) esta enzima actúa en ambas cadenas sentido/antisentido (miRNA/miRNA*). Finalmente, el miRNA metilado es transportado al citoplasma gracias a la acción de una proteína exportina HST (HASTY), donde finalmente los miRNA maduros son procesado por la enzima DICER y AGO. Ambas proteínas están asociadas con el complejo enzimático de silenciamiento inducido por RNA (RISC, RNA-Induced Silencing Complex) para dar inicio a procesos de silenciamiento (D'Ario *et al.*, 2017; Ricaño-Rodríguez *et al.*, 2014).

Los siRNAs tienen precursores de ssRNA de diferentes orígenes, los cuales por diferentes vías dan origen a un dsRNA. Por lo que exceptuando por la síntesis de algunos precursores todo el proceso de biogénesis de los siRNAs tiene lugar en el citoplasma. Es decir los a diferencia de los miRNAs los siRNAs no tienen que ser exportados al citoplasma, pero los dsRNAs precursores son igualmente metilados por la enzima HEN1 y procesados por DICER y AGO en el complejo RISC, el cual llevara a cabo la degradación del RNA mensajero (Carthew *et al.*, 2009; Guleria *et al.*, 2011; Kamthan *et al.* 2015; Won *et al.*, 2014).

En plantas se reconocen dos tipos de siRNAs: los siRNAs primarios y los siRNAs secundarios. Los siRNAs primarios se subdividen en siRNAs de horquilla (hp-siRNAs, hairpin siRNAs), siRNAs heterocromáticos (hetsiRNAs, heterochromatic

siRNAs) y siRNAs anti-sentido natural (natsiRNAs, natural antisense siRNAs), estos últimos pueden tener una estructura cis o trans dependiendo del alineamiento de los precursores. Los siRNAs secundarios pueden subdividirse en siRNAs de fase secundaria (phasiRNAs, siglas del inglés phased secondary small interfering RNAs), siRNAs de acción en trans (trans-activating siRNA o tasiRNAs) y siRNAs activado epigenéticamente (easiRNAs, del inglés epigenetically activated siRNAs) (Borges *et al.*, 2015; Xie y Yu, 2015).

Los hp-siRNAs comúnmente poseen una longitud de entre 21-24nt y participan únicamente en el proceso de PTGS. Estos provienen de un precursor que posee una estructura similar a la de los miRNA es decir una estructura de dsRNA. Sus precursores son moléculas de ssRNA que presentan de dos regiones complementarias dispuestas de forma invertida que son separadas por una región no complementaria (estructura tallo-asa). Al igual que los miRNAs son sintetizados la RNA POLII, sin embargo, la longitud de la doble hebra es mayor a la de un miRNA por lo que un hp-siRNA puede dar lugar a varios siRNAs gracias a la acción de las enzimas DLC1, DCL2, DCL3 y DCL4 (Borges *et al.*, 2015; Guleria *et al.*, 2011).

Los hetsiRNAs tienen una longitud de 24nt tienen como precursor dsRNAs los que se originan por la acción de la enzima RDR2 (RNA-DEPENDENT RNA POLYMERASE II), que usa como sustrato ssRNA no codificante expresado por la acción de la RNA POLV a partir de genes específicos para hetsiRNAs. Este dsRNA es procesado por la enzima DCL3 lo que finalmente da lugar a diferentes hetsiRNAs. Posteriormente, los het-siRNAs son reconocidos por la enzima AGO4 y participan en la metilación de DNA mediado por metiltransferasas (Borges *et al.*, 2015; Xie y Yu, 2015).

Los siRNAs secundarios, tienen en común que todas sus vías de síntesis usan como molde hebras de ncRNA o mRNA sintetizadas por la RNA POLII. El proceso de biogénesis inicia por la acción de las enzimas AGO, las cuales une un miRNAs de 22nt o hasta dos miRNA de 21b a estos precursores. Posteriormente inicia la síntesis de dsRNA por la actividad de la RDR6 (RNA-DEPENDENT RNA POLYMERASE 6) y la SGS3 (SUPRESOR OF GENE SILENCINING3); después,

estos dsRNAs son procesados por las enzimas DCL3, DCL2 o DCL4 para dar origen a los siRNAs secundarios de 24nt, 22nt o 21nt respectivamente (Borges *et al.*, 2015; Deng *et al.*, 2018; Fei *et al.*, 2013; Vázquez y Hohn 2012).

Los pha-siRNA tienen como precursores mRNAs que son sometidos al mecanismo de biogénesis de sRNAs secundarios para dar origen a phasiRNAs de 21nt, 22nt y 24nt. Los ta-siRNAs son un caso particular de pha-siRNA, la diferencia entre estos dos radica en su biogénesis. Pues su precursor son ncRNA (genes TAS) con una longitud de 21nt o 22nt (Borges *et al.*, 2015; Fei *et al.*, 2013; Komiyama, 2017).

Los vsiRNAs por su parte son los siRNAs cuyos precursores son transcritos producidos por el virus. Dentro de esta categoría se encuentran los denominados vsiRNAs primarios, que son los siRNAs expresados por genes virales o bien aquellos que se forman por la acción directa de proteínas DCL sobre el dsRNA de origen viral. Pero además se encuentran los vsiRNAs secundarios cuyo origen es similar al de los siRNAs secundarios endógenos. En este caso las proteínas RDR usan como hebra molde un ssRNA de origen viral al cual se le unió un miRNA. (Guleria *et al.*, 2011; Sarkies *et al.*, 2014; Zhu *et al.*, 2012).

1.8 Diseño de construcciones con tecnología RNAi

La tecnología de RNAi busca imitar y aprovechar el proceso de biogénesis y funcionalidad de los sRNAs (Ortiz-Quintero, 2009). Por lo tanto, el diseño de construcciones de RNAi se rige por diversos procesos. Se ha reportado que la correcta unión de los siRNAs y los microRNAs en el complejo RISC depende de la presencia de grupos fosfato en el extremo 5' resultante del mecanismo de biogénesis, este grupo fosfato es crucial para la degradación de los mRNA pues la distancia del grupo fosfato a la cadena guía determina posición del sitio de unión en el mRNA. Además, se conoce que la estabilidad de la estructura secundaria de la unión sRNAs/RNA depende del grado de homología y tipo de sustitución (mismatch) que existe entre ellos, estabilidad que puede cuantificarse mediante la energía libre de Gibbs (ΔG). En este sentido un valor positivo de energía libre de Gibbs implica que no se unirán miRNA/mRNA, mientras que una energía libre de Gibbs negativa

implica la unión de ambos. Por lo tanto, incrementos de la energía libre de Gibbs se interpretan como reducción a la estabilidad mientras que incrementos negativos (reducción) se interpretan como un incremento en la estabilidad de los miRNAs (Dai *et al.*, 2018; Kertesz *et al.*, 2007; Mückstein *et al.*, 2006). Otro factor que se tiene que considerar, para la eficiencia de un sistema de silenciamiento, es el tamaño del sRNAs, pues este determina las posibles rutas en los que puede participar como en PTGS, en TGS o la generación de siRNAs secundarios. Por lo anterior también es importante considerar el tipo de precursor, ya que existen precursores como los hp-siRNAs que dan lugar a varios siRNAs (Medina-Hernández *et al.*, 2013; Yang *et al.*, 2011).

Además del gasto energético en la síntesis de siRNAs, estos procesos derivan en el silenciamiento de múltiples genes afectando negativamente procesos biológicos importantes. Finalmente se debe considerar los promotores involucrados en la transcripción de genes relacionados a la biogénesis de sRNAs. Debido a que la regulación negativa de un sRNAs se verá reflejada en un menor nivel de silenciamiento de su gen objetivo (Ricaño-Rodríguez *et al.*, 2014).

La tecnología de silenciamiento génico se ha hecho más eficiente por el desarrollo de nuevas tecnologías como la de secuenciación masiva de ácidos nucleicos y la generación de bases de datos de miRNAs, mRNAs y DNA, que nos permite predecir los sitios que son posibles objetivos de un miRNA (Ellwanger *et al.*, 2011; Holen, 2006; Hsu *et al.*, 2007). Sin embargo, conforme avanzamos en la comprensión del papel de los sRNAs en la regulación genética de las plantas, se llega a la comprensión que el diseño eficiente de una construcción de silenciamiento génico basado en RNAi debe de considerar la homología que el o los sRNAs generados por la construcción tienen con un gen objetivo, si tiene otros posibles objetivos (objetivos alternos) y sí la interacción de estos sRNAs con los objetivos alternos puede dar lugar a sRNAs secundarios o bien puede derivar en la metilación del DNA genómico (Raja *et al.*, 2008). Estas características pueden considerarse como deseables o no dependiendo del objetivo de nuestra construcción. Por ejemplo, se ha reportado que el reconocimiento de múltiples regiones blanco puede potenciar el

efecto de las construcciones diseñadas para la protección de plantas frente a diferentes virus (Medina *et al.*, 2013; Shweta y Jawaid, 2014).

Si bien, en la actualidad hay más entendimiento sobre los mecanismos de silenciamiento génico como: que enzimas participan, rutas de biogénesis, el papel que juega la estructura en su mecanismo de acción aún desconocemos mucho sobre las interacciones que existen entre diferentes sRNAs durante el proceso de infección viral. Pero sobre todo existe poca información en desarrollar una construcción que pueda ser eficiente en el control de infecciones mixtas begomovirales; por lo que el estudio in silico de estrategias para potenciar la eficiencia de los mecanismos de silenciamiento génico para el control de diferentes begomovirus de una región, es viable ya que existe un amplio número de herramientas bioinformáticas como miRANDA y RNA hybrid y bases de datos como mirBASE (Baig y Kha, 2013; Tousi *et al.*, 2017). Por lo que, en este trabajo, se plantea un análisis in silico de miRNAs de tomate que puedan ser blanco de regiones de los genomas de *Squash Leaf Curl Virus* (SLCuV), *Tomato Chino la Paz Virus* (ToChLPV), *Pepper Golden Mosaic Virus* (PepGMV) y *Tomato Yellow Leaf Curl virus* (TLCV), todos ellos reportados en BCS y que posteriormente estos miRNAs puedan ser evaluados en un sistema de RNAi como mecanismo de control a infecciones mixtas begomovirales.

2. ANTECEDENTES

La comunidad científica ha hecho un gran esfuerzo por entender el papel de los RNAs pequeños en el proceso de infección/defensa durante las interacciones bióticas. En este sentido, a inicios de la primera década del 2000 surgieron un gran número de trabajos que buscaban entender el papel de los sRNAs tanto en el desarrollo de las plantas, como en la defensa viral posteriormente surgieron los primeros estudios relacionados a las interacciones planta/geminivirus por ejemplo Bisaro y colaboradores (2006), documentaron como algunas proteínas de los geminivirus, interactúan de manera negativa en la eficiencia del sistema de silenciamiento génico y a la caracterizar los sRNAs derivados de geminivirus. (Carthew *et al.*, 2009; D'Ário *et al.*, 2017; Li *et al.*, 2017; Rodrigo *et al.*, 2010).

Las interacciones entre miRNAs de origen vegetal y virus del género BGVs fue una investigación de Amin y colaboradores (2011) quien empleando heterólogos de miRNA de *A. thaliana* detectaron a miRNAs cuya acumulación incrementa durante infecciones begomovirales en *Nicotiana benthamiana*. Además, una de sus aportaciones importantes es proponer la posible participación de algunos miRNAs (como el mir159 y el mir160) en el desarrollo de síntomas asociados a infecciones virales. Así, como explicar los mecanismos por los que la acumulación de estos miRNAs afecta el desarrollo de la planta y se resalta el hecho de que es difícil comprender que pese a que los BGVs poseen mecanismos para inhibir las rutas de biogénesis de sRNAs, aun así se induce la acumulación de miRNAs específicos.

El trabajo como el de Yang y colaboradores (2011) reportan que durante el proceso de infección (BGV/betasatelite) se induce la expresión de vsiRNAs y siRNAs en plantas de *N. benthamiana*. Para la identificación de los siRNAs se llevó a cabo la secuenciación de siguiente generación en la plataforma Solexa (actualmente Illumina). En este estudio se determinó que la mayoría de los sRNAs generados son vsiRNAs y siRNAs de 22 nt, y se sugiere que estos podrían derivar del mecanismo biogénesis de siRNAs secundarios y para ser utilizados por la planta como sistema de defensa frente a los geminivirus. Además, infieren que la estructura de la mayoría

los siRNAs de 21 nt encontrados en este estudio son procesados por AGO1, por lo que a su vez, se asocian a la defensa viral. Finalmente, en este reporte se sugiere que los vsiRNAs y siRNAs 24 nt podrían ser procesados por AGO4 y podrían estar asociados a la metilación del DNA viral o vegetal.

Posteriormente Miozzi y colaboradores (2013) utilizando la plataforma SOLEXA/Illumina secuenciaron los sRNAs de plantas de tomate y *N. benthamiana* que fueron infectados por *Tomato yellow leaf curl Sardinia virus* (TYLCSV), con su trabajo de análisis bioinformáticos encuentran la posible interacción entre los sRNAs derivados de los hospederos y los vsiRNAs. Aunque parte de estos resultados fueron validados por hibridación, sugieren la necesidad de más trabajos que profundicen en el tema para caracterizar el papel de estas moléculas en el proceso de infección de TYLCSV.

Otro trabajo importante es el de Rogans y colaboradores (2016), quienes usaron plataforma de illumina HiSeq2000 para secuenciar los vsiRNAs de plantas de tapioca (*Manihot esculenta* Crantz) infectada por *South Africa cassava mosaic virus* (SACMV). En este estudio ellos encontraron que durante el proceso de infección de este begomovirus existen patrones diferentes de expresión de vsiRNAs en plantas susceptibles respecto a plantas resistentes a begomovirus aunque ellos resaltan el papel que tiene el tiempo o la etapa de infección en estos patrones de expresión. Pese a que en este trabajo no se consideran el papel de los miRNAs ni siRNAs de origen endógeno (hospedero), observan que las plantas susceptibles permiten la replicación de ADN viral por lo que sugieren que la acumulación de vsiRNAs en esta variedad se debe a un ineficiente silenciamiento por la maquinaria de la planta. Aunque también mencionan que estos resultados contrastan con el papel conocido de C4 en afectar la maquinaria de silenciamiento. Finalmente reportan un incremento en los vsiRNAs de 24pb que coinciden con la IR durante la etapa de recuperación, sin embargo, a diferencia de otros trabajos no lograron encontrar evidencias de metilación.

Recientemente Piedra-Aguilera y colaboradores (2019) reportaron la aplicación herramientas de secuenciación de siguiente generación para caracterizar el

transcriptoma, metiloma y sRNAoma de tomate durante una infección con TYLCV. En este trabajo se determinó que no hay una correlación directa entre la abundancia de vsRNAs de 24 nt y la metilación del genoma viral. Aunque, se encontraron evidencias de metilación en regiones promotoras y el efecto que esta tiene en la diferencia de expresión de los genes C1/C4 y V1/V2, por lo que se llegó a establecer la pregunta si este fenómeno es consecuencia del mecanismo de defensa de las plantas al tratar de silenciar el virus o bien si son cambios necesarios mediados por los vsRNAs para completar el ciclo de infección.

Actualmente, el papel de los diferentes tipos de sRNAs (virus y hospedero) es más claro durante el proceso infectivo del virus y en el proceso de defensa de las plantas frente a los mismos. Por lo que el uso de construcciones a partir de los componentes virales es una estrategia factible para el control de las infecciones begomovirales. En el trabajo publicado por Medina y colaboradores (2013) usaron dos construcciones basadas en la IR de dos BGVs (PepGMV y ToChLPV) para reducir la severidad de los síntomas de infección por PepGMV en *N. benthamiana* con ambas construcciones de forma independiente. También reportan que la expresión de los componentes virales es dependiente de la construcción usada; por lo que es posible que la protección con construcciones heterólogas se logre por una vía diferente a la conferida por la construcción homóloga.

Otras estrategias han optado por usar a los miRNAs que se sobre-expresan en plantas resistentes a la infección por BGVs, por ejemplo, en el trabajo de Ali y colaboradores (2013) reportaron una reducción significativa del título viral en plantas de *N. benthamiana*, al realizar los retos infectivos con los virus CLCuBuV y CLCuKoV. Esto mediante el uso de una construcción de miRNA (miRNA artificial) diseñada a partir del mir169 de la planta algodón (*Gosypium hirsutum* L.). Demostrando que una construcción podría brindar una protección frente a ambos virus debido al nivel de complementariedad entre el miRNA y su objetivo. Sin embargo, esta estrategia también presenta algunas limitantes pues se ha documentado que la regulación de algunos miRNAs conlleva tanto a la aparición de

síntomas relacionados a infecciones virales como el aumento de la severidad de los mismos (Ali *et al.*, 2013). Además, otros trabajos de investigación han optado por apoyarse en las bases de datos y en la homología de los miRNAs para realizar estudios *in silico*, con la finalidad de predecir que estos pueden ser usados en un sistema de silenciamiento contra BGVs o en las construcciones que activan mecanismos de RNAi, sin afectar algún gen involucrado o crucial para el desarrollo del hospedero (Tousi *et al.*, 2017).

Respecto a la predicción de objetivos de miRNAs Wang y colaboradores (2004) hacen la predicción de los objetivos de 83 miRNA de *A. thaliana* por métodos computacionales, posteriormente logran comprobar la existencia de 16 objetivos de estos miRNAs mediante hibridación (Northern blot). Otro punto importante es que la mayoría de los emparejamientos predichos miRNA/mRNA en los que existen errores de alineamiento, tienen una estructura secundaria con una menor energía libre que las reportadas para la interacción de miRNA/mRNA en animales, los cuales derivan en la represión de la traducción del mRNA objetivo, por lo que sugieren la existencia de un mecanismo similar en plantas.

Por otro lado, Ellwanger y colaboradores (2011), establecieron un conjunto de características para el reconocimiento y unión de un miRNA con su objetivo. Por ejemplo postularon reglas para la predicción *in silico* de los sitios de unión de un miRNA basados en conocimiento experimental, pues demostraron que mediante el uso de secuencias tipo miRNA core seed (región semilla) de 6nt se logra predecir el 67% de los sitios objetivo verdaderos.

Las investigaciones entre la interacción miRNA y mRNA, permitieron el surgimiento de diferentes herramientas bioinformáticas que comenzaron a considerar más factores para la predicción de verdaderos objetivos. Uno de estos trabajos fue el de Dai y colaboradores (2011) en el que crean un sistema de puntuación considerando diferentes factores (biológicos y termodinámicos) y en el 2017 este esquema de puntuación fue modificado por el autor para mejorar las predicciones (Dai *et al.*, 2011; Dai *et al.*, 2018).

Las predicciones *in silico* se han aplicado para el diseño de construcciones contra BGVs. Por ejemplo, Naqvi y colaboradores (2011) usando herramientas bioinformáticas identificaron varios miRNAs de tomate que podrían unirse al genoma de *Tomato leaf curl virus* (ToLCV), aunque sus predicciones no fueron validadas les permiten especular y explorar la evolución conjunta de los mecanismos de RNAi de los virus/planta y la posible existencia de nuevas interacciones entre hospedero y virus. Ya que, se encontró que los miRNAs del hospedero podrían interactuar con la mayoría de los ORFs virales.

Mientras que Baing y colaboradores (2013) hacen la predicción de miRNAs que pueden unirse al genoma de *Cotton leaf curl Multan virus* (CLCuMuV) y su betasatellite asociado *Cotton leaf curl Multan betasatellite* (CLCuMB) que infectan algodón. En este trabajo se encontró que múltiples miRNAs con potencial de unión a la IR y los ORFs del genoma del virus, principalmente a C1 también encuentran sitios blanco en la región de solapamiento de C1/C4. También resaltan que alguno de los miRNAs tendría sitios blancos en múltiples ORFs con niveles de especificidad (complementariedad) en cada sitio. Finalmente analizan los efectos que podrían tener la sobre expresión de estos miRNAs en la defensa contra el virus (por el silenciamiento de genes y afecciones en el ciclo de replicación viral); sin embargo no analiza el efecto que esto tendría en el desarrollo de la planta. Más tarde estas predicciones *in silico* fueron usadas por Akmal y colaboradores (2017), quienes crean una construcción de miRNA para sobre-expresar miR398 y miR2950 simultáneamente. Al retar con CLCuMuV y CLCuMB a plantas de algodón transformadas con esta construcción obtuvieron como resultado la reducción de la severidad de los síntomas asociados a estos virus.

De mismo modo Tousi y colaboradores (2017) hicieron la predicción *in silico* de miRNAs de tomate que se unen al genoma de ToLCV. En este estudio se predijo que los miR159, miR403 y miR156 del tomate tienen potencial de unirse al genoma viral. Los dos primeros con la capacidad de unirse a distintos ORFs, mientras que el miR156 parecía tener complementariedad con la IR. Posteriormente evaluaron la acumulación de estos miRNAs en dos variedades de plantas de tomate. Una de

estas variedades era resistente a la infección y la otra variedad susceptible. Ellos reportaron que el mir159 y el mir403 se sobre-expresan durante la infección en la variedad susceptible, mientras que el miR156 se sobre-expresa en la variedad resistente. Estos resultados les llevan a concluir que el mir156 tiene un papel importante en la metilación del genoma viral lo que lleva una disminución de la expresión génica del virus. Además, especulan sobre la posible preferencia de unión del mir156 con el virus, pues esto explicaría como pese a sobre-expresarse este miRNA durante la infección, su gen objetivo (hospedero) no se regula de forma negativa y por el contrario se puede sobre-expresar. Este resultado les lleva a sugerir que el mir156 tiene un papel importante en la metilación del genoma viral lo que a su vez podría dar lugar a la disminución de la expresión génica del virus. También, comentan sobre la posible preferencia de unión del mir516 con el virus, pues esto explicaría cómo el gen objetivo (Gen SPL 6 del hospedero) no se regula de forma negativa y por el contrario se sobre-expresa pese a la acumulación de este miRNA. Además, postulan que el mir159 y mir403 podrían estar involucrados en el desarrollo de síntomas similares a los síntomas virales, ya que los genes objetivos de estos miRNAs están relacionados al desarrollo de las hojas. En este sentido, los autores resaltan el hecho de que pueden existir interacciones complejas entre el virus y su hospedero, por lo que se indican que se requieren más estudios para elucidar qué mecanismos del silenciamiento están participando en las interacciones virus/hospedero.

El estado del arte de este tema, deja claro que el análisis *in silico* de predicción del objetivo verdadero, es un paso esencial y debes ser previo al diseño de construcciones basadas en miRNAs del hospedero. Para así aprovechar el potencial de unirse al virus y lograr disminuir el título viral presente en el hospedero, dando lugar al mejoramiento de la salud del hospedero y disminuir significativamente los síntomas ocasionados por el virus; y a la par y de manera crucial evitar que el microRNA seleccionado intervenga en el desarrollo/ crecimiento del hospedero.

3. JUSTIFICACIÓN

La tecnología de RNAi ofrece la oportunidad de afrontar el problema de las infecciones begomovirales; Sin embargo, la eficiencia de esta estrategia depende de muchos factores, debido a lo complejo del sistema de silenciamiento génico y a las interacciones virus hospedero. En la actualidad existen dos estrategias basadas en construcciones de RNAi para el control de infecciones virales: i) Construcciones de RNAi basadas regiones del genoma viral (dsRNAs homólogos o heterólogos a los BGVs) y ii) Construcciones de RNAi basadas en miRNAs de la planta con potencial para unirse al genoma viral (amiRNAs homólogos a miRNAs del hospedero). Los estudios han demostrado que tanto las construcciones de RNAi basadas en regiones del genoma de los BGVs como aquellas basadas en miRNAs del hospedero pueden ayudar a reducir la severidad de síntomas asociados a enfermedades ocasionadas por BGVs. Sin embargo, también hay reportes que han demostrado que los sRNAs expresados por ambos tipos de construcciones pueden derivar en efectos negativos sobre el desarrollo de la planta. Además varios estudios han demostrado que el uso de análisis *in silico* para predecir las interacciones de los siRNAs con su secuencia objetivo en los virus o en el hospedero provee mayor precisión en la generación de las construcciones usadas para resolver el problema de las infecciones provocadas por BGVs. Por lo que, en el caso de la identificación *in silico* de los miRNA presentes en el tomate que influyen el silenciamiento de BGVs por RNAi están relacionados con i) el reconocimiento de secuencias blanco en regiones conservadas intra o inter especies, ii) el reconocimiento de secuencias homologas y iii) la sobre-expresión de estos miRNAs no tengan efectos adversos en el desarrollo de la planta. El conjunto de los miRNAs seleccionados, nos permitirá el diseño de construcciones más eficientes para una protección de amplio espectro contra los BGVs presentes en BCS.

4. HIPÓTESIS

Si los miRNAs de plantas juegan un papel importante en la regulación genética y en la defensa contra begomovirus; entonces identificar miRNAs *in silico* que no tenga efectos significativos en el desarrollo de *Solanum Lycopersicum*, pero que sea blanco de *Tomato Chino la Paz Virus* (ToChLPV), *Tomato Yellow Leaf Curl virus* (TYLCV), *Pepper Golden Mosaic Virus* (PepGMV), *Squash Leaf Curl Virus* (SLCuV), permitirá generara información a priori de miRNAs con potencial para utilizarse en un sistema de silenciamiento eficiente y de amplio espectro.

5. OBJETIVOS

5.1 Objetivo general

Identificación *in silico* los miRNAs de *S. lycopersicum* que pueden unirse al genoma de los begomovirus ToChLPV, TYLCV, SLCuV y PepGMV reportados en el estado, con el potencial de utilizarse posteriormente en un sistema de silenciamiento eficiente y de amplio espectro.

5.2 Objetivos específicos

- Caracterizar el genoma de los BGVs de interés y sus variaciones como especie.
- Establecer relaciones filogenéticas entre ToChLPV y los virus caracterizados en BCS.
- Analizar complementariedad de los miRNAs de *S. lycopersicum* con los genes de los BGVs.
- Determinar *in silico* los miRNAs con mayor capacidad de reconocer sitios blancos para activar el silenciamiento de BGVs.

6. MATERIALES Y MÉTODOS

6.1 Selección de especies, aislamientos y variantes de los BGVs

Debido a que el presente trabajo se enfoca en problemas ocasionados por BGVs en Baja California Sur (BCS). Se descargaron de GenBank únicamente las secuencias de los genomas de los BGVs cuya presencia se confirmó previamente por nuestro grupo de trabajo los cuales fueron: *Tomato Chino la Paz Virus* (ToChLPV), *Tomato Yellow Leaf Curl Virus* (TYLCV), *Pepper Golden Mosaic Virus* (PepGMV), *Squash Leaf Curl Virus* (SLCuV) (Tabla I). Posteriormente se descargaron los genomas asociados a BGVs que fueron reportadas en el continente Americano y reconocidas por el ICTV (Tabla I). Del mismo modo se descargaron los genomas de variantes del medio oriente de SLCuV. Los archivos de las secuencias de los genomas se dividieron en las secuencias de los ORFs presentes de forma individual, así como de las proteínas codificadas. Todas las secuencias se agrupan por especies para su análisis. Posteriormente se aisló la secuencia de la IR de cada genoma, las cuales también se agruparon por especie. Los BGVs analizados se les asignó una identificación en base a su acrónimo (ICTV), localidad y/o región donde se aisló y número de acceso en el GenBank (NCBI). Las regiones se expresan con la abreviación del país separada por dos puntos de la abreviación de la localidad. Los países se abreviaron de la siguiente manera: Costa Rica (CR), Egipto (EG), Estados Unidos (US), Guatemala (GU), Israel (IL), Jordania (JO), México (MX), Nicaragua (NI), Palestina (PL) y República Dominicana (Do). Finalmente se realizó el alineamiento de las secuencias con ClustalW en la paquetería MEGA X V10.0.4 (Kumar *et al.*, 2018).

Tabla I. Secuencias seleccionadas de los begomovirus usados en este estudio, corresponde a aquellas secuencias reportadas como especies por el Comité Internacional de Taxonomía de Virus (ICTV).

Especie	Descripción	Número de acceso NCBI	Primera referencia	Año
ToChLPV	MX:BCS*	HM459852	Cardenas-Conejo, Y	2010
	MX:BCS:LP	AY339618	Holguin-Pena, R.J.	2006
	MX:BCS:EC	AY339619	Holguin-Pena, R.J.	2006
	MX:Sin	DQ347948	Gregorio-Jorge, J. Ambriz.	2006
	MX:Sin2	DQ347949	Gregorio-Jorge, J.	2006
	MX:SLP	JN676150	Banuelos-Hernandez, B	2012
TYLCV	MX:BCS*	HM459851	Cardenas-Conejo, Y.	2010
	Do	AF024715	Nakhla, M.K.	1994
	US:FL	AY530931	Yang, Y.	2004
	MX:Cul	DQ631892	Brown, J.K.	2006
	US:TX	EF110890	Isakeit, T.	2007
	US:AZ	EF210554	Idris, A.M.	2007
	MX:Son	EF210555	Idris, A.M.	2007
	MX:Sin	EF523478	Kon, T.	2007
	US:CA	EF539831	Rojas, M.R.	2007
	MX:Sin:Gu	FJ609655	Gamez-Jimenez, C.	2009
	US:Poa	GU322423	Melzer, M.J.	2010
	US:Wai	GU322424	Melzer, M.J.	2010
	GU:Sal	GU355941	Salati, R.	2010
	US:KLK	HM988987	Melzer, M.J.	2010
	MX: SLP	JQ303121	Banuelos-Hernandez, B.,	2011

Especie	Descripción	numero de ac	Primera referencia	año
SLCuV	Mx:BCS *	MF187211	Vargas, M.	
	US:AZW	AF256203	Brown, J.K.	2000
	Jo:Mal	EF532620	Al-Musa, A.	2008
	LB	HM368373	Sobh, H	2012
	IL	HQ184436	Abudy, A.	2010
	JO:Sarv	JX131281	Anfoka, G.H	2012
	JO:Hor	JX444577	Haj Ahmad, F.	2013
	PL	KC441465	Ali-Shtayeh, M.S.	2014
	EG:ism	KC895398	El Attar, A.K.	2013
	US:IV	DQ285016	Brown, J.	2005
PepGMV	MX:Tam*	U57457	Torres-Pacheco	1996
	CR	AF149227	De La Torre, R.	1998
	US:Mos	AY928512	Brown, J.K.	2005
	US:Dis	AY928514	Brown, J.K.	2005
	CR:(US:Ser)	AY928516	Brown, J.K.	2005
	US:Mt	EF210556	Idris, A.	2007
	MX:Gua1	GU128148	Carrillo-Tripp, J.	2009
	MX:Gua2	GU128149	Carrillo-Tripp, J.	2009
	MX:SLP1	GU564595	Ambriz-Granados, S.	2010
	CR:Tur 2	JN688724	Vargas, J.A.	2011
	CR:Tur	JN688725	Vargas, J.A.	2011
	CR:Cer	JN688726	Vargas, J.A.	2011
	CR:Tab	JN688727	Vargas, J.A.	2011
	CR:LG	JQ743486	Vargas, J.A.	2012

6.2 Comparación nucleotídica para búsqueda de divergencia

Se realizó la comparación por pares de las secuencias de cada una de las variantes analizadas para determinar las regiones más conservadas. Para ello se utilizó como referencia las secuencias de la variante reportada en el estado. En el caso de Tomato chino se tomó como referencia la variante MX: BCS [HM459852] para el análisis de ToChLPV debido a que es la variante que se reportó en infección mixta con TYLCV. Una vez calculado el porcentaje de identidad para cada uno los elementos de los virus seleccionados, por especie se graficó el porcentaje de identidad de los aislamientos de BCS. Para ello se realizó el alineamiento de los

genomas por especie para después calcular la distancia filogenética con el programa SimPlot V3.5.1 con el modelo de parámetros Kimura-2 en regiones de 200 nt.

6.3 Diversidad nucleotídica

La diversidad de nucleotídica (π) se estimó mediante la prueba de neutralidad de Tajima para cada uno de los ORFs de cada especie, usando los parámetros con valores predeterminados con la paquetería de MEGA X V10.0.4 (Kumar *et al.*, 2018).

6.4 Presión de selección

El análisis de presión de selección se estimó mediante la aprobación de tres hipótesis alternas todas las secuencias de cada ORF por especie, mediante el test-Z por codones utilizando el modelo de parámetros de Kimura-2 y los modelos de sustitución (sinónima [α_s] y no sinónima [α_n]) con un soporte estadístico de 1500 réplicas de “bootstrap” para la estimación de la variación de la diferencia entre α_s y α_n , las cuales consistieron en evaluar la neutralidad [sustituciones sinónimas (δ_s) = sustituciones no sinónimas (δ_n)], presión positiva [$\delta_n > \delta_s$] y presión purificadora o negativa [$\delta_s > \delta_n$]. Además, se calculó la cantidad de fenómenos de inserción/delección, sustituciones sinónimas y no sinónimas mediante el programa SNAP (<https://www.hiv.lanl.gov>) utilizando, para cada especie, el alineamiento de cada uno de los ORF por especie, tomando como base la acumulación de estos eventos, codón por codón a lo largo de la secuencia analizada, para todas las posibles comparaciones pareadas. Finalmente, los valores obtenidos se graficaron en el programa SigmaPlot V14.0.

6.5 Análisis filogenético

Los genomas se alinearon con ClustalW con los parámetros predeterminados de la paquetería de MEGA X V10.0.4 (Kumar *et al.*, 2018). Posteriormente el alineamiento se editó manualmente tomando como criterio la eliminación de las regiones menos

informativas. Se estableció el mejor modelo de ajuste de sustitución de nucleótidos con el programa “find best-fit substitution model” con los parámetros predeterminados de la paquetería MEGA X v10.0.4. El análisis filogenético se llevó a cabo con el método de máxima verosimilitud (ML) usando el modelo de parámetros de Tamura-3 y 1500 réplicas de “Bootstrap” de soporte estadístico. El árbol inicial fue obtenido automáticamente por el método BioNJ. El resto de parámetros fueron los predeterminados.

6.6 Predicción de sitios semilla

Se descargaron las secuencias de miRNAs maduros de tomate de la base de datos miRBase (<http://www.mirbase.org>). Se usó el software miRanda V3.3a para la predicción de posibles sitios objetivos de dichos miRNAs en los diferentes ORFs y la IR de cada especie. Como límite de confianza se permitió un error de complementariedad en la región semilla la cual se consideró entre las posiciones 2-8 nt de la región 5' del miRNA, de acuerdo a lo reportado por Tousi *et al.*, (2017). Se estableció un listado de referencia de los miRNAs que reconocieron un sitio blanco después del análisis de predicción de los sitios semilla en el genoma de los BGVs. Posteriormente se realizó otro análisis con el servidor psRNATarget (<http://plantgrn.noble.org/psRNATarget/>) utilizando la lista de miRNAs previamente generada, así como las secuencias de los ORFs e IRs de los BGVs analizados. Para este análisis se eligió el esquema V2 del 2017 (Dai *et al.*, 2018) y se usaron los parámetros predeterminados. Los miRNA que encontraron una región objetivo en alguna de las secuencias se denominaron miRNA candidatos, los cuales se agruparon por familia de miRNA, especie de reconocimiento de sitio blanco y, gen objetivo.

6.7 Especificidad de los miRNAs candidatos con los genes objetivo en los BGVs y de los genes del hospedero

Con el fin determinar la eficiencia de los miRNAs candidatos en un sistema de silenciamiento de BGVs se realizó un filtro donde se eliminaron aquellos miRNAs

que silenciaron a menos de la mitad de los aislamientos en una sólo especie. Se identificaron los genes objetivos del hospedero y tipo de proceso al que responde (por ejemplo, que regule la expresión del miRNA) o el que afecta la expresión de los miRNAs restantes. Tomando como base la información disponible de reportes previos. También se consideró la cantidad de proteínas que podría afectar y la cantidad de genes que silencia en los BGVs; así como la estabilidad de unión con su secuencia objetivo. Finalmente, a partir de estos criterios se seleccionaron los miRNAs que podrían ser usados para el desarrollo de construcciones sin afectar el desarrollo de las plantas de manera significativa.

7. RESULTADOS

7.1 Comparación genómica de especies caracterizadas en BCS con los aislados reconocidos por el ICTV

Al comparar el DNA-A de ToChLPV [HM459852], con los demás aislamientos reportados de BCS presentan una identidad del 92% al 96%; siendo el aislamiento de El carrizal el de mayor identidad (Tabla II). Mientras que los aislamientos de TYLCV mostraron un porcentaje de identidad entre 97%-99% con el aislamiento de BCS [HM459651], siendo el aislamiento del estado de Sonora el de mayor identidad. Los aislamientos reportados de SLCuV mostraron una identidad del 96%-99% con el aislamiento de BCS [MF187211], siendo el aislamiento de Jordania (Mal) el de mayor identidad. Los aislamientos de PepGMV mostraron una identidad entre 92%-94% con el aislamiento de Tamaulipas [U57457]; siendo el aislamiento de US (mos) el de mayor identidad (Tabla II).

Los resultados del análisis de la región intergenica mostraron que la región es altamente conservada dentro de los BGVs reportados en la región (92%-99%); siendo PepGMV el virus con menor identidad en esta región (92.35%). Además, el estudio mostró que el porcentaje de identidad en la región derecha de la IR de todos los aislamientos es menor que la región izquierda de la IR (Anexo 1). El análisis de los ORFs para el gen C1 de los aislamientos de ToChLPV y PepGMV son los menos conservados, con una identidad promedio del 92.35% - 93.92%; mientras que TYLCV y SLCuV mostraron una identidad mayor al 97% en promedio, en el gen C2 los aislamientos de TYLCV fueron los que mostraron el menor porcentaje de identidad en promedio (92%), mientras que para el gen C3 los aislamientos de SLCuV fueron las que mostraron menor identidad en promedio (94.72%). Para C4 todas las especies mostraron una identidad promedio mayor al 97%. Para V1 los aislamientos de PepGMV fueron las menos conservadas con un promedio de 93.98% (Tabla II).

US:AZ:W	97.2	97.4	99	97.6	97.7	97	97.5	95.5	96.8	92	96.7	97.6	-	-
Jo:Mal	96.9	97.9	99	97.2	99.5	99.2	98.3	95.5	88.6	67.2	98.5	97.2	-	-
LB	99.09	98.8	98.8	99.6	100	100	97.8	97.8	98.7	96	99.7	99.6	-	-
IL	98.83	99.1	98.7	99.6	99.5	52.6	97.8	96.3	98	95.6	99.6	99.6	-	-
JO:Sarv	98.94	99.3	98.6	99.6	100	100	98.3	97.8	98.4	95.2	99.6	99.6	-	-
JO:Hor	98.98	99.1	98.5	99.2	99.8	100	98	97.8	98.7	96	99.3	99.2	-	-
PL	98.94	99.1	98	100	99.8	99.2	97.8	96.3	97.9	93.6	99.7	100	-	-
EG:ism	97.2	94.8	95.2	100	99	97.7	97.3	96.3	92.1	49.2	99.6	100	-	-
US:IV	98.13	98.4	95.1	98.4	98.7	98.5	98	97	97.9	94.4	98	98.4	-	-
Promedio	98.2	98.2	98.5	98	97.9	99	99.3	93.8	97.9	96.7	96.3	86.6	99	99
PepGMV														
MX:Tam	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CR	92.14	90.5	91.8	92	93.3	88.8	93.2	89.4	-	-	93	97.2	-	-
US:Mos	95.26	94.6	94.6	94	97.4	96.9	96.7	93.2	-	-	95.1	98.4	-	-
US:Dis	94.81	93.8	95.1	94.3	94.6	89.7	95	93.9	-	-	94.8	98	-	-
CR:(US:Ser)	92.6	90.6	92	92	93.1	91.4	94.7	93.9	-	-	93.4	98	-	-
US:Mt	94.93	93.2	93.8	94.3	98.5	99.2	98	96.2	-	-	95.2	97.6	-	-
MX:Gua1	94.8	95.2	95.7	96.3	94.1	89.7	94	88.6	-	-	94.6	98.4	-	-
MX:Gua	94.58	94.3	95.1	94.8	94.6	90.5	94.2	91.7	-	-	94.4	98.4	-	-
MX:SLP1	94.92	94.9	95.2	94.8	96.2	94	94.5	90.2	-	-	94.4	97.6	-	-
CR:Tur	92.52	90.3	92	92.6	94.6	92.2	94.2	92.4	-	-	93.1	97.2	-	-
CR:Tur	92.6	90.9	92.2	93.4	94.6	92.2	94.2	93.2	-	-	92.9	97.2	-	-
CR:Cer	92.94	91.2	92.3	93.4	94.9	93.1	94.7	93.2	-	-	93.5	97.6	-	-
CR:Tab	92.87	90.9	92.2	93.1	94.6	92.2	94.5	93.2	-	-	93.7	97.6	-	-
CR:LG	92.6	90.3	92.3	93.1	94.1	91.4	93.2	91.7	-	-	93.5	97.6	-	-
Promedio	93.6	92.4	93.4	93.7	95	92.4	94.7	92.4			94	97.8		

+Proteínas que muestran en un incremento significativo en la similitud de la proteína con respecto a la similitud del gen.

*Proteínas que muestran en una disminución significativo en la similitud de la proteína con respecto a la similitud del gen.

Las gráficas de similitud muestran que el genoma de los aislamientos de ToChLPV en las regiones de ORFs e IR presentan diferentes porcentajes de similitud. En la

mayoría de los aislamientos las regiones con mayor divergencia fueron: la región de 0 a 200 nt (IR) donde la similitud cae hasta el 86%, la región de 1100 nt a 1500 nt (C2/C3) donde la similitud varía entre el 85%-92% y la región de 2000 nt a 2500 nt (C1 y C4) donde la similitud decae hasta un 45% (Fig. 3a). Mientras que la similitud de TYLCV es estable en todas sus regiones, manteniéndose sobre el 94% (Fig. 3b). En el caso de PepGMV, la similitud con los aislamientos reportados en América es superior al 90%, excepto en la región C1/C4 donde para la mayoría de los aislamientos decae hasta un 85% (Fig. 3c). La similitud de SLCuV se mantiene sobre el 94%, excepto en la región de 900 nt a 1100 nt (C2) en todos los aislamientos y la región de 2000 nt a 2500 nt (C1/C4) para los aislamientos de Egipto y Jordania:Mal en las que la similitud decae hasta un 80% (Fig. 3d).

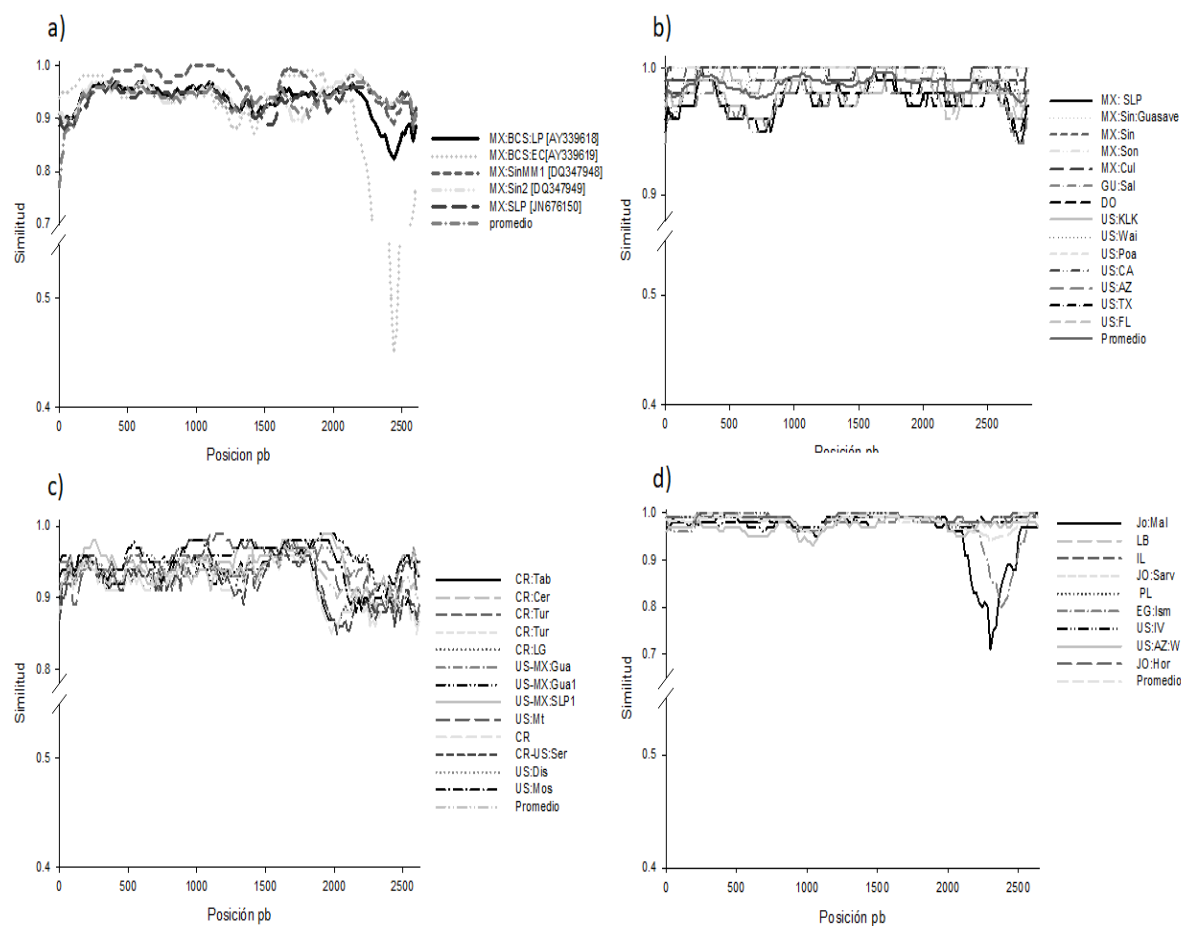


Figura 3. Gráficas por especie de similitud de BGVs de referencia comparadas con

aislamientos de BCS. Porcentaje de similitud calculado para 200 nt por cada 20nt (gapstrip: Kimura-2). Porcentaje de similitud por posición en los genomas de BGVs analizados. a) aislamientos de ToChLPV comparado con MX:BCS HM459852, b) aislamientos de TYLCV de América comparados con Mx:BCS HM459851, c) aislamientos de PepGMV reportadas en América respecto comparados con Mx:Tam U57457) y d) aislamientos de SLCuV reportadas para américa y medio oriente comparados con MX:BCS MF187211.

7.2 Homología de las proteínas de los Begomovirus aislados de BCS

En relación a las diferencias entre las proteínas codificadas por estos genes (Tabla II) se determinaron diferencias significativas entre el porcentaje de la identidad del gen y de la proteína en C4 y V1/CP de ToChLPV; para PepGMV en V1/CP y para SLCuV en C4. En estos casos el porcentaje de identidad siempre fue mayor para la proteína de CP respecto al gen mientras que el porcentaje de identidad de C4 fue menor para la proteína que para el gen.

7.3 Diversidad nucleotídica

Con base en los análisis de diversidad nucleotídica del genoma A de los BGVs (Tabla III), se encontró mayor variabilidad dentro de las variantes de ToChLPV ($\pi=0.061$) con respecto a los otros BGVs detectados en la región, tales como PepGMV, TYLCV y SLCuV. Los aislamientos de PepGMV presentaron un $\pi=0.054$, los aislamientos de SLCuV y TYLCV presentaron valores de $\pi=0.019$ y $\pi=0.013$, respectivamente. De acuerdo al análisis de diversidad de los ORFs (Tabla III), el gen con mayor diversidad fue C1 de ToChLPV. Los genes de cada especie mostraron valores diferentes de diversidad, siendo la región C1 la que mostró la mayor diversidad en ToChLPV y PepGMV, mientras que para SLCuV y TYLCV fueron los genes C4 y AV1, respectivamente. Los resultados muestran una relación positiva para el gen C4 y C1 [$\pi(C4)/\pi(C1)>1$] entre TYLCV y SLCuV, mientras que entre los aislamientos de ToChLPV fueron negativos [$\pi(C4)/\pi(C1)<1$]. Para PepGMV no se calculó esta relación, ya que el gen C4 no está caracterizado en la mayoría de los genomas de los BGVs aislados. Mientras que en todas las especies la relación entre los genes C2 y C3 fue aproximadamente de uno [$\pi(C2)/\pi(C3)\cong 1$].

Las gráficas de similitud muestran la disminución de la similitud en la región C1/C4 que es una región de solapamiento de secuencias codificantes con diferentes marcos de lectura para dos proteínas.

Tabla III. Tabla de diversidad nucleotídica del genoma y genes por especie

ToChLPV			SLCuV		
n=6	π	D	n=10	π	D
V1	0.035315*	0.577846	V1	0.014903	-0.98635
C1	0.072095+	-0.18845	C1	0.018597	-1.43253
C2	0.057537	0.523211	C2	0.010606*	-1.58688
C3	0.051462	0.419044	C3	0.013882	-1.85203
C4	0.048837	-0.81408	C4	0.042482+	-1.84771
Total	0.061413	0.063634	Total	0.019511	-1.55154
TYLCV			PepGMV		
n=15	π	D	n=14	π	D
V1	0.01542+	0.224956	V1	0.050814	-0.67277
V2	0.006892*	-0.06183	C1	0.054819+	-0.02322
C1	<u>0.010588</u>	-1.30438	C2	0.046858*	-0.34238
C2	0.012278	-0.58208	C3	0.050814	-0.67277
C3	0.013075	-0.56981	C4	-	-
C4	<u>0.01312</u>	-1.38654	Total	0.054488	-0.11488
Total	0.01391	-0.76174			

La diversidad de nucleótidos (π), de la prueba de neutralidad de Tajima (D). (+) los valores de diversidad más altos para cada especie. (*) los valores de diversidad más bajos. (u) valores de interés resaltados. Los valores negativos=menor heterocigosidad; los valores positivos = más heterocigosidad.

7.4 Presión de selección

Se realizó un análisis de presión de selección en las secuencias codificantes (ORFs) de las especies de BGVs, para determinar el efecto positivo, negativo o neutro. Los análisis del test-Z basado en los codones de los genes de BGVs indicaron que la presión de selección fue negativa para la mayoría de los casos.

El análisis de distribución de mutaciones sinónimas y no sinónimas mostró que en ToChLPV, PepGMV y TYLCV existe una mayor acumulación de mutaciones sinónimas en C1; sin embargo, el número total de sustituciones sinónimas es mayor

en ToChLPV y PepGMV; mientras que para SLCuV la mayor acumulación son mutaciones no sinónimas (Fig. 4).

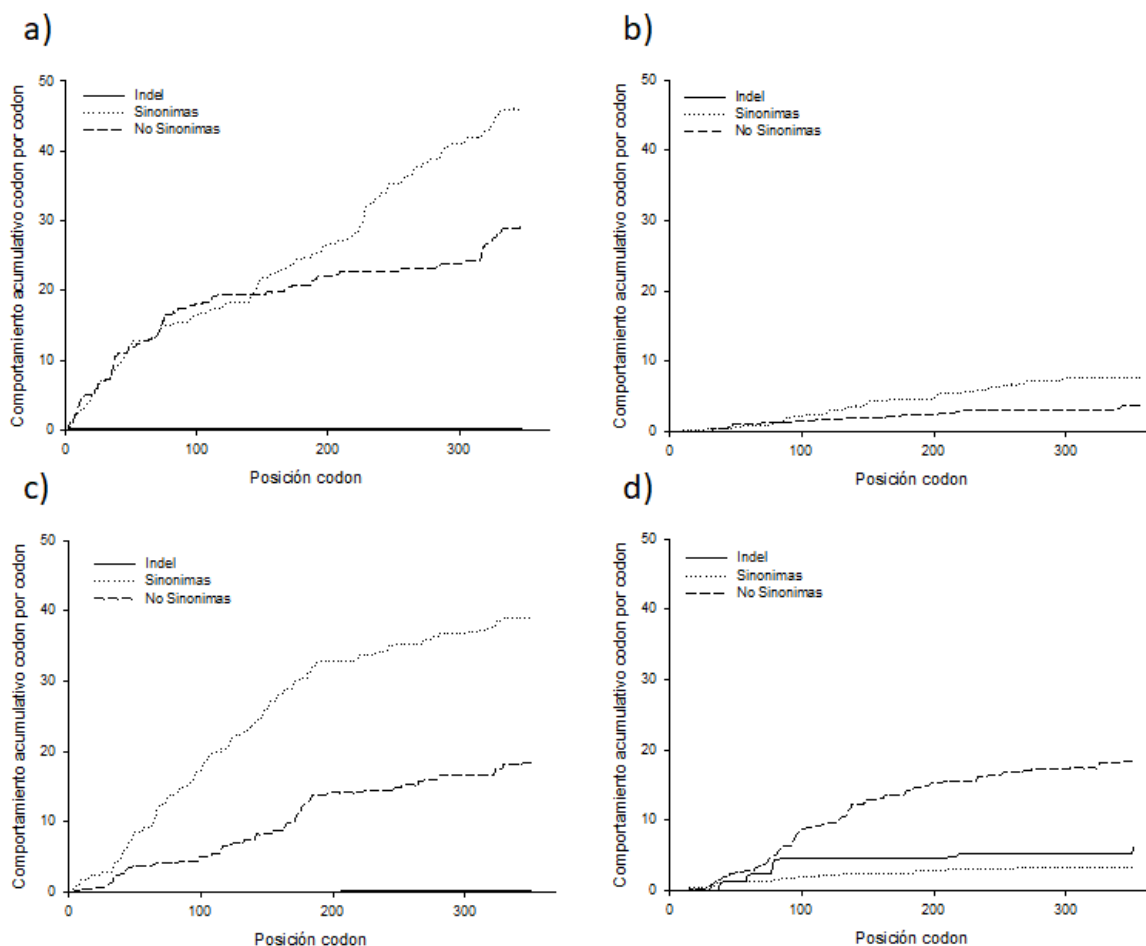


Figura 4. Distribución de sustituciones sinónimas y no sinónimas en el gen C1. Acumulación de sustituciones sinónimas y no sinónimas, fenómenos de inserción-delección respecto a la posición en el gen en codones de las cuatro especies de BGVs: a) ToChLPV, b) TYLCV, c) PepGMV y d) SLCuV.

El número de sustituciones sinónimas y no sinónimas en C2 y C3 son similares, siendo mayor para ToChLPV y PepGMV que en TYLCV y SLCuV (Fig.5 y Fig. 6).

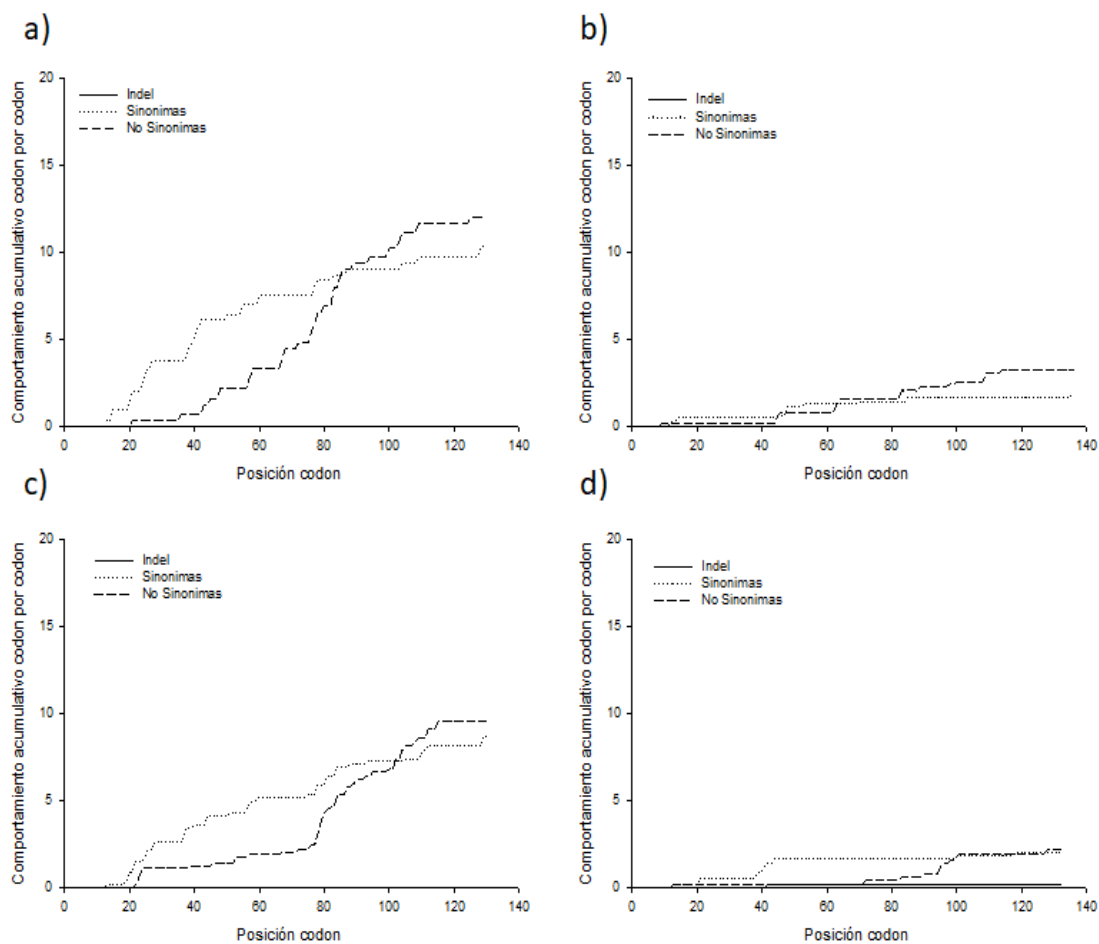


Figura 5. Distribución de sustituciones sinónimas y no sinónimas en el gen C2. Acumulación de sustituciones sinónimas y no sinónimas, fenómenos de inserción-delección respecto a la posición en el gen en codones de las cuatro especies de BGVs: a) ToChLPV, b)TYLCV, c) PepGMV y SLCuV.

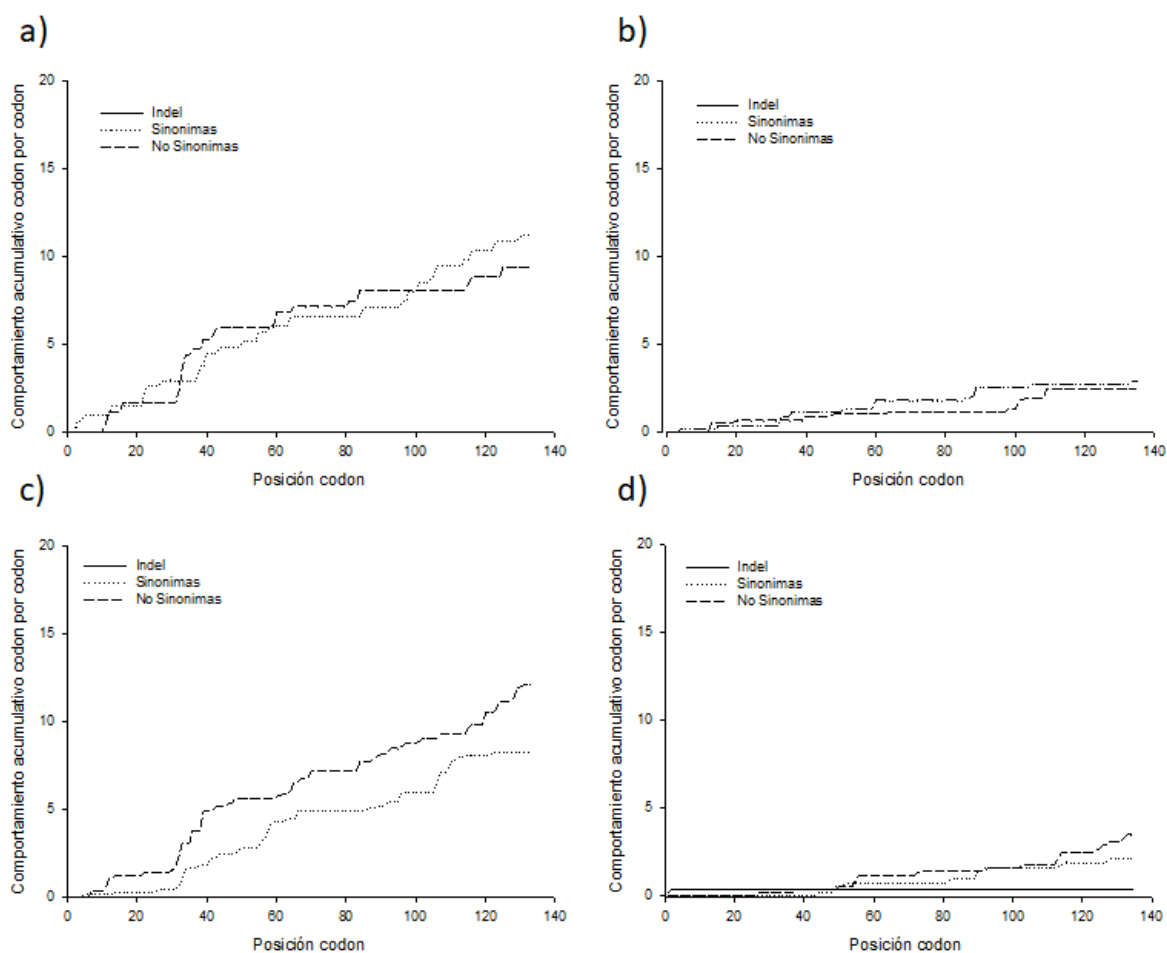


Figura 6. Distribución de sustituciones sinónimas y no sinónimas en el gen C3. Acumulación de sustituciones de sinónimas y no sinónimas, fenómenos de inserción-delección respecto a la posición en el gen en codones de las cuatro especies de BGVs: a)ToChLPV, b)TYLCV, c) PepGMV y d)SLCuV.

Para el gen V1 el número de sustituciones sinónimas es mayor al de las no sinónimas en todos los BGVs analizados; sin embargo, en ToChLPV y PepGMV es más de dos veces mayor comparado con TYLCV y SLCuV (Fig. 7).

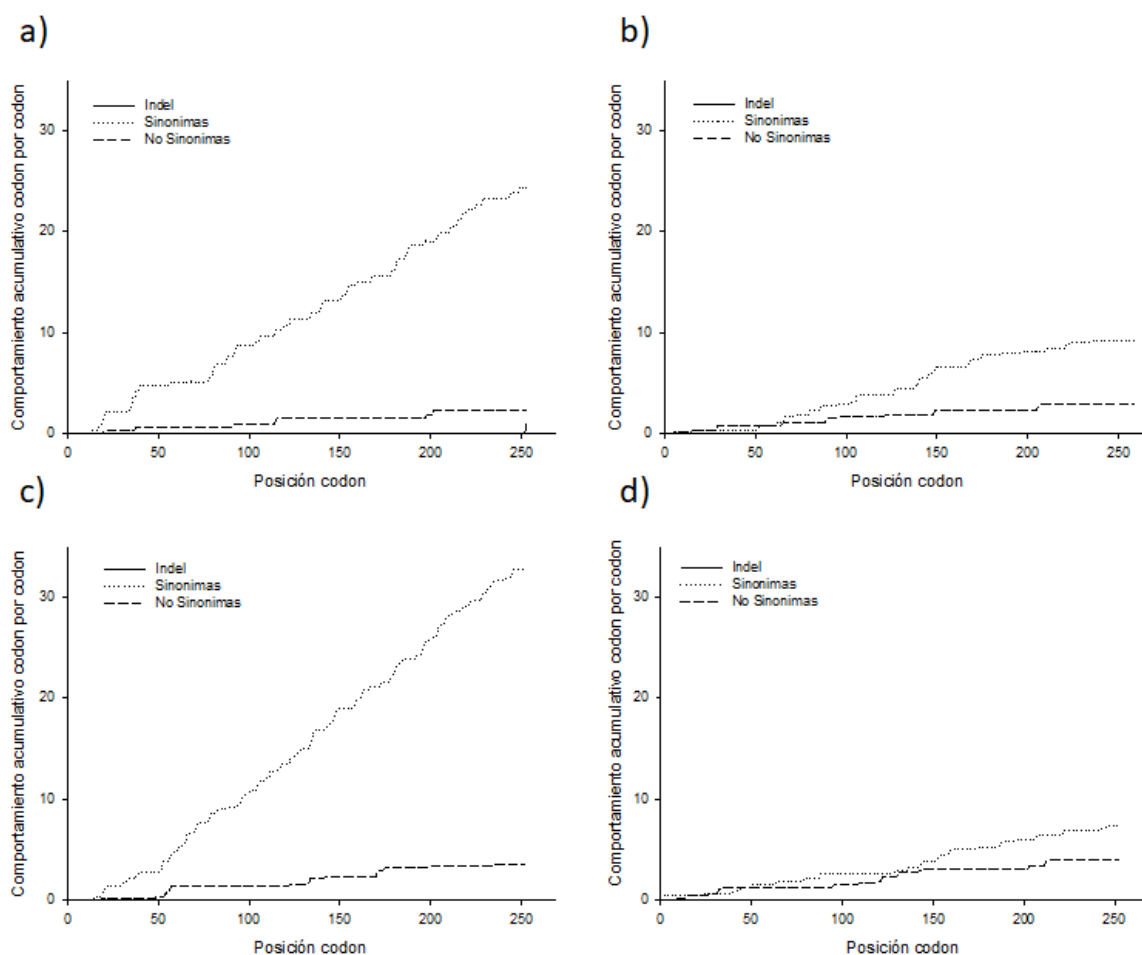


Figura 7. Distribución de sustituciones sinónimas y no sinónimas en el gen V1. Acumulación de sustituciones sinónimas y no sinónimas y, fenómenos de inserción-delección respecto a la posición en el gen en codones de las cuatro especies de BGVs: a)ToChLPV, b)TYLCV, c) PepGMV y d)SLCuV.

En el caso del gen codificante para C4 el número de sustituciones no sinónimas fue mayor que las sustituciones sinónimas para ToChLPV y TYLCV mientras que para SLCuV el número de sustituciones sinónimas y no sinónimas es similar (Fig. 8).

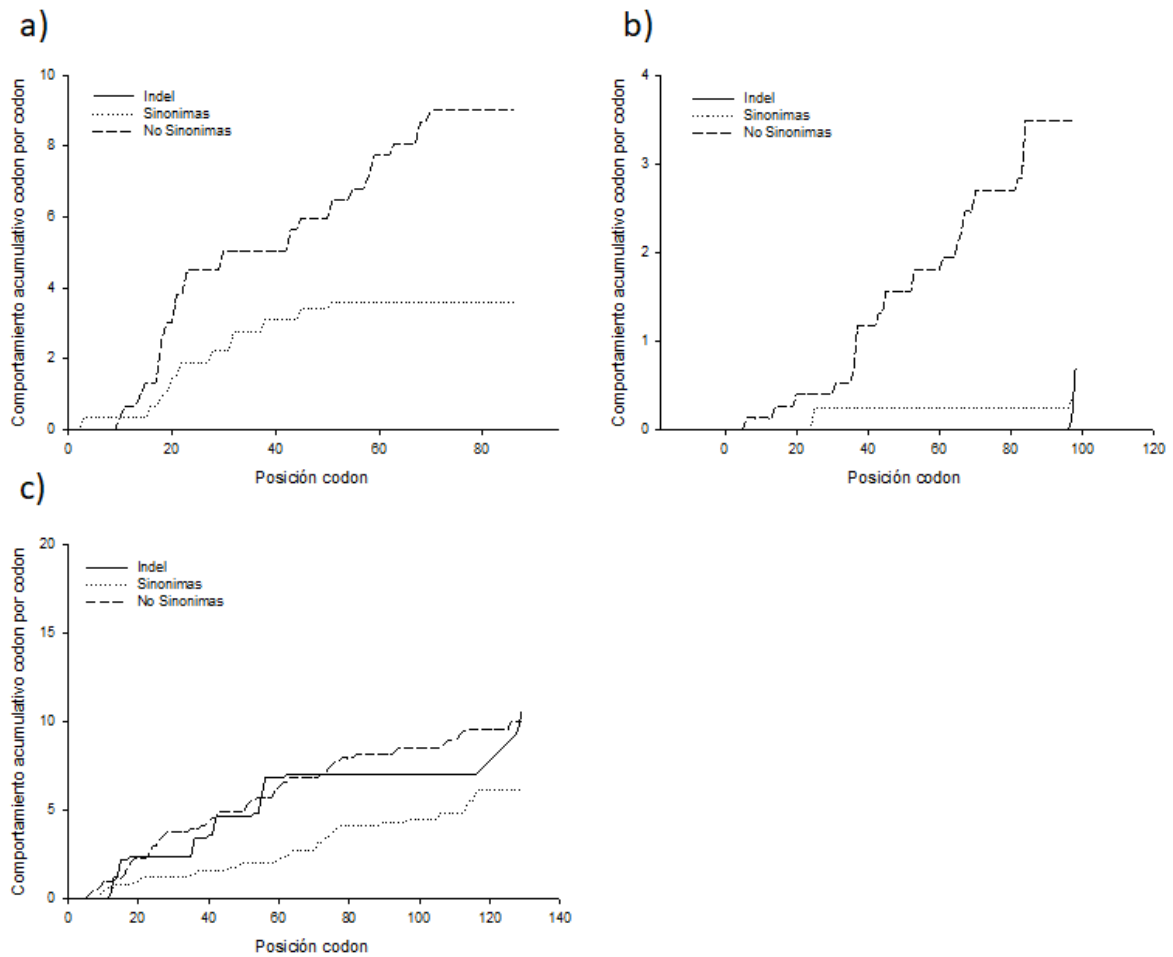


Figura 8. Distribución de sustituciones sinónimas y no sinónimas en el gen C4. Acumulación de sustituciones sinónimas y no sinónimas y, fenómenos de inserción-delección respecto a la posición en el gen en codones de tres de las cuatro especies de BGVs: a) ToChLPV, b) TYLCV, c) SLCuV.

Finalmente, con TYLCV el gen V2 mostró que los valores de sustituciones se mantuvieron bajos (Fig. 9).

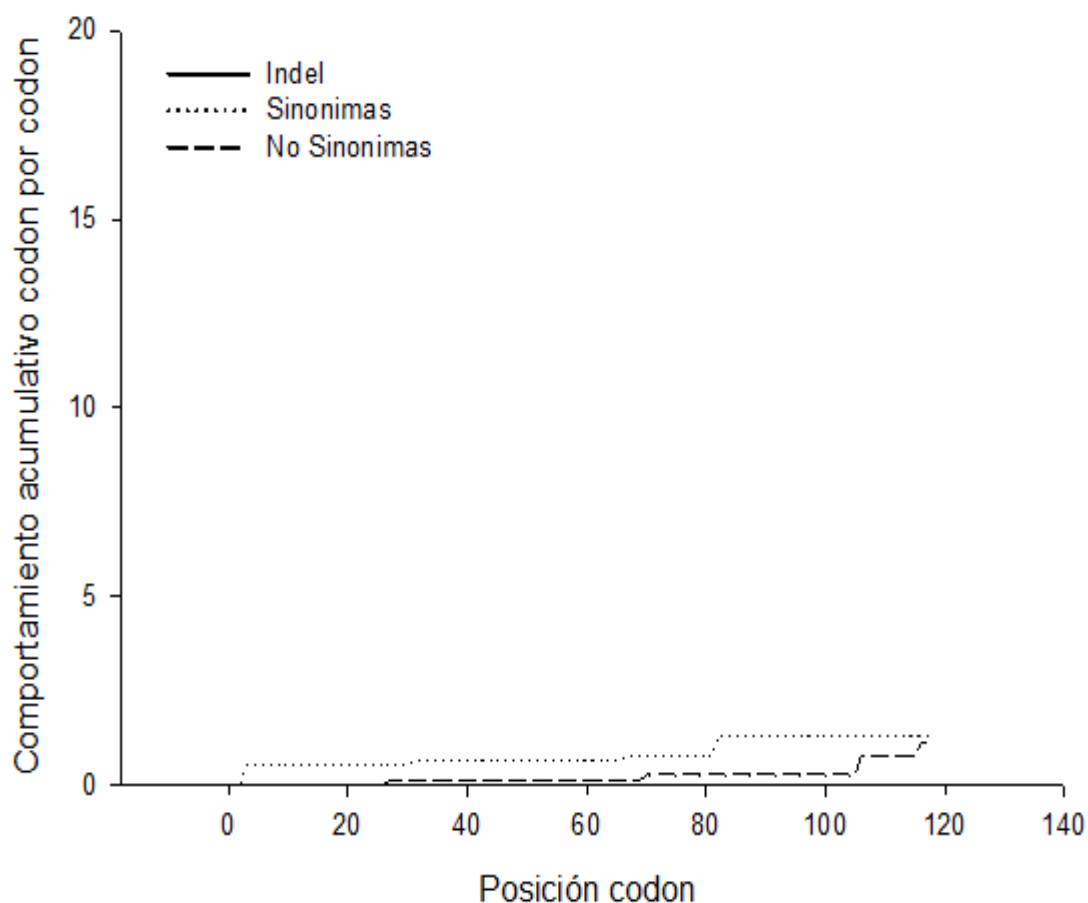


Figura 9. Distribución de sustituciones sinónimas y no sinónimas en el gen V2. Acumulación de sustituciones sinónimas y no sinónimas y, fenómenos de inserción/delección respecto a la posición en el gen en codones de V2 de TYLCV.

En la Tabla IV, se muestra la probabilidad de no rechazar las diferentes hipótesis nula de neutralidad estricta ($dN=dS$) proponiendo como hipótesis alternativa la presión positiva ($dN>dS$) o negativa ($dN<dS$). Los valores de probabilidad inferiores a 0.05 se consideran significativos al nivel del 5%. La estadística de prueba ($dN - dS$) se muestra en la columna estadística ("stat"). dS y dN son los números de sustituciones sinónimas y no sinónimas por sitio, respectivamente.

Tabla IV. Probabilidad de rechazar la hipótesis de presión de selección.

ToChLPV	dN-dS Neutra		dN-dS Positiva		dS- dN Negativa	
	<i>prob</i>	<i>stat</i>	<i>prob</i>	<i>stat</i>	<i>prob</i>	<i>stat</i>
AV1	0	-7.95	1	-7.89	0	8.13
C1	0	-0.97	1	-9.28	0	9.42
C2	0	-3.04	1	-2.96	0	3.09
C3	0	-3.5	1	-3.6	0	3.58
C4	0.42	-0.8	1	-0.74	0.22	0.76

TYLCV						
AV1	0	-4.2	1	-4.22	0	4.21
V2	0.22	-1.24	1	-1.25	1	1.3
C1	0	-4.84	1	-4.81	0	4.86
C2	0.24	-1.19	1	-1.27	0.11	1.21
C3*	0.08	-1.78	1	-1.77	0.05	1.69
C4*	0.09	1.73	0.05	1.66	1	-1.55

PepGMV						
AV1	0	-10.97	1	-10.95	0	11.87
C1	0	-10.3	1	-9.63	0	9.98
C2	0	-3.1	1	-3.5	0	3.28
C3	0	-2.72	1	-2.64	0.1	2.49

SLCuV						
AV1	0	-4.06	1	-4.02	0	4.22
C1	0.02	2.31	0.01	2.26	1	-2.2
C2*	0.07	-1.83	1	-1.82	0.05	1.7
C3*	0.13	-1.51	1	-1.61	0.06	1.6
C4*	0.02	-2.45	1	-2.47	0.01	2.47

Se muestra la probabilidad de no rechazar las diferentes hipótesis nula de neutralidad estricta ($dN=dS$) proponiendo como hipótesis alternativa la presión positiva ($dN>dS$) o negativa ($dN<dS$). Los valores de probabilidad inferiores a 0.05 se consideran significativos al nivel del 5%. La estadística de prueba ($dN - dS$) se muestra en la columna Estadística. dS y dN son los números de sustituciones sinónimas y no sinónimas por sitio, respectivamente.

7.5 Relación filogenética entre ToChLPV y BGVs presentes en BCS

En el análisis filogenético realizado a partir de los genomas completos de los BGVs seleccionados, estableció la base para que se definieran dos grupos discretos, los cuales están conformados por los BGVs monopartitas y bipartitas. En el

dendrograma (Fig. 10) el clado de los monopartitas quedó constituido por las variantes de TYLCV, y se dividió en dos subclados, uno formado por las variantes de Guatemala, República Dominicana, San Luis Potosí, Texas y Arizona, mientras que el otro estaba formado por las variantes de Sonora, Sinaloa, Culiacán y otras variantes de Estados Unidos. Este último subclado fue donde la variante de BCS. [TYLCV-BCS] se agrupó (Bañuelos-Hernández *et al.*, 2012).

El clado de los begomovirus Bipartitas se agrupó por especie. El subclado formado por las variantes de ToChLPV fue el que mostró una mayor distancia filogenética con respecto a los clados de las otras especies. La variante BCS:LP [AY339618] es el que tiene mayor distancia con el resto de variantes. El subclado de PepGMV formó dos subclados, uno conformado por los aislamientos de PepGMV de México y Estado Unidos y otro por los BGVs aislados en Costa Rica.

Las variantes de SLCuV se agruparon en un clado o grupo monofilético definido, la distancia entre variantes dentro de este clado fue la menor respecto a otros clados. Sin embargo, se formaron dos subclados, uno formado por las variantes de Estados Unidos y el otro está formado por las variantes de Medio Oriente. En este último se agrupó la variante de BCS.

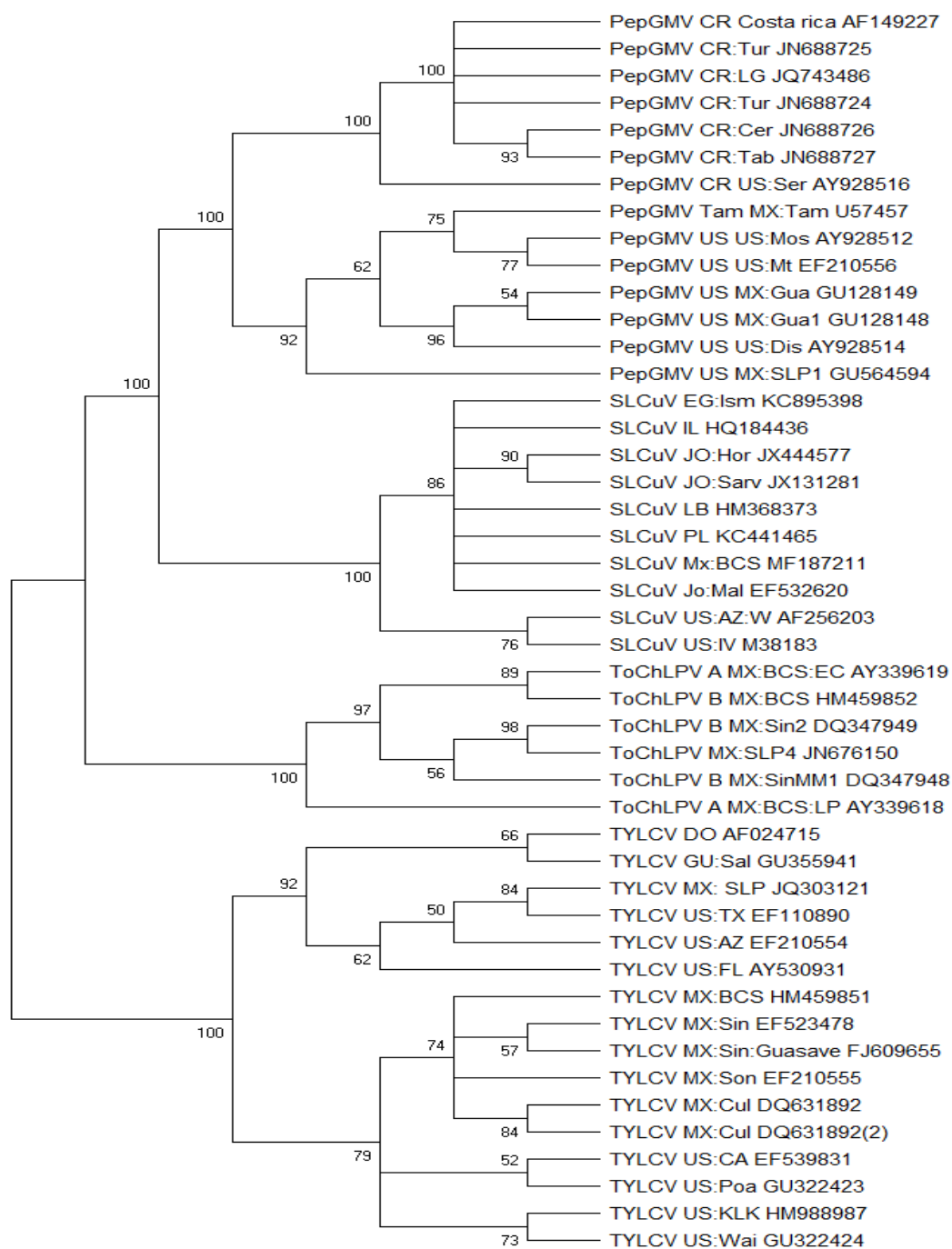


Figura 10. Dendrograma construido a partir de la mayoría de los aislamientos reportados en América de los BGVs. La historia evolutiva se infirió utilizando el método de Máxima verosimilitud basado en el modelo de parámetros de Tamura-3, con 1500 réplicas de “bootstrap” para soporte estadístico. Costa Rica (CR), Egipto (EG), Estados Unidos (US), Guatemala (GU), Israel (IL), Jordania (JO), México (MX), Nicaragua (NI), Palestina (PL) y República Dominicana (Do).

7.6 Predicción de sitios semilla

Respecto a la búsqueda específica *in silico* de los miRNAs de tomate con potencial para el silenciamiento inter-específico de especies del género *Begomovirus*, se encontraron distintas familias de miRNAs que se unen a diferentes sitios blanco (SB) que son regiones con alta complementariedad dentro de un gen (Tabla V). Algunos de estos miRNAs cumplen con los criterios establecidos y presentan potencial para silenciar a todos los integrantes de una de las especies analizadas, otros miRNAs que podrían brindar una protección cruzada con algunos de los BGVs de BCS, así como contra otros BGVs encontrados en otras partes de México, USA y Centroamérica. El análisis de los SB por especie, de todos los aislamientos, muestra que algunos genes tienen SBs que son reconocidos por más de un integrante de una familia de miRNA. Debido a la distribución desigual que presentan los miRNAs; ninguno logró reconocer a los cuatro BGVs reportados en la región; pero mir403 si posee complementariedad con PepGMV, TYLCV y ToChLPV (Tabla V).

Tabla V. Genes de BGVs reconocidos por miRNAs

miRNA y sus sitios semilla por variante								
# Acceso	Descripción	IR	C1	C2	C3	C4	V1	V2
ToChLPV								
HM459852	MX:BCS	mir482d mir1918					miR164a-3p	
AY339618	MX:BCS:LP	mir482d	mir9471(a,b)-5p	miR1918	miR1918 miR9478-5p		miR164a-3p	
AY339619	MX:BCS:EC	mir482d mir1918	mir9471(a,b)-5p miR482e-5p					
DQ347948	MX:Sin	mir482d mir1918		miR403-5p miR1918	miR403-5p miR1918 miR9478-5p			
DQ347949	MX:Sin2	mir482d mir1918		miR403-5p	miR403-5p miR9478-5p			
JN676150	MX:SLP	mir482d mir1918	miR482(d,e)-5p		miR9478-5p			

TYLCV								
HM459851	MX:BCS		miR395(a,b) miR9477-3p	miR403-5p	miR403-5p		miR395b	

AF024715	DO	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
AY530931	US:FL	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
DQ631892	MX:Cul	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
EF110890	US:TX	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
EF210554	US:AZ	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
EF210555	MX:Son	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
EF523478	MX:Sin	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
EF539831	US:CA	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
FJ609655	MX:Sin:Gu	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
GU322423	US:Poa	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
GU322424	US:Wai	miR395(a ,b) miR9477-3p	mir403 miR166c- 5p	miR166c- 5p miR403- 5p	miR395b
GU355941	GU:Sal	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
HM988987	US:KLG	miR395(a ,b) miR9477-3p	miR403- 5p	miR403- 5p	miR395b
JQ303121	MX: SLP	miR395(a ,b) miR9477-3p			miR395b
----- PepGMV					
U57457	MX:Tam		miR9478- 5p miR403-5p	miR403-5p miR9478- 5p	miR394-5p
AF149227	CR				miR156e- 3p miR394-5p
AY928512	US:Mos		miR9478- 5p	miR9478- 5p	miR394-5p
AY928514	US:Dis		miR9478- 5p	miR5302b- 3 miR9478- 5p miR397	miR394-5p
AY928516	CR:(US:Ser)				miR394-5p
EF210556	US:Mt		miR9478- 5p		miR171c
GU128148	MX:Gua1		miR9478- 5p	miR5302b- 3 miR9478- 5p miR397	miR394-5p miR164a- 3p
GU128149	MX:Gua		miR9478- 5p	miR5302b- 3 miR9478- 5p miR397	miR9472- 3p
GU564594	MX:SLP1				miR394-5p

JN688724	CR:Tur			miR156e-3p miR394-5p
JN688725	CR:Tur			miR156e-3p miR394-5p
JN688726	CR:Cer			miR156e-3p miR394-5p
JN688727	CR:Tab			miR156e-3p miR394-5p
JQ743486	CR:LG			miR156e-3p miR394-5p

SLCuV				
MF187211	Mx:BCS	miR395(a,b) miR482(d,e)-5p	miR395(a,b) miR482(d,e)-5p	
AF256203	US:AZ:W	miR395(a,b) miR9471a-5p miR482(d,e)-5p	miR482(d,e)-5p	miR164a-3p
EF532620	Jo:Mal			
HM368373	LB	miR482(d,e)-5p	miR482(d,e)-5p	
HQ184436	IL	miR482(d,e)-5p	miR482(d,e)-5p	
JX131281	JO:Sarv	miR482(d,e)-5p	miR482(d,e)-5p	
JX444577	JO:Hor	miR482(d,e)-5p	miR482(d,e)-5p	
KC441465	PL	miR482(d,e)-5p	miR482(d,e)-5p	
KC895398	EG:ism	miR482(d,e)-5p	miR482(d,e)-5p	
M38183	US:IV	miR482(d,e)-5p	miR395(a,b) miR482(d,e)-5p	

7.7 Análisis por miRNA

Se encontraron 14 familias de miRNAs de tomate con alta complementariedad en al menos un sitio del ADN-A de los begomovirus (en las 4 especies de BGVs de BCS). Los miRNAs más representativos fueron mir395 y mir403 (Fig. 11).

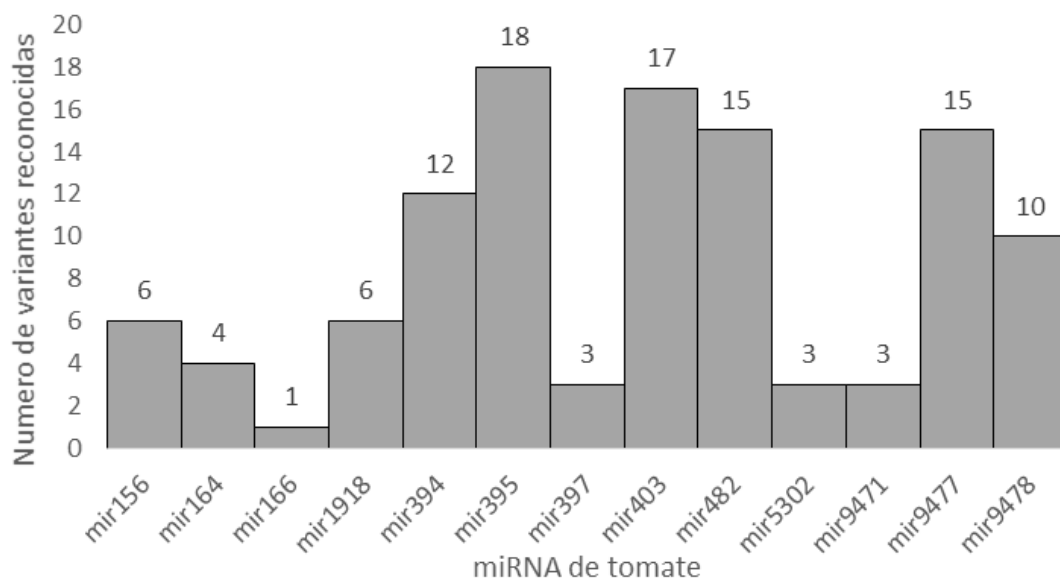


Figura 11. Representación global del número de sitios reconocidos por los miRNAs en SLCuV, ToChLPV, PepGMV y TYLCV. Todos los BGVs presentan al menos un sitio blanco (SB) reconocido por cada uno de los miRNA. La gráfica considera como un solo caso aquellos miRNA que encontraron más de un sitio blanco (SB) entre los ORFs y la IR.

Al realizar este análisis por especie se encontraron seis familias de miRNA (mir164, mir1918, mir403, mir482, mir9471 y mir9478) con diferentes sitios blanco (SB) en diferentes regiones genómicas del ToChLPV (Fig. 12a). Los miRNAs mir1918 y mir482 fueron las dos familias que se encontraron como sitios blancos en los seis aislamientos de ToChLPV, mientras que para el resto se encontraron en al menos dos aislamientos.

Del mismo modo cuatro de estas familias (mir395, mir403, mir166 y mir9477) mostraron tener SB en los diferentes genes y en la IR en los diferentes aislamientos de TYLCV (Fig. 12b), donde el mir395 y mir9477 se unen a los 15 aislamientos; el mir403 reconocen los SB en 14 aislamientos y mir166 de un aislamiento. Por otro lado, las familias de miRNAs; mir156, mir164, mir394, mir397, mir403 mir5302 y mir9478, tuvieron SB en los aislamientos de PepGMV (Fig. 12c), siendo el mir394 el que tiene mayor cantidad de sitios (13-14) con los BGVs aislados de BCS, mientras que el mir403 solo se podría unir a un aislamiento. Finalmente, las cuatro familias de miRNAs; mir482, mir164, mir395 y mir9471, mostraron poder unirse a los SB en el genoma de SLCuV (Fig. 12d).

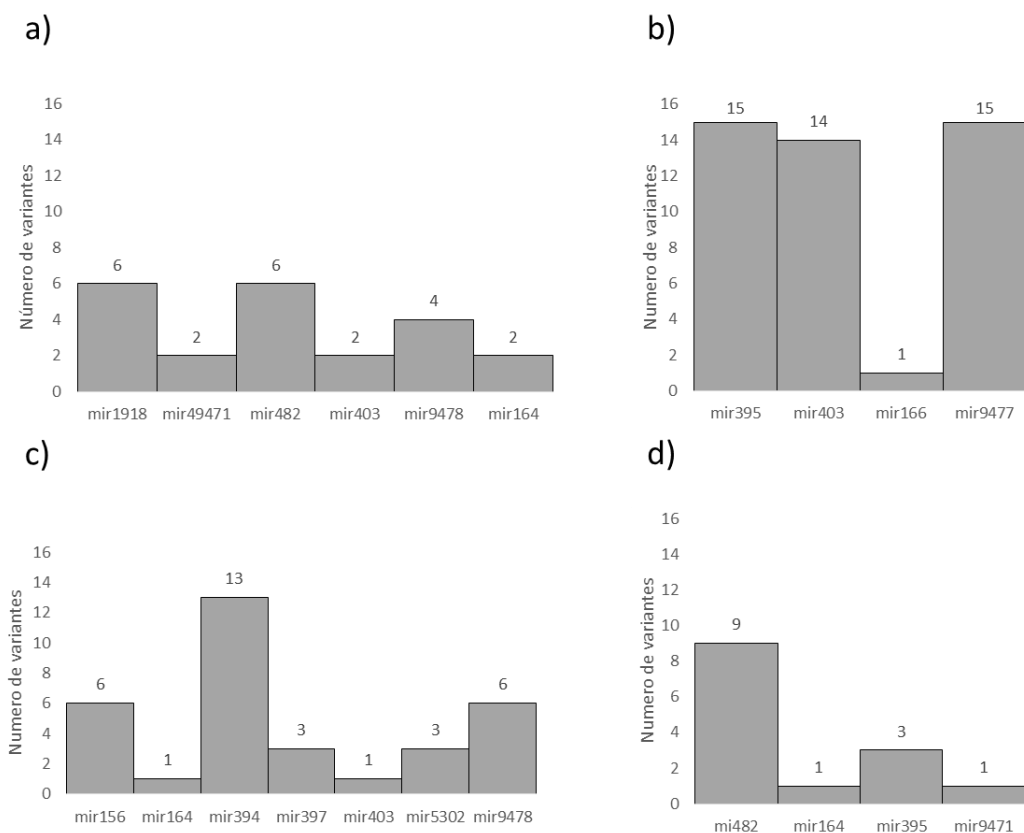


Figura 12. Representación gráfica del número de aislamientos reconocidos por los miRNAs por cada una de los virus. Número de aislamientos por BGVs con al menos un sitio blanco reconocido por cada uno de los miRNA. La gráfica considera como un solo caso aquellos miRNAs que encontraron más de un sitio blanco entre los ORFs y la IR. a) ToChLPV (6 aislamientos analizados), b) TYLCV (15 aislamientos), c) PepGMV (14 aislamientos) y d) SLCuV (10 aislamientos).

7.8 Análisis por cada secuencia viral

Al realizar este análisis por gen (IR, C1, C2, C3, C4, V1), se encontraron diferentes SB para los diferentes miRNAs (Fig. 13). Los sitios SB se encontraron principalmente en el gen C1 donde existen 46 sitios para todos los aislamientos de las especies de BGVs analizados de este sitios 26 comparte con el gen C4. También se encontraron 35 SB reconocidos por los miRNA para el gen C3 dentro de los aislamientos de BGVs analizados, de los cuales 26 que se encontraron en la región

del gen C3 están en común con el gen C2. Por su parte en la IR y V1 se encontraron 11 y 36 SB, respectivamente.

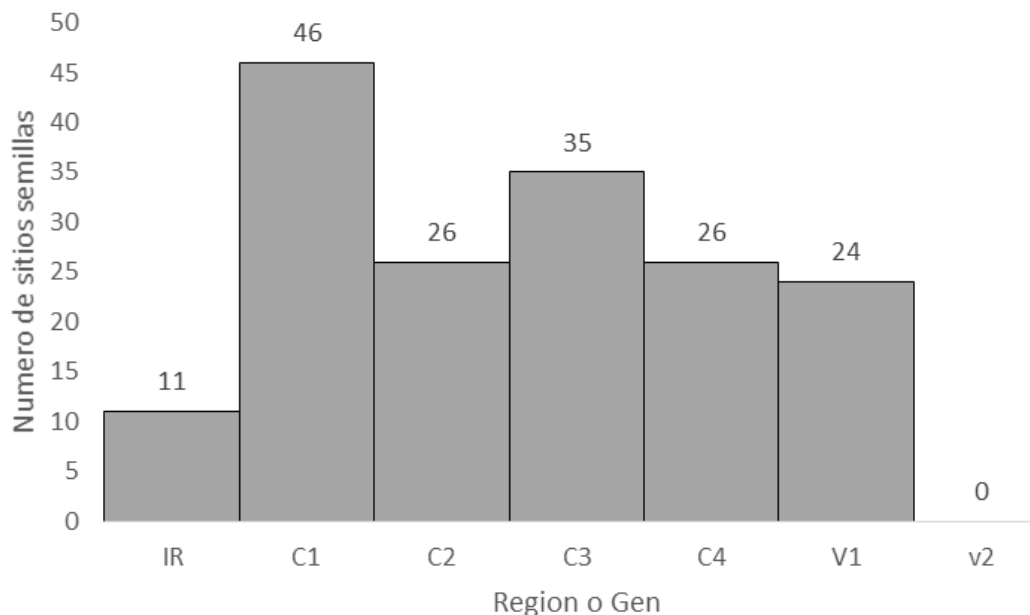


Figura 13. Número sitios blanco reconocido por los miRNA en los diferentes componentes dentro de los diferentes aislamientos de los BGVs. Número total de sitios blanco predichos in silico para todas los aislamientos (de todas las especies) de BGVs analizados. Esta gráfica considera como dos eventos diferentes si un mismo miRNA reconoce una región dos veces o bien a dos regiones diferentes en un mismo aislamiento.

El análisis después se llevó a cabo por especie y se determinó que para los aislamientos de BGVs de ToChLPV la región con mayor número de SB fue la IR en el cual mir482d y mir1918 reconocen dos sitios en los aislamientos de BGVs analizadas, dando un total de 11 sitios. En el gen C3 se encontraron 8 SB reconocidos por diferentes familias de miRNAs (Fig.14a).

En los aislamientos de TYLCV se determinó que el C1 tiene la mayor cantidad de SB, de los cuales dos familias de miRNAs (miR395a y el miR9477-3p) que reconocen dos sitios en los 15 aislamientos de TYLCV (Fig.14b).

Para PepGMV, en el gen V1 los miR156e-3p, mir171, miR394-5p y miR164a-3p reconocen 33 SB para 14 aislamientos de BGVs (Fig. 14c).

Mientras que para SLCuV el gen C1 tuvo la mayor cantidad de SB, 12 sitios reconocidos por el miR395a, el miR395b, el miR9471a-5p, el miR482d-5p y el miR482e-5p que tienen los SB en nueve de los 10 aislamientos de SLCuV (Fig 14d).

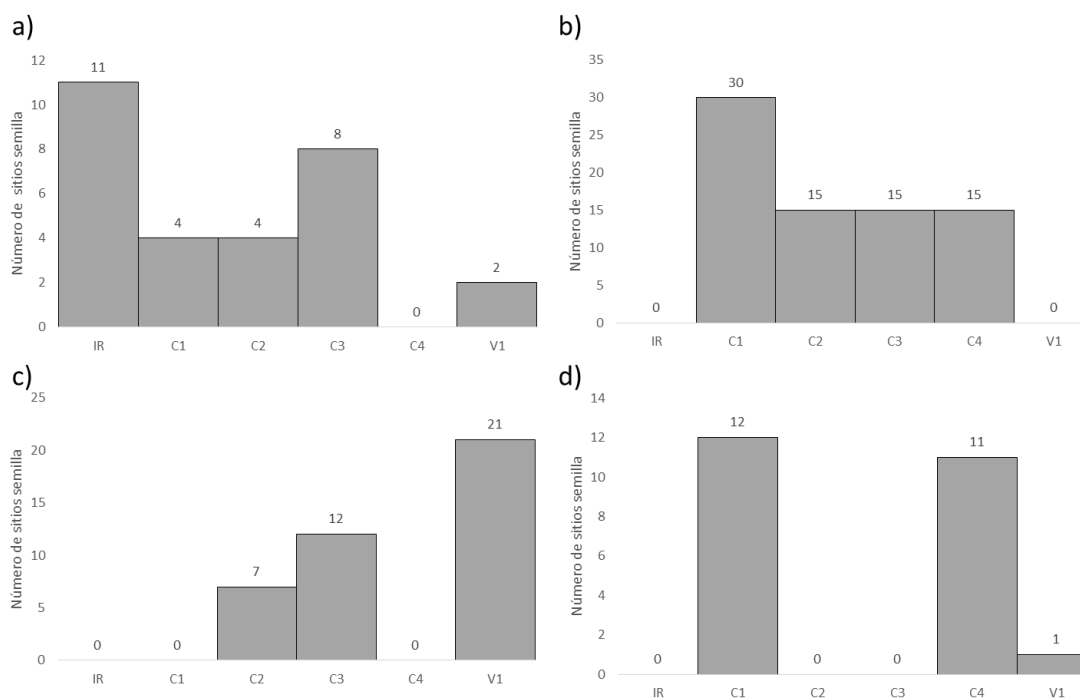


Figura 14. Representación del número de sitios blanco por gen y por begomovirus. Número total de sitios blanco *in silico* para los aislamientos de cada especie analizada: a) ToChLPV (6 aislamientos analizados), b) TYLCV (15 aislamientos), c) PepGMV (14 aislamientos) y d) SLCuV (10 aislamientos). Esta gráfica considera como dos eventos diferentes si un mismo miRNA reconoce una región dos veces o bien a dos regiones diferentes en un mismo aislamiento.

7.9 Eficiencia *in silico* de los miRNAs con potencial de silenciamiento

Se encontraron cuatro miRNAs que se unen a un sitio blanco en una sola especie en más de la mitad de los aislamientos: mir394, mir395, mir1918 y mir9477 (Tabla VI). De la misma forma se identificaron cuatro miRNAs que se unen a un sitio blanco en más de una especie con complementariedad en más de la mitad de los aislamientos de una especie: mir164, mir403, mir482 y mir9478 (Tabla VII).

No se encontró ningún miRNA con la capacidad de reconocer a los aislamientos de las cuatro especies de BGVs aquí estudiadas, ni ningún miRNA con la capacidad de unirse a todos los aislados de PepGMV.

Tabla VI. Comparación de genes blanco en el hospedero y BGVs para miRNA con los genes blanco en una sola especie

miRNA	Hospedero		Virus	
	Gen o Familia de genes	Regulación	Gen	tipo de silenciamiento
mir394	LCR	Estrés biótico y abiótico	V1 PepGMV	Degradación
mir395	SULTR2, ATPS	Estrés biótico	C1/C4 TYLCV	Degradación
mir1918	Proteína RING finger	Estrés biótico	IR ToChLPV	No clara
mir9477	Aciltransferasa	Procesos metabólicos	C1 TYLCV	Degradación

Gen o familia de genes silenciado(s) por los miRNAs en el hospedero, el tipo de regulación en el hospedero. También se muestra el gen silenciado y la especie (BGVs) así como el tipo de silenciamiento (degradación, transcripción u otro). Así, como la energía libre de Gibbs (ΔG) de la unión del miRNA con los genes del hospedero y con los genes virales.

Tabla VII. Comparación de genes blanco en hospedero y virus para miRNA con múltiples especies blanco de BGVs

	Hospedero			Virus		
	Gen	Δ G	Regulación	Gen	Δ G	silenciamiento
mir164	NAC		Regulación floración	V1 ToChLPV V1 SLCUV V1 PepGMV C2/C3 TYLCV C2		Degradación
mir403	Ago2		No clara	ToChLPV C2 PepGMV C1/C4		Degradación
mir482	FRG 3		Estrés biótico	SLCUv IR ToChLPV C2, C3		Degradación
mir947 8	*		*	PepGMV C2 ToChLPV		Degradación

Gen o familia de genes silenciado(s) por los miRNAs en el hospedero, el tipo de regulación en el hospedero. También se muestra el gen silenciado y la especie (BGVs) así como el tipo de silenciamiento (degradación, transcripción u otro). Así, como la energía libre de Gibbs (ΔG) de la unión del miRNA con los genes del hospedero y con los genes virales.

8. DISCUSIÓN

Las cuatro especies de BGVs presentes en BCS fueron analizadas en el presente estudio, las cuales han sido previamente caracterizadas, tres BGVs son bipartitas (ToChLPV, SLCuV y PepGMV) y uno monopartita (TYLCV). Para aumentar la probabilidad de éxito en el diseño una construcción de RNAi para proteger cultivos de tomate frente a los BGVs de una región es indispensable considerar la diversidad de BGVs en la región de interés. También, es posible que con una o varias construcciones se busque hacer frente a ciertas especies de BGVs en específico, bien por razones prácticas o bien por la posible interacción de estos en un hospedero. En este sentido una construcción podría estar diseñada para hacer frente a una variante específica de una especie o bien a varias variantes o a toda una especie.

El objetivo de este estudio fue encontrar miRNAs que cumplan los requisitos para ser usados en un sistema de protección contra BGVs mediante el uso de construcciones de RNAi y que a la par reconozcan las variantes actuales y emergentes de la misma especie o frente diferentes especies reportadas en regiones cercanas. Ya que estas pueden ser introducidas al estado por fenómenos naturales o antropogénicos como el comercio de plantas o bien desplazamiento del insecto vector.

Usar como criterio la selección de secuencias reportadas en el ICTV como variantes por el ICTV permitió reducir el número de politomias, lo que se reflejó en la conservación de la estructura del árbol filogenético obtenido y dio lugar a que el análisis fuera robusto. El dendograma generado en este trabajo mostro que todas las variantes de PepGMV presentes en México tienen un ancestro común posterior al fenómeno que dio origen a la especie; el cual además, es diferente al del ancestro que presentan las variantes de Costa Rica lo cual es consistente con reportes previos (Lotrakul *et al.*, 2000). Por otra parte el árbol filogenético muestra que las variantes de TYLCV presentes en México pueden haber sido originados por dos eventos de introducción, al igual que para PepGMV (Dufy y Holmes, 2007). En el caso de SLCuV pese a que todas las variantes mostraron ser filogenéticamente más

cercanas que el resto de las especies, es decir poseen un ancestro común más reciente que para el resto de BGVs de BCS, SLCuV parece estar filogenéticamente más relacionado a las variantes de medio oriente que a las variantes de América. Esto último había sido previamente reportado (Medina-Hernández *et al.*, 2019) y aunque se requieren más estudios para comprobar el origen geográfico de la especie este resultado indica que es posible que exista intercambio de material genético entre BCS y medio oriente. En el caso de las variantes de ToChLPV el dendograma muestra que todas las variantes de esta especie son más cercanas filogenéticamente que entre las variantes de otras especies como TYLCV. Estos resultados reflejan la complejidad de los fenómenos de dispersión de los virus y la variabilidad/diversidad que puede existir aun dentro de los virus de en una misma región aun si pertenecen a una sola especie. Por lo que para la selección de miRNAs para ser usados en una construcción de RNAi se decidió considerar a las variantes presentes en las regiones cercanas, esto tomando en cuenta futuros fenómenos de dispersión. Además, se consideró que para una especie las variantes genéticas al tener diferentes orígenes filogeográficos podrían no ser reconocidos por los mismo miRNAs.

Una primera aproximación para la búsqueda de regiones conservadas dentro de cada especie consistió en la comparación de la secuencias de todas las variantes analizadas mediante alineamientos locales.

Al comparar los genomas analizados, con un porcentaje de identidad >97% como criterio de corte, se determinó que las variantes de TYLCV aisladas en BCS muestran un contraste con los resultados obtenidos en el análisis filogenético. La razón de esta diferencia se puede deber a que el criterio de análisis se enfocó en regiones específicas (alineamientos locales o restringidos) para así calcular el porcentaje de identidad, mientras en el análisis filogenético el criterio principal se basó en el alineamiento de toda la secuencia del genoma.

En el caso de ToChLPV se determinó un comportamiento opuesto, dando lugar a una mayor divergencia en el análisis filogenético. Por otro lado, para las especies de SLCuV (>96%) y PepGMV (<95%) el análisis de las regiones y el alineamiento

de los genomas mostró un comportamiento congruente con la proximidad filogenética entre las variantes de estas especies.

Análisis similares en los que se compararon cada ORF, proteína y las IRs por separado, revelaron que dentro cada especie existen ORFs más conservados (mayor porcentaje de identidad) que el resto de las regiones no codificantes. Además, en las gráficas de similitud se determinó que dentro de las secuencias de los ORFs existen regiones más conservadas entre las variantes de las especies analizadas. Estos resultados concuerdan con lo que ha sido reportado previamente en los que se caracterizó a los BGVs de una región geográfica, indicando que estas variaciones son consecuencia de las fuerzas evolutivas que derivan en diferentes grados de presión de selección sobre cada gen (Medina *et al.*, 2019; Seal *et al.*, 2006). Ambos análisis mostraron una mayor variabilidad en la región correspondiente a los genes C1/C4 (Tabla II, Fig. 3), lo cual concuerda con estudios realizados en otros BGVs que indican que esta región suele ser muy variable (Hosseinzadeh *et al.*, 2014; Xie *et al.*, 2013). La región correspondiente al gen V1 fue una de las más conservadas, lo que concuerda con lo descrito por Brown *et al.*, 1996. Estos análisis permiten identificar regiones en las que existe menor variación con respecto a una secuencia de referencia.

Al análisis de la diversidad nucleotídica de cada uno de los ORFs y las IRs por separado (Tabla II) mostró que TYLCV, SLCuV, PepGMV tienen una menor diversidad (Π) que ToChLPV. En las variantes de PepGMV existe una mayor variación de secuencias. El valor negativo para el estadístico D calculado para el DNA-A y la mayoría de los ORF de estas las especies analizadas en el estado de BCS puede indicar que en estos BGVs la baja diversidad sea producto de un cuello de botella evolutivo (Gutiérrez *et al.*, 2002; Hosseinzadeh *et al.*, 2014; Lima *et al.*, 2017; Melgarejo *et al.*, 2013; Ramesh *et al.*, 2007; Sánchez-Campos *et al.*, 2013; Xie *et al.*, 2013).

La diversidad calculada para ToChLPV fue la más alta y pareciera enfrentarse a una situación evolutiva diferente a la de las otras 3 especies; los genes C1 y C2 de este BGVs fueron los que mostraron mayor diversidad, por lo que este debería ser un

criterio a considerar para evitar miRNAs que flanquean estas regiones debido al número de variaciones. El resto de los genes de ToChLPV mostraron un valor de diversidad similar a la diversidad calculada para todos los genes de PepGMV. Pese a que los valores de diversidad calculados para ToChLPV y PepGMV son varias veces mayor que los mostrados por TYLCV y SLCuV no podemos hablar de una alta diversidad dentro del género pues valores similares (>0.04) de diversidad nucleotídica fueron calculados para otros BGVs (Hosseinzadeh *et al.*, 2014; Lima *et al.*, 2017).

La conservación del gen V1 se puede explicar por la importancia que tiene la secuencia de aminoácidos en la estructura y función de la proteína de la cápside (CP). Del mismo modo que se puede explicar la baja diversidad de C2 y C3 a partir de su papel en la replicación viral y la importancia que tiene su estructura para llevar a cabo esta función (Gutiérrez *et al.*, 2002).

Las diferencias en los valores de identidad y diversidad nucleotídica para C1 y C4 de TYLCV y SLCuV como se muestra se debe a la acumulación de mutaciones en la región de solapamiento de estos dos genes. Además, las mutaciones en esta región (C1/C4) tienen efectos diferentes en las proteínas que cada uno codifica pues tienen marcos de lectura diferentes y desfasados por 2 nucleótidos. Todo esto deriva en la acumulación de mutaciones no sinónimas (nS) en C4 lo que a su vez afecta el porcentaje de similitud para la proteína codificada. En cambio para C1 la mayoría de las mutaciones son mutaciones sinónimas (S) lo que deriva en un mayor porcentaje de similitud entre las proteínas Rep de cada especie.

Por otro lado, las variantes de ToChLPV muestran justo lo contrario una menor acumulación de mutaciones en la región C1/C4 respecto al resto de C1. La diferencia en estos valores para los genes C1 y C4, indican que para una misma especie puede haber diferentes tipos de presión de selección. En las 4 especies la mayoría de los genes se encontró bajo una presión negativa pues la mayoría de las mutaciones son sinónimas lo que conserva las propiedades de las proteínas que codifican. Sin embargo para TYLCV la presión de selección es positiva para C4 mientras que es negativa para C1. Como ya se explicó pese a estar presentes en

la misma región estos genes tienen marcos de lectura diferentes. Es posible que como se reportó para otros virus (Xie *et al.*, 2013) la mayoría de las mutaciones se encuentre en el tercer nucleótido de C1 y que corresponde al primer nucleótido de C4 lo que explicaría estos resultados. Este fenómeno también tiene un sentido evolutivo pues se ha reportado que las variaciones en la proteína C4 tienen un papel importante en el reconocimiento de nuevas variantes (Sánchez-Campos *et al.*, 2013) mientras que cambios en la proteína de la replicación (Rep) puede derivar en efectos negativos en el ciclo de replicación viral (Gutierrez C., 2002, Ramesh *et al.*, 2007). Melgarejo y colaboradores dieron una explicación biológica al gran número mutaciones no sinónimas del gen C4 de algunos virus, debido a la participación de C4 en proceso de infección (secuestrando el mecanismo de silenciamiento) los cambios en esta proteína podrían ser cruciales en el reconocimiento de nuevos hospederos (Melgarejo *et al.*, 2013). Esto último contribuiría la diseminación del virus y jugaría un papel importante en virulencia del mismo. De hecho de los virus estudiados en este trabajo TYLCV es el virus con mayor número reportes en México, también posee un gran número de especie de hospederos y el que se encuentra más ampliamente distribuido. Por lo que su baja diversidad o poca variación podría estar vinculado a su “éxito” como virus. Por otro lado gran diversidad calculada para el resto de los virus que se traduce en mutaciones sinónimas podría ser parte de una estrategia de los virus para evitar los mecanismos de silenciamiento innato de los hospederos (Lima *et al.*, 2017).

Los miRNAs seleccionados tienen complementariedad en regiones altamente conservadas en los genomas de los BGVs. Además, varios miRNAs seleccionados tienen sitios blancos (SB) en las regiones más conservadas de los genes con alta variabilidad. Al respecto, se determinó que un incremento en la variabilidad en cualquiera de las regiones en el genoma de los BGVs analizados, podría incrementar el número de miRNAs que podrían reconocer posibles SB, lo cual influye de manera negativa en la cantidad de variantes de cada especie de BGV que podrían ser reconocidos por los miRNAs. Esto podría sugerir que si los distintos aislamientos de cada especie de BGV tuvieran un índice alto de mutaciones

(diversidad nucleotídica), por ende, habría un aumento en la cantidad de posibles SB para los diferentes miRNAs, ya que habría un mayor número de combinaciones viables (secuencias de reconocimiento ligeramente distintas); lo cual se traduce en cada una de las variaciones puede ser objetivo de distintos miRNAs. Por ejemplo dos de las cuatro familias de miRNAs (el miR395 y el miR9477) tienen la capacidad de reconocer SB en todas las variantes de TYLCV analizadas. Mientras que en SLCuV otro virus con baja variabilidad el miR482 tiene sitios blanco en la mayoría de las variantes. En el caso de PepGMV no se encontró ningún miRNA que reconociera todas las variantes de este BGV. En el caso de ToChLPV, el virus con mayor variabilidad, se encontró que dos familias de miRNAs (mir482 y el mir1918) tienen un sitio blanco en una región conservada en la IR en la mayoría de las variantes esta especie. Es importante mencionar que mir1918 tiene sitios diana en los genes C2 y C3 de las variantes de ToChLPV de Sinaloa y de BCS. Mientras que el mir482 sitios blanco en la variante de San Luis Potosí y sitios blanco en la mayoría de los aislamientos de SLCuV. De hecho en todos los BGVs analizados algunos miRNAs podrían unirse a SB de dos genes de una misma especie o una región compartida por dos genes como C1/C4 y C2/C3. En este sentido el silenciamiento de dos SB por un mismo miRNA implica que este miRNA tendrá un mayor impacto en el ciclo de replicación viral. Por otro lado, mir9478 tiene sitios blancos en el gen C3 de cuatro de las variantes analizadas de ToChLPV. Aunque existen estudios que muestran que el silenciamiento de este gen podría no tener gran impacto en el ciclo viral (Tripathi *et al.*, 2018), además, este mismo miRNA tiene sitios blancos en todas las variantes de PepGMV de México y Estados Unidos. Finalmente mir9471, mir164 y mir403 reconocen dos de las variantes de ToChLPV cada uno por lo que no sería conveniente usarlos como sistema de protección frente a esta especie. En el caso de TYLCV quizá uno de los mejores candidatos es el mir9477 que reconoce una región en C1 en todos los aislamientos analizados por lo tanto podría afectar la expresión de este gene; el mir395 tienen sitio blanco en la región de solapamiento de C1/C4 en todos los aislamientos analizados; mientras el mir403 podría reconocer la región C2/C3 en 14 de los 15 aislamientos. El reporte de Piedra-

Aguilera *et al.* (2019) acerca de la expresión de vi-siRNAs de TYLCV durante una infección en tomate determinaron que dos de los sitios donde se alinearon una gran cantidad de los vi-siRNAs fue en las regiones C2/C3 y C1/C4. Además, en este reporte se determinó que durante la infección existe un incremento en la expresión de los genes V1/V2 (75% de las lecturas). En este sentido es posible que la expresión de estos vi-siRNAs sea resultado de la actividad de los miRNA de tomate, es decir, que sean vi-siRNAs secundarios de origen viral. En los análisis realizados, se sugiere que tanto mir395, mir403 y el mir9477 se podrían unir a C1/C2 y C3/C4, respectivamente. Estos miRNAs tienen el potencial de dirigir la degradación de genes virales. Además, estos miRNAs tienen un tamaño de 22nt, 21nt y 21nt pares de bases, respectivamente; que es el tamaño reportado de los miRNAs que participan en la biogénesis de siRNAs secundarios. Se ha reportado que la biogénesis de siRNAs secundarios está dirigida por un miRNA de 22nt o dos miRNAs de 21nt (Cuperus *et al.*, 2010). Finalmente debido a que el mir166 solo reconoce una variante no es un candidato a ser usado en el silenciamiento de TYLCV ni de ninguno de los virus analizados.

Para hacer frente a PepGMV se podría usar el miR394-5p en un sistema de silenciamiento pues tiene sitios diana en la región de V1 en todas variantes analizadas; por otro lado todas las variantes de Costa Rica analizadas tienen sitios blancos para miRNAs de la familia mir156 en AV1. Mientras que los aislados de México tienen sitios blancos para el mir9478 en la región de C2. Finalmente el miR397 y el miR5302 reconocen muy pocas variantes de PepGMV y de otros virus por que se descartan como candidatos a ser usados en un sistema de silenciamiento.

Para SLCuV el único buen candidato a ser usado en un sistema de silenciamiento es el mir482 pues tiene un sitio blanco en la región de C1/C4 de la mayoría de las variantes de esta especie. La familia de mir971 no son buenos candidatos pues solo reconocen una variante, el mir395 reconoce pocas variantes de este virus pero podría reconocer la mayoría de las variantes de TYLCV.

Respecto a esto se sabe que la sobreexpresión de mir394 en *A. thaliana* tiene efectos sobre el proceso de desarrollo de las plantas y en algunos experimentos lo relacionan a la sobre expresión de este microRNA con la tolerancia a estrés abiótico, además también se asocia con la susceptibilidad a estrés biótico (hongos patógenos) (Ayubov *et al.*, 2019; Rajwansh *et al.*, 2019). Por otro lado estimamos que este microRNA tendría una unión estable con el gen AV1 del PepGMV lo que reduciría el número de partículas virales pero tendría un efecto “bajo” sobre la replicación del genoma viral y no afectaría la expresión del resto de proteínas virales. Por lo que podría considerarse como un mal candidato para la defensa contra PepGMV pese a que los efectos de su sobre expresión podrían no ser muy grandes. Del mismo modo, la sobre expresión del mir395 en *A. thaliana* muestra tener efectos negativos en el desarrollo de la raíz y hojas; pero, únicamente en condiciones de deficiencia de sulfato (Singh *et al.*, 2012), como ya se mencionó este tiene potencial de unirse con los genes C1 y C4 los cuales participan activamente en la replicación del virus (Kumar *et al.*, 2017). Además, la región en la que este miRNA reconoce su SB en el genoma de TYLCV es la región que mayor potencial de unión de sRNAs virales. Estas características lo hace un miRNA candidato para ser usado en una construcción para conferir resistencia a TYLCV, sin embargo, se debe requerir más análisis para determinar su posible papel en la generación de siRNAs secundarios y su relación con el mir403. En este sentido un ensayo para determinar la eficiencia podría servir a la vez para determinar los mecanismos de silenciamiento viral en los que participa.

En plantas de tomate la sobre expresión del mir403 afecto negativamente el tiempo de floración, la viabilidad de las semillas y la morfología de las hojas (Zhang *et al.*, 2015). En el análisis realizado se encontró que el mir403 podría regular el gen AC2 de TYLCV, PepGMV y ToChLPV. Sin embargo, los efectos negativos de la sobre expresión de este miRNA en el desarrollo del tomate, dan lugar a descartarlo como candidato para ser usado en una construcción contra los BGVs. En cambio el mir482 regula genes relacionados a la defensa contra hongos y en algunas especies la sobre expresión de este miRNA reduce la susceptibilidad a algunos hongos (Zhu

et al., 2013). Además, este miRNA puede regular los genes AC1, C2 y C3 lo que puede afectar la replicación y transcripción del begomovirus por lo que este puede ser un candidato a ser usado en un sistema de defensa a begomovirus. Finalmente, el papel del mir9478 en el desarrollo del tomate no es del todo claro pero en este trabajo encontramos que puede unirse en los genes C2 y C3 de algunos begomovirus. Por lo tanto son necesarios trabajos para dilucidar el papel de este miRNA en el desarrollo del tomate y en la defensa contra estos begomovirus.

En lo referente al mir1918 se reportó que la sobre expresión de este miRNA incrementa la susceptibilidad del tomate a hongos patógenos (*Phytophthora infestans*) (Luan *et al.*, 2016) por lo que la sobreexpresión de este miRNA en campo podría ser contraproducente brindando protección contra ToChLPV pero dejando vulnerable a los cultivos contra el ataque de hongos, debido a esto no se considera como un buen candidato para un sistema de defensa contra begomovirus. Por otro lado, el mir9477 tiene una actividad directa en la regulación de aciltransferasas y metiltransferasas en plantas de tomate, pero aún se desconoce el efecto de la sobre expresión de este miRNA en el desarrollo de estas plantas (Candar-Cakir *et al.*, 2016). Debido a que se desconocen los efectos negativos en el desarrollo y por los análisis realizados en este estudio, este miRNA puede ser un candidato de estudio para ser usado en la defensa contra los BGVs en plantas de tomate.

En conjunto, los análisis bioinformáticos realizados revelaron que varias regiones de los genomas de SLCuV, TYLCV, PepGMV y ToChLPV poseen una alta probabilidad de ser objetivos de varios miRNAs. Estos miRNAs podrían unirse en las regiones conservadas de los ORFs. Estos ORFs codifican para los genes C1, C2, C4 y V1. Este análisis *in silico* brinda evidencias *a priori*, de que algunos de los miRNA candidatos podrían tener consecuencias negativas en el desarrollo de las plantas, lo cual es criterio de exclusión para ser usados en sistemas de silenciamiento génico mediado por sRNAs. Sin embargo, los resultados obtenidos abren la posibilidad para establecer la base de miRNAs candidatos para ser usados en construcciones que los sobre-expresen y evaluar sus eficiencias de protección

contra BGVs. Además, así se podría determinar el efecto de la sobre expresión de estos miRNAs en el desarrollo de la planta de tomate.

9. CONCLUSIONES

- Los estudios de diversidad genética, los análisis filogenéticos y el análisis in *silico* son herramientas indispensables, para una selección específica de los miRNAs y de los sitios blancos, los cuales influyen en el diseño de una construcción para las metodologías de RNAi y el silenciamiento de una amplia gama de BGVs que afectan a los cultivos.
- En begomovirus con alta variabilidad genética, existen regiones suficientemente conservadas como para ser un sitio blanco de algunos miRNAs, lo cual nos indica que estos miRNA juegan un papel importante en la relación hospedero/virus.
- Proponemos como candidatos a mir395 y mir9477 ya que presentan objetivos blancos en TYLCV, SLCuV, PepGMV, ToChLPV, además de que para estos miRNAs no se han reportaron efectos secundarios para el hospedero.
- mir395 y mir9477 presentan potencial para ser empleados en la activación de un sistema de protección como el silenciamiento génico o RNAi, ante la presencia de infecciones mixtas begomovirales.

10. LITERATURA CITADA

Abrahamian, P., H. Sobh, R. Seblani, Y. Abou-Jawdah. 2015. Co-infection of two criniviruses and a begomovirus enhances the disease severity in cucumber. *Eur. J. Plant Pathol.* 142(3):521-530.

Akmal, M., M. S.Baig, J. A.Khan. 2017. Suppression of cotton leaf curl disease symptoms in *Gossypium hirsutum* through over expression of host-encoded miRNAs. *J. Biotechnol.* 263:21-29.

Ali, I., I. Amin, R. W.Briddon, S. Mansoor. 2013. Artificial microRNA-mediated resistance against the monopartite begomovirus Cotton leaf curl Burewala virus. *Virol. J.* 10:1-9.

Amin, I., B. L. Patil, R. W.,Briddon, S. Mansoor, C. M. Fauquet. 2011. A common set of developmental miRNAs are upregulated in *Nicotiana benthamiana* by diverse begomoviruses. *Virol. J.* 8:1-10.

Ascencio-Ibáñez, J. T., Z. I. Monsalve-Fonnegra, M. B. Pruna-Camacho, R. Díaz-Plaza, R. F. Rivera-Bustamante. 1999. Los Geminivirus. *Revista Mexicana de Fitopatología* 17(2):113-127.

Ayubov, M. S., M. H. Mirzakhmedov, V. R. Sripathi, Z. T.Buriev, K. A. Ubaydullaeva, D. E. Usmonov, I. Y. Abdurakhmonov. 2019. Role of MicroRNAs and small RNAs in regulation of developmental processes and agronomic traits in *Gossypium* species. *Genomics.* 111(5):1018-1025.

Baig, M. S., J. A. Khan. 2013. Identification of *Gossypium hirsutum* miRNA targets in the genome of cotton leaf curl multan virus and betasatellite. *Indian J. Biotechnol.* 12(3):336-342.

Bañuelos-Hernández, B., J. A. Mauricio-Castillo, Y. Cardenas-Conejo, R. G. Guevara-González, G. R. Arguello-Astorga. 2012. A new strain of tomato severe leaf curl virus and a unique variant of tomato yellow leaf curl virus from Mexico. *Arch. Virol.* 159(9):1835-1841.

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

Bojórquez, R. M. C., J. G. Gallego, P. S. Collado. 2013. Propiedades funcionales y beneficios para la salud del licopeno. *Nutrición Hospitalaria.* 28(1):6-15.

Borah, B. K., F. Zarreen, G. Baruah, I. Dasgupta. 2016. Insights into the control of geminiviral promoters. *Virology* 495:101-111.

- Borges, F., R. A. Martienssen. 2015. The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16(12):727-741.
- Brodersen, P., J. A. Cutts, A. D. Howard, B. H. Lewis, S. W. Squyres, K. E. Herkenhoff, O. Voinnet. 2008. Widespread Translational Inhibition by Plant miRNAs and siRNAs. *Science.* 320(5880):1185-1190.
- Brown, J. K., F. M. Zerbini, J. Navas-Castillo, E. Moriones, R. Ramos-Sobrinho, J. C. F. Silva, E. Fiallo-Olivé, R. W. Briddon, C. Hernández-Zepeda, A. Idris, V. G. Malathi, D. P. Martin, R. Rivera-Bustamante, S. Ueda, A. Varsani. 2015. Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Arch. Virol.* 160:1593-1619.
- Camejo, D., W. Torres. 2000. La salinidad y su efecto en los estadios iniciales del desarrollo de dos cultivares de tomate (*S. lycopersicum*). *Cultivos Tropicales.* 21(2):23-26.
- Candar-Cakir, B., E. Arican, B. Zhang. 2016. Small RNA and degradome deep sequencing reveals drought-and tissue-specific micrnas and their important roles in drought-sensitive and drought-tolerant tomato genotypes. *Plant Biotechnol. J.* 14(8):1727-1746.
- Carthew, R. W., E. J. Sontheimer. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell.* 136(4):642-655.
- Chen, H., A. A. Arsovski, K. Yu, A. Wang. 2016. Genome-Wide Investigation Using sRNA-Seq, Degradome-Seq and Transcriptome-Seq Reveals Regulatory Networks of microRNAs and Their Target Genes in Soybean during Soybean mosaic virus Infection. *Plos One*, 11(3):e0150582. <https://doi.org/10.1371/journal.pone.0150582>
- Cruz Bojórquez R. M., J. González-Gallego, P. Sánchez Collado. 2013. Propiedades funcionales y beneficios para la salud del licopeno. *Nutrición Hospitalaria.* 28(1):6-15.
- Cuellar, W. J., M. Galvez, S. Fuentes, J. Tugume, J. Kreuze. 2015. Synergistic interactions of begomoviruses with Sweet potato chlorotic stunt virus (genus Crinivirus) in sweet potato (*Ipomoea batatas* L.). *Mol. Plant Pathol.* 16(5):459-471.
- Cuperus, J. T., A. Carbonell, N. Fahlgren, H. Garcia-Ruiz, R. T. Burke, A. Takeda, J. C. Carrington. 2010. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. *Nat. Struct. Mol. Biol.* 17(8):997.
- D'Ario, M., S. Griffiths-Jones, M. Kim, M. D'Ario, S. Griffiths-Jones, M. Kim. 2017. Small RNAs: Big Impact on Plant Development. *Trends Plant Sci.* 22(12):1056-1068.

Dai, X., P. X. Zhao. 2011. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 39(suppl_2):W155-W159.

Dai, X., Z. Zhuang, P.X. Zhao. 2011. Computational analysis of miRNA targets in plants: current status and challenges. *Brief. Bioinformatics.* 12(2):115-121.

Dai, X., Z. Zhuang, Z., P. X. Zhao. 2018. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46(W1):W49-W54.

Darré, M. 2019. Factores de pre y poscosecha que afectan el contenido de compuestos antioxidantes en hortalizas. Tesis (Doctorado). La Plata, Buenos Aires, Argentina. Universidad Nacional de la Plata. Facultad de Ciencias Agrarias y Forestales. 178 p.

Deng, P., S. Muhammad, M. Cao, L. Wu. 2018. Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants. *Plant Biotechnol. J.* 16(5):965-975.

Duffy, S., E. C. Holmes. 2007. Multiple introductions of the Old World begomovirus Tomato yellow leaf curl virus into the New World. *Appl. Environ. Microbiol.* 73(21):7114-7117.

Ellwanger, D. C., F. A. Büttner, H. W. Mewes, V. Stümpflen. 2011. The sufficient minimal set of miRNA seed types. *Bioinformatics.* 27(10):1346-1350.

Fei, Q., R. Xia, B. C. Meyers. 2013. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell.* 25(7):2400-2415.

García-Ruiz, H., M. T. García-Ruiz, S. M. Gabriel-Peralta, C. B. Miravel-Gabriel, K. El-Mounadi. 2016. Mecanismos, aplicaciones y perspectivas del silenciamiento génico de virus en plantas. *Revista Mexicana de Fitopatología.* 34(3):1-22.

Goykovic Cortés, V., G. Saavedra del Real. 2007. Algunos efectos de la salinidad en el cultivo del tomate y prácticas agronómicas de su manejo. *Idesia.* 25(3):47-58.

Guleria, P., M. Mahajan, J. Bhardwaj, S. K. Yadav. 2011) Plant Small RNAs: Biogenesis, mode of action and their roles in abiotic stresses. *Genomics Proteomics Bioinformatics.* 9(6):183-199.

Gutierrez, C. 2002. Strategies for geminivirus DNA replication and cell cycle interference. *Physiol. Mol. Plant Pathol.* 60(5):219-230.

Hanley-Bowdoin, L., E. R. Bejarano, D. Robertson, S. Mansoor. 2013. Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nat. Rev. Microbiol.* 11:777-788.

Hanley-Bowdoin, L., S. B. Settlage, B. M. Orozco, S. Nagar, D. Robertson. 2000. Geminiviruses - models for plant DNA replication, transcription and cell cycle regulation. *Crit. Rev. Biochem. Mol. Biol.* 35:105-140.

Holen, T. 2006. Efficient prediction of siRNAs with siRNA rules 1.0: An open-source JAVA approach to siRNA algorithms. *RNA.* 12(9):1620-1625.

Holguín-Peña, R. J., R. Vazquez-Juarez, R. F. Rivera-Bustamante. 2005. A new begomovirus causes tomato leaf curl disease in Baja California Sur, Mexico. *Plant Dis.* 89(3):341-341.

Holguín-Peña, R. J., G. R. Arguello-Astorga, J. K. Brown, R. F. Rivera-Bustamante. 2006. A new strain of Tomato chino La Paz virus associated with a leaf curl disease of tomato in Baja California Sur, Mexico. *Plant Dis.* 90(7):973-973.

Holguín-Peña, R. J., L. G. Hernández-Montiel, H. Latisnere-Barragán. 2010. Identificación y distribución geográfica de *Bemisia tabaci* Gennadius y su relación con enfermedades begomovirales en tomate (*Solanum lycopersicum* L.) de Baja California, México. *Revista Mexicana de Fitopatología.* 28(1):58-60.

Holguín-Peña, R.J., R. Vázquez-Juárez, R. R. F. Rivera Bustamante. 2004. Rango de Hospedantes, Incidencia y Filogenia del Virus del Mosaico Dorado del Chile (PepGMV) en Baja California Sur, México. *Revista Mexicana de Fitopatología.* 22(2):206-215.

Hosseinzadeh, M. R., M. Shams-Bakhsh, S. K. Osaloo, J. K. Brown. 2014. Phylogenetic relationships, recombination analysis, and genetic variability among diverse variants of tomato yellow leaf curl virus in Iran and the Arabian Peninsula: Further support for a TYLCV center of diversity. *Arch. Virol.* 159(3):485-497.

Hsu, P. W. C., L. Z. Lin, Hsu, S. D. Hsu, J. B. K., Hsu, H. D. Huang. 2007. ViTa: Prediction of host microRNAs targets on viruses. *Nucleic Acids Res.* 35(SUPPL. 1):381-385.

Inoue-Nagata, A. K., L. C. Albuquerque, W. B. Rocha, T. Nagata. 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage ϕ 29 DNA polymerase. *J. Virol. Methods.* 116(2):209-211.

Inoue-Nagata, A. K., M. F. Lima, R. L. Gilbertson. 2016. A review of geminivirus diseases in vegetables and other crops in Brazil: current status and approaches for management. *Hortic. Bras.* 34(1):8-18.

Kamthan, A., A. Chaudhuri, M. Kamthan, A. Datta. 2015. Small RNAs in plants: recent development and application for crop improvement. *Front. Plant Sci.* 6:1-17.

- Kertesz, M., N. Lovino, U. Unnerstall, U. Gaul, E. Segal. 2007. The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39(10):1278-1284.
- Khalid, S., M. Zia-ur-Rehman, U. Hameed, F. Saeed, F. Khan, M. S. Haider. 2017. Transmission specificity and coinfection of mastrevirus with begomovirus. *Int. J. Agric.Biol.* 19(1):105-113.
- Khraiwesh, B., J. K. Zhu, J. Zhu. 2012. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta Gene Regul. Mech.* 1819(2):137-148.
- Kumar, S., A. Srivastava, A. Kumari, R. Raj, M. Jaidi, S. K. Raj. 2017. Begomovirus disease management measures, now and then. In: Saxena S., A. Tiwari (eds.). *Begomoviruses: occurrence and management in Asia and Africa.* Springer. Singapore. pp 71-92.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, K. Tamura. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35(6):1547-1549.
- Lebsky, V., J. Hernández-González, G. Arguello-Astorga, Y. Cardenas-Conejo, A. Poghosyan. 2011. Detection of phytoplasmas in mixed infection with begomoviruses: A case study of tomato and pepper in Mexico. *Bull. Insectology.* 64 (Supplement): S55-S56.
- Li, S., C. Castillo-González, B. Yu, B., X. Zhang. 2017. The functions of plant small RNA s in development and in stress responses. *Plant Journal.* 90(4):654-670.
- Lima, A., J. C. Silva, F. N. Silva, G. P. Castillo-Urquiza, F. F. Silva, Y. M. Seah, F. M. Zerbini. 2017. The diversification of begomovirus populations is predominantly driven by mutational dynamics. *Virus Evol.* 3(1):vex005 (1-14).
- Lotrakul, P., R. A. Valverde, R. De La Torre, J. Sim, J., A. Gomez. 2000. Occurrence of a strain of Texas pepper virus in Tabasco and Habanero pepper in Costa Rica. *Plant Dis.* 84(2):168-172.
- Luan, Y., J. Cui, W. Wang, J. Meng. 2016. MiR1918 enhances tomato sensitivity to *Phytophthora infestans* infection. *Sci. Rep.* 6:35858. doi:10.1038/srep35858.
- Medina-Hernández, D., M. G. Caamal-chan, J. Holgu. 2019. Molecular characterization and phylogenetic analysis of a Squash leaf curl virus isolate from Baja California Sur, Mexico. *PeerJ* 7:e6774. doi: 10.7717/peerj.6774.

- Medina-Hernández, D., R. F. Rivera-Bustamante, F. Tenllado, R. J. Holguín-Peña. 2013. Effects and effectiveness of two RNAi constructs for resistance to pepper golden mosaic virus in *Nicotiana benthamiana* plants. *Viruses*. 5(12):2931-2945.
- Medina-Hernández, D., M. Vargas-Salinas, E. O. Rueda-Puente, R. J. Holguín-Peña. 2019. Seasonal Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) MEAM1 Species and Impact on Incidence of Begomoviral Diseases in Baja California Sur. *J. Econ. Entomol.* 112(3):1055-1061.
- Melgarejo, T. A., T. Kon, M. R. Rojas, L. Paz-Carrasco, F. M. Zerbini, R. L. Gilbertson. 2013. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. *J. Virol.* 87(10):5397-5413.
- Mendez-Lozano, J., I. Torres-Pacheco, C. M. Fauquet, R. F. Rivera-Bustamante. 2003. Interactions between geminiviruses in a naturally occurring mixture: Pepper huasteco virus and Pepper golden mosaic virus. *Phytopathology*. 93(3):270-277.
- Mückstein, U., H. Tafer, J. Hackermüller, S. H. Bernhart, P. F. Stadler, I. L. Hofacker. 2006. Thermodynamics of RNA–RNA binding. *Bioinformatics*. 22(10):1177-1182.
- Naqvi, A. R., N. R. Choudhury, S. K. Mukherjee, Q. M. R. Haq. 2011. In silico analysis reveals that several tomato microRNA/microRNA* sequences exhibit propensity to bind to tomato leaf curl virus (ToLCV) associated genomes and most of their encoded open reading frames (ORFs). *Plant Physiol. Biochem.* 49(1):13-17.
- Noriega, D., A. Valencia, B. Villegas. 2016. RNA de interferencia (RNAi): una tecnología novedosa con potencial para el control de insectos plaga. *Rev. U.D.C.A. Act. & Div. Cient.* 19(1):25-35.
- Ortiz-Quintero, B. 2009. RNA de interferencia: Origen y aplicación en el silenciamiento de genes. *Revista de Investigación Clínica*. 61(5):412-427.
- Piedra-Aguilera, Á., C. Jiao, A. P. Luna, F. Villanueva, M. Dabad, A. Esteve-Codina, A. G. Castillo. 2019. Integrated single-base resolution maps of transcriptome, sRNAome and methylome of Tomato yellow leaf curl virus (TYLCV) in tomato. *Sci. Rep.* 9(1):2863.
- Pradhan, B., V. V. Tien, N. Dey, S. K. Mukherjee. 2017. Molecular Biology of Geminivirus DNA Replication. *Avid Science*. <https://www.avidscience.com/wp-content/uploads/2017/05/VR-17-01.pdf>
- Raja, P., B. C. Sanville, R. C. Buchmann, D. M. Bisaro. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. *J. Virol.* 82(18):8997-9007.

- Raja, P., J. N. Wolf, D. M. Bisaro. 2010. RNA silencing directed against geminiviruses: Post-transcriptional and epigenetic components. *Biochim. Biophys. Acta Gene Regul. Mech.* 1799(3-4):337-351.
- Rajwanshi, R., K. J. Devi, G. R. Sharma, B. Lal. 2019. Role of miRNAs in Plant-Microbe Interaction. In: Kumar M., A. Muthusamy, V. Kumar, N. Bhalla-Sarin (eds.). *In vitro Plant Breeding towards Novel Agronomic Traits*. Springer, Singapore. pp 167-195.
- Ramesh, S. V., A. K. Mishra, S. Praveen. 2007. Hairpin RNA-Mediated Strategies for Silencing of Tomato Leaf Curl Virus AC1 and AC4 Genes for Effective Resistance in Plants. *Oligonucleotides*. 17(2):251-257.
- Rentería-Canett, I., B. Xoconostle-Cázares, R. Ruiz-Medrano, R. F. Rivera-Bustamante. 2011. Geminivirus mixed infection on pepper plants: synergistic interaction between PHYVV and PepGMV. *Viol. J.* 8(1):104-117.
- Reveles-Torres, L. R., R. Velásquez-Valle, J. A. Mauricio-Castillo, S. Salas-Muñoz. 2012. Detección de infecciones mixtas causadas por Begomovirus y Curtovirus en plantas de Chile para secado en San Luis Potosí, México. *Revista Mexicana de Fitopatología*. 30:1-6.
- Ricaño-Rodríguez, J., E. A. Zavala-González, M. Ramírez-Lepe. 2014. Silenciamiento génico en plantas: Mecanismos moleculares del RNA de interferencia y aplicaciones biotecnológicas. *Rev. Fitotec. Mex.* 37(4):339-350.
- Rodrigo, G., J. Carrera, A. Jaramillo, S. F. Elena. 2011. Optimal viral strategies for bypassing RNA silencing. *J. R. Soc. Interface*. 8(55):257-268.
- Rodríguez-Negrete, E. A., J. J. Morales-Aguilar, G. Domínguez-Duran, G. Torres-Devora, E. Camacho-Beltrán, N. E. Leyva-López, J. Méndez-Lozano. 2019. High-throughput sequencing reveals differential Begomovirus species diversity in non-cultivated plants in Northern-Pacific Mexico. *Viruses*. 11(7):594-618.
- Rogans, S. J., F. Allie, J. E. Tirant, M. E. C. Rey. 2016. Small RNA and methylation responses in susceptible and tolerant landraces of cassava infected with South African cassava mosaic virus. *Virus Res.* 225:10-22.
- Sánchez-Campos, S., A. Martínez-Ayala, B. Márquez-Martín, L. Aragón-Caballero, J. Navas-Castillo, E. Moriones. 2013. Fulfilling Koch's postulates confirms the monopartite nature of tomato leaf deformation virus: A begomovirus native to the New World. *Virus Res.* 173(2):286-293.

- Sarkies, P., E. A. Miska. 2014. Small RNAs break out: The molecular cell biology of mobile small RNAs. *Nat. Rev. Mol. Cell Biol.* 15(8):525-535.
- Seal, S. E., F. VandenBosch, F., M. J. Jeger. 2006. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. *CRC Crit. Rev. Plant Sci.* 25(1):23-46.
- Sharma, H. J., S. K. Sharma, N. B. Singh. 2019. Genome complexity of begomovirus disease and a concern in agro-economic loss. *J. Appl. Biol. Biotechnol.* 7(3):78-83.
- Shweta, J. A. Khan. 2014. In silico prediction of cotton (*Gossypium hirsutum*) encoded microRNAs targets in the genome of Cotton leaf curl Allahabad virus. *Bioinformatics.* 10(5):251-255.
- Singh, K., A. Talla, A., W. Qiu. 2012. Small RNA profiling of virus-infected grapevines: evidences for virus infection-associated and variety-specific miRNAs. *Funct. Integr. Genomics.* 12(4):659-669.
- States, U., B. Faso. 2017. List of begomovirus isolates as of Jan. 19th, 2017 Species name Abutilon golden mosaic Yucatan virus Abutilon mosaic Bolivia virus Isolate. https://talk.ictvonline.org/ictv_wikis/geminiviridae/m/files_gemini/7195/download
- Tousi, N., O. Eini, R. Ahmadvand, A. Carra, L. Miozzi, E. Noris, G. P. Accotto. 2017. In silico prediction of miRNAs targeting ToLCV and their regulation in susceptible and resistant tomato plants. *Australas. Plant Pathol.* 46(4):379-386.
- Tripathi, A., K. Goswami, M. Tiwari, S. K. Mukherjee, N. Sanan-Mishra. 2018. Identification and comparative analysis of microRNAs from tomato varieties showing contrasting response to ToLCV infections. *Physiol. Mol. Biol. Plants.* 24(2):185-202.
- Vaca-Vaca, C., J. F. Betancur-Pérez, K. López-López. 2012. Distribución y diversidad genética de Begomovirus que infectan tomate (*Solanum lycopersicum* L) en Colombia. *Revista Colombiana de Biotecnología.* 14(1):60-76.
- Vazquez, F., T. Hohn. 2013. Biogenesis and Biological Activity of Secondary siRNAs in Plants. *Scientifica.* <https://doi.org/10.1155/2013/783253>
- Walter N. L., A. Kvarnheden. 2014. Mixed infection by two West African tomato-infecting begomoviruses and ageratum leaf curl Cameroon betasatellite in tomato in Cameroon. *Arch. Virology.* 159(11):3145-3148.
- Wang, X. J., J. L. Reyes, N. H. Chua, T. Gaasterland. 2004. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.* 5(9):R65.

- Watanabe, Y. 2011. Overview of Plant RNAi. In: Kodama H., Komamine A. (eds.) RNAi and Plant Gene Function Analysis. Methods in Molecular Biology (Methods and Protocols). Humana Press. Totowa, New Jersey. Vol. 744 pp 1-11.
- Won, S. Y., R. E. Yumul, X. Chen. 2014. Small RNAs in Plants. In: Howell S. H. (ed.). Molecular Biology. Springer, Switzerland, A. G. pp 95-127.
- Xie, M., B. Yu. 2015. siRNA-directed DNA methylation in plants. *Curr. Genomics*. 16(1):23-31.
- Xie, Y., L. Zhao, X. Jiao, T. Jiang, H. Gong, B. Wang, X. Zhou. 2013. A recombinant begomovirus resulting from exchange of the C4 gene. *J. Gen. Virol.* 94(8):1896-1907.
- Xu, X. Z., Q. P. Liu, L. J. Fan, X. F. Cui, X. P. Zhou. 2008. Analysis of synonymous codon usage and evolution of begomoviruses. *J. Zhejiang Univ. Sci. B.* 9(9):667-674.
- Yang, J. S., E. C. Lai. 2011. Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Molecular Cell*. 43(6):892-903.
- Yang, X., Y. Wang, W. Guo, Y. Xie, Q. Xie, L. Fan, X. Zhou. 2011. Characterization of small interfering RNAs derived from the geminivirus/betasatellite complex using deep sequencing. *PLoS ONE* 6(2):e16928. <https://doi.org/10.1371/journal.pone.0016928>
- Zerbini, F. M., R. W. Briddon, A. Idris, D. P. Martin, E. Moriones, J. Navas-Castillo, A. Varsani. 2017. ICTV Virus Taxonomy Profile: Geminiviridae. *J. Gen. Virol.* 98:131-133.
- Zhang, C., Z. Xian, W. Huang, Z. Li. 2015. Evidence for the biological function of miR403 in tomato development. *Sci. Hortic.* 197:619-626.
- Zhang, W., N. H. Olson, T. S. Baker, L. Faulkner, M. Agbandje-McKenna, M. I. Boulton, R. McKenna. 2001. Structure of the maize streak virus geminate particle. *Virology*. 279(2):471-477.
- Zhu, H., H. Guo. 2012. The role of virus-derived small interfering RNAs in RNA silencing in plants. *Sci. China Life Sci.* 55(2):119-125.
- Zhu, Q. H., L. Fan, Y. Liu, H. Xu, D. Llewellyn, I. Wilson. 2013. miR482 regulation of NBS-LRR defense genes during fungal pathogen infection in cotton. *PLoS ONE* 8(12):e84390. <https://doi.org/10.1371/journal.pone.0084390>