



**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
IZTACALA**

**VARIACIÓN EN LA ESTRUCTURA DEL
GENOMA DE CLOROPLASTO DE
Mammillaria (CACTACEAE).**

TESIS

Que para obtener el grado académico de:

LICENCIADO EN BIOLOGÍA

Presenta:

VICTOR ABDERRAM CÓRDOBA NAVARRETE

Directora de tesis:

DRA. SOFÍA SOLÓRZANO LUJANO



Los Reyes Iztacala, Estado de México

Diciembre 2021



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS ACADÉMICOS

Al jurado revisor de esta Tesis: Dra. Sofía Solórzano Lujano, directora de esta tesis, al Dr. Ángel Salvador Arias Montes, Dra. Fabiola Soto Trejo, Dra. Patricia Dolores Dávila Aranda y la Dra. Sandra Luz Gómez Acevedo por revisar y ayudarme con sus observaciones al desarrollo de este proyecto.

Esta Investigación fue realizada gracias al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica de la UNAM (PAPIIT-DGAPA IN228619): Caracterización de la estructura del genoma completo de cloroplasto de *Mammillaria* y sus especies cercanas. También, de este proyecto, se me otorgó una beca de seis meses para la obtención de los resultados y escritura de esta tesis.

Al Dr. Ángel Salvador Arias Montes porque me proporcionó varias de las muestras de los tejidos de las especies que se analizaron en este estudio. Otras muestras de tejidos y de ADN fueron obtenidas de proyectos anteriores de PAPIIT (IN220711, IN222216, IN217208) ya concluidos y que estuvieron bajo la responsabilidad de la Dra. Sofía Solórzano Lujano.

Gracias a la Unidad de Síntesis y Secuenciación del Instituto de Biotecnología, UNAM, Cuernavaca, Morelos, donde se sintetizaron los oligonucleótidos utilizados en este trabajo.

A la M. en C. Laura Margarita Márquez Valdelamar y a la M. en C. Nelly María López Ortiz, técnicas del Laboratorio de Secuenciación Genómica de la Biodiversidad y de la Salud del instituto de Biología de la UNAM por su apoyo en la secuenciación Sanger de las muestras analizadas.

A la FES Iztacala y a la UNAM por darme la oportunidad de superarme y concluir mis estudios profesionales.

AGRADECIMIENTOS PERSONALES

Agradezco a mis padres por apoyarme en todo lo que he necesitado, por quererme y procurarme, sobre todo por mostrarme que se puede ser fiel a tus ideales, admiro su tenacidad.

A mi hermana por ser mi amiga, por divertirse conmigo y de alguna manera comprenderme tan profundamente.

A mi novia por tantos grandes momentos, por ser tan maravillosa y apoyarme siempre, por quererme tanto y enseñarme tanto en tan poco tiempo.

A mis amigos por haberme acompañado durante este gran trayecto, por pasarla bien conmigo, por ser tan geniales y auténticos.

A mis profesores de la carrera, por transmitirme su conocimiento y sobre todo su pasión por la ciencia y por la vida.

Una vez más a mi tutora, la Dra. Sofía Solórzano, por su guía, por sus grandes consejos y su sabiduría, por su paciencia, su tiempo y su genuino interés.

Finalmente quiero dar gracias a la vida, a lo maravilloso del mundo, a la biología, a la ciencia y a la naturaleza, agradezco el poder para que cada hombre pueda construirse y dar significado a su vida.

ÍNDICE

RESUMEN	1
INTRODUCCIÓN	3
HIPÓTESIS	15
OBJETIVOS	15
MÉTODOS	16
RESULTADOS.....	22
DISCUSIÓN	28
CONCLUSIONES.....	36
LITERATURA CITADA.....	37
ANEXOS	46

RESUMEN

La genómica comparada permite inferir procesos evolutivos a distintas escalas jerárquicas taxonómicas, a partir de la comparación de la estructura, definida como el número y tipo de genes, y cómo estos se arreglan relativamente entre sí. En particular, en la familia Cactaceae se han documentado varias estructuras en el genoma completo de cloroplasto (ADNcp). De esas estructuras la que destaca es la que carece de un bloque de genes repetido en tándem pero con distinta orientación. Este bloque de genes se conoce comúnmente como inversos repetidos (IRs) y está ausente en dos especies de cactáceas columnares (*Carnegiea gigantea* y *Lophocereus schottii*). En contraste, en las otras especies que cuentan con el ADNcp sí tienen IRs, pero con distinta estructura. En particular en el género *Mammillaria* se hallaron tres tipos de estructuras del ADNcp en siete especies pertenecientes a tres subgéneros, indicando que este género tiene una diversidad estructural en su ADNcp y que el tipo de estructura no es concordante con alguna clasificación taxonómica infragenérica previamente propuesta. En esta tesis se investigó si las tres estructuras de ADNcp documentadas en *Mammillaria* están presentes en otras especies de este género. Para ello, se muestrearon 15 series taxonómicas agrupadas en ocho subgéneros del género *Mammillaria*. Para realizar el estudio se utilizaron los oligonucleótidos diseñados previamente para amplificar por PCR las regiones flanco de los IRs, así como otros loci que varían entre las tres estructuras. Estos productos de PCR se secuenciaron con Sanger y se procesaron para corroborar su identidad con las bases de acceso libre <https://www.ncbi.nlm.nih.gov/>. Además, las secuencias de los loci *ycf2-trn^{CAU}-rpl23* se utilizaron para determinar las relaciones filogenéticas de las especies aquí estudiadas. Las secuencias de las especies *Anacampseros filamentosa*, *C. gigantea*, *Celosia argentea*, *L. schottii* y *Portulaca oleracea* se usaron como grupos externos. Sus secuencias son de acceso libre en la página <https://www.ncbi.nlm.nih.gov/>. Como resultado de esta tesis se encontró la estructura 1 solamente en la especie ya reportada previamente (*M. albilflora*, subgénero *Krainzia*, serie *Herrerae* + *Pectiniferae*). En contraste, la estructura 2,

que se documentó previamente en los subgéneros *Mammillaria* y *Krainzia*, en esta tesis se reconoció en otros subgéneros y series, siendo un total de cinco subgéneros que agrupan cuatro series: I. subgénero *Cocheimea* (*M. poselgeri*, sin serie), II. subgénero *Dolichotele* (*M. baumii*, sin serie), III. subgénero *Krainzia* (*M. hernandezii*, serie *Longiflorae*), IV. subgénero *Mammillaria* (*M. geminispina*, serie *Leucocephalae*), (*M. mystax*, serie *Polydrae*), (*M. rhodanta* serie *Rhodantae*) y V. subgénero *Mammilliopsis* (*M. senilis*, sin serie). La estructura 3, que previamente se documentó en el subgénero *Pellospherma*, en esta tesis también se encontró en dos subgéneros que agrupan a seis series: I. subgénero *Oehmea* (*M. beneckeii*, sin serie) y II. subgénero *Mammillaria* (*M. wiesingeri*, serie *Heterochlorae*), (*M. densispina*, serie *Leptocladode*), (*M. uncinata*, serie *Mammillaria*), (*M. rekoi*, serie *Polycanthae*), (*M. schiedeana*, serie *Proliferae*), (*M. sphacelata*, serie *Sphacelatae*). Además, en las especies del subgénero *Mammillaria* que agrupa a tres series: *M. bocasana* (serie *Stylothelae*), *M. decipiens* (serie *Decipientes*) y *M. lasiacantha* (serie *Lasiacanthae*) y del subgénero *Chilita* (*M. dioica*, sin serie) no se encontró ninguna de las tres estructuras reportadas. Por otra parte, la filogenia mostró que las especies del subgénero *Mammillaria* que tienen la estructura 2 se agrupan, excluyendo a las que tienen la estructura 3. Los resultados de esta tesis confirman una diversidad estructural en el ADNcp del género *Mammillaria*, la cual aún no ha sido completamente descrita, como lo indica el hecho de que cuatro especies aquí estudiadas no amplificaron para alguna de las tres estructuras previamente descritas. La forma en la que se agrupan las especies según su la estructura de su ADNcp no es concordante con las clasificaciones taxonómicas infragenericas propuestas para *Mammillaria*. Sin embargo cabe destacar que las relaciones filogenéticas aquí obtenidas indican que la estructura 2 del ADNcp, es más frecuente en especies del subgénero *Mammillaria*.

INTRODUCCIÓN

La genómica comparada estudia la similitud y diferencias entre los genomas completos de las especies a un nivel molecular, incluyendo sus características funcionales y estructurales (Touchman, 2010). Además, la genómica comparada se encarga de rastrear el origen, cambio y pérdida a través de los linajes evolutivos de los genomas. Esta información permite comprender procesos evolutivos, tales como la mutación, recombinación, transferencia horizontal de genes y selección natural, que rigen los cambios de las características del genoma (Xia, 2013). La genómica comparada también analiza los genomas completos nucleares y los de orgánulos celulares como la mitocondria y el cloroplasto (Eguarte *et al.*, 2003). En particular, el cloroplasto es el orgánulo donde ocurre el proceso de la fotosíntesis. Durante este proceso los electrones son transportados por la membrana tilacoidal, con la finalidad de formar moléculas de ATP y NADPH, las cuales se utilizan para la síntesis de glucosa a partir de CO₂ en el ciclo de Calvin (Taiz & Zeiger, 2002). En particular, el genoma completo de cloroplasto (ADNcp, de aquí en adelante) ha sido de gran ayuda para entender procesos fisiológicos en plantas y algas (Odintsova & Yurina, 2006).

Los genes que se encuentran en el ADNcp se encargan de la fotosíntesis, por ejemplo el gen *rbcL* que codifica para la subunidad grande de la enzima Rubisco (Odintsova & Yurina, 2006). Estos genes se dividen en verdaderos genes, secuencias de ADN que se traducen a proteínas, así como regiones no codificantes particularmente los intrones y los espacios intergénicos, ARN de transferencia (ARNt) y ARN ribosomal (ARNr) (Shaw *et al.*, 2007). Los genes del ADNcp también están involucrados en otras funciones: la transcripción de su propio ADNcp, que codifican para las subunidades de la ARNp (*rpoC*) (Jiao & Guo, 2014); y otros relacionados con el metabolismo fotosintético participando en la síntesis de aminoácidos, ácidos grasos y pigmentos (Odintsova & Yurina, 2006).

Sobre el origen y la evolución del ADNcp sabemos que este tiene un origen cianobacteriano, además que el contenido y número de genes y la estructura del

genoma se ha conservado durante los últimos 800 millones de años, pues en el alga verde *Mesostigma viridae*, se documentó que el 81% de sus genes se encuentran en Embryophyta y en otras algas rojas y verdes (Lemieux *et al.*, 2000).

En algas, el ADNcp es mucho más heterogéneo en el tamaño y composición genética que en plantas terrestres. En Chlorophyta, el ADNcp contiene entre 100,000 hasta más de 200,000 pb (Lemieux *et al.*, 2016), incluso alcanza más de 520,000 pb en el alga *Floydiela terrestris* (Broudard *et al.*, 2010). Debido a que el ADNcp se encarga de procesos tan específicos, en particular la transcripción del mismo ADN y la fotosíntesis, su secuencia nucleotídica suele ser muy conservada, siendo los ARNr los más conservados con 98% de similitud entre diferentes especies de algas (Odintsova & Yurina, 2006).

Las pruebas que ayudan a discernir el origen y evolución del ADNcp están ligadas al origen del cloroplasto. Estas pruebas son la presencia de los mismos genes en algas verdes y en plantas terrestres (Lemieux *et al.*, 2000), las similitudes físicas entre las algas cianobacterias y los cloroplastos, así como sus paredes hechas con una doble membrana (Whatley, 1981). Además, las evidencias moleculares muestran la similitud de las secuencias entre el ADN nuclear de cianobacterias y el ADNcp de otras algas unicelulares (Palmer & Delwiche, 1998).

Los cloroplastos de las células vegetales son endosimbiontes semiautónomos derivados de cianobacterias de vida libre (Gray & Doolittle, 1982). Hace más de 1,000 millones de años un organismo eucarionte fagocitó o fue parasitado por una cianobacteria (McFadden & van Dooren, 2004). Este fue el evento endosimbiótico primario, que generó con el tiempo, a Rhodophyta, Chlorophyta, Glaucophyta y Embryophyta (Jensen & Leiter, 2014). Actualmente, la teoría endosimbiótica ya no se pone en duda pues hay las suficientes pruebas para aclarar el origen de los cloroplastos (Odintsova & Yurina, 2006).

La reducción del genoma es una de las tendencias que apoyan la teoría endosimbiótica y que tienen implicaciones evolutivas. Aproximadamente el 90% de los genes de las cianobacterias ancestrales fueron transferidos al genoma nuclear del organismo huésped (Clegg *et al.*, 1994). Esto provocó que el cloroplasto

cediera su autonomía al núcleo de la célula (de Vries & Archibald, 2018). Este proceso de transferencia de genes del ADNcp al genoma del núcleo continúa ocurriendo. Por ejemplo, el gen *rbcS* que codifica para la parte pequeña de la proteína Rubisco, fue transferido al genoma nuclear mientras este mismo gen continúa presente en el ADNcp en algas no fotosintéticas (Bottomley & Bohnert, 1982).

Actualmente, se han generado más de 3,400 secuencias de genomas completos de cloroplasto de especies de Embryophyta, las cuales están depositadas en el Genbank (NCBI, 2020). Esta acumulación de secuencias ha permitido conocer más a fondo las características del ADNcp de los taxones de plantas terrestres, por ejemplo, la variación en su longitud, número de genes, estructura y función (Odintsova & Yurina, 2006).

En Embryophyta el ADNcp es una molécula típicamente circular que se encuentra dentro del cloroplasto, está compuesto por una doble cadena de ADN y tiene una longitud de entre 120,000 y 217,000 pares de bases (pb) (Palmer, 1985). Sin embargo, existen excepciones como la especie parásita *Epifagus virginiana* que, tiene tan solo 70,000 pb, debido a que no realiza fotosíntesis y fotorespiración (Wolfe *et al.*, 1992).

En Embryophyta el número de genes varía desde 42 (*Epifagus virginiana*, familia Orobanchaceae, orden Lamiales), (Wolfe *et al.*, 1992) hasta más de 130 en el género *Camellia* (familia Theaceae, orden Ericales) (Huang *et al.*, 2014).

De acuerdo con su estructura, el ADNcp de Embryophyta es circular y se divide en cuatro secciones, dos regiones únicas, una de ellas más larga ("Large Single Copy", LSC por sus siglas en inglés), que la otra ("Small Single Copy", SSC) y dos regiones duplicadas pero en sentidos contrarios llamadas Inversos Repetidos ("Inverted Repeats", IRA y el IRB), separados entre sí por las dos regiones únicas (Palmer & Thompson, 1982). Cada uno de los IRs mide típicamente entre 12,000 y 25,000 pb (Palmer & Thompson, 1982), aunque algunos pueden tener menos de 1,000 pb. Por ejemplo la especie *Pinus thunbergii* que pertenece a las gimnospermas presenta 495 pb (Wakasugi *et al.*, 1994), mientras que en

angiospermas como las cactáceas *M. albiflora* tienen 1348 pb y *M. pectinifera* con 1544 pb entre ambos IRs (Solórzano *et al.*, 2019). El número y posición de los genes dentro de los IRs puede ser muy variable incluso entre especies cercanas (Brázda *et al.*, 2018). Los IRs en Embryophyta contienen los cuatro genes que contienen el ARNr, los cuales son indispensables para la síntesis proteica y la replicación del genoma (Palmer, 1985).

Los IRs se encuentran en la mayoría de grupos de plantas terrestres (Lemieux *et al.*, 2016). Durante la evolución del ADNcp los IRs, aparentemente, se han perdido varias veces de modo independiente (Turmel *et al.*, 2017). La formación de los IRs es resultado de genes que forman parte de la composición del ADNcp, ya sea de LSC o de SSC.

En Embryophyta, se ha documentado en varios linajes la ausencia de los IRs lo que se ha interpretado como una pérdida. En las gimnospermas, las familias Pinaceae (orden Pinales) y Cupressaceae (orden Cupressales), cuentan con varias especies en las que faltan los IRs (Chaw *et al.*, 2018). También están ausentes en algunos miembros de angiospermas, por ejemplo en la subfamilia Papilionoideae (familia Leguminosae, orden Fabales) (Lavin *et al.*, 1990), en un grupo de plantas parásitas no fotosintéticas en la familia Orobanchaceae (orden Lamiales) (Wicke *et al.*, 2013) y en algunas especies de la familia Cactaceae (orden Caryophyllales), como *C. gigantea* (Sanderson *et al.*, 2015) y *L. schottii* (verificado en el GenBank con el código de acceso NC_041727.1)

La genómica comparada, además de poder rastrear y comprobar el origen del cloroplasto, también nos permite identificar procesos evolutivos a nivel molecular y estructural a través de los cambios moleculares que se presentan entre los taxones comparados (Xia, 2013). La velocidad de cambio de la secuencia del ADNcp es más lenta que la del genoma nuclear (Xu *et al.*, 2015), pero más rápida que la de los genes del genoma de la mitocondria, lo cual permite realizar estudios evolutivos por debajo del nivel de orden y familia como con el genoma mitocondrial (Wolfe *et al.*, 1987).

Los espaciadores intergénicos y los intrones muestran ser un buen marcador de variación molecular, como demuestra Huang *et al.* (2014), quienes compararon las regiones homólogas de ADNcp de especies cercanas del género *Camellia* (familia Theaceae, orden Ericales), observando una variabilidad muy alta en regiones no codificantes.

Los IRs muestran ser una fuente de información muy valiosa para conocer las relaciones evolutivas de las plantas, pues contienen secuencias con diferencias muy marcadas entre especies cercanas (Brázda *et al.*, 2018). La principal fuente de variación estructural en el ADNcp es la expansión, contracción y pérdida de los IRs, lo que indica diferentes procesos evolutivos (Mower *et al.*, 2018). La pérdida de un segmento entero de uno de los IRs es una mutación significativa que puede definir un grupo monofilético (Downie & Palmer, 1992).

La mayoría de los estudios de genómica comparada sobre organismos que tienen cloroplastos están dirigidos a angiospermas por ser el grupo más numeroso de Embryophyta, con más de 30,000 especies vivientes y del que se puede obtener mayor información (Jiao & Guo, 2014). Las angiospermas también han sido estudiadas en cuanto a sus relaciones evolutivas, al ser el grupo más diversificado de plantas superiores como lo propuesto por Downie & Palmer (1992) que sugirieron que la estructura del ADNcp y sus rearrreglos pueden ser una herramienta adecuada para reconstruir la filogenia de niveles taxonómicos orden y familia.

Dentro de las angiospermas uno de los órdenes más diversos es Caryophyllales, el cual consiste de 39 familias y más de 12,500 especies (Hernández-Ledesma, *et al.*, 2015). Muchos de las familias de este orden se caracterizan por presentar adaptaciones inusuales como carnivoría, halofitismo y múltiples orígenes de fotosíntesis tipo C4 y CAM (Yao *et al.*, 2019).

En su estudio Crawley & Hilu (2012) muestrearon 30 de las 40 familias de Caryophyllales utilizando el genes *matK* con su intrón *trnK*, obteniendo seis árboles parsimoniosos y confirmando el origen monofilético para el orden. Por su parte, Yao *et al.* (2019), incluyeron genomas completos y datos sobre la estructura

del ADNcp de representantes de las 40 familias de Caryophyllales. Los autores reconocieron como procesos evolutivos la pérdida de genes e intrones, así como eventos de pseudogenización e inversión de intrones; además observaron la expansión, contracción y pérdida de IRs, encontrando una correlación positiva entre la tasa de sustitución de bases y los cambios en la estructura del ADNcp.

Las especies de Caryophyllales, hasta hoy analizadas, tienen una estructura típica de ADNcp, el cual está dividido en cuatro regiones con dos IRs (Yao *et al.*, 2019). En particular, esta tesis se enfoca en aquellas especies de la familia Cactaceae, que son originarias del continente americano, siendo una de las familias más representativas en los climas áridos y semiáridos, aunque también se distribuye en las zonas tropicales y templadas de América (Bravo-Hollis & Sánchez-Mejorada, 1978). Esta familia consiste de más de 2,000 especies (Anderson, 2001). En México se registran 63 géneros que agrupan aproximadamente 669 especies, de las cuales 518 son endémicas a este país (Guzmán, 2007). Los miembros de la familia Cactaceae tienen un origen reciente y su tiempo de origen calculado es de 35 millones de años aproximadamente, durante el periodo geológico del Eoceno (Arakaki *et al.*, 2011).

Esta investigación se enfocó en el estudio de varias especies de la familia Cactaceae, entre las que se encuentra, el saguaro (*Carnegiea gigantea*), una especie carente de la estructura cuadripartita, pues perdió los IRs (Sanderson *et al.*, 2015). Esta especie habita en el desierto Sonorense, alcanza los 16 metros de altura y puede vivir hasta 175 años. Esta especie tiene un cambio poblacional muy lento como es típico de la mayoría de las especies de Cactaceae. En contraste, en siete especies del género *Mammillaria* se documentaron tres tipos distintos de IRs, los cuales difieren en extensión, número y tipo de genes que los componen, además, en la región LSC se documentaron otros rearrreglos génicos e inversiones de bloques de hasta seis genes (Solórzano *et al.*, 2019).

El género *Mammillaria* cuenta con alrededor de 155 especies, de las cuales 113 son endémicas de México (Hunt, 2006). Estas cactáceas se caracterizan por tener tallos globosos-aplanados o cilíndricos, de hábitos terrestres o rupícolas.

Presentan tubérculos en forma de mamilas, dispuestos en series espiraladas (Bravo-Hollis & Sánchez-Mejorada, 1991). Este género diversificó hace aproximadamente 7.6 - 5.33 millones de años, durante el periodo Mioceno (Arakaki *et al.*, 2011). Hernández-Hernández *et al.* (2014), propusieron que la aridificación de América del Norte pudo ser el proceso que promovió la diversificación del género.

Cerca del 37% de las especies endémicas de *Mammillaria* presentan una distribución restringida, por lo que son consideradas microendémicas. Por tanto, debido a la pérdida y fragmentación de los hábitats, así como a la extracción ilegal de plantas, 109 especies microendémicas se encuentran en alguna categoría de riesgo en la NOM-059-SEMARNAT-2010 (DOF, 2020). A nivel mundial, la Unión Internacional para la Conservación de la Naturaleza (IUCN) incluye 163 especies de este género en algún grado de riesgo (IUCN, 2020). Por su parte la CITES (Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres), la cual representa la normatividad en materia de comercio internacional de especies amenazadas, incluye 165 especies en el Apéndice II y una (*M. pectinifera*), en el Apéndice I (CITES, 2020). La importancia de estas normativas reside en su carácter legal, cuyo propósito es regular la protección y aprovechamiento de las especies amenazadas. Sin embargo, existen inconsistencias en estas listas debido a que la delimitación y número de especies de *Mammillaria* ha cambiado a través del tiempo generando una inestabilidad taxonómica para este grupo, lo cual dificulta las acciones de manejo de recursos naturales por parte de las agencias gubernamentales, tales como los sectores académicos y privados (Arias, 2005).

Los cambios constantes en la clasificación taxonómica en las especies de *Mammillaria* se debe principalmente a su gran diversidad y complejidad morfológica tanto a nivel infra como supra genérico (Arias, 1998), lo cual ha dado como resultado que existan varias clasificaciones infragenéricas (Bravo-Hollis & Sánchez-Mejorada 1991; Lüthy, 1995; Hunt, 2006). La clasificación infragenérica más reciente de *Mammillaria* fue propuesta por Hunt (2006), y se basa en

caracteres morfológicos como la forma de las ramas, características de flores y frutos, número de espinas, y características de las semillas como el hilum y la testa. Esta clasificación reconoce 165 especies de *Mammillaria* agrupadas en ocho subgéneros, de los cuales dos, *Krainzia* y *Mammillaria* están divididos en un nivel inferior de clasificación llamado serie. En total se reconocen 15 series. Por su parte, otro sistema de clasificación de *Mammillaria* es el propuesto por Lüthy (1995, 2001) quien reconoce cinco subgéneros y 22 series (cuadro 1).

Cuadro 1: comparación de la clasificación infragenérica propuestas por Hunt (2006) y Lüthy (1995).

Hunt (2006)		Lüthy (1995, 2001)	
Subgénero	Serie	Subgénero	Serie
<i>Mammillaria</i>	<i>Decipentes</i>	<i>Mammillaria</i>	<i>Bombycina</i>
	<i>Heterochlorae</i>		<i>Decipientes</i>
	<i>Lasiacanthae</i>		<i>Heterochlorae</i>
	<i>Leptocladodae</i>		<i>Lasiacanthae</i>
	<i>Leucocephalae</i>		<i>Leptocladodae</i>
	<i>Mammillaria</i>		<i>Leucocephalae</i>
	<i>Polycanthae</i>		<i>Mammillaria</i>
	<i>Polydrae</i>		<i>Polyacanthae</i>
	<i>Proliferae</i>		<i>Polyedrae</i>
	<i>Sphacelatae</i>		<i>Rhodantae</i>
	<i>Supertextae</i>		<i>Proliferae</i>
	<i>Stylotaelae</i>		<i>Sphacelatae</i>
	<i>Rhodantae</i>		<i>Supertextae</i>
<i>Cocheimea</i>		<i>Cocheimea</i>	<i>Ancistracanthae</i>
			<i>Bartschella</i>
			<i>Cocheimea</i>
<i>Pellosperma</i>		<i>Pellosperma</i>	<i>Herrerae</i>
			<i>Longiflorae</i>
			<i>Pectiniferae</i>

			<i>Pellospherma</i>
			<i>Zephyranthoides</i>
<i>Dolichotele</i>		<i>Dolichotele</i>	
<i>Oehmea</i>		<i>Oehmea</i>	
<i>Chilita</i>			
<i>Krainzia</i>	<i>Herrerae + Pectiniferae</i>		
	<i>Longiflorae</i>		
<i>Mammilliopsis</i>			

Las clasificaciones infragenéricas están basadas principalmente en la morfología de estructuras florales; sin embargo, en Cactaceae estas estructuras pueden observarse solo una vez al año o cuando existe una alta variación fenotípica dentro de una misma especie debido a las condiciones locales (Duminil & Di Michele, 2009).

Debido a la dificultad que conlleva la obtención de los criterios taxonómicos morfológicos, se ha propuesto el uso de marcadores moleculares para intentar resolver la filogenia y taxonomía de la familia Cactaceae, en particular en el género *Mammillaria* (Wallace, 1995).

Algunos estudios realizados con marcadores moleculares (*matK*, *trnK-matK*, *trnL-trnF*, *rbcL*, *rpoB*, *rpl16*, *trnK*, *rpoC1*, *trnK-psbA*, *rpl20-rps12*, *trnL-trnF*, *trnT-trnL*, *psbA-trnH*), concluyeron que el género *Mammillaria* no es monofilético sino parafilético. Sin embargo, no se pudieron resolver sus relaciones filogenéticas con especies de otros siete géneros de la tribu *Cacteae* (Hunt 2006): *Coryphantha*, *Encephalocarpus*, *Escobaria*, *Mammilloya*, *Neolloydia*, *Ortegocactus* y *Pelecyphora* (Butterworth *et al.*, 2002; Butterworth & Wallace, 2004; Nyffeler, 2002; Crozier, 2005; Bárcenas *et al.*, 2011;). En otro estudio utilizando el gen *matK*, considerado como marcador universal porque ha funcionado para el 95% de taxones de Embryophyta, mostró que en Cactaceae la resolución no es suficiente, pues solamente los autores lograron identificar a nivel de género en clados basales, como Opuntoidae (Yesson *et al.*, 2011).

Por otro lado, con el espaciador trnL-trnF, se encontró que las relaciones entre 21 especies del género *Mammillaria* no se resolvieron completamente o tienen soporte bajo, particularmente las del subgénero *Mammillaria* que resultó ser polifilético (Harpke, 2006). Este resultado contrasta con el obtenido por Crozier (1995), quien a partir del análisis con 10 marcadores de cloroplasto (los intrones rpl16, trnK, rpoC1, los espaciadores intergenicos trnK-psbA, rpl20-rps12, trnL-F, trnT-trnL y los genes *matK*, *rbcL*, y *rpoB*), estableció que solo las especies del subgénero *Mammillaria* podrían incluir entidades taxonómicas que representan un linaje evolutivo natural, en tanto que las de los otros subgéneros pueden estar relacionados con otros géneros. En consecuencia en el género *Mammillaria* (sensu Hunt, 2006) podrían agruparse entidades evolutivas distintas conformando actualmente a este género como un grupo no natural.

Uno de los resultados frecuentes de los estudios filogenéticos de la familia Cactaceae o enfocados en miembros de *Mammillaria*, es la formación de politomías en las ramas terminales, debido a la falta de resolución molecular de los marcadores utilizados. Todo lo anterior refleja la ausencia de marcadores moleculares que resuelvan las relaciones filogenéticas dentro y fuera de la familia Cactaceae. Es por esto que adquieren relevancia los estudios de genómica comparada pues se podrían no solo entender las relaciones filogenéticas de la Cactaceae, sino que se podrían contestar preguntas sobre especiación en un género en particular, procesos evolutivos de divergencia dentro de la familia y con el resto de los miembros de Caryophyllales.

Por otro lado, Solórzano *et al.* (2019) encontraron, mediante estudios de secuenciación masiva, una gran diversidad estructural del ADNcp en *Mammillaria*, representado por tres estructuras identificadas, estas varían el contenido y orden de los genes del ADNcp, particularmente en el contenido de sus IRS (figuras 1, 2 y 3).

Estructura 1 IRB

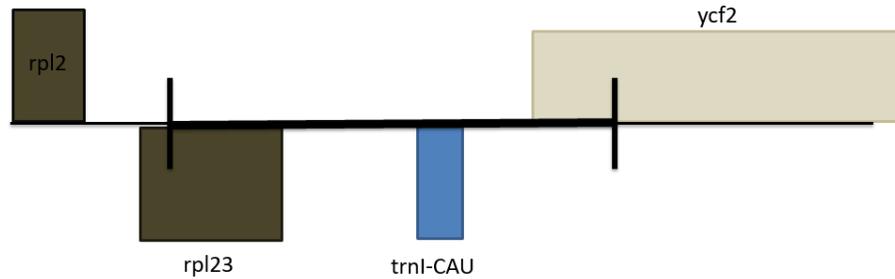


Figura 1. Los inversos repetidos de la de la estructura 1 del genoma del cloroplasto identificado en *Mammillaria*, miden ~670 pb y se componen solo de tres genes. El gen *ycf2* penetra solo parcialmente con 265 pb en su extremo 5'. En este trabajo se utilizó a *Mammillaria albiflora* como representante de esta estructura. Tomado y modificado de Solórzano *et al.* (2019).

Estructura 2 IRB

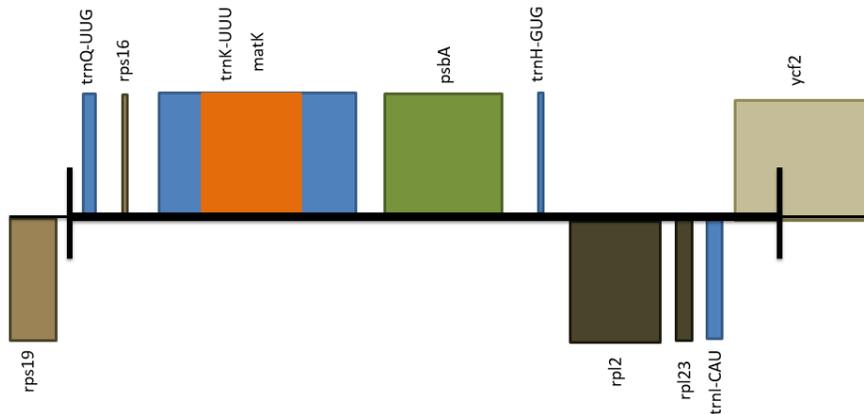


Figura 2. Los inversos repetidos de la estructura 2 del genoma del cloroplasto identificado en *Mammillaria*, miden ~7,200 pb y se componen de 11 genes y un espacio intergénico. En este trabajo se utilizó a *Mammillaria supertexta* como representante de esta estructura, que a diferencia de otras especies encontradas con la misma estructura carece del gen *rpl23* en sus IRs. Tomado y modificado de Solórzano *et al.* (2019).

Estructura 3 IRB

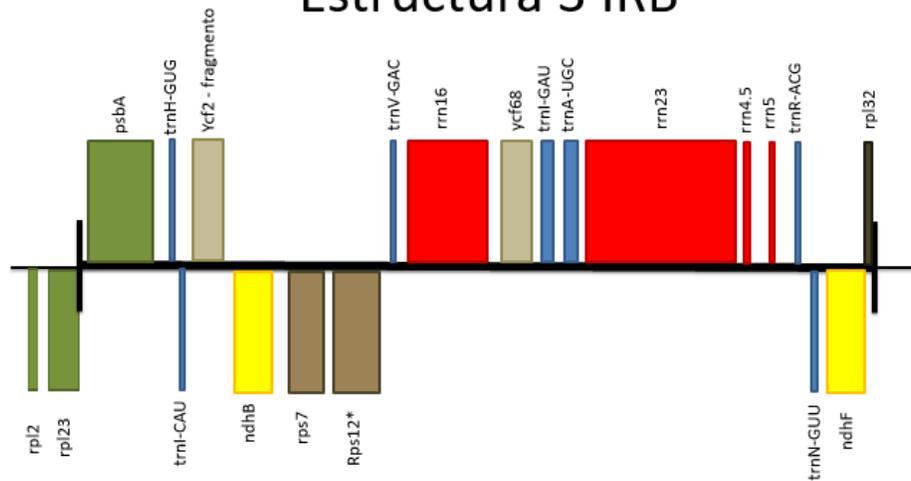


Figura 3. Figura 2. Los inversos repetidos de la estructura 3 del genoma del cloroplasto identificado en *Mammillaria*, miden ~14,100 pb y se componen de 22 genes. El gen *rpl23* está ausente en el IRB por lo que estos IR son divergentes. En este trabajo se utilizó a *Mammillaria zephyranthoides* como representante de esta estructura. Tomado y modificado de Solórzano *et al.* (2019).

Por otro lado, hasta la fecha, ninguno de los estudios basados en marcadores de ADN enfocados en el género *Mammillaria* ha logrado resolver sus relaciones filogenéticas. Una de las consecuencias de la falta de una filogenia resuelta es que ninguna de las clasificaciones taxonómicas infragenéricas propuestas para *Mammillaria* (Lüthy, 1995, 2001; Hunt, 2006) (figura 1) está confirmada por alguna filogenia. En esta tesis se siguió la clasificación infragenérica más reciente que es la de Hunt (2006), con la finalidad de delimitar el muestreo taxonómico.

HIPÓTESIS

Ho: La estructura del genoma de cloroplasto en las especies de los ocho subgéneros y 15 series reconocidos del género *Mammillaria*, son iguales a nivel de series y de subgéneros.

Ha: La estructura del genoma de cloroplasto en las especies de los ocho subgéneros y 15 series reconocidos del género *Mammillaria* es iguales a nivel de serie de un mismo subgénero, pero difiere entre los ocho subgéneros.

OBJETIVOS

Objetivo general

Determinar, a partir de las secuencias de ADN de los flancos de los inversos repetidos, la estructura que poseen en su genoma del cloroplasto las especies de los ocho subgéneros y las 15 series del género *Mammillaria* (Cactaceae), para corroborar su concordancia con la clasificación infragenérica *sensu* Hunt (2006).

Objetivos particulares

1. Caracterizar la estructura de un conjunto de 21 especies pertenecientes a los ocho subgéneros y 15 series de *Mammillaria* (Cactaceae), a partir de las secuencias flanco de los inversos repetidos del genoma de cloroplasto.
2. Determinar las relaciones filogenéticas de las 21 especies estudiadas, con la finalidad de verificar si las estructuras del genoma de cloroplasto se reflejan en un árbol filogenético.

MÉTODOS

Muestreo Taxonómico.

Debido a que el género *Mammillaria* incluye alrededor de 163 especies, se seleccionaron 21 de ellas para representarlo a nivel de subgénero y de series taxonómicas. Para esta representación taxonómica se siguió la clasificación de Hunt (2006) (cuadro 2). Muestras de tejido de 17 especies fueron obtenidas del Jardín Botánico del Instituto de Biología de la UNAM, mientras que otras de ADN genómico de seis especies provienen de proyectos previos concluidos y anteriores a esta tesis que se encuentran disponibles en el Laboratorio de Ecología Molecular y Evolución, UBIPRO, FES IZTACALA, UNAM.

Cuadro 2: Lista de especies estudiadas agrupadas de acuerdo a los ocho subgéneros y 15 series del género *Mammillaria* según Hunt (2006), junto a su voucher del Jardín Botánico o del proyecto del que fueron obtenidas. *: El subgénero no incluye el nivel taxonómico de series. **: Especies utilizadas en el análisis filogenético de *ycf2-rpl23*. La última columna contiene la referencia del número de acceso de las especies utilizadas para el análisis filogenético del locus *matK*, que puede ser consultado en la página <https://www.ncbi.nlm.nih.gov/>.

Subgénero	Series	Especie	Voucher de Jardín Botánico	Voucher proyecto PAPIIT	Código de acceso para <i>matK</i>
1.- <i>Chilita</i>	*	1.- <i>Mammillaria dioica</i> **	S. Arias 1805 (MEXU)	-	FN997457
2.- <i>Cochemiea</i>	*	2.- <i>M. poselgeri</i> **	H. Sánchez-Mejorada 4317 (MEXU)	-	
3.- <i>Dolichothele</i>	*	3.- <i>M. baumii</i> **	S. Arias 1992 (MEXU)	-	FN997100
4.- <i>Krainzia</i>	1.- <i>Herrerae</i> + <i>Pectiniferae</i>	4.- <i>M. albiflora</i>	-	IN220711	
	2.- <i>Longiflorae</i>	5.- <i>M. hernandezii</i> **	-	IN217208	MH037655

5.- <i>Mammillaria</i>	3.- <i>Lasiacanthae</i>	6.- <i>M. lasiacantha</i>	S. Arias 2078 (MEXU)	-	
	4.- <i>Stylothelae</i>	7.- <i>M. bocasana</i>	H. Sánchez- Mejorada 3290 (MEXU)	-	
	5.- <i>Proliferae</i>	8.- <i>M. schiedeana</i> **	S. Arias 717 (MEXU)	-	FN997179
	6.- <i>Sphacelatae</i>	9.- <i>M. sphacelata</i> **	-	IN217208	
	7.- <i>Decipientes</i>	10.- <i>M. decipiens</i>	S. Arias 2043 (MEXU)	-	
	8.- <i>Leptocladodae</i>	11.- <i>M. densispina</i>	D. Aquino 359 (MEXU)	-	
	9.- <i>Heterochlorae</i>	12.- <i>M. wiesingeri</i>	S. Aquino 2256 (MEXU)	-	
	10.- <i>Rhodantae</i>	13.- <i>M. rhodantha</i> **	U. Guzmán 1391 (MEXU)	-	
	11.- <i>Polycanthae</i>	14.- <i>M. rekoii</i>	-	IN217208	
	12.- <i>Supertextae</i>	15.- <i>M. supertexta</i> **	-	IN222216	FN997397
	13.- <i>Leucocephalae</i>	16.- <i>M. geminispina</i> **	D. Aquino 279 (MEXU)	-	FN997280
	14.- <i>Mammillaria</i>	17.- <i>M. uncinata</i>	D. Aquino 289 (MEXU)	-	
	15.- <i>Polydrae</i>	18.- <i>M. mystax</i> **	D. Aquino 23813 (MEXU)	-	
6.- <i>Mammillopsis</i>	*	19.- <i>M. senilis</i> **	S. Arias 1890 (MEXU)	-	FN997161
7.- <i>Oehmea</i>	*	20.- <i>M. beneckeii</i> **	JLSVM 6 (MEXU)	-	FN997206
8.- <i>Phellosperma</i>	*	21.- <i>M.</i> <i>zephyranthoides</i> **	-	IN222216	FN997203

Extracción de ADN genómico total.

Para cada especie (cuadro 2), se obtuvo una muestra de tejido de aproximadamente 30 mg y se pulverizó en un mortero de porcelana. El ADN genómico total (ADNg), se extrajo utilizando el DNeasy plant minikit de QIAGEN® (QIAGEN, Alameda, E.U.A.), siguiendo las instrucciones del fabricante. Para visualizar el ADNg se realizaron electroforesis horizontales en geles de agarosa al 0.8%, teñidos con bromuro de etidio y se corrieron a 112V durante 20 minutos. Las bandas fueron reveladas con luz ultravioleta (UV). Las muestras de ADNg se guardaron a -20°C durante todo el proceso de su análisis molecular.

Amplificación por PCR y secuenciación Sanger

En esta tesis se utilizaron las condiciones de PCR establecidas por Solórzano *et al.* (2019) para amplificar los flancos a los IRs y tres rearrreglos genómicos localizados en la región LSC. Se incluyeron como controles positivos a *M. albiflora* de la estructura 1, *M. supertexta* de la estructura 2 y *M. zephyranthoides* de la estructura 3. Además se diseñaron nuevos oligonucleótidos con la plataforma gratuita en línea del programa primer3 4.1.0 (Untergasser *et al.*, 2012) para corroborar otras inversiones detectadas en la región LSC.

Para diseñar estos nuevos oligonucleótidos se usaron como referencias las secuencias del genoma completo de *M. albiflora* (número de acceso: MN517610.1) y *M. zephyranthoides* (número de acceso: MN517611).

En total se utilizaron cinco loci (cuadro 3) para examinar las tres estructuras: el locus *rps19-trnICAU* proviene del flancos del IRB de la estructura 1 (figura 1), el locus *ycf2-rpl23* flanquea el IRB de la estructura 2 (figura 2), y el locus *rpl2-trnICAU*, que incluye a los genes *ycf2*, *trnI* y *rpl23* y que se encuentran dentro de los IRB de las tres estructuras.

Cuadro 3: Nombre de los loci amplificados por PCR y los oligonucleótidos usados. La temperatura de alineamiento y el tamaño aproximado del producto de PCR amplificado. Los oligonucleótidos marcados con (*) fueron diseñados en este trabajo, el resto fueron obtenidos de Solórzano *et al.* (2019)

Nombre del locus	Pareja de oligonucleótidos	Tm	Tamaño del amplificado	Procedencia
<i>rps19-trnI^{CAU}</i>	2AF / rps19-trnIR*	57°C	1300 pb	Flanco de IR de estructura 1
<i>ycf2-rpl23</i>	11F / 11R	57°C	200 pb	Flanco de IR de estructura 2
<i>rpl2-trnI^{CAU}</i>	3F / 3R	57°C	800 pb	Flanco de IR de estructura 2
<i>trnP^{UUG}-rpl20</i>	MzeprnPF* / Mzeprpl20R*	57°C	850 pb	Rearreglo de la LSC de estructura 3
<i>trnE^{UUC}-psbZ</i>	trnE-psbZ4F* / trnE-psbZ4R*	56°C	600 pb	Rearreglo de la LSC de estructura 3

Sin embargo, en el caso de las especies que presentan la estructura 3, se incluyen dentro del locus los genes *trnH* y *psbA* (figura 3). Asimismo, en esta última estructura, se incluyeron dos loci: *trnP^{UUG}-rpl20* y *trnE-psbZ*, ambos se encuentran dentro del LSC. Esto se hizo con el fin de observar si hay concordancia en la estructura 3, ya que Solórzano *et al.* (2019) menciona que existen dos importantes inversiones génicas en esta zona de la estructura 3, en comparación de la 2 y 1.

Las amplificaciones por PCR se realizaron en el termociclador Veriti™ (Applied biosystems, CA, E.U.A.). El ciclo general de PCR fue el siguiente: 95° de desnaturalización inicial por tres minutos, seguido de 34 ciclos a 95° por 45 segundos, alineamiento durante 45 segundos usando la temperatura requerida por la pareja de oligonucleótidos, 72° por 45 segundos y un ciclo de extensión final a 72° C por cinco minutos. El volumen final de cada reacción fue de 12.5 µl y contenía las concentraciones finales de: Buffer de PCR 10x, MgCl₂ 2.5 mM,

dNTPs 10 mM, oligonucleótido (primer) F y R 10 pM, BSA 10x, Taq polimerasa 0.5 u y 1 µl de ADNg extraído de las muestras. Los productos amplificados por PCR se revelaron por electroforesis horizontal en un gel de agarosa al 1.1% teñidos con bromuro de etidio. El gel se corrió a 90V por 30 minutos y posteriormente se visualizó con luz ultravioleta. Los productos de PCR amplificados se secuenciaron con método Sanger en el equipo modelo 3500XL de 24 capilares de Life Technology (Applied biosystems, CA, E.U.A.).

Análisis de secuencias.

Las secuencias de cada uno de los loci se visualizaron y editaron con el software Chromas 2.6.5 (Technelysium, Australia). Posteriormente los archivos individuales se convirtieron a formato FASTA y se visualizó el tamaño y contenido de las secuencias. Además, se ensamblaron las secuencias correspondientes con el software BioEdit 7.0.5.3 (Hall, 1999).

Las secuencias en formato FASTA se organizaron por locus y posteriormente se alinearon con base en el programa MEGA X (Kumar *et al.*, 2018), en el cual se utiliza el algoritmo ClustalW (Thompson *et al.*, 1994), para comparar la similitud y el contenido de las secuencias entre el mismo locus de varias especies.

Análisis filogenético.

Una vez obtenido los resultados de la amplificación de los loci en las especies estudiadas (cuadro 4, sección resultados), se buscó un locus para realizar el análisis filogenético de secuencias. El locus elegido fue *ycf2-rpl23*, debido a que amplificó para la mayoría de especies estudiadas con buena calidad y para las tres estructuras, además por tener más diferencias entre bases que el resto de loci. Además de este locus se hizo una comparación con el locus *matK*, debido a que es uno de los más usados para análisis filogenéticos además de su fácil obtención. Finalmente se concatenaron estos dos loci para hacer un tercer análisis (*ycf2-rpl23/matK*).

Para el análisis filogenético del locus *ycf2-rpl23*, de la familia Cactaceae se utilizaron 13 de las 21 especies de *Mammillaria* analizadas en este trabajo (cuadro

2) y *C. gigantea* (número de acceso en el GenBank: *KT164771*) y *L. schotti* (*NC_041727*). Como grupos externos se incluyeron, junto con otras tres especies también del orden Caryophyllales: *Celosia argentea* (Amaranthaceae) (*NC_041294*), *Anacampseros filamentosa* (Anacampserotaceae) (*MK397913*) y *Portulaca oleracea* (Portulacaceae) (*NC_036236*). A las especies *C. gigantea*, *L. schotti* y a los grupos externos, excepto *C. argentea*, se les realizó un inverso repetido con el programa MEGA X (Kumar, 2018), para poder comparar las secuencias, debido a la orientación de los loci en los IRs.

En cuanto al locus *matK* se realizó una búsqueda en el GenBank de las especies elegidas para el análisis del locus *ycf2-rpl23*. En total se localizaron nueve especies que estaban disponibles en el GenBank (cuadro 2). Además de *C. gigantea* (*KF437837*) y *L. schotti* (*AY015309*), también se incluyeron *A. filamentosa* (*MK397913*), *C. argentea* (*GQ434274*) y *P. oleracea* (*HQ620836*) como grupos externos.

Las 13 secuencias de ADN de *ycf2-rpl23* y las nueve de *matK* se alinearon utilizando el programa MAFFT v7.310 (Kato & Daron, 2013). Posteriormente las secuencias se concatenaron con en el programa Mesquite (Maddison & Maddison, 2019), dando así 13 secuencias concatenadas para su análisis filogenético.

Para obtener los modelos de sustitución nucleotídica de los análisis *ycf2-rpl23*, *matK* y *ycf2-rpl23/matK* concatenados, se utilizó el programa MEGA X (Kumar, 2018) que, analiza 24 modelos de sustitución con el criterio de Información Akaike (AIC).

Para los tres análisis se realizó una reconstrucción filogenética de Máxima Verosimilitud con en el programa RAXML-HPC v8.2.10 (Stamatakis, 2014) y se utilizó el modelo GTR + G con 1000 repeticiones para obtener el árbol filogenético.

RESULTADOS

Los resultados muestran que de los cinco loci estudiados, cuatro amplificaron para más de una especie, lo que ayuda a conocer parte de la estructura del ADNcp de cada una de ellas (cuadro 4; anexos 1 a 5).

Los resultados de esta tesis mostraron que la estructura 1 no se presentó en las especies aquí estudiadas de *Mammillaria*, y solo amplificó en el control positivo *M. albilflora*. La estructura 2 se encontró en especies de cuatro series y cinco subgéneros. La estructura 3 se presentó en especies de cinco series y dos subgéneros.

La estructura 2 se verificó con el locus *ycf2-rpl23* y se encontró en los subgéneros 1. *Cocheimea* (*M. poselgeri*, sin serie), 2. *Dolichotele* (*M. baumii*, sin serie), 3. *Krainzia* (serie Longiflorae (*M. hernandezii*)), 4. *Mammillaria* (serie *Leucocephalae* (*M. geminispina*), serie *Polydrae* (*M. mystax*), serie *Rhodantae* (*M. rhodanta*)) y 5. *Mammilliopsis* (*M. senilis*, sin serie).

La estructura 2 también amplificó para *M. beneckeii*, *M. schiedeana*, *M. sphacelata* y *M. zephyranthoides*, sin embargo, estas amplificaron también para el locus utilizado para verificar la estructura 3 (*rpl2-trnI^{CAU}*). Por lo que se considera que estas especies pertenecen a la estructura 3, ya que ésta tiene el mismo locus que se puede encontrar en la estructura 2, sin embargo no se encuentran en los flancos de los IR (figura 2 y 3).

La estructura 3 se encontró en los subgéneros 1. *Oehmea* (*M. beneckeii*, sin serie) y 2. *Mammillaria* (serie *Hetecochlorae* (*M. wiesingeri*), serie *Leptocladode* (*M. densispina*), serie *Mammillaria* (*M. uncinata*), serie *Polycanthae* (*M. rekoii*), serie *Proliferae* (*M. schiedeana*), y serie *Sphacelatae* (*M. sphacelata*)).

Para identificar las especies de *Mammillaria* que contienen la estructura 3 se utilizó el locus *trnI^{CAU}-rpl2* que está basado en la estructura 2, pues todas las especies de *Mammillaria* amplificaron para este locus. Sin embargo, al verificar el tamaño del amplificado, en *M. beneckeii* (subgénero *Oehmea*, sin serie), *M.*

densispina (subgénero *Mammillaria*, serie *Leptocladode*), *M. rekoii* (subgénero *Mammillaria*, serie *Polycanthae*), *M. schiedeana* (subgénero *Mammillaria*, serie *Proliferae*), *M. sphacelata* (subgénero *Mammillaria*, serie *Sphacelatae*), *M. uncinata* (subgénero *Mammillaria*, serie *Mammillaria*) y *M. wiesingeri* (subgénero *Mammillaria*, serie *Heterochlorae*), el gel de agarosa mostró que los fragmentos de ADN en estas especies son notablemente más grandes, midiendo 1300 pb, comparado con el resto de las especies que solamente alcanzan 900 pb (anexo 3). En particular en *M. senilis* (subgénero *Mammillopsis*), se encontró una banda doble, lo que implica que los amplificadores fueron productos inespecíficos, lo cual dio paso a que el resultado fuera descartado debido a las limitaciones técnicas requeridas para la secuenciación Sanger.

Dos amplificadores del locus *rpl2-trn^{CAU}* no fueron secuenciados para las especies *M. senilis*, *M. densispina*, *M. rekoii* y *M. sphacelata*, debido a una baja calidad del amplificado. Tampoco se amplificó este locus para *M. geminispina*, ya que hay certeza de que posee la estructura 2.

En lo que respecta a los dos rearrreglos génicos de la estructura 3, se diseñaron oligonucleótidos basados en la LSC del ADNcp para observar si hay concordancia con la estructura 3, ya que Solórzano *et al.* (2019) menciona que existen dos importantes inversiones génicas en esta zona de dicha estructura, a diferencia de la 2 y 1. En el caso del locus *trn^{EUUC}-psbZ*, únicamente amplificó para el control positivo *M. zephyranthoides*, mientras que el locus *trn^{PUUG}-rpl20* amplificó para *M. beneckeii* (subgénero *Oehmea*) y *M. baumii* (subgénero *Phellosperma*).

Finalmente, *M. bocasana* (serie *Stylothelae*), *M. decipiens* (serie *Decipientes*) y *M. lasiacantha* (serie *Lasiacanthae*) del subgénero *Mammillaria*, así como *M. dioica* del subgénero *Chilita*, solo amplificaron para el locus *rpl2-trn^{CAU}*. Sin embargo, ellos amplificaron con el tamaño de banda más corto por lo que no fue posible identificarlos con ninguna de las estructuras del ADNcp del género *Mammillaria*.

Cuadro 4. Especies analizadas y ampliaciones de los flancos de los inversos repetidos de las tres estructuras del genoma de cloroplasto de *Mammillaria* y de dos rearrreglos génicos ubicados en la región LSC. La lista taxonómica sigue la clasificación infragenérica de Hunt (2006) En la primera columna se presentan los ocho subgéneros, seguidos de las series y las especies analizadas; se presenta el tamaño en pares de bases (pb), del fragmento amplificado. El símbolo astérico (*), indica que el producto amplificado mostró productos inespecíficos de PCR que fueron descartados. El guion (-) indica que el producto amplificado dio un resultado negativo al compararlo con un control. Las letras NA indican que no se hicieron las repeticiones necesarias para eliminar el resultado dudoso.

Taxón			Tamaño (pb) de loci usados para verificar la estructura de los inversos repetidos (las tres primeras columnas) y de dos rearrreglos génicos localizados en la región larga (dos últimas columnas).				
			<i>rps19-trnI</i> ^{CAU}	<i>ycf2-rpl23</i>	<i>rpl2-trnI</i> ^{CAU}	<i>trnP</i> ^{UUG} - <i>rpl20</i>	<i>trnE</i> ^{UUC} - <i>psbZ</i>
Subgénero	Serie	Especie					
1.- <i>Chilita</i>		1.- <i>Mammillaria dioica</i>	N/A	900	1000	-	-
2.- <i>Cocheimea</i>		2.- <i>M. poselgeri</i>	N/A	900	900	-	-
3.- <i>Dolichotele</i>		3.- <i>M. baumii</i>	N/A	900	900	890	-
4.- <i>Krainzia</i>	1.- <i>Herrerae</i> + <i>Pectiniferae</i>	4.- <i>M. albiflora</i>	1100*	-	800	-	-
	2.- <i>Longiflorae</i>	5.- <i>M. hernandezii</i>	N/A	900	900	-	-
5.- <i>Mamillaria</i>	3.- <i>Decipientes</i>	6.- <i>M. decipiens</i>	N/A	-	900	-	-
	4.- <i>Heterochlorae</i>	7.- <i>M. wiesingeri</i>	N/A	-	1300	-	-
	5.- <i>Lasiacanthae</i>	8.- <i>M. lasiacantha</i>	N/A	-	900	-	-
	6.- <i>Leptocladodae</i>	9.- <i>M. densispina</i>	N/A	-	1300	-	-
	7.- <i>Leucocephalae</i>	10.- <i>M. geminispina</i>	N/A	900	N/A	-	-
	8.- <i>Mammillaria</i>	11.- <i>M. uncinata</i>	N/A	-	1300	-	-
	9.- <i>Polycanthae</i>	12.- <i>M. rekoii</i>	N/A	-	1300*	-	-
	10.- <i>Polydrae</i>	13.- <i>M. mystax</i>	N/A	900	1000	-	-
	11.- <i>Proliferae</i>	14.- <i>M. schiedeana</i>	N/A	900	1300	-	-
	12.- <i>Rhodanthae</i>	15.- <i>M. rhodantha</i>	N/A	900	900	-	-

	13.- <i>Sphacelatae</i>	16.- <i>M. sphacelata</i>	N/A	900	1300*	-	-
	14.- <i>Stylothelae</i>	17.- <i>M. bocasana</i>	N/A	-	900	-	-
	15.- <i>Supertextae</i>	18.- <i>M. supertexta</i>	-	900	N/A	-	-
6.- <i>Mammillopsis</i>		19.- <i>M. senilis</i>	N/A	900	-*	-	-
7.- <i>Oehmea</i>		20.- <i>M. beneckeii</i>	N/A	900	1300	850	-
8.- <i>Phellosperma</i>		21.- <i>M. zephyranthoides</i>	-	900	1300	850	1100

En cuanto al análisis filogenético, la topología del árbol del locus *matK* (figura 5), no fue congruente con las otras dos, Sin embargo, si lo fue con el árbol del locus *ycf2-rpl23* y el árbol concatenado de *ycf2-rpl23/matK* (figura 6).

La filogenia obtenida en los tres análisis presenta una topología poco resuelta. Esto se refleja también en los bajos valores de soporte bootstrap que indican poca resolución filogenética entre las 18 especies analizadas. La topología de los tres árboles no es concordante con las estructuras del ADNcp de *Mammillaria*, pues en las ramas más basales se encuentran especies cuya estructura, según los resultados encontrados en este trabajo, pertenecen a la estructura 2 y 3. Además se encontraron especies que no coincidieron con ninguna de las tres estructuras. Sin embargo, para los tres árboles se puede observar que las especies *M. supertexta*, *M. mystax*, *M. rhodantha* y *M. geminispina* que se encuentran juntas formando un grupo pertenecen a la estructura 2 según lo encontrado en este trabajo y el subgénero *Mammillaria* en la clasificación de Hunt (2006).

En los tres análisis *C. argentea*, *P. oleracea* y *A. filamentosa*, que funcionan como grupos externos, formaron un grupo bien diferenciado y hermano al formado por *C. gigantea* y *L. schottii* que pertenecen a Cactaceae.

En el análisis del locus *ycf2-rpl23*, los valores de soporte resultaron ser los más altos de las tres filogenias, con valores que van desde el 47.97% hasta el 25.93% entre los nodos. En el análisis del locus *matK* la filogenia muestra una topología poco resuelta y politomías y los valores de soporte varían entre 13.32% y 10.74%. Finalmente en el análisis de los locus concatenados *ycf2-rpl23/matK*, no se muestran politomías como en el caso de la filogenia obtenida únicamente con *matK*. En este caso, los valores de soporte van desde 16.75% hasta 7.61%. Cabe señalar que los valores en las especies del grupo externo y el grupo de *C. gigantea* y *L. schottii* son mayores que en el análisis de *matK*, pero en las especies de *Mammillaria*, resultaron ser los más bajos en los tres análisis.

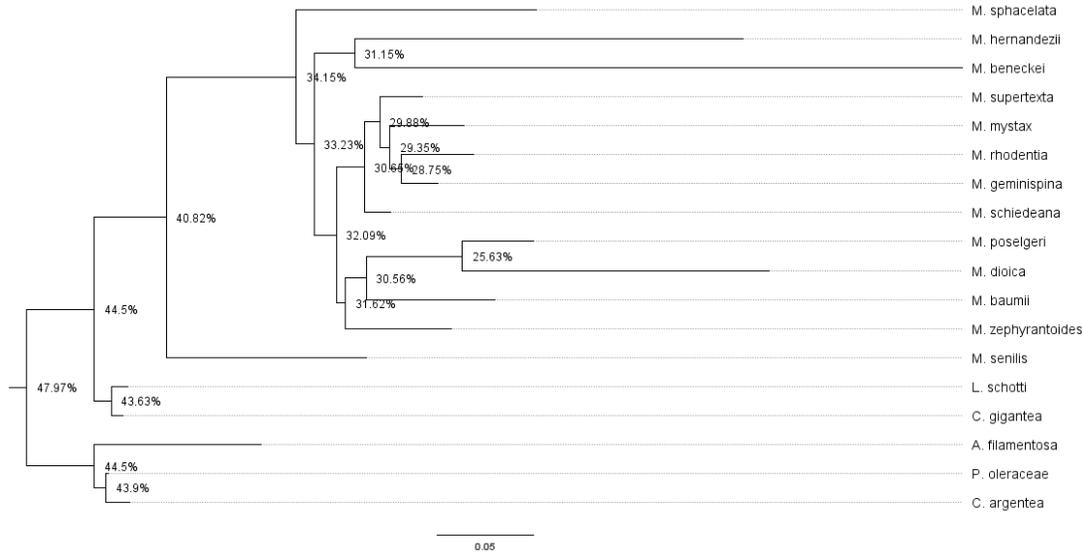


Figura 4. Árbol filogenético obtenido con Máxima Verosimilitud basado en las secuencias del locus *ycf2-rpl23* de 13 especies de *Mammillaria*, usando como grupos externos a otras especies del orden Caryophyllales *L. schotti* y *C. gigantea* (Cactaceae), *A. filamentosa* (Anacamperotaceae), *C. argentea* (Amaranthaceae) y *P. oleracea* (Portulacaceae). El análisis se realizó bajo el modelo GTR + G. El soporte estadístico bootstrap se indica con los valores de porcentaje de 1000 réplicas. La barra en la parte inferior indica 0.05 sustituciones por nucleótido.

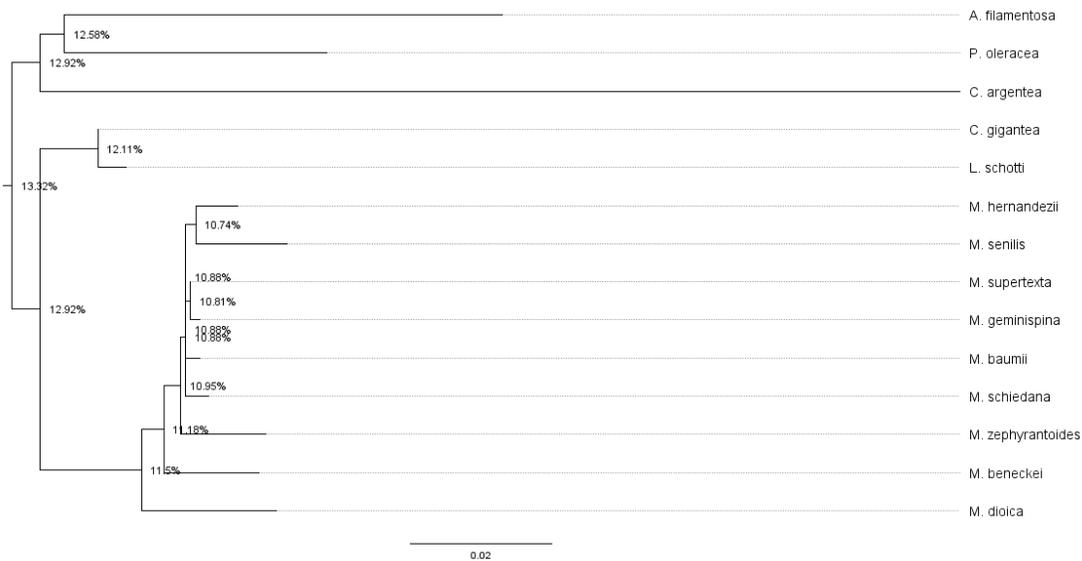


Figura 5. Árbol filogenético de Máxima Verosimilitud basado en las secuencias del locus *matK* de nueve especies de *Mammillaria*, usando como grupos externos a otras especies del orden Caryophyllales *L. schotti* y *C. gigantea*, *A. filamentosa*, *C. argentea* y *P. oleracea*. El análisis se realizó bajo el modelo GTR + G. El soporte estadístico bootstrap se indica con los valores de porcentaje de 1000 réplicas. La barra en la parte inferior indica 0.05 sustituciones por nucleótido.

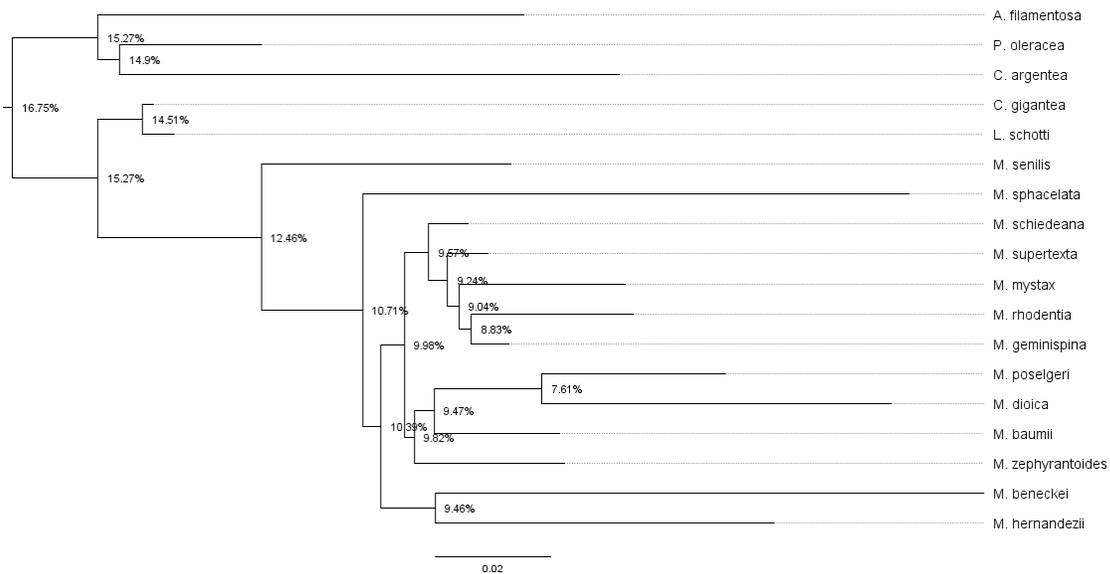


Figura 6. Árbol filogenético de Máxima Verosimilitud basado en las secuencias concatenadas de los loci *ycf2-rpl23* y *matK* de 13 especies de *Mammillaria*, usando como grupos externos a otras especies del orden Caryophyllales *L. schotti*, *C. gigantea*, *A. filamentosa*, *C. argentea* y *P. oleracea*. El análisis se realizó bajo el modelo GTR + G. El soporte estadístico bootstrap se indica con los valores de porcentaje de 1000 réplicas. La barra en la parte inferior indica 0.05 sustituciones por nucleótido.

DISCUSIÓN

Los resultados de esta tesis muestran que las estructuras de ADNcp de las 21 especies representantes de los ocho subgéneros y 15 series de *Mammillaria* aquí estudiadas, no son concordantes con los niveles taxonómicos de subgénero y serie enunciados en la clasificación infragenérica para *Mammillaria* propuesta por Hunt (2006). Este resultado es concordante con lo reportado por Solórzano *et al.* (2019) para siete especies de *Mammillaria* que, tampoco fueron concordantes con esta clasificación.

La clasificación para todos los miembros de la familia Cactaceae más extendida es la propuesta por Hunt (2006), la cual está basada en caracteres morfológicos, sin

embargo, diversos autores han cuestionado esta clasificación pues lo encontrado por Butterworth *et al.* (2002), Butterworth & Wallace (2004), Crozier (2005) y (Bárcenas *et al.*, 2011) sugieren que *Mammillaria* es un clado polifilético a varios niveles, dificultando la tarea de reconocer grupos “naturales” dentro de las clasificaciones de Luthy (1995) y Hunt (2006), debido a las convergencias morfológicas producto de la homoplasia (Sanderson & Donoghue, 1989). Los resultados de este trabajo sobre las relaciones del subgénero *Dolichotele* (representado por *M. baumii*) que presenta la estructura 2 al igual que tres series del subgénero *Mammillaria* y del subgénero *Oehmea* (representado por *M. benecke*), que presenta la estructura 3 al igual que cinco series del subgénero *Mammillaria*, apoyan el origen polifilético del subgénero *Mammillaria* y su relación cercana con los subgéneros *Dolichotele* y *Oehmea*, que propone Harpke *et al.* (2006) y Butterworth & Wallace (2004). Estos autores han propuesto cambios a la taxonomía del grupo. La separación de taxones tan diversos como *Mammillaria* por su aparente origen polifilético ya se ha observado en angiospermas, tal es el caso de la familia Leguminosae (orden Fabales) donde existe la subfamilia Papilionoideae, cuyos IRs están ausentes en seis tribus de un total de 25 que pertenecen a esta subfamilia. En este caso se sugirió que, desde una perspectiva parsimoniosa, un ancestro relacionado con la tribu *Milletieae*, que no presenta esta delección de los IRs, heredó esa pérdida a los descendientes de las seis tribus de la subfamilia. Con este resultado se demostró que la estructura del ADNcp puede apoyar la evidencia basada en caracteres morfológicos utilizada en la clasificación de Papilionoideae (Lavin *et al.*, 1990). Debido a esto se demostró que se puede utilizar la estructura del ADNcp, en particular de los IRs, como un método para confirmar las relaciones evolutivas de diferentes grupos vegetales, tal como se plantea en este trabajo con la verificación de las estructuras y su relación con la taxonomía.

La diversidad estructural del ADNcp es concordante con lo reportado para *Mammillaria* por Solórzano *et al.* (2019), pues en este trabajo se encontró que 14 especies de *Mammillaria*, pertenecen a dos de las estructuras que se propusieron por Solórzano *et al.* (2019). También se encontró que cuatro especies no

amplificaron para ninguna de las tres estructuras conocidas. En el trabajo de Solórzano *et al.*, (2019), se documentó que de las siete especies sometidas a una secuenciación masiva para observar su ADNcp, se obtuvieron tres estructuras diferentes. Esta variación es inusual en un género y más aún en una familia. Lo más común es lo encontrado, por ejemplo, en el género *Camellia* (familia Theaceae, orden Ericales), en el cual como resultado de un estudio de secuenciación masiva de ocho especies, se encontró que todas presentan una estructura idéntica (Huang *et al.*, 2014).

La variación estructural de *Mammillaria* también se observa en *C. gigantea* y *L. schotti* que tiene una estructura del ADNcp diferente al resto de los miembros de Caryophyllales, al carecer totalmente de IRs (Sanderson *et al.*, 2015). Por tanto, se observa que la expansión, contracción y pérdida de IRs es un fenómeno extendido en la subfamilia Cactoideae, a la cual pertenecen los cactus columnares y los globosos pequeños del género *Mammillaria*. En particular la contracción de IRs puede observarse en *M. albiflora*, pues Solórzano *et al.* (2019) reporta que sus IRs miden ~670 pb. La pérdida, contracción y expansión de los IRs ya ha demostrado anteriormente la diferenciación de dos taxones, como en las especies de Polygonaceae (Caryophyllales) y en parte de correspondientes a la familia Plumbaginaceae (Caryophyllales). En estas familias se observa una expansión de los IRs. Respecto a esta última familia, la especie *Limonium tenellum* no cuenta con la expansión de los IRs, lo cual la separa taxonómicamente del resto de especies analizadas de este nivel taxonómico. Dicha separación se confirmó con los métodos filogenéticos basados en secuencias de ADN (Yao *et al.*, 2019).

En Caryophyllales existen reportes de cambios estructurales importantes como en *Pisoniella arborescens* (Nyctaginaceae), *Drosophyllum lusitanicum* (Drosophyllaceae), las cuales presentan una contracción de los IRs. Sin embargo al comparar el número de variaciones en la estructura del ADNcp de *Mammillaria* y *Carnegiea*, de la subfamilia Cactoideae, con otras especies de la familia Cactaceae, los resultados no ratifican la diversidad estructural, pues las especies *Weingartia kergliana*, *Blossfeldia liliputiana*, *Maihuenia poeppigii*, *Pereskia*

sacharosa, *Pereskiopsis diguetii* y *Opuntia decumbens* muestran una estructura de los IRs más cercana a la típica de Caryophyllales, pues no presentan contracciones ni expansiones de los IRs (Yao *et al.*, 2019). En otros ordenes de angiospermas también se ha registrado una alta diversidad estructural, como en Geraniaceae (orden Geraniales), en la cual se reportan expansiones, contracciones y pérdidas de IRs (Ruhlman & Jahnsen, 2018).

Los resultados del presente trabajo reportan por primera vez la presencia de la estructura 2 propuesta por Solórzano *et al.* (2019), para las siguientes especies: *M. geminispina*, *M. mystax*, *M. rhodantha*, *M. senilis*, *M. poselgeri*, *M. baumii* y *M. hernandezii*. Además, en el caso de la especie *M. baumii* solamente una parte del LSC del ADNcp amplificó y se encontró la presencia de la estructura 3, por lo que existe la posibilidad de que exista una estructura del ADNcp diferente a la encontrada previamente o una combinación de estas dos estructuras. La estructura 3 está presente en las especies *M. beneckeii*, *M. densispina*, *M. schiedeana*, *M. sphacelata*, *M. uncinata*, *M. wisingeri* y *M. rekoii*. Este resultado está basado en la diferencia de tamaño del amplificado del locus *rpl2-trnI^{CAU}*. La amplificación inconsistente de los loci podría deberse a una mayor diversidad estructural que la registrada por Solórzano *et al.* (2019), al existir especies que no amplificaron para las estructuras propuestas.

En la subfamilia Cactoideae la variación estructural es aparentemente un proceso extendido según el trabajo de Solórzano *et al.* (2019), que encontró tres estructuras identificadas en nueve especies de *Mammillaria*, así mismo, en este trabajo se encontraron varias estructuras, además de especies cuyo ADNcp no corresponde con ninguna de las estructuras buscadas, por lo que existe la posibilidad de estructuras no identificadas aun para *Mammillaria*, además, en el trabajo de Sanderson *et al.*, (2015) también parece existir variación estructural pues la estructura encontrada en *C. gigantea* y *L. schottii* muestra que sus IRs fueron perdidos (como comprobamos en la base de datos del GenBank, código de acceso NC_041727.1). En contraste, la variación encontrada, utilizando los marcadores comúnmente utilizados como *matK*, *trnK*, *rbcL*, *rpoB* y algunos otros

para *Mammillaria* y géneros afines, ha demostrado ser insuficiente para la resolución filogenética a nivel de género y de especies, pues se presentan politomías en los nodos de las especies y se presenta un bajo soporte para varios de sus clados (Butterworth & Wallace, 2004; Crozier, 2005; Barcenas, 2011). De hecho, el bajo soporte también puede observarse en la filogenia realizada con *matK* en este trabajo, como se detalla más adelante. Sin embargo en el estudio de Chincoya *et al.* 2020, se utilizó la genómica comparativa para encontrar las secuencias de ADN con altos niveles de variación molecular para nueve especies de *Mammillaria*, entre las cuales se encuentran *M. albiflora*, *M. supertexta* y *M. zephyrantooides*, las cuales se incluyen en el presente trabajo. La alta variación encontrada en las especies es concordante con lo hallado en la familia Geraniaceae (orden Geraniales), para la cual los cambios estructurales del ADNcp se han correlacionado con la aceleración en la tasa de sustitución de los nucleótidos, con el aumento de secuencias repetidas (Guisinger *et al.*, 2008) y con la estabilidad del ADNcp (Blazier *et al.*, 2014). Sin embargo, este no es el caso en *M. zephyrantooides*, ya que a pesar de tener la estructura de ADNcp con los IRs más grandes, tiene el menor número de secuencias repetidas, que son fuentes importantes de variación. Esto concuerda con lo reportado en muchas especies de angiospermas donde se encontró que la tasa de sustitución de los genes localizados en copias dentro de los IRs, es 3.7 veces más lenta que la de los genes que tienen una sola copia (Zhu *et al.*, 2016). Es decir, que en el caso de *M. zephyrantooides* se reduce la tasa de sustitución y se preserva la homogeneidad de las secuencias dentro de los IRs (Ruhlman & Jansen, 2014). La aceleración de sustitución de nucleótidos se registró principalmente en genes que codifican para las proteínas ribosomales y en las unidades de la ARNp que codifican en el cloroplasto (Weng *et al.*, 2014).

La razón por la que los marcadores más utilizados como *matK*, el intron 2 del *trnK*, *rbcL*, *rpoB*, *rpl16*, *rpoC1*, *trnK-psbA*, *rpl20-rps12*, *trnL-F* y *trnT-trnL*, no logran ordenar a las especies de Cactaceae resolviendo sus relaciones evolutivas, puede deberse a la baja tasa de sustitución del ADNcp, como resultado del cambio poblacional lento, pues las cactáceas pueden vivir varias décadas y hasta cientos

de años (Steenbergh & Lowe, 1977). También se puede deber a su reciente diversificación. A partir de los estudios moleculares de Arakaki, *et al.* (2011), los autores sugirieron que los eventos de diversificación para los taxones más numerosos de Cactaceae tales como *Opuntia*, *Mammillaria*, *Coryphantha* y *Phyllocactaeae*, ocurrieron en el periodo Mioceno tardío (hace 5 - 10 millones de años). Sin embargo, los resultados reportados por Chincoya *et al.* (2020), indican que se puede observar una alta variación fuera de los marcadores más comunes y que dicha variación en el mismo genoma ha ocurrido en diferentes procesos evolutivos. Uno de ellos está basado en la velocidad de sustitución de bases dentro y fuera de los IRs, que parece ser muy heterogéneo en las distintas regiones del ADNcp. El otro proceso es de tipo estructural, ya que el tiempo de replicación del ADNcp (Mizuno *et al.*, 2013), el “splicing” y la estabilidad del genoma afectan los procesos de replicación para ciertos genes (Leach, 1994).

El resultado obtenido en este trabajo sobre la especie *M. hernandezii* apoya la discrepancia encontrada por Solórzano *et al.* (2019), en relación a la cercanía de las especies del subgénero *Krainzia*. Así, *M. hernandezii*, (subgénero *Krainzia*), amplificó para la estructura 2, al igual que *M. solisoides* (subgénero *Krainzia*) en el trabajo de Solórzano *et al.* (2019), ambas especies presentan la estructura 2 y el mismo subgénero. Para Hunt (2006) *M. solisoides* es una subespecie de *M. pectinifera*, la cual tiene la estructura 1 según lo reportado por Solórzano *et al.* (2019) que proponen a *M. solisoides* como especie independiente y en acuerdo con Arias, *et al.* (2012).

Debido a que *M. bocasana*, *M. decipiens*, *M. dioica* y *M. lasiacantha* no amplificaron para ninguno de los loci de los flancos de los IRs o de LSC, se pueden manejar dos hipótesis. La primera es que estas especies podrían tener otras estructuras del ADNcp, aún no documentadas para *Mammillaria*. La segunda hipótesis, es que ellas cuenten con alguna de las estructuras del ADNcp ya documentadas, pero que tengan rearrreglos o mutaciones en los genes donde se anclan los oligonucleótidos durante la PCR. La diversidad estructural potencial observada en *Mammillaria*, permite considerar el diseño de nuevos estudios y

herramientas como los que optan por el análisis la estructura del ADNcp y otras partes del genoma, además de los análisis filogenéticos con secuencias de bases, para validar propuestas filogenéticas y taxonómicas para este y otros taxones de la familia Cactaceae.

En cuanto al análisis filogenético, el locus propuesto para resolver las relaciones filogenéticas fue *ycf2-rpl23*, el cual se utilizó para observar la estructura del ADNcp. Así, al compararlo con el marcador *matK*, los valores de soporte de las repeticiones de la topología resultan mayores a *matK* que resulta ser el locus más comúnmente utilizado para resolver las relaciones filogenéticas en Embryophyta (Yesson *et al.*, 2011).

El locus *ycf2-rpl23*, incluye la región *trnI-ycf2* encontrada por Chincoya *et al.* 2020, la cual representa una de las 20 regiones con más nivel de variación encontrados en su estudio. Su nivel de variación es concordante con los resultados encontrados en este estudio, pues aunque no logró resolver con suficiente nivel de variación y soporte un árbol filogenético, si demostró una eficiencia mayor al árbol basado en *matK* al compararlos.

Los tres árboles obtenidos con los loci *ycf2-rpl23*, *trnK* y *ycf2-rpl23/trnK* mostraron valores de bootstrap inferiores al 50%, por lo que entran en la categoría de “sin soporte” (Zomlefer *et al.*, 2001). En trabajos posteriores debería ser resuelto con otros marcadores que tengan un mejor soporte estadístico y se encuentren en todas especies de *Mammillaria*, también podría utilizarse una mezcla de varios marcadores.

M. geminispina, *M. mystax*, *M. rhodantha* y *M. supertexta* forman un grupo monofilético en el árbol de consenso del locus *ycf2-rpl23* y del árbol concatenado *ycf2-rpl23/matK*, lo que concuerda con la filogenia de Butterworth & Wallace (2004). Estas especies se encuentran en un solo clado, lo que indica su cercanía, excluyendo a otras también consideradas por Butterworth & Wallace (2004) que están en este trabajo. Además todas estas especies tienen la estructura 2 del ADNcp propuesto para *Mammillaria* (cuadro 4) y al subgénero *Supertexta* según Hunt (2006). De esta manera se puede observar que esta parte del árbol obtenido

es congruente con la estructura del ADNcp y con la clasificación taxonómica de Hunt (2006). Sin embargo, *M. sphacelata* no se agrupa con estas especies, aunque pertenece al mismo subgénero, *Mammillaria*.

En el árbol de consenso de *ycf2-rpl23* y en el concatenado *ycf2-rpl23/matK*, las especies *M. geminispina*, *M. rhodantha*, *M. supertexta* (subgénero *Mammillaria*) y *M. senilis* (subgénero *Mammillopsis*) forman un grupo monofilético. Todas ellas presentan la estructura 2. Esto coincide con lo encontrado por Crozier (2005), quien en su árbol de consenso para Cactaceae, en el cual incluyó a 44 especies de *Mammillaria*, siete de ellas también aquí estudiadas, muestra que *M. geminispina*, *M. rhodantha*, *M. senilis* y *M. supertexta* forman parte de una rama que genera un grupo monofilético, mientras que *M. beneckeii*, *M. dioica* y *M. sphacelata* forman otro, al igual que en la topología del árbol *ycf2-rpl23* y en el árbol concatenado *ycf2-rpl23/matK* donde se observan como especies excluidas del grupo núcleo del árbol que contiene a las especies del subgénero *Mammillaria* y la estructura 2 del ADNcp. Estas especies también tienen estructuras diferentes a la estructura 2, ya que *M. beneckeii* y *M. sphacelata* tienen la estructura 3 y *M. dioica* no tiene una estructura conocida. Es así como las especies del subgénero *Mammillaria* utilizadas en este estudio parecen tener la estructura 2 del ADNcp, mientras que las asociadas a la estructura 3, pertenecen a otros subgéneros, excepto por la posición de *M. senilis* en los árboles *ycf2-rpl23* y el concatenado *ycf2-rpl23/matK*.

Además, en el caso de *M. senilis*, el locus *rpl2-trn^{CAU}* (anexo 3), presentó una banda doble de amplificación lo que significa que existe más de una copia del locus de diferente tamaño o una región de inicio del locus (Hauge & Litt, 1993). En otros casos la banda no amplificó de manera consistente, ni con la calidad necesaria para mandarlo a secuenciar (anexo 2 y 3).

Los resultados de esta tesis difieren con la idea de Crozier (2005) de que solo las especies pertenecientes al subgénero *Mammillaria* representan un linaje monofilético en el género *Mammillaria*, como muestran las estructuras 2 y 3 que están presentes en especies del subgénero *Mammillaria* y en otros subgéneros,

sugiriendo un origen polifilético de este subgénero. Además, existen especies cuya estructura no está identificada dentro y fuera del subgénero *Mammillaria*, lo que aumenta la incertidumbre sobre el origen del subgénero.

Es posible que la estructura del genoma completo de cloroplasto sea un marcador que permita esclarecer no solo la filogenia sino también la taxonomía dentro de *Mammillaria*, para lo que se sugiere extender el muestreo taxonómico e implementar una estrategia de multiarreglos para disminuir los costos y el tiempo de la obtención de los resultados.

CONCLUSIONES

Los resultados de esta tesis, indican que los loci aquí estudiados sirven para identificar arreglos génicos en los flancos de los IRs y de la región LSC.

Tres especies del subgénero *Mammillaria* y una del subgénero *Chilita* no presentaron ninguna de las tres estructuras identificadas, lo que permite sugerir que se presenten altos niveles de mutación en las secuencias de ADN que no permitieron el anclaje de los oligonucleótidos, o incluso que pueda haber nuevas estructuras que no han sido aún identificadas y que podrían descubrirse en estudios futuros.

Los marcadores utilizados no resolvieron con suficiente resolución las relaciones entre las especies estudiadas como para tener un resultado concluyente, sin embargo, el árbol obtenido muestra que las especies que tienen la estructura 2 y que pertenecen al subgénero *Mammillaria* se agrupan, sugiriendo una relación filogenética.

LITERATURA CITADA.

Anderson, E. F. 2001. The cactus family. Portland, Oregon, U.S.A. 776. Timber Press.

Arakaki, M., Christin, P.-A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., Springs, E., Moore, M. J. & Edwards, E. J. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences*, 108, (20), 8379-8384.

Arias, S. 1998. La sistemática de Cactáceas en México, breve recuento y perspectivas. *Boletín de la Sociedad Botánica Mexicana*, 63, 153-165.

Arias, S., Guzmán, U., Mandujano, M. C., Soto, M., & Golubov, J. 2005. Las especies mexicanas de cactáceas en riesgo de extinción. I. Una comparación entre los listados NOM-059-ECOL-2001 (México), La lista Roja (UICN) y CITES. *Cactáceas y Suculentas Mexicanas*, 50, (4), 100-125.

Arias, S., S. Gama, B. Vázquez & L. U. Guzmán. 2012. Flora del Valle de Tehuacán-Cuicatlán. Fascículo 95, Cactaceae Juss. México. Instituto de Biología. Universidad Nacional Autónoma de México.

Bárcenas, R. T., Yesson, C., & Hawkins, J. A. 2011. Molecular systematics of the Cactaceae. *Cladistics*, 27, (5), 470-489.

Blazier, J. C., Jansen, R. K., Mower, J. P., Govindu, M., Zhang, J., Weng, M. L. & Ruhlman, T. A. 2016. Variable presence of the inverted repeat and plastome stability in *Erodium*. *Annals of Botany*, 117, (7), 1209–1220.

Bottomley, W., & Bohnert, H. J. 1982. The biosynthesis of chloroplast proteins. En: Nucleic acids and proteins. En plants II, Berlin, Alemania, Heidelberg. 531-596. Springer.

Bravo-Hollis, H & Sánchez-Mejorada, H 1978. Las cactáceas de México: Vol. 1. D.F. México. 744. Universidad Nacional Autónoma de México.

Bravo-Hollis, H & Sánchez-Mejorada, H. 1991. Las cactáceas de México: Vol. 3. D.F. México. 564. Universidad Nacional Autónoma de México.

Brázda, V., Lýsek, J., Bartas, M., & Fojta, M. 2018. Complex Analyses of Short Inverted Repeats. En All Sequenced Chloroplast DNAs. *BioMed. Research International*, 2018, (1), 1-10.

Brouard, J. S., Otis, C., Lemieux, C., & Turmel, M. 2010. The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the Chlorophyceae. *Genome Biology and Evolution*, 2010, (2), 240-256.

Butterworth, C. A., & Wallace, R. S. 2004. Phylogenetic studies of *Mammillaria* (Cactaceae)-insights from chloroplast sequence variation and hypothesis testing using the parametric bootstrap. *American Journal of Botany*, 91, (7) 1086-1098.

Butterworth, C. A., Cota-Sanchez, J. H., & Wallace, R. S. 2002. Molecular systematics of tribe Cacteeae (Cactaceae: Cactoideae): a phylogeny based on rpl16 intron sequence variation. *Systematic Botany*, 27, (2), 257-270.

Chaw, SM., Wu, SC. & Sudianto, E. 2018. Evolution of gymnosperm plastid genomes. *Advances in botanical research*. 85, 195-222.

Chincoya, D. A., Sanchez-Flores, A., Estrada, K., Díaz-Velásquez, C. E., González-Rodríguez, A., Vaca-Paniagua, F., Dávila, P., Arias, S. & Solórzano, S. (2020). Identification of High Molecular Variation Loci in Complete Chloroplast Genomes of *Mammillaria* (Cactaceae, Caryophyllales). *Genes*, 11(7), 830.

CITES, 2020. Lista de especies CITES. Secretaría CITES, Ginebra, Suiza, y Cambridge, Reino Unido. En: <http://checklist.cites.org/#/es>. Consultado el 28 de noviembre de 2020.

Clegg M.T., Gaut B.S., Learn G.H. & Morton B.R. 1994. Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 91, (15), 6795-6801.

Crawley, S. S., & Hilu, K. W. 2012. Caryophyllales: Evaluating phylogenetic signal in trnK intron versus *matK*. *Journal of Systematics and Evolution*, 50, (5), 387-410.

Crozier, B. S. 2005. Systematics of Cactaceae Juss: phylogeny, cpDNA evolution, and classification, with emphasis on the genus *Mammillaria*. Haw. Tesis Doctoral. Universidad de Austin, Texas.

de Vries, J., & Archibald, J. M. 2018. Plastid Autonomy vs Nuclear Control Over Plastid Function. En *Advances in Botanical Research*, 85, 1-28. E.U.A. Academic Press.

DOF (Diario Oficial de la Federación) 2020. Proyecto de Modificación del Anexo Normativo III, Lista de especies en riesgo de la Norma Oficial Mexicana NOM-059-SEMARNAT-2010, Protección ambiental-Especies nativas de México de flora y fauna silvestres-Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio-Lista de especies en riesgo. México, 30 de diciembre de 2010. Recuperado de: https://dof.gob.mx/nota_detalle_popup.php?codigo=5173091, consultado el 13 de abril de 2020.

Downie, S. R., & Palmer, J. D. 1992. Use of chloroplast DNA rearrangements in reconstruction plant phylogeny. Soltis. En *Molecular systematics of plants*. 14-35. Boston, MA. E.U.A. Springer.

Duminil, J., & Di Michele, M. 2009. Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystems*, 143, (3), 528-542.

Eguiarte, E. A., Castillo, A., Souza, V. 2003. Evolución Molecular y genómica en angiospermas. *Interciencia*, 28, (3), 141-147.

Gray, M. W., & Doolittle, W. F. (1982). Has the endosymbiont hypothesis been proven?. *Microbiological reviews*, 46, (1), 1-42.

Guisinger, M. M., Kuehl, J. V., Boore, J. L., & Jansen, R. K. (2008). Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (47), 18424-18429.

Guzmán, U., Arias, S., & Dávila, P. 2007. Catálogo de cactáceas mexicanas. 1ra reimpresión México. Universidad Nacional Autónoma de México y Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO).

Hall TA. 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 41 (41), 95–98.

Harpke, D., Peterson, A., Hoffmann, M. H., & Röser, M. 2006. Phylogenetic evaluation of chloroplast trnL–trnF DNA sequence variation in the genus *Mammillaria* (Cactaceae). *Schlechtendalia*, (14), 7-16.

Hauge, X. Y., & Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Human Molecular Genetics*, 2 (4), 411-415.

Hernández-Hernández, T., Brown, J. W., Schlumpberger, B. O., Eguiarte, L. E., & Magallón, S. 2014. Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist*, 202 (4) 1382-1397.

Hernández-Ledesma, P., Berendsohn, W. G., Borsch, T., Von Mering, S., Akhiani, H., Arias, S., Castañeda-Noa, I., Eggl, I., Eriksson, R., Flores-Olvera, H., Fuentes-Bazán, S., Kadereit, G., Klak, C., Korotkova, N., Nyffeler, R., Ocampo, G., Ochoterena, H., Oxelman, G., Rabeler, R. K., Sánchez, A., Schlumperberger, B. O. & Uotila, P. 2015. A taxonomic backbone for the global synthesis of species diversity in the angiosperm order Caryophyllales. *Willdenowia*, 45(3), 281-383.

Huang, H., Shi, C., Liu, Y., Mao, S. Y., & Gao, L. Z. 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *Evolutionary biology*, 14 (1), 151.

Hunt, D., Taylor, N., & Charles, G. 2006. The New Cactus Lexicon, Vols. I & II: Descriptions and Illustrations of the Cactus Family. Milborne Port, Inglaterra. DH Books.

IUCN Red List of Threatened Species. Versión 2020-2. www.iucnredlist.org. Consultado el 28 de noviembre de 2020.

Jensen, P. E., & Leister, D. 2014. Chloroplast evolution, structure and functions. *F1000prime reports*, 6, (40), 6-40.

Jiao, Y., & Guo, H. 2014. Prehistory of the Angiosperms: characterization of the ancient genomes. En *Advances in Botanical Research*, 69, 223-245 Academic Press.

Katoh, K.; Daron, M.S. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology Evolution*, 30, 772–780.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* 35 (6), 1547-1549.

Lavin, J., Doyle, J., & Palmer, J. D. 1990. Evolutionary Significance of the Loss of the Chloroplast-DNA Inverted Repeat in the Leguminosae Subfamily *Papilionoideae*,” *Evolution*, 44 (2), 390.

Leach, D.R. 1994. Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *BioEssays*, 16 (12), 893–900.

Lemieux, C., Otis, C., & Turmel, M. 2000. Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature*, 403 (6770), 649-652.

Lemieux, C., Otis, C. & Turmel, M. 2016. Comparative Chloroplast Genome Analyses of Streptophyte Green Algae Uncover Major Structural Alterations in the Klebsormidiophyceae, Coleochaetophyceae and Zygnematophyceae. *Frontiers in Plant Science*, 2016 (7), 697.

Lüthy, J. M. 1995. Taxonomische untersuchung der gattung *Mammillaria* Haw. Ph.D. dissertation, Universität Bern, Bern.

Lüthy, J. M. 2001. A revised classification of the "primitive" mammillarias. *Journal of the Mammillaria Society*, 41, 6-7.

Maddison, W. P. & Maddison D.R. 2019. Mesquite: a modular system for evolutionary analysis. Version 3.61. Recuperado de <http://www.mesquiteproject.org>. Consultado el 20 de noviembre de 2020.

McFadden, G. I. & van Dooren, G. G. 2004. Evolution: Red algal genome affirms a common origin of all plastids. *Current Biology*. 14 (13), 514-516.

Mizuno, K., Miyabe, I., Schalbetter, S. A., Carr, A. M. & Murray, J. M. 2013. Recombination-restarted replication makes inverted chromosome fusions at inverted repeats. *Nature*, 493 (4731), 246–249.

Mower, J. P., & Vickrey, T. L. 2018. Structural Diversity Among Plastid Genomes of Land Plants. En *Advances in botanical research*. 263-292, E.U.A. Academic Press.

National Center for Biotechnology Information. 2020. Genbank (base de datos). Recuperado de: <https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles/Plants>. Consultado el 19 de enero de 2020.

Nyffeler, R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from *trnK/matK* and *trnL-trnF* sequences. *American Journal of Botany*, 89 (2), 312-326.

Odintsova, M. S., & Yurina, N. P. 2006. Chloroplast genomics of land plants and algae. In *Biotechnological applications of photosynthetic proteins: biochips, biosensors and biodevices* 57-72, Boston, MA. E.U.A. Springer.

Palmer, J. D. 1985. Comparative organization of chloroplast genomes. *Annual review of genetics*, 19 (1), 325-354.

Palmer, J. D., & Delwiche, C. F. 1998. The origin and evolution of plastids and their genomes. In *Molecular systematics of plants II*. 375-409. Springer, Boston, MA.

Palmer, J. D. & Thompson, W. F. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost, *Cell*, 29, (2), 537–550.

Ruhlman, T. A., & Jansen, R. K. 2014. The plastid genomes of flowering plants. En *Chloroplast biotechnology*, Totowa, 3-38, NJ E.U.A. Humana Press,

Ruhlman, T. A., & Jansen, R. K. 2018. Aberration or analogy? The atypical plastomes of Geraniaceae. En *Advances in botanical research*, 85, 223-262. E.U.A. Academic Press.

Sanderson, M. J., Copetti, D., Búrquez, A., Bustamante, E., Charboneau, J. L., Eguiarte, L. E. Kumar. S., Oh Lee, H., Lee, J., McMahon, M., Steele, K., Wing, R., Yang. T., Derrick. Z. & Wojciechowski. M. F. 2015. Exceptional reduction of the plastid genome of saguaro cactus (*Carnegiea gigantea*): Loss of the *ndh* gene suite and inverted repeat. *American Journal of Botany*, 102 (7), 1115-1127.

Sanderson, M. J., & Donoghue, M. J. 1989. Patterns of variation in levels of homoplasy. *Evolution*, 43 (8), 1781-1795.

Shaw, J., Lickey, E. B., Schilling, E. E., & Small, R. L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American journal of botany*, 94 (3), 275-288.

Solórzano, S., Chincoya, D. A., Sanchez-Flores, A., Estrada, K., Díaz-Velásquez, C. E., González-Rodríguez, A., Vaca-Paniagua, F., Davila, P. & Arias, S. 2019. De Novo Assembly Discovered Novel Structures in Genome of Plastids and Revealed Divergent Inverted Repeats in *Mammillaria* (Cactaceae, Caryophyllales). *Plants*, 8 (10), 392.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, (30) 1312–1313.

Steenbergh, W. H. & Lowe. C. H. 1977. Ecology of the Saguaro. II. Reproduction, Germination, Establishment, Growth and Survival of the Young Plant. National

Park Service Scientific Monograph Series No. 8, U.S., Washington, D.C. Government Printing Office.

Taiz L. & Zeiger E. 2002. *Plant Physiology*. 3^a edición. MA. E.U.A. 690. Sinauer Associates, Sunderland.

Thompson, J. D., Higgins, D. G., Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. 22 (22). 4673–4680.

Touchman, J. 2010. Comparative Genomics. *Nature Education Knowledge*. 3 (10).13.

Turmel, C., Otis, & Lemieux, C. 2017. Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophyceean green algae. *Scientific Reports*, 7 (1), 1-14.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M & Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acid Research*. 40 (15), 115.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T., & Sugiura, M. 1994. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proceedings of the National Academy of Sciences*, 91 (21), 9794-9798.

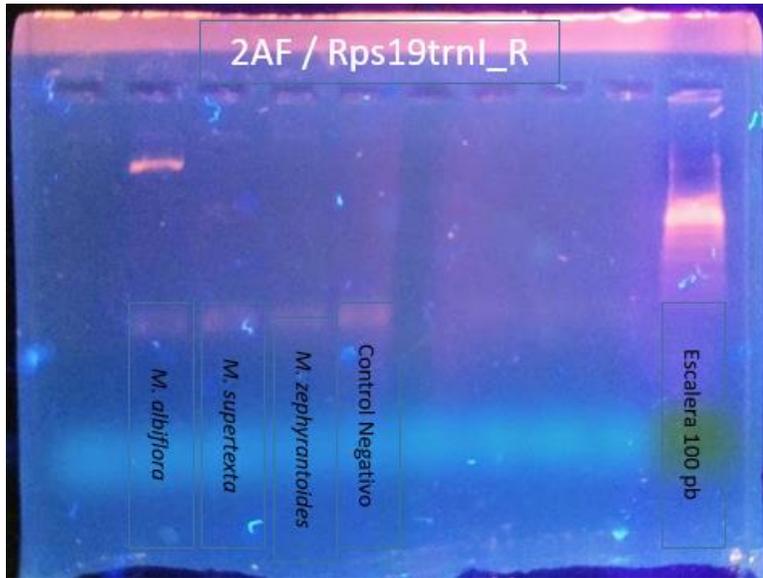
Wallace, R. S. 1995. Molecular systematic study of the Cactaceae: using chloroplast DNA variation to elucidate cactus phylogeny. *Bradleya*, 1995 (13), 1-12.

Weng, M.-L., Blazier, J. C., Govindu, M., & Jansen, R. K. 2014. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Molecular Biology and Evolution*, 31 (3), 645–659.

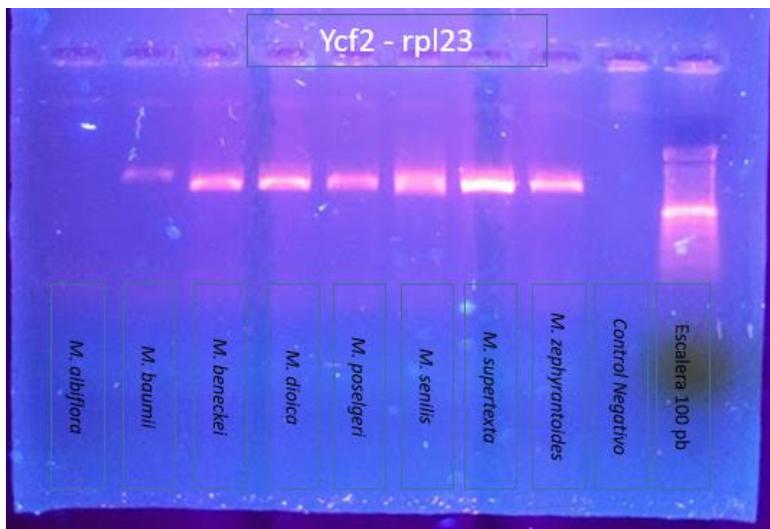
- Whatley, J. M. 1981. Chloroplast evolution--ancient and modern. *Annals of the New York Academy of Sciences*, 1981 (361), 154-165.
- Wicke, S. 2013. Genomic evolution in Orobanchaceae. En *Parasitic Orobanchaceae* 267-286. Berlin, Heidelberg. Springer.
- Wolfe, K. H., Li, W. H., & Sharp, P. M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences*, 84 (24), 9054-9058.
- Wolfe, K. H., Morden, C. W., & Palmer, J. D. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the National Academy of Sciences*, 89 (22), 10648-10652.
- Xia, X. 2011. Comparative genomics. In *Handbook of Statistical Bioinformatics*. Berlin, Heidelberg, Alemania. 567-600. Springer.
- Xu, J. H., Liu, Q., Hu, W., Wang, T., Xue, Q., & Messing, J. 2015. Dynamics of chloroplast genomes in green plants. *Genomics*, 106 (4), 221-231.
- Yao, G., Jin, J. J., Li, H. T., Yang, J. B., Mandala, V. S., Croley, M., & Soltis, D. E. 2019. Plastid phylogenomic insights into the evolution of Caryophyllales. *Molecular phylogenetics and evolution*, 2019 (134), 74-86.
- Yesson, C., Bárcenas, R. T., Hernandez, H. M., Ruiz-Maqueda, M. de la L., Prado, A. Rodriguez, V. M. & Hawkings, J. A. 2011. DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. *Molecular Ecology Resources*, 11 (5), 775-783.
- Zhu, A., Guo, W., Gupta, S., Fan, W. & Mower, J. P. 2016. Evolutionary dynamics of the plastid inverted repeat: The effects of expansion, contraction, and loss on substitution rates, *New Phytologist*, 209 (4), 1747–1756.
- Zomlefer, W., Williams, N., Whitten, W. & Judd, W. 2001: Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae) with emphasis on *Zigadenus*: Evidence from ITS and trnI–F sequence data. *American Journal of Botany*. 88: 1657–1669.

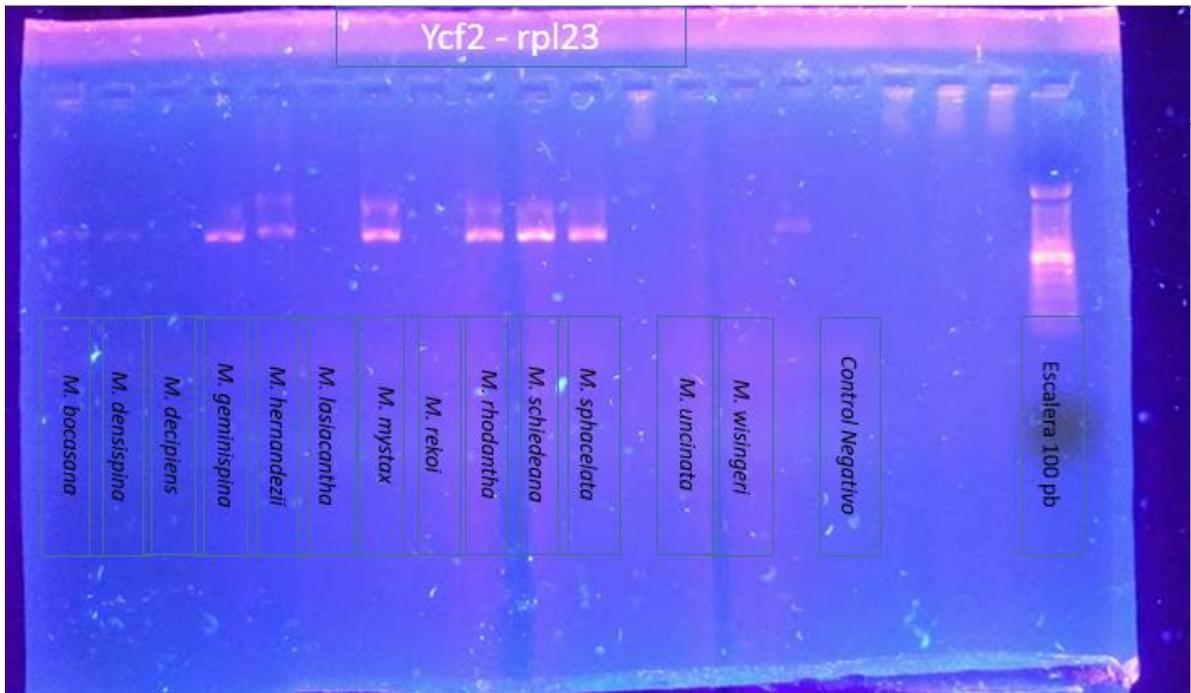
ANEXOS

Anexo 1: amplificados del locus *rps19-trnI^{CAU}*.

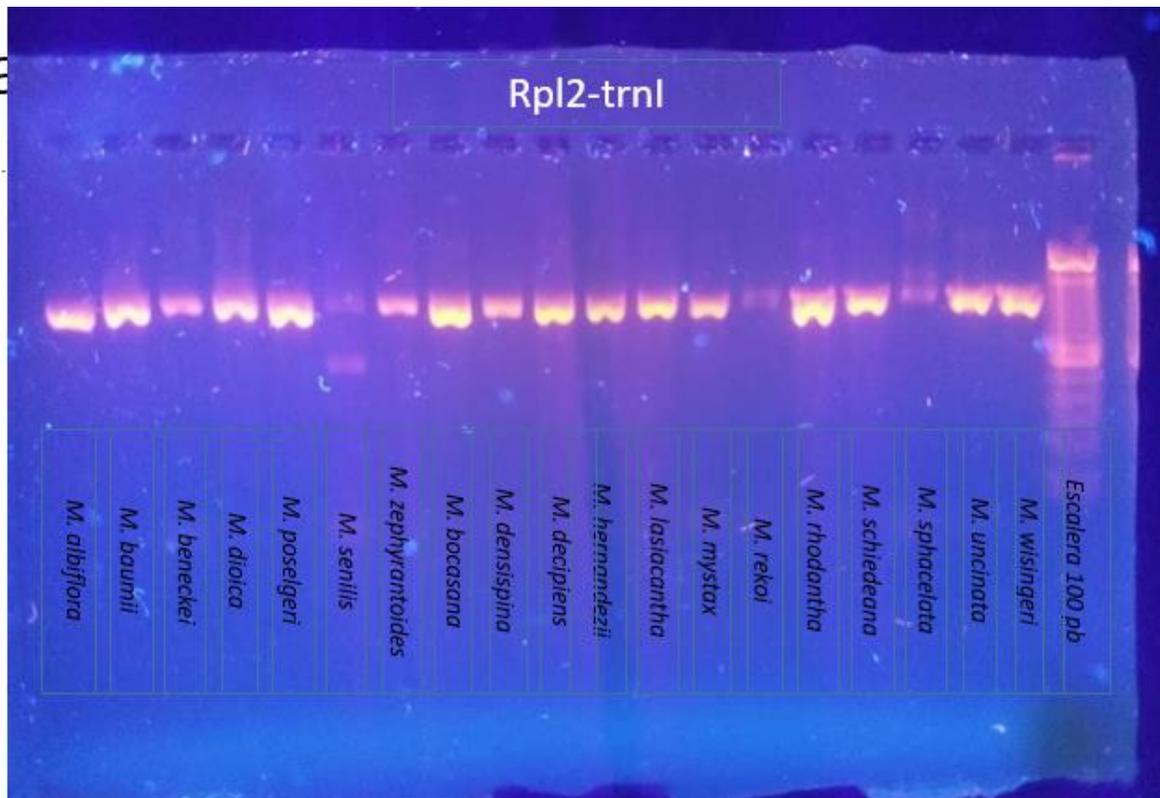


Anexo 2: Amplificados del locus *yfc2-rpl23*.

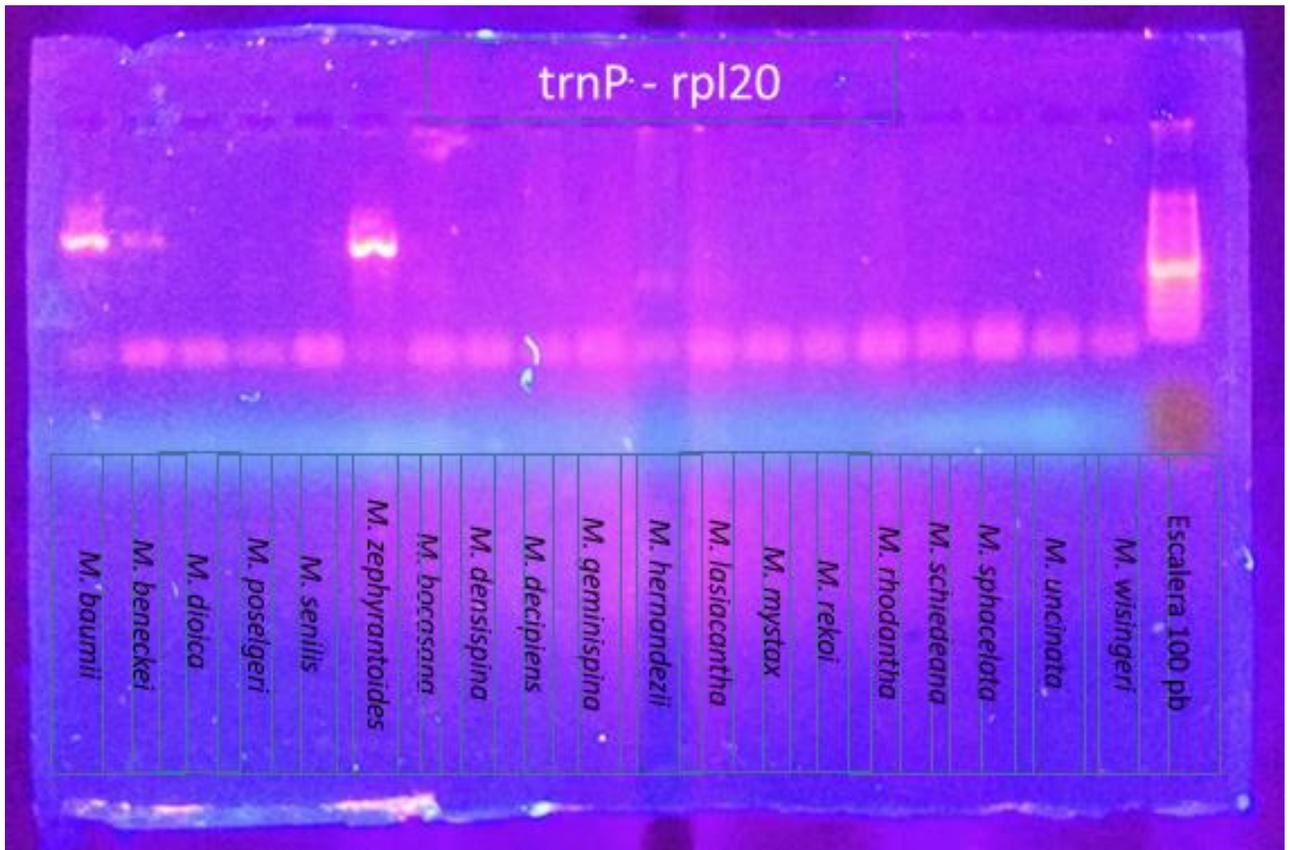
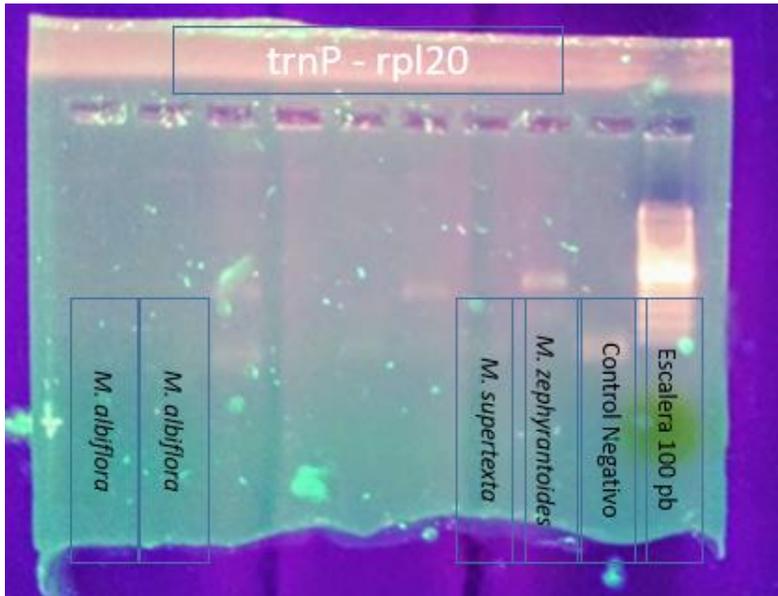




Anexo 3: Amplificados del locus *rpl2-trnI*^{CAU}.



Anexo 4: Amplificados del locus *trnP^{UUU}-rpl20*.



Anexo 5: Amplificados del locus *trnE^{UUC}-psbZ*.

