



Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

**ESTUDIO DE REGIONES ADN SATÉLITE EN TRES
ESPECIES DE LA SECCIÓN *Rigidae* DEL GÉNERO
*Agave***

Tesis que presenta

LAURA ANGÉLICA ESPINOSA BARRERA

En opción al título de

DOCTOR EN CIENCIAS

(Ciencias Biológicas: Opción Biotecnología)

Mérida, Yucatán, México
2018

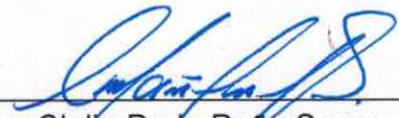
CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.
POSGRADO EN CIENCIAS BIOLÓGICAS



RECONOCIMIENTO

Por medio de la presente hago constar que el trabajo de tesis de doctorado de **Laura Angélica Espinosa Barrera** titulado "Estudio de regiones ADN satélite en tres especies de la sección *Rigidae* del género *Agave*" fue realizado en el Laboratorio de Marcadores Moleculares y Genómica Funcional de la Unidad de Biotecnología, del Centro de Investigación Científica de Yucatán, A.C., en la línea de Agrobiotecnología, bajo la dirección del Dr. Lorenzo Felipe Sánchez Teyer, dentro de la opción Biotecnología perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Atentamente:



Dra. Clelia De la Peña Seaman
Directora de Docencia

Mérida, Yucatán, México, a 30 de octubre de 2018.

DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en la sección de Materiales y Métodos Experimentales, los Resultados y Discusión de este documento proviene de las actividades de experimentación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que a razón de lo anterior y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información, le pertenecen patrimonialmente al Centro de Investigación Científica de Yucatán, A.C., y en el mismo tenor, reconozco que si derivaren de este trabajo productos intelectuales o desarrollos tecnológicos, en lo especial, estos se registrarán en todo caso por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.

Firma:  _____

Laura Angélica Espinosa Barrera

Este trabajo se llevó a cabo en el Laboratorio de Marcadores Moleculares y Genómica Funcional de la Unidad de Biotecnología del Centro de Investigación Científica de Yucatán, en el que participé bajo la dirección del Dr. Lorenzo Felipe Sánchez Teyer.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por la beca otorgada con el número 391947.

Al Centro de Investigación Científica de Yucatán, A.C., por prestarme sus instalaciones para realizar este proyecto.

A la Unidad de Biotecnología del CICY, por las instalaciones proporcionadas para la realización del proyecto.

Al Dr. Lorenzo Felipe Sánchez Teyer, por permitirme formar parte de su laboratorio y brindarme las facilidades para la elaboración de este trabajo.

Al Dr. José Alberto Narváez Zapata, por su apoyo en el asesoramiento del artículo.

A mi comité tutorial conformado por los Dres. Rosa María Escobedo Gracia-Medrano, Germán Carnevali Fernández-Concha y Jorge Rubio Piña, por sus consejos, sugerencias y comentarios para la mejora de este trabajo.

A mi comité de revisión de tesis y examen de grado, conformado por el Dr. Lorenzo Felipe Sánchez Teyer, el Dr. José Alberto Narváez Zapata, la Dra. Rosa María Escobedo Gracia-Medrano, el Dr. Germán Carnevali Fernández-Concha, el Dr. Jorge Rubio Piña, la Dra. Elizabeth Ortiz Vázquez y el Dr. Luis Sáenz Carbonell, muchas gracias por sus aportaciones.

Al Dr. Antonio Rescalvo Morales, por instruirme la técnica de hibridación fluorescente *in situ* (FISH).

A la M. en C. Adriana Quiroz Moreno, por su apoyo técnico.

A mi compañera de estudios M. en C. Sara Hernández Castellano, por su apoyo en la técnica de Southern Blot.

A la IQ. Silvia Andrade Canto, por su apoyo en la edición fotográfica de este trabajo

A la M en C. Fabiola Escalante Erosa por su apoyo técnico.

A las doctoras María Tamayo Ordoñez y Yahaira Tamayo Ordoñez, por su apoyo técnico y observaciones durante los seminarios de grupo.

Al Dr. Jorge Humberto Ramírez Prado, por su apoyo en los análisis bioinformáticos en este proyecto.

Al Dr. Luis Carlos Rodríguez Zapata, por proporcionarme material y reactivos para la realización de la técnica de Southern Blot.

A la M. en C. Ángela Ku González, por su apoyo técnico en el microscopio confocal.

A mis compañeros de grupo Zamaria De la Torre Espinosa, Gerardo Campos Rivero, María José García Castillo y Eduardo Hernández Navarro, por sus sugerencias en los experimentos del laboratorio y seminarios.

DEDICATORIAS

Con cariño para André y Lalo

ÍNDICE

INTRODUCCIÓN.....	1
CAPÍTULO I. ANTECEDENTES	3
1.1. Definición y características del ADN satélite.....	3
1.2. Mecanismos de origen y evolución del AND satélite.....	4
1.3. Técnicas para el aislamiento de ADN satélite.....	5
1.4. El ADN satélite en el estudio de especies filogenéticamente relacionadas	6
1.5. La familia Agavaceae.....	8
1.6. Características fisiológicas y morfológicas de <i>Agave</i>	9
1.7. Cariotipo de <i>Agave</i>	10
1.8. La sección Rigidiae del género <i>Agave</i>	11
1.9. Importancia del género <i>Agave</i>	12
1.10. ADN satélite en <i>Agave tequilana</i>	13
JUSTIFICACIÓN	15
HIPÓTESIS	16
OBJETIVO GENERAL	16
OBJETIVOS ESPECÍFICOS	16
ESTRATEGIA EXPERIMENTAL.....	17
CAPÍTULO II	19
AISLAMIENTO Y CARACTERIZACIÓN DE REGIONES ADN SATÉLITE EN A.	

<i>angustifolia</i> , <i>A. fourcroydes</i> y <i>A. tequilana</i>	19
2.1 INTRODUCCIÓN	19
2.2. MATERIALES Y MÉTODOS	21
2.2.1. Material biológico.....	21
2.2.2. Extracción de ADN.....	21
2.2.3. Diseño de cebadores.	21
2.2.4. Amplificación de fragmentos por PCR	21
2.2.5. Clonación y secuenciación.	222
2.2.6. Análisis de secuencia.....	22
2.2.7. Hibridación Southern Blot.....	22
2.3 RESULTADOS	24
2.3.1. Diseño de cebadores para la amplificación de ADN satélite en <i>Agave</i>	24
2.3.2. Análisis de secuencias de ADN satélite en especies de <i>Agave</i>	25
2.3.3. Análisis e identificación de motivos repetidos en secuencias de ADN satélite de <i>Agave</i>	34
2.3.4. Obtención de secuencias consenso de ADN satélite de las especies de <i>Agave</i>	36
2.3.5. Southern Blot y posible organización de dos subfamilias satélites en <i>Agave</i>	41
2.3.6. Diseño de cebadores para las subfamilias A y B de ADN satélite en <i>A. tequilana</i> , <i>A. fourcroydes</i> y <i>A. angustifolia</i>	42
2.4 DISCUSIÓN.....	44

2.4.1. Aislamiento e identificación de ADN satélite en <i>Agave</i>	44
2.4.2. Identificación de motivos que comprenden las secuencias ADN satélite en <i>Agave</i>	46
2.4.3. Uso de información generada por ADN satélite en plantas	47
CAPÍTULO III	49
MAPEO FÍSICO DE ADN SATÉLITE EN LAS ESPECIES <i>A. tequilana</i>, <i>A. fourcroydes</i> y <i>A. angustifolia</i>	49
3.1. INTRODUCCIÓN.....	49
3.2. MATERIALES Y MÉTODOS.	50
3.2.1. Obtención de preparaciones con células en metafase de <i>Agave</i>	50
3.2.2. Marcaje de la sonda con Digoxigenina (dig-11-dUTP).....	50
3.2.3. Hibridación in situ fluorescente de ADN satélite en especies de <i>Agave</i>	51
3.2.4. Fotografía de células en metafase y análisis de las señales	52
3.3. RESULTADOS	54
3.3.1. Mapeo físico de ADN satélite en tres especies de la sección <i>Rigidiae</i> del género <i>Agave</i>	54
3.4. DISCUSIÓN.....	60
3.4.1. Estudios de ADN satélites implicados en los centrómeros	60
3.4.2. Estudios de ADN satélites implicados en regiones subteloméricas y teloméricas.....	61
3.4.3. Análisis citogenético comparativo de ADN satélites en plantas cercanamente relacionadas.....	62

DISCUSIÓN GENERAL	64
CONCLUSIONES.....	67
PERSPECTIVAS	68
BIBLIOGRAFÍA.....	69

LISTADO DE FIGURAS

Figura 1.1. Hibridación <i>in situ</i> fluorescente (FISH) de Tail en <i>Leymus racemus</i> , TR9 de <i>Pisum sativum</i> y pcR1 de <i>Oryza sativa</i>	3
Figura 1.2. Campo de cultivo de <i>A. tequilana</i> var. Azul.	9
Figura 1.3. Células en metafase después del proceso de “squash”	11
Figura 2.1. Electroforesis de los productos de amplificación para los cebadores ATsat102-113 (A) y ATsat-109 (B)..	255
Figura 2.2. Fenograma realizado a partir de fragmentos amplificados en las especies de <i>A. tequilana</i> , <i>A. fourcroydes</i> , <i>A. angustifolia</i> con el cebador ATsat-102-113.	277
Figura 2.3. Fenograma realizado a partir de fragmentos amplificados por PCR en las especies <i>A. tequilana</i> , <i>A. fourcroydes</i> , <i>A. angustifolia</i> con el cebador ATsat-109.	299
Figura 2.4. Representación de los motivos repetidos internos en la especie <i>A. angustifolia</i>	344
Figura 2.5. Representación de los motivos repetidos internos en la especie <i>A. fourcroydes</i>	355
Figura 2.6. Representación de los motivos repetidos internos en la especie <i>A. tequilana</i>	355
Figura 2.7. Distancias genéticas entre las secuencias ADN satélite de las especies de Agave de este estudio.	377
Figura 2.8. Diversidad y generación de secuencias consenso de ADN satélite para el grupo 1 en especies de <i>Agave</i>	388
Figura 2.9. Diversidad y generación de secuencias consenso de ADN satélite para el grupo 2 en especies de <i>Agave</i>	399
Figura 2.10. Southern Blot para las subfamilias 1 y 2 de satélites.....	411

Figura 2.11. Electroforesis de los fragmentos generados por los cebadores Tipo I y Tipo II en las especies de Agave.. 433

Figura 3.1. Estrategia experimental para la hibridación *in situ* de ADN satélite en especies de Agave. 533

Figura 3.2. Fotografías de preparaciones de células en metafase de *A. tequilana*, *A. fourcroydes* y *A. angustifolia*..... 544

Figura 3.3. Hibridación *in situ* fluorescente en especies de Agave con la sonda I (AACHE102-554)..... 566

Figura 3.4. Hibridación *in situ* fluorescente en especies de Agave con la sonda II (AT102-531).. 577

LISTADO DE TABLAS

Tabla 1.1. Ejemplos de ADN satélites en plantas reportados en la base de datos del PlantSat (Macas <i>et al.</i> , 2002).....	6
Tabla 1.2. El ADN satélite y las aplicaciones que se le han conferido en plantas.....	7
Tabla 1.3. Secciones de los subgéneros <i>Littea</i> y <i>Agave</i>	12
Tabla 1.4. Aplicaciones y usos de las especies del género <i>Agave</i>	13
Tabla 1.5. Secuencias candidatas de ADN satélite en <i>Agave tequilana</i>	14
Tabla 2.1. Características de los cebadores sintetizados a partir de las secuencias candidatas a ADN satélite de <i>A. tequilana</i>	24
Tabla 2.2. Motivos consenso generados a partir del cebador AteqSat102 que se encuentran presentes en regiones de ADN satélite reportadas en el PlantSat (Macas <i>et al.</i> , 2002).....	31
Tabla 2.3. Motivos consenso generados a partir del cebador AteqSat109 que se encuentran presentes en regiones de ADN satélites reportados en el PlantSat (Macas <i>et al.</i> , 2002).....	33
Tabla 2.4. Identificación de secuencias consenso con el PlantSat, el Genbank, con secuencias de TSA (Transcriptome Shotgun Assembly) y con secuencias de una clona BIBAC de <i>A. tequilana</i>	40
Tabla 2.5. Cebadores diseñados a partir de las secuencias del Tipo I y II de AteqSat102.....	42
Tabla 3.1. Identificación de señales en cromosomas de 3 especies de agave Sonda I AACHE102-554. 58	
Tabla 3.2. Identificación de señales en cromosomas de 3 especies de agave. Sonda II AT102-531.....	58

RESUMEN

En el presente estudio se realizó el análisis de ADN satélites en tres especies de *Agave* (*A. angustifolia* Haworth, *A. tequilana* Webber y *A. fourcroydes* Lemaire). La Auto PCR se usó inicialmente para generar productos de PCR repetidos en *A. tequilana* y la información de la secuencia se usó para construir diferentes conjuntos de cebadores. El conjunto de cebadores ATsat-102-113 fue capaz de amplificar productos de PCR repetidos, incluidos probables multímeros que no exhibieron el perfil de escalera típico, en todas las accesiones de *Agave* analizadas. El análisis de la secuencia de estos amplicones identificó dos subfamilias de ADN satélites que comparten seis motivos cortos conservados pero exhiben una gran distancia genética y diferentes rasgos de secuencia entre ellos. Los monómeros obtenidos tienen una longitud superior a 100 pb y son ricos en porcentajes de adenina / timina, como muchos satélites de plantas. Estas subfamilias tipo satélite de *Agave* tienen diferentes tasas de homogeneización y la subfamilia 2 está más conservada. Los monómeros tipo satélite probablemente se transcriben y exhiben una fuerte relación con las características del retrotransposón LTR-Gypsy y con las secuencias de *A. tequilana* TSA, algunas de las cuales muestran dominios relacionados con elementos móviles en ubicaciones diferentes a las secuencias similares a satélites identificadas. El análisis FISH de estas subfamilias similares a satélites solo puede ser confiable determinado en las accesiones diploides (*A. angustifolia* var. *Marginata* y *A. tequilana* 'Azul'). Las señales observadas generalmente mostraron hibridación subteloamérica con señales intersticiales.

ABSTRACT

Repeated DNA analysis was conducted in three *Agave* species (*A. angustifolia* Haworth, *A. tequilana* Webber and *A. fourcroydes* Lemaire). Self-priming PCR was initially used to generate repeated PCR products and their sequence information was then used to build different primers sets. The AT102-113 primer set was able to amplify repeated PCR products, including probably multimers without becoming a typical ladder profile, in all *Agave* accessions analyzed. Sequence analysis of these amplifications identified two subfamilies of repeated DNA which share six conserved short motifs but exhibit a high genetic distance and different sequence traits between them. Consensus monomers obtained have a length higher than 100 bp and are rich in Adenine/Thymine percentage, as many satellites of plants. These *Agave* satellite-like subfamilies have different homogenization rates and the subfamily 2 is more conserved. The satellite-like monomers are transcribed and exhibited a strong relation with transposable element signatures of the LTR/Gypsy families and with *A. tequilana* TSA sequences, some which exhibits transposon-related domains in different locations to the satellite-like sequences identified. FISH analysis of these satellites-like subfamilies only may be reliable determined in the diploid (*A. angustifolia* Haworth var. Marginata and *A. tequilana* 'Azul') accessions. The signals observed were generally showed subtelomeric hybridization with interstitial signals.

INTRODUCCIÓN

El genoma nuclear de las células eucariotas está compuesto de secuencias únicas (genes) y diferentes familias de ADN repetido. Las secuencias repetidas pueden estar altamente representadas en el genoma (30-60%) (Sharma y Raima, 2003; Ribeiro *et al.*, 2017) y provocar cambios genéticos que desempeñan un papel importante en la variabilidad, evolución y divergencia de las plantas (Heslop-Harrison, 2003). De igual manera, las secuencias repetidas contribuyen al tamaño del genoma y a la estructura de los cromosomas influyendo en la variación del genoma (Flavell, 1986; Neumann *et al.*, 2006). Las secuencias repetidas se pueden encontrar organizadas en tándem o dispersas (Kishii, Nagaki y Tsujimoto, 2001; Schmidt y Heslop-Harrison, 1998).

El ADN satélite es una secuencia repetida en tándem que contiene un motivo mayor a 100 pb siendo el constituyente principal de la heterocromatina, se localiza en regiones subtroméricas y centroméricas en los cromosomas de las plantas (Plohl *et al.*, 2008). El estudio del ADN satélite proporciona información que permite comprender la evolución de la heterocromatina (Samoluk *et al.*, 2017), los procesos que dan origen a los ADN satélites (Plohl *et al.*, 2012), las proporciones que ocupan en el genoma (Richard *et al.*, 2013) y cómo evoluciona el genoma de algunas especies (Mehrotra *et al.*, 2014).

El género *Agave* está ampliamente representado en México ya que un gran número de especies se encuentran distribuidas a lo largo de todo el país (Ochoa *et al.* 2006) y son consideradas plantas emblemáticas. Entre las aplicaciones comerciales que tienen estas especies están la obtención de bebidas alcohólicas, como el mezcal y el tequila, la producción de fibras y de compuestos como la inulina o sapogeninas (Narvaez-Zapata y Sánchez-Teyer, 2009).

Los agaves presentan diferentes niveles de ploidía, que van desde diploides hasta octaploides (Castorena-Sánchez *et al.*, 1991; Palomino *et al.*, 2008), tienen un cariotipo bimodal con 25 cromosomas pequeños y 5 grandes (Castorena-Sánchez *et al.*, 1991) y la cantidad de ADN 1Cx varían desde 4.08 pg hasta 9.2 pg (Palomino *et al.*, 2004).

La información que se ha generado sobre los agaves en cuestión de la evolución de sus genomas es poca clara, por lo tanto múltiples preguntas quedan por responder para comprender los procesos evolutivos que presentan las especies poliploides de *Agave*. Recientemente, en este género se ha enfocado al estudio del ADN satélite para dilucidar las relaciones evolutivas de las especies poliploides. En *Agave tequilana* se aislaron

secuencias candidatas a regiones de ADN satélite mediante la técnica de auto-PCR (Espinosa-Barrera, 2014, tesis de maestría). Con este antecedente, en el presente trabajo se realizó un análisis filogenético para determinar la relación ancestral entre tres especies de la sección *Rigidae* del género *Agave*, a través de un análisis comparativo de secuencias de ADN satélite y su localización en el genoma.

Nuestros resultados sugieren que existen dos familias de ADN satélites, las cuales pertenecen a cada uno de los dos pares de cebadores ATsat102-113 y ATsat-109, respectivamente, diseñados a partir de secuencias candidatas en *A. tequilana*. Los análisis del perfil de la electroforesis, y el de secuencias para tres especies de agave arrojan que existe una variación entre ellas; además, para el cebador ATsat-102-113 se ubicaron dos subfamilias con un tamaño distinto pero presentes en las tres especies de este estudio. También se encontró la presencia de motivos cortos dentro de las secuencias obtenidas y compartidos con otras secuencias de ADN satélites depositadas en la base de datos PlantSat (Macas *et al.*, 2002). El mapeo físico de dos sondas distintas en cromosomas de tres especies diploides de *Agave* mostró una distribución diferencial a pesar de ser consideradas especies cercanas, siendo *A. angustifolia* var. *Marginata* la que mostró una gran cantidad de señales en sus cromosomas.

La información que proporcionan las secuencias repetidas que son compartidas puede establecer relaciones de parentesco y con las subfamilias que divergen es probable inferir los procesos evolutivos que han sufrido las especies diploides y poliploides del género *Agave*.

CAPÍTULO I. ANTECEDENTES

1.1. Definición y características del ADN satélite

El ADN satélite es una secuencia repetida no codificante organizada en tándem que se ha definido como un componente especie o género específico de genomas eucariotas (Hemleben *et al.*, 2007). Esta secuencia repetida se caracteriza por contener un tamaño de motivo mayor a 100 pb (Macas *et al.*, 2002), contiene motivos cortos repetidos dentro de su secuencia, un alto porcentaje de Adeninas/Timinas (AT%), está presente en la heterocromatina y se encuentra localizado en las regiones subteloméricas o centroméricas de los cromosomas (Heslop-Harrison, Brandes y Schwarzacher, 2003). La longitud completa del arreglo en tándem del ADN satélite puede medir de 1Kb hasta 100 Mb y puede ocupar un 20% del ADN nuclear en plantas (Plohl *et al.*, 2008).

El ADN satélite se encuentra en constante cambio, motivo por el cual ha sido empleado como marcador molecular para estudios en la evolución de cariotipos de especies cercanas (Lim, *et al.* 2000; Pires *et al.* 2004; Zhang *et al.* 2015). La variabilidad genética del ADN satélite puede originar la formación de diversas familias de estas secuencias, las cuales pueden ser específicas de una especie y estar presentes en las especies de un mismo género o incluso familia (Quesada del Bosque *et al.*, 2014); este modelo es la Teoría de la librería de ADN satélites (Fry y Salsler, 1977) y trata de explicar la ocurrencia de los perfiles de ADN satélites en especies cercanamente relacionadas como una consecuencia del cambio en número de copias dentro de la librería (genoma).

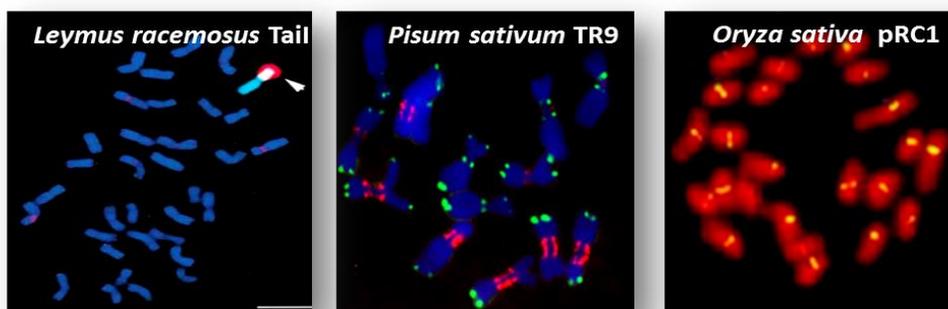


Figura 1.1. Hibridación *in situ* fluorescente (FISH) de Tail en *Leymus racemosus*, TR9 de *Pisum sativum* y pRC1 de *Oryza sativa*. La localización de estos ADN satélites es distinta, así como puede encontrarse en una sola posición de un solo cromosoma (cabeza de flecha) puede estar presente en la región subtelomérica (en color verde) o centromérica (en color amarillo) de los cromosomas. Fotos de Dong *et al.*, 1998, Kubis *et al.* 1998 y Kishii, *et al.*, 1999.

La presencia de más de un ADN satélite es común en el genomas de plantas (Macas *et al.*, 2002). Por ejemplo, en el género *Musa* se identificaron los satélites maTR_CL18 y maTR_CL33 en 9 accesiones y mostraron secuencias conservadas con una alta homología entre las especies de este género (Cizkova *et al.*, 2013). También el género *Phaseolus* contiene el satélite *Khipu* (Richard *et al.*, 2013), y recientemente identificaron que este último tiene adyacente a otro satélite denominado *Jumper* que solo está presente en *P. vulgaris* pero de diferente tamaño, distribución y probablemente diferente mecanismo de amplificación (Ribeiro *et al.*, 2017).

A pesar de que al ADN satélite se le considera una región no codificante en algunos organismos incluidos vertebrados, invertebrados y plantas se ha reportado que el ADN satélite es temporalmente transcrito en una etapa del desarrollo en algunos tipos de célula, tejido u órgano (Ugarkovic, 2005). En arroz (*Oryza sativa*) el ADN satélite centrómero, CentO, fue evaluado para observar su transcripción a través del procesamiento de siRNA sugiriendo que este satélite puede estar participando en la modificación epigenética de la cromatina en esta especie (Lee *et al.*, 2006).

Se ha observado que los transcritos del ADN satélite pueden generarse por choque térmico en humanos, salamandras y grillos, lo que sugiere que los transcritos de ADN satélite pueden ser precursores de ARN que tienen la capacidad de adoptar estructuras secundarias tipo cabeza de martillo y funcionar como ribozimas (Trofimova *et al.*, 2015; Ugarkovic, 2005).

1.2. Mecanismos de origen y evolución del AND satélite

El origen del ADN satélite en plantas es desconocido, sin embargo se han propuesto algunos mecanismos. Estos incluyen al cruzamiento desigual, eventos de conversión de genes, desplazamiento de replicación y la transposición (Garrido-Ramos, 2015; Plohl *et al.* 2008).

El cruzamiento desigual resulta del apareamiento entre repeticiones no equivalentes en regiones de ADN formadas por unidades de repetición, de modo que produce una deleción en un cromosoma recombinante y una duplicación correspondiente en el otro (Lewin, 2008).

Los eventos de conversión de genes se dan a través del intercambio de cadenas para formar genes; en este mecanismo las secuencias de los genes se comparan directamente

y son homogeneizadas por las enzimas que detectan las bases apareadas inapropiadamente (Lewin 2008).

El desplazamiento en la replicación sucede cuando la cadena hija se desliza hacia atrás una unidad de repetición al aparearse con la cadena molde. Cada evento de deslizamiento agrega una unidad de repetición a la cadena hija y la duplicación de esta cadena hija en el siguiente ciclo genera un ADN dúplex con un número mayor de repeticiones (Lewin, 2008).

De igual manera, se ha reportado que los retroelementos, con su maquinaria de copiado e inserción en el genoma, pueden enriquecer las secuencias de estas regiones en tándem (Tek, Song, Macas, & Jiang, 2005). Por último, ha sido reportado que el ADN satélite puede derivarse de la región intergénica (IGS) del ADN ribosomal de especies pertenecientes a la familia Solanaceae (Jo *et al.* 2011).

1.3. Técnicas para el aislamiento de ADN satélite

El ADN satélite fue descubierto en plantas a través de la ultracentrifugación de gradiente de densidades por CsCl, en el cual se observó, junto a una cantidad mayor de ADN, la presencia de bandas adicionales (Hemleben, 1990). El arreglo en tándem del ADN satélite ha permitido su aislamiento en diferentes especies de plantas y entre las técnicas para dicho aislamiento se incluyen la restricción enzimática, la auto-PCR, la amplificación con cebadores específicos y búsqueda en bibliotecas genómicas (Buntjer y Lenstra, 1998; Grebenstein *et al.* 1996; Nouzová *et al.* 1999), entre otras. En la actualidad existe una base de datos para los ADN satélites de plantas denominada PlantSat que facilita la identificación y caracterización de los motivos repetidos de estas regiones (Macas, Meszaros y Nouzová, 2002) (Tabla 1.1).

El aislamiento del ADN satélite ha permitido conocer que son parte importante del centrómero, regiones subteloméricas y que están involucrados en la formación y mantenimiento de la heterocromatina (Čížková *et al.*, 2013; Matyasek, Fulnecek, Leitch, & Kovarik, 2011). El descubrimiento de estas secuencias repetidas en especies relacionadas permite conocer su distribución, su variabilidad y su relación con otras regiones repetidas como los retroelementos (Biscotti, Olmo y Heslop-Harrison, 2015; Clarkson *et al.* 2005; Yoong Lim *et al.* 2006).

Tabla 1.1. Ejemplos de ADN satélites en plantas reportados en la base de datos del PlantSat (Macas *et al.*, 2002), en la cual existe más de un satélite por especie.

Especie	Tipo satélite	Tamaño (pb)	Referencias
<i>Arabidopsis thaliana</i>	Arabidopsis thaliana_500	500	Simoens <i>et al.</i> , 1988
<i>Brassica oleracea</i>	340	342	Harrison y Heslop-Harrison, 1995
<i>Cucumis sativus</i>	180	180	Hemleben <i>et al.</i> , 1986
<i>Erianthus spp.</i>	EACIR1	367	Simoens <i>et al.</i> , 1988.
	Eracsi228	352	
	Eracsi 294	185	
<i>Oryza sativa</i>	TrsA	354	Ohtsub y Ohtsub, 1994
	TrsB	159	
<i>Pisum sativum</i>	PisTR A	212	Neuman <i>et al.</i> , 2001
	PisTR B	50	
<i>Secale cereale</i>	D1100	1047	Houben <i>et al.</i> , 1996
	E3900	3894	
<i>Nicotiana tabacum</i>	NTS9	90	Matzke <i>et al.</i> , 1998.
<i>Solanum brevidens</i>	A	155	Preisner <i>et al.</i> , 1994
	B	228	
	C	33	
<i>Zea mays</i>	CentC	155	Ananiev <i>et al.</i> , 1998
	TR1	330	

1.4. El ADN satélite en el estudio de especies filogenéticamente relacionadas

Las secuencias de ADN satélite se encuentran en la heterocromatina y entre las funciones que se les ha atribuido, de gran importancia el estudio de esta secuencia, son que forman parte de la estructura del centrómero de los cromosomas y amortiguan las mutaciones de las secuencias (Heslop-Harrison, Brandes y Schwarzacher, 2003). De igual manera, se ha observado que pueden participar como sitio para el anclaje de proteínas y en la formación de ARN pequeño de interferente ARNip (regulación epigenética) (Ugarkovic, 2005). Las regiones de ADN satélite también se han utilizado para el estudio de origen y evolución de poliploides (Clarkson *et al.* 2004; Lim *et al.* 2006).

Debido a la rápida evolución de estas regiones, a consecuencia de las mutaciones acumuladas a lo largo de la evolución, puede originar diferentes polimorfismos de utilidad

para el análisis filogenético de diversidad y especiación (Tek et al., 2005; Wolfgruber et al., 2009). El alto número de copias y elevado nivel de polimorfismo reportado en las regiones de ADN satélite ha sido utilizado para el desarrollo de marcadores específicos de cromosomas en diferentes especies de plantas (Sharma y Raina, 2005).

En rábano (*Raphanus sativus* L.) se determinó que los ADN satélites son las regiones repetidas más abundantes dentro del genoma de esta especie y se encontraron tres ADN satélites nuevos; CL1 presente en el centrómero, CL25 en la región subtelomérica y CL43 en el telómero, siendo el CL1 y CL25 nuevos marcadores para los cromosomas (He et al., 2015). En *Phaseolus microcarpus* Mart., se identificó un nuevo ADN satélite denominado “*jumpe*” y se encuentra adyacente a otro “*kipu*”; se determinó que el primero puede ser originado por eventos de retrotransposición, debido a la gran cantidad de elementos repetidos que contiene el genoma (54%) y además resultó ser específico para esta especie (Ribeiro et al., 2017).

Tabla 1.2. El ADN satélite y las aplicaciones que se le han conferido en plantas.

Taxón	Aplicación	Información proporcionada	Bibliografía
<i>Aegilops speltoides</i>	Marcador cromosomal	Spelt1 es específico de la especie <i>Aegilops speltoides</i> y se posiciona en el centrómero del cromosoma 4.	Salina <i>et al.</i> , 1998 Shaima y Raima, 2005
<i>Leymus racemus</i>	Marcador cromosomal	Se identificó a Tai1 en la región subtelomérica del cromosoma 3.	(Kishii <i>et al.</i> , 1999)
<i>Pisum sativum</i>	Análisis de filogenia	Psat27 fue específico para la especie <i>P. sativum</i> cuando fue comparado con otras especies de leguminosas.	(Neumann, Nouzová, & Macas, 2001)
<i>Nicotiana</i>	Análisis de filogenia	Subfamilias de ADN satélites pertenecientes a HRS60 han logrado identificar nuevas secuencias en diferentes secciones del género.	(S-h Jo et al., 2011; Lim et al., 2004; Matyasek et al., 2011)
Solanaceae	Análisis de filogenia	Identificación de secuencias satélite específicos para 3 especies diferentes del género.	(S-h Jo et al., 2011)
<i>Solanum tuberosum</i>	Marcador cromosomal	St24 fue identificado en los cromosomas 2 de esta especie y es específico del centrómero.	(Gong et al., 2012)

Los estudios de localización y mapeo de los ADN satélite para estudios citogenéticos se han logrado a través de técnicas de hibridación *in situ*; con esta metodología se puede visualizar la posición de una secuencia del ácido nucleico (ADN o ARN) en los cromosomas, núcleos y tejidos (Ohmido *et al.*, 2010). La hibridación se realiza entre el ADN cromosómico y secuencias de ácidos nucleicos marcados, conocida como sonda (Herrera, 2007).

Así mismo, para especies del género *Arabidopsis*, el estudio de ADN satélites en el centrómero evidenció que la ausencia y presencia de estas secuencias en los cromosomas pueden explicar la divergencia, y la conservación de motivos (pequeñas secuencias) que permiten establecer la relación entre especies de un grupo taxonómico cercano (Heslop-Harrison, Brandes y Schwarzacher, 2003).

En 2011, Zhao *et al.*, llevaron a cabo un estudio comparativo en los cromosomas de tres variedades de la especie *Cucumis sativus*: *C. sativus* var. *xishuangbannensis*, *C. sativus* var. *hardwickii* y *C. sativus* var. *sativus*; los autores notaron que a partir de la organización y distribución de ADN repetido en los cromosomas se pueden inferir la relación ancestral que presentan las variedades, encontrando que *C. sativus* var. *sativus* es más cercana a *C. sativus* var. *xishuangbannensis* que a *C. sativus* var. *hardwickii* (Zhao *et al.*, 2011). Un estudio más reciente en especies cultivadas y silvestres del género *Cucumis*, reveló la divergencia de la estructura cromosómica a través de secuencias repetidas (Zhang *et al.* 2015).

El empleo de regiones de ADN satélite para inferir relaciones evolutivas entre especies, o variedades relacionadas, es una herramienta que puede revelar los patrones evolutivos del ADN repetido como consecuencia de la poliploidía en el genoma de las especies de *Agave*. Así mismo, el estudio de regiones de ADN satélite puede contribuir a explicar el papel que desempeña esta región en especies que han sufrido un incremento del tamaño de su genoma, y en un futuro pueden ser empleadas como marcadores cromosómicos para esclarecer relaciones filogenéticas entre especies cercanas o clasificadas en diferentes taxón.

1.5. La familia Agavaceae

Las especies de agaves pertenecen a la familia Agavaceae que se distribuye principalmente en América en ambientes áridos y semiáridos. Esta familia consta de 9

géneros y exhibe una estructura básica similar de acuerdo a su roseta, flores e inflorescencias (Good-Avila, Souza, Gaut, & Eguiarte, 2006).

Agave es el género más grande de la familia Agavaceae, con aproximadamente 208 especies (Good-Avila et al., 2006). En nuestro país se han reportado 186 taxones, de los cuales 129 (69%) son endémicas. El género alcanza su máximo nivel de riqueza y diversidad de especies en México, por lo cual es considerado centro de origen y diversificación del grupo (García-Mendoza, 2007). Es posible que la riqueza de especies se deba a los hábitats heterogéneos que se presentan en nuestro país. Los agaves son abundantes en las provincias de las serranías meridionales del centro de México, Sierra Madre Occidental, Altiplano mexicano, península de Baja California y Sierra Madre Oriental (García-Mendoza, 2007).



Figura 1.2. Campo de cultivo de *A. tequilana* var. Azul, Valladolid Yucatán. Fotografía Laura Espinosa Barrera.

1.6. Características fisiológicas y morfológicas de *Agave*

Los agaves son plantas xerófitas que presentan modificaciones o especializaciones morfológicas como una estrategia para sobrevivir en ambientes desérticos. Son plantas perennes, con hojas dispuestas en espiral y arregladas en rosetas en el ápice de un tallo,

el cual puede ser corto o largo y erecto. Las hojas por lo general son suculentas, fibrosas y carnosas, su forma varía de lineal a lanceolada u ovalada, y los márgenes exhiben una gran diversidad morfológica pero casi siempre tienen una espina al final del ápice (Gentry, 1982; Eguiarte, Souza y Silva-Montellano, 2000).

Las especies del género *Agave* presentan un sistema de reproducción semélparos o monocápicos. Aquellas plantas que son semélparos mueren después de reproducirse. Presentan ciclos de vida largos que van desde los 5 hasta 20 años. Tienen reproducción asexual por medio de estolones o bulbillos, los cuales se pueden formar en la base de la planta, así como a lo largo del escapo floral, facilitando el rápido aumento en el tamaño de sus poblaciones (Gentry, 1982; Eguiarte, Souza y Silva-Montellano, 2000; Garcia-Mendoza, 2007). Su escapo floral puede llegar a medir hasta 15 veces el tamaño de la planta, por lo que son visibles para una gran diversidad de polinizadores, como los murciélagos, polillas, abejas, colibríes, entre otros, lo que permite mayor dispersión del polen para la reproducción sexual y semillas (Schaffer y Schaffer, 1979; Gentry, 1982).

1.7. Cariotipo de *Agave*

En los agaves, el número de cromosoma básico (x) y el valor haploide (n) consta de 30 cromosomas. En *Agave* el nivel de ploidía reportado varía desde diploides ($2x$) hasta octaploides ($8x$) (Banerjee y Sharma, 1987). Los procesos de hibridación, poliploidía y reproducción vegetativa son considerados como una estrategia evolutiva importante en este género. Estos procesos implican cambios rápidos del genoma debido a la hibridación y al origen de la poliploidía (autopoliploidía y alopoliploidía) (Soltis y Soltis, 2012). Los agaves contienen un cariotipo bimodal que consta de 5 cromosomas largos y 25 pequeños (Castorena-Sánchez *et al.*, 1991).

De las 208 especies que pertenecen al género *Agave*, en solamente 27 de estas se han analizado el tamaño del genoma, y la mayoría son diploides con una cantidad $2C$ de ADN que va desde 6pg en *A. fourcroydes* hasta 9.6 pg en *A. desertii* (Bennett y Leitch, 2005; Palomino *et al.* 2008). En el 2008, Robert *et al.*, determinaron la cantidad de ADN en el genoma nuclear de *Agave*, con un promedio monoploide del tamaño genómico de $1Cx$ es de ~7.5pg (Robert *et al.*, 2008).

Se han realizado trabajos de hibridación *in situ* fluorescente en especies de *Agave*, entre estos destaca el de Robert *et al.*, (2008) en el cual se determinó la ubicación de los genes de ADN ribosomal 5S y 35S y el incremento proporcional de señales de estos genes con

respecto al nivel de ploidía. Posteriormente, Gómez-Rodríguez *et al.*, (2013) determinaron que el gen ribosomal 18S se encuentra en cromosomas grandes acrocéntricos y el gen ribosomal 5S en cromosomas pequeños metacéntricos o submetacéntricos de la especie *A. tequilana* (Gomez-Rodriguez, Rodriguez-Garay, Palomino, Martínez, & Barba-Gonzalez, 2013).

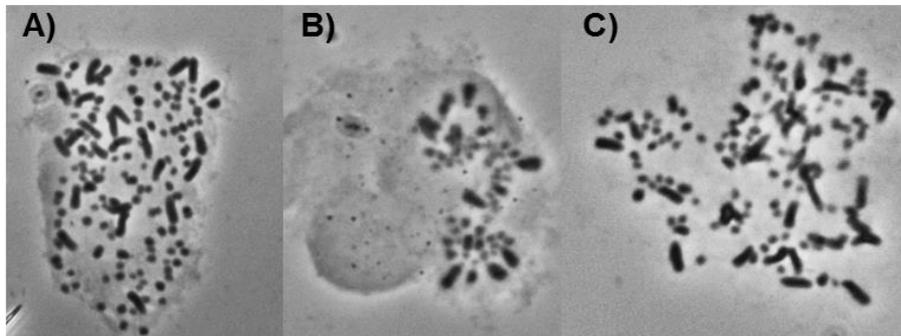


Figura 1.3. Células en metafase después del proceso de “squash”. A, C) Cromosomas de la especie *A. fourcroydes* var. Yaax ki, esta especie es un pentaploide ($2x=5n=150$). B) Cromosomas de *A. tequilana*, esta especie es diploide ($2x=2n=60$). Fotografías Laura Espinosa Barrera.

1.8. La sección *Rigidae* del género *Agave*

El género *Agave* está constituido por más de 208 especies (Good-Avila et al., 2006), que de acuerdo a la morfología de la inflorescencia Gentry (1982) dividió en dos subgéneros: *Littaea* que es polifilético y *Agave* que es monofilético. A su vez, estos subgéneros se encuentran divididos en secciones; para el subgénero *Littaea* tiene ocho secciones con 54 especies y el subgénero *Agave* cuenta con doce secciones con 136 especies en México (Tabla 1.3) (Gentry, 1982).

En la sección *Rigidae* del género *Agave* pertenecen distintas especies que han sido empleadas debido a su alto contenido en fibra y azúcares, es un grupo mayormente tropical y en general la producción de inflorescencias se beneficia de tormentas en el verano y el otoño (Gentry, 1982; Narváez-Zapata y Sánchez-Teyer, 2009). Según Gentry (1982), este grupo se caracteriza por su morfología foliar: hojas angostas y rígidas, dispuestas en forma de espiral, y por su morfología floral: sus flores son verdosas a amarillentas con tépalos pequeños que se marchitan pronto y que están dispuestas en panículos pequeños abiertos y con brácteas pequeñas, producen cápsulas ovoides y

anchas con semillas grandes. Algunas de las especies son fáciles de reconocer, mientras que otras presentan una gran variedad de formas tales como el complejo *Agave angustifolia* Haw. Otros grupos difíciles de caracterizar a nivel de especie son los cultivados (para fibra y licores), la mayoría de los cuales son clones y algunos son híbridos estériles como *Agave fourcroydes*. Las 12 especies que pertenecen a esta sección son: *A. aktites* Gentry, *A. breedlovei* Gentry, *A. decipiens* Baker, *A. karwinskii* Zucc., *A. panamana* Trelease, *A. stringens* Trelease, *A. angustifolia* Gentry, *A. datylio* Zucc., *A. fourcroydes* Lem., *A. macroacantha* Zucc., *A. rhodacantha* Trelease y *A. tequilana* Weber.

Otro problema que enfrenta la taxonomía de este grupo es que algunas de las especies han sido clasificadas con base en el uso que se les da y no por el conocimiento de su morfología y biología (Gentry, 1982), por lo que es necesario el desarrollo de marcadores genéticos que permitan una clasificación molecular de las especies que comprenden esta sección. En este estudio se propone estudiar al ADN satélite para evaluar las relaciones filogenéticas de la sección *Rigidae* del género *Agave*.

Tabla 1. 3.Secciones de los subgéneros *Litteae* y *Agave*. El subgénero *Litteae* consta de 8 secciones y *Agave* consta de 12.

Subgénero <i>Litteae</i>	Subgénero <i>Agave</i>
<i>Amolae</i>	<i>Americanae</i>
<i>Choritepalae</i>	<i>Campaniflorae</i>
<i>Filiferae</i>	<i>Crenatae</i>
<i>Marginatae</i>	<i>Deserticolae</i>
<i>Parvoroflorae</i>	<i>Ditepalae</i>
<i>Polycephalea</i>	<i>Hiemiflorae</i>
<i>Striatae</i>	<i>Marmoratae</i>
<i>Urceolatae</i>	<i>Parryanae</i>
	<i>Rigidae</i>
	<i>Salminae</i>
	<i>Sisalanae</i>
	<i>Umbelliflorae</i>

1.9. Importancia del género *Agave*

El género *Agave* es considerado de importancia ecológica y económica ya que ha sido empleado para la fabricación de ropa, sogas, alimentos, entre otros productos.

Actualmente, en México varias especies de este género son aprovechadas como materia prima en procesos tradicionales y agroindustriales (Tabla 1.4). Por ejemplo, el subgénero *Agave*, en particular las secciones *Rigidae* y *Sisalanae*, fueron seleccionadas para cultivarlas, ya que a partir de ellas se producía casi todo el suministro de fibra natural, la cual se extraía del henequén y sisal. Esta fibra es empleada para la fabricación de cordeles, tejidos, sacos, redes, muebles, sandalias, canastos, entre otros. En Yucatán, el cultivo de *Agave fourcroydes* (henequén) fue una de las industrias mejor desarrolladas.

A partir de *A. fourcroydes*, se obtienen productos naturales bioactivos de alto valor agregado como sapogeninas esteroidales con propiedades antiinflamatorias, antiparasitarias o hemolíticas; aplicables en medicina tradicional (Castorena-Sánchez *et al.*, 1991; Narváez-Zapata y Sánchez-Teyer, 2009).

En la industria de bebidas, el *Agave* es utilizado para la producción de mezcal y el tequila. Con la especie *Agave tequilana*-variedad azul (NOM-006-SFCI-1994) se produce el tequila, un producto de exportación con denominación de origen y de gran importancia económica para México, además de los fermentos y destilados como el pulque y el mezcal que son actividades con relativa importancia económica para distintas regiones del país (Robert *et al.*, 2008). Así mismo, muchas de las especies del género *Agave* también son utilizadas como plantas de ornato (Eastmond y Robert, 2000).

Tabla 1. 4. Aplicaciones y usos de las especies del género *Agave*.

Especie	Aplicación	Bibliografía
<i>Agave tequilana</i>	Bebidas alcohólicas	Valenzuela <i>et al.</i> , 1997; Valenzuela <i>et al.</i> , 2003
<i>Agave sisalana</i>	Fibra natural	Banerje y Sharma, 1988
<i>Agave fourcroydes</i>	Fibra natural	Boulanger <i>et al.</i> , 1985
<i>Agave angustifolia</i>	Planta ornamental	Banerje y Sharma, 1988; Eastmond y Robert, 2000.
<i>Agave angustifolia</i>	Bebidas alcohólicas	Moreno-Salazar <i>et al.</i> , 2007

1.10. ADN satélite en *Agave tequilana*.

Hasta el momento no hay información publicada sobre la presencia de regiones de ADN satélite en el genoma de especies del género *Agave*. El único estudio es el realizado por, Espinosa-Barrera *et al.*, 2014 (tesis de maestría) cuyo objetivo fue el aislamiento y

localización del ADN satélite en *A. tequilana* (Tabla 1.5). Para identificar el ADN satélite se aislaron fragmentos por auto-PCR, obteniendo seis secuencias.

Por el alineamiento múltiple de secuencias se encontró que son diferentes en tamaño y composición. Así mismo, la localización de una secuencia en particular (AteqSat 105) se encontró en núcleos de *A. tequilana* por medio de hibridación *in situ* fluorescente y se observaron múltiples señales de esta región en el genoma de *A. tequilana* lo que sugiere que en esta especie diploide pueden estar presentes subfamilias de ADN satélite. También, se encontró que comparten motivos con otras secuencias de ADN satélite de especies pertenecientes a la clase Lillioopsida.

Tabla 1. 5. Secuencias candidatas de ADN satélite en *Agave tequilana*.

Secuencia	Tamaño (pb)	%AT	Enzima de restricción
AteqSat 101	404	57	<i>HindIII, BsiI,</i>
AteqSat 102	320	54	<i>HindIII, BsiI, NcoI</i>
AteqSat 105	271	45	<i>SalI, XbaI, TaqI, MseI</i>
AteqSat 109	214	49	<i>BsiI, SmaI, XmaI</i>
AteqSat 110	213	50	<i>Eco57I, ScaI</i>
AteqSat 113	188	57	<i>XbaI, AsuII BclI</i>

Debido a que hasta el momento es desconocida la relación filogenética de especies del género *Agave* se propone el estudio del ADN satélite para establecer los patrones evolutivos y el empleo de estas regiones como un marcador cromosómico y filogenético que a futuro permita optimizar la clasificación taxonómica de especies de la sección *Rigidae*.

JUSTIFICACIÓN

El ADN satélite es una región repetida en cientos de copias de ADN no-codificante, que es el principal constituyente estructural de la heterocromatina; su evolución es importante para la organización del genoma en plantas (Plohl et al., 2008). Debido a sus características, y a la variabilidad en la secuencia de estas regiones ha sido empleada en especies de interés agronómico para establecer las relaciones de parentesco con otras especies (Hemleben, Kovarik, Torres-Ruiz, Volkov, & Beridze, 2007; Neumann, Koblížková, Navrátilová, & Macas, 2006). Uno de los usos que se le ha conferido al ADN satélite es como marcador molecular específico en el genoma lo que ha permitido la caracterización de especies híbridas, el estudio taxonómico y evolutivo de plantas de interés comercial como banana, trigo, arroz y patata (Ananiev et al.1998; Čížková et al. 2013; Jo et al. 2011; Salina et al. 1998).

Uno de los principales modelos para el estudio de regiones repetidas en plantas es el género *Nicotiana*, en el que se han encontrado varias familias de ADN satélite en diversas secciones del género, estos reportes han sido complementados con la hibridación *in situ* fluorescente de estas secuencias en los cromosomas de estas especies para un análisis de los cariotipos. Por ejemplo, en la sección *Alatae* del género *Nicotiana* se ha descrito que las regiones repetidas pueden determinar que las especies de esta sección se encuentran en un proceso lento de homogenización (Lim et al. 2006).

En la familia Solanaceae se ha realizado un estudio del ADN satélite, particularmente localizado en la región intergénica (IGS) del ADN ribosomal, para conocer las relaciones entre 3 géneros de esta familia y se encontró que son específicos para cada uno del género que fue evaluado; en este análisis se concluyó que los ADN satélites se han fijado junto con el ADN ribosomal en las especies de los géneros de la familia Solanacea y que pueden ser empleados para determinar las relaciones ancestrales entre dichos géneros (S-h Jo et al., 2011).

Dadas las evidencias que se tienen sobre el estudio de regiones de ADN satélite en plantas, se sugiere que pueden ser empleadas para inferir la relación filogenética y ciertos patrones evolutivos en el genoma poliploide que permitan contar con un marcador cromosomal y molecular para apoyar la clasificación taxonómica entre las especies pertenecientes a la sección *Rigidae* del género *Agave*.

HIPÓTESIS

La caracterización de las secuencias de ADN satélite en algunas especies de la sección *Rigidae* del género *Agave* permitirá observar una diferencia en la secuencia de esta región entre dichas especies.

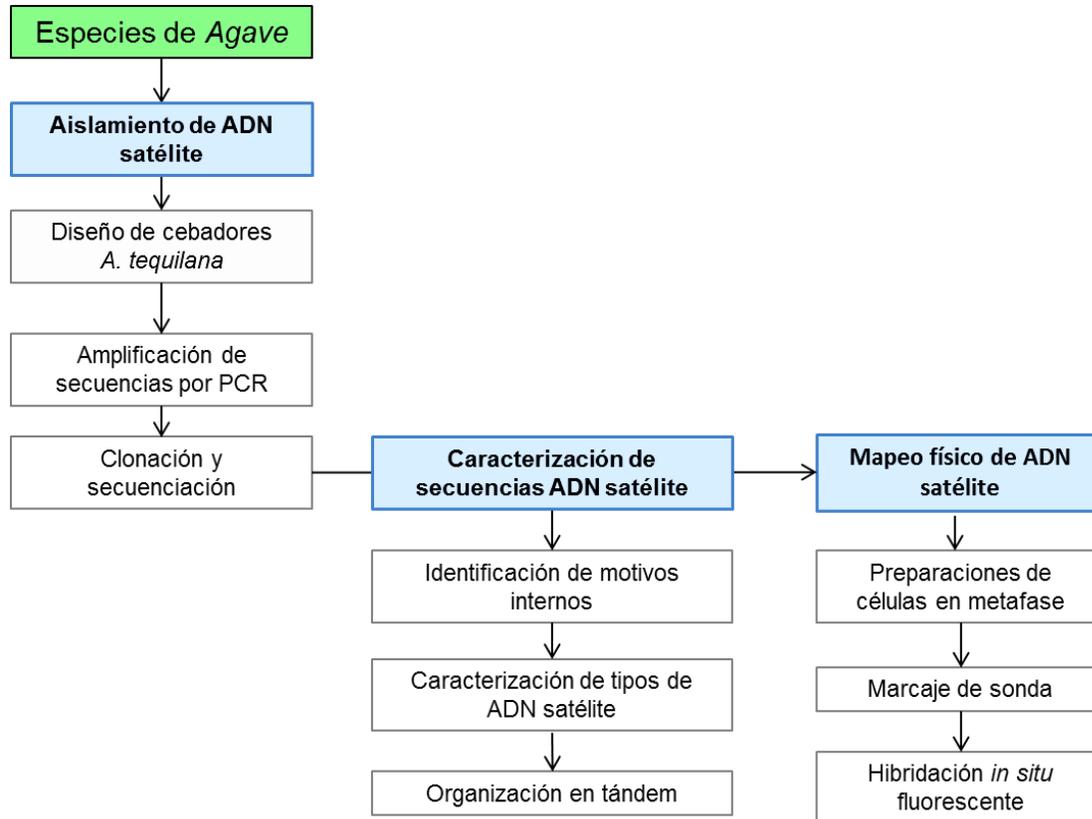
OBJETIVO GENERAL

Determinar la variabilidad y localización de las secuencias de ADN satélite por medio de secuenciación e hibridación *in situ* para un análisis filogenético en tres especies de la sect. *Rigidae* del género *Agave*.

OBJETIVOS ESPECÍFICOS

1. Aislar las regiones ADN satélite en tres de las especies de la sección *Rigidae* del género *Agave*.
2. Caracterizar las secuencias de ADN satélite de tres especies de la sección *Rigidae* del género *Agave*.
3. Realizar el mapeo físico del ADN satélite en tres especies de la sección *Rigidae* por hibridación *in situ* fluorescente (FISH).

ESTRATEGIA EXPERIMENTAL



CAPÍTULO II AISLAMIENTO Y CARACTERIZACIÓN DE REGIONES ADN SATÉLITE EN *A. angustifolia*, *A. fourcroydes* Y *A. tequilana*

2.1. INTRODUCCIÓN

Las regiones satélites se encuentran altamente representadas en el genoma de eucariotas y se ha descrito que estas regiones están en constante cambio lo que permite, durante la evolución, que ciertas secuencias se fijen y otras se eliminen del genoma (Garrido-Ramos, 2015; Plohl *et al.*, 2008). Estas secuencias se amplifican rápido y producen repeticiones específicas a un género, especie o incluso un cromosoma individual dando lugar a la aparición de familias o subfamilias de ADN satélite (Kishii, Nagaki y Tsujimoto, 2001).

El ADN satélite es una de las regiones más intrigantes del genoma debido a su gran abundancia, el cambio evolutivo y el impacto que pueden tener en los procesos de la especiación (Ferree y Prasad, 2012). Varios mecanismos han sido propuestos para explicar esta rápida expansión de ADN satélite, incluyendo el mecanismo de entrecruzamiento desigual y duplicación segmental (Flavell *et al.*, 1974; Hourcade *et al.*, 1973; Ma y Jackson, 2006) pero hasta el momento se desconoce el mecanismo por el cual estas regiones han sido mantenidas durante la evolución en el genoma de las especies de plantas.

Así mismo, el ADN satélite es considerado como una secuencia repetida en tándem con gran tamaño y pueden ser específicas para una especie o género en particular (Hemleben *et al.*, 2007) y también se ha reportado que existen ADN satélites preservados por períodos evolutivos largos y pueden contener motivos conservados (Meštrović *et al.*, 2015; Plohl *et al.*, 2012). Un ejemplo de esto fue el trabajo realizado en los géneros de *Helictotrichon* y *Avena*, y especies de las tribus *Andropogenoae* y *Oryzaceae* en donde fueron detectadas regiones de ADN satélite que se distribuían de manera diferente con cada grupo taxonómico (Grebenstein *et al.*, 1996). Otro estudio realizado fue el de Lim *et al.*, (2006) en especies del género *Nicotiana*, en el que se compararon las secuencias de ADN satélite de especies con 0.2, 1 y 5 millones de años; se observó que las secuencias pertenecientes a las especies de 5 millones de años, en comparación a aquellas pertenecientes a 0.2 y 1 millones de años, presentan cambios y amplificación de nuevos motivos de ADN satélite (Lim *et al.*, 2006).

Las secuencias repetidas están compuestas por diferentes contenidos de AT (adenina/timina) y CG (citosina/guanina); su alto contenido en AT se ha utilizado para la

caracterización del ADN satélite y por lo tanto, los cambios en las secuencias repetidas pueden ser especialmente útiles para contribuir a la comprensión de la evolución del genoma.

En estudios previos realizados por Espinosa-Barrera (2014, tesis de maestría) se logró la obtención de seis secuencias candidatas a partir de la auto PCR (Buntjer y Lestra, 2001) en la especie *A. tequilana*. Estas secuencias candidatas son el inicio de la búsqueda de regiones ADN satélites en especies de agave, pues a partir de éstas se realizó el diseño de cebadores y su posterior caracterización molecular. Se obtuvieron más de 100 secuencias que fueron analizadas a través de programas bioinformáticos y se observaron dos grupos de secuencias compartidas en todas las especies. Se sugiere que estos grupos pueden ser dos subfamilias de ADN satélites en agave.

2.2. MATERIALES Y MÉTODOS

2.2.1 Material biológico

En este estudio se emplearon 5 accesiones de tres especies del género *Agave*, que son *A. tequilana* (AT), *A. fourcroydes* var. Kitam ki (KK), *A. fourcroydes* var. Yaax ki (YK), *A. angustifolia* var. Chelem ki (CHE), y *A. angustifolia* var. Marginata (MAR). *Agave tequilana* ($2n=2x=60$ cromosomas), *Agave angustifolia* var. Marginata ($2n=2x=60$), *A. angustifolia* var. Chelem ki ($2n=6x=180$), *A. fourcroydes* var. Kitam ki ($2n=3x=90$) y *A. fourcroydes* var. Yaax ki ($2n=5x=150$) que se encuentran en el Jardín Botánico Regional “Roger Orellana” del Centro de Investigación Científica de Yucatán (CICY).

2.2.2 Extracción de ADN

Para la extracción de ADN se utilizó tejido de las hojas maceradas con nitrógeno líquido y el protocolo de CTAB, reportado por Doyle y Doyle (1990).

2.2.3. Diseño de cebadores

A partir de las secuencias obtenidas en el trabajo de maestría de Espinosa-Barrera (2014), se diseñaron los cebadores para regiones ADN satélite, con el programa DNAMAN versión 3.0 y OligoAnalyzer (<https://www.idtdna.com/calc/analyzer>). Las características que se consideraron para el diseño de oligonucleótidos son la temperatura de alineamiento, contenido de GC, la complementariedad, la especificidad, el tamaño de oligonucleótidos y el tamaño de fragmento amplificado.

2.2.4. Amplificación de fragmentos por PCR

Una vez diseñados los cebadores se determinaron las condiciones adecuadas para la mezcla de reacción de la PCR empleando: 100ng de ADN genómico, 100 pM de cada cebador (AtSat-102 y AtSat-109), 1.5 mM de $MgCl_2$, 0.2 mM de dNTP's, 1X buffer, y 0.05 U de DNA Taq polymerase, y se ajustó a 25 μ l con agua destilada ultrapura. Las condiciones de PCR fueron: un primer paso de desnaturalización de 5 min a 95°C, seguido de 30 ciclos de 2 min a 95°C, 1.5 min a 56 °C y 1 min a 72 °C de elongación, después un paso de elongación a 10 min at 72 °C. Los fragmentos fueron observados en un gel de agarosa (1%).

2.2.5. Clonación y secuenciación

Los fragmentos de interés fueron clonados en el vector pGEM T-easy (PROMEGA) y transformados por choque térmico en células competentes de *E. coli* DH10B. Las transformantes se seleccionan por resistencia a 50 µg/ml de ampicilina. Se realizó la ligación del fragmento en el vector pGEM T-easy como indica el fabricante. Para la extracción de plásmidos se utilizó el método de lisis alcalina de Birnboim & Doly (1979).

La verificación de la transformación se realizó mediante amplificación del inserto usando oligonucleótidos universales M13 REV (-24) 5'-AACAGCTATGACCARG y M13 FWD (-40) 5'-GTTTTCCCAGTCACGAC, que se posicionan en los extremos del sitio múltiple del vector. Los fragmentos clonados se enviaron a secuenciar a MACROGEN (<http://www.macrogen.com/>), en Seúl, Corea. Los alineamientos para editar la secuencia del plásmido pGEM-T easy Vector systems y la caracterización de la secuencia de cada fragmento clonado se realizó con el programa BioEdit (Hall, 1999).

2.2.6. Análisis de secuencia

Las secuencias de nucleótidos se alinearon y compararon con la base de datos del GenBank (<http://www.ncbi.nlm.nih.gov/>) y se utilizó el programa BioEdit Sequence Alignment Editor 5.0.6 (Hall, 1999).

Una vez obtenidas las secuencias de ADN satélite en las especies de *A. tequilana*, *A. angustifolia* y *A. fourcroydes* se llevó a cabo el análisis comparativo de las mismas, con el programa MEGA 6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

2.2.7. Hibridación Southern Blot

El Southern Blot se llevó a cabo mediante la hibridación de la sondas de la subfamilia 1 y 2, AACHE102-554 y AT102-531, respectivamente, previamente marcadas por PCR con Digoxigenina (Dig-11-dUTP, Sigma). Se usaron 50 ng de ADN como molde para la amplificación de la sonda, ATsat102-113 a 10 µM, nucleótidos 2 mM sin marcar (dATP, TTP, dCTP y dGTP en 100 mM, nucleótidos marcados 1 mM (digoxigenina-11-dUTP, MgCl₂ 50 mM, buffer de PCR 1X y 0.2 U de Taq polimerasa. El producto de PCR generado con el cebador ATsat-102-113 fue transferido a una membrana de Hybond-N+ mediante transferencia por capilaridad y la membrana se fijó con rayos ultravioleta y se guardaron a 4 °C hasta su uso. La temperatura de hibridación fue de 42 °C durante toda la noche y los lavados de astringencia fueron a 65 °C.

Para el revelado de las membranas se agregó una solución con el anticuerpo anti-digoxigenina conjugado con fosfatasa alcalina (Roche TM) (dilución de 1/10000) y se incubó por 30 min a temperatura ambiente, se lavaron dos veces por 15 min con solución de malato y 0.3% Tween 20, y finalmente con 20 ml de Tris-NaCl pH 9.5 por 5 min.

La detección quimioluminiscente se realizó con CSPD (Roche TM) siguiendo el protocolo del fabricante. Las membranas fueron expuestas a películas de rayos X por 20min y para el revelado de las películas se usó revelador y fijador Kodak ®. Por último, se lavaron las películas de rayos X con agua y se realizó el análisis de señales.

2.3. RESULTADOS

2.3.1. Diseño de cebadores para la amplificación de ADN satélite en *Agave*

Se diseñaron y sintetizaron dos pares de cebadores, a partir de secuencias candidatas a satélites obtenidas previamente en *A. tequilana*, denominados ATsat-102-113 y ATsat-109 (Tabla 2.1). De igual manera, se determinó la temperatura de alineamiento y las condiciones adecuadas para la amplificación de la región ADN satélite mediante una PCR de gradiente de temperaturas, utilizando como molde de amplificación el ADN genómico de *A. tequilana* (AT) ($2n=2x=60$), *A. fourcroydes* var. Kitam ki (KK) ($2n=3x=90$), *A. fourcroydes* var. Yaax ki (YK) ($2n=5x=150$), *A. angustifolia* var. Chelem ki (CHE) ($2n=6x=180$), y *A. angustifolia* var. Marginata (MAR) ($2n=2x=60$). La temperatura de alineamiento óptima fue de 56°C (Figura 2.1).

Tabla 2.1. Características de los cebadores sintetizados a partir de las secuencias candidatas a ADN satélite de *A. tequilana*.

Cebador	Secuencia 5' A 3'	TM (°C)	%GC
ATsat-102-113 FWD	ACTAGTTGATGAGTCCTGA	59	42
ATsat-102-113 REV	CGAGCCAGCCAGCCG	61	80
ATsat-109 FWD	AGACATGACTCGCTTCCC	59	55
ATsat-109 REV	TCAGACAAACAAGTTACCC	63	42

Los fragmentos de amplificación con los cebadores ATsat-102-113 y ATsat-109 observados en la electroforesis (Figura 2.1) para las especies *A. angustifolia*, *A. fourcroydes* y *A. tequilana*, mostraron un patrón de bandeo distinto para dichas especies; se determinaron fragmentos intensos a partir de ~200 pb para todas las especies y también se observaron fragmentos a lo largo de todo el carril con menor intensidad. En los fragmentos amplificados con el cebador ATsat-102-113 la especie *A. angustifolia* var. Chelem ki mostró mayor cantidad de estos. A pesar de las diferencias generadas entre cebadores, estas también se observaron en los tamaños de fragmentos amplificados por especie e incluso entre las variedades de las mismas especies.

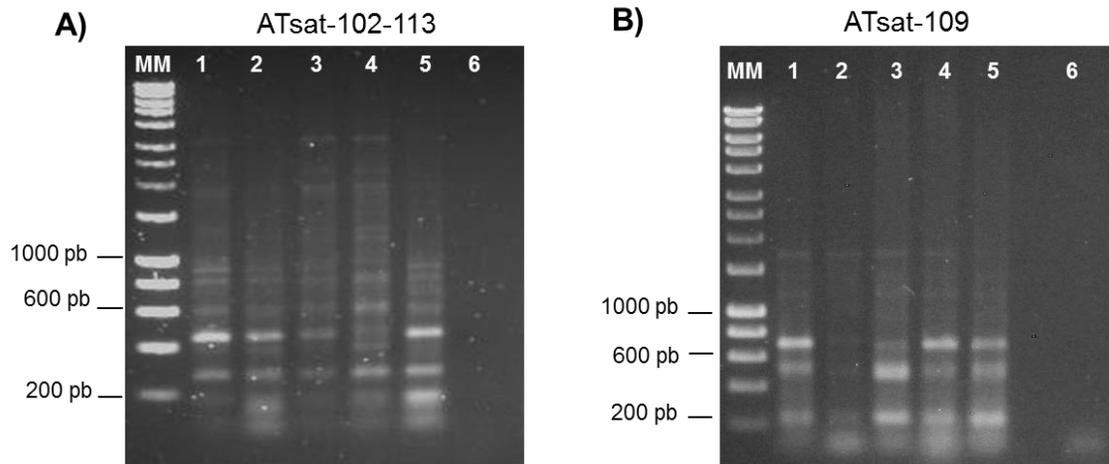


Figura 2. 1. Electroforesis de los productos de amplificación para los cebadores ATsat102-113 (A) y ATsat-109 (B) (1) *A. tequilana* ($2n=2x=60$), (2) *A. fourcroydes* var. Kitam ki ($2n=3x=90$), (3) *A. fourcroydes* var. Yaax ki ($2n=5x=150$), (4) *A. angustifolia* var. Chelem ki ($2n=6x=180$), (5) *A. angustifolia* var. Marginata ($2n=2x=60$) y (6) testigo positivo (agua) La electroforesis se realizó en un gel al 1% de agarosa y el marcador de peso molecular es de 1Kb.

El perfil de múltiples bandas y el tamaño observado en la electroforesis sugiere que estos fragmentos se encuentran en tándem característico del ADN satélite. Además de esto, el patrón de bandeo similar sugieren una posible homogenización de subfamilias de ADN satélite que se están compartiendo entre las diferentes especies de *Agave*. La intensidad particular de ciertas bandas podría estar relacionada con la abundancia de ciertas familias de ADN satélites observada en las especies de *Agave*.

2.3.2. Análisis de secuencias de ADN satélite en especies de *Agave*

Las secuencias de los fragmentos amplificados con los cebadores ATsat-102-113 y ATsat-109 para *A. angustifolia*, *A. fourcroydes* y *A. tequilana* fueron analizadas en Blast (NCBI) lo que permitió determinar los motivos consenso, encontrándose pequeños motivos en común con otras secuencias de ADN satélite contenida en la base de datos del PlantSat (Macas et al., 2002). Se obtuvieron más secuencias con el cebador ATsat-102-113 (63) que con el ATsat-109 (27); se determinó la similitud de las secuencias y se construyó un fenograma (Figura 2.2.)

En el fenograma resultante de secuencias ATsat-102-113 (Figura 2.2) se distinguieron cinco grupos. En el grupo I con 30 secuencias, de las cuales 7 pertenecen a la especie *A. angustifolia*, 11 de *A. angustifolia* var. Chelem ki y 6 de *A. angustifolia* var. Marginata; 8 secuencias propias de la especie de *A. fourcroydes*, 5 de *A. fourcroydes* var. Kitam ki y

las restantes de *A. fourcroydes* var. Yaax ki; para la especie *A. tequilana* se encontraron en el grupo 5 secuencias.

El grupo II consta de 18 secuencias, de las cuales 10 son de *A. angustifolia* var. Marginata y 2 de *A. angustifolia* var. Chelem ki; en tanto que para 2 corresponden a la especie *A. fourcroydes* var. Kitam ki y para *A. fourcroydes* var. Yaax ki también fueron 2. En el caso de *A. tequilana* se encontraron 2 secuencias en este grupo.

El grupo III conformado por el menor número de secuencias con un total de 15, de las cuales 3 fueron de *A. angustifolia* y estas se dividieron en 1 secuencias para *A. angustifolia* var. Marginata y 2 secuencias para *A. angustifolia* var. Chelem ki; para la especie de *A. fourcroydes* se observaron 8 secuencias en total, 5 fueron de *A. fourcroydes* var. Kitam ki y 3 fueron para *A. fourcroydes* var. yak ki; por último se encontraron 4 secuencias pertenecientes a la especie *A. tequilana*.

Los grupos I (azul) y II (morado) de ATsat-102-113 presentaron dos subgrupos con una gran cantidad de secuencias relacionadas. El 76% de las secuencias del grupo I (30 secuencias) comprende una secuencia de aproximadamente 140 pb y es compartida por las especies *A. angustifolia*, *A. fourcroydes* y *A. tequilana*. Otro caso similar al grupo I, es el grupo II, en el que se encontraron dos subgrupos que contienen secuencias de las 3 especies, anteriormente mencionadas, siendo el subgrupo con el mayor número de secuencias y comprende una secuencia de 107 pb, en este grupo las secuencias agrupadas fueron mayormente de *A. angustifolia* var. Marginata.

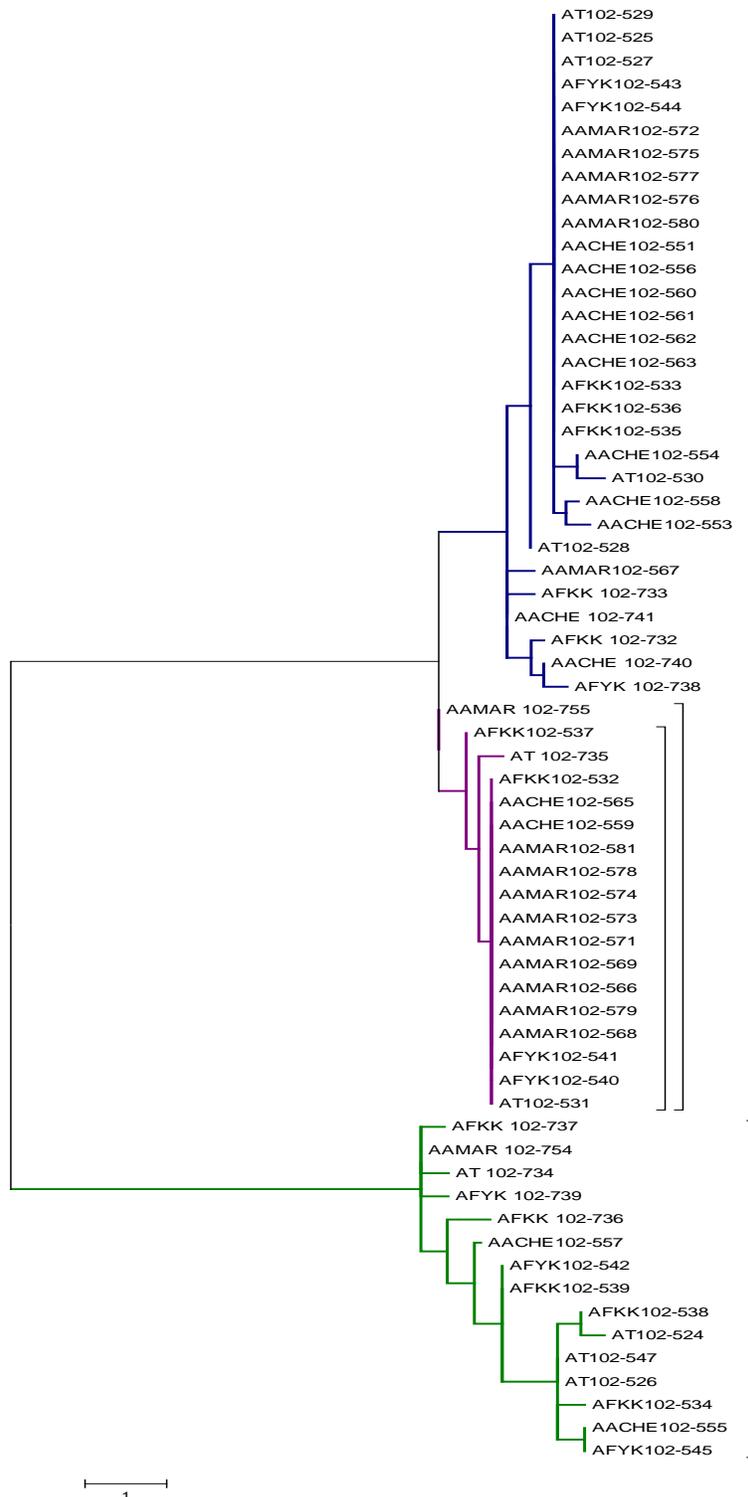


Figura 2.2. Fenograma de secuencias amplificadas con el cebador ATsat-102-113 en las especies de *A. tequilana*, *A. fourcroydes*, *A. angustifolia*. El análisis fue realizado en 63 secuencias con el Método de Máxima verosimilitud con un muestreo de 1000 réplicas. Los grupos I y II (azul y morado, respectivamente) son los que cuentan con un mayor número de secuencias.

El fenograma generado para el cebador ATsat-109 (Figura 2.3) mostró seis grupos. En el grupo I se ubicaron 12 secuencias todas pertenecientes a la especie *A. angustifolia*, en sus dos variedades: 6 de *A. angustifolia* var. Chelem ki y 6 de *A. angustifolia* var. Marginata. El grupo II está constituido de 2 secuencias idénticas de *A. tequilana*. El grupo III contiene 3 secuencias, una de *A. tequilana* y 2 de *A. fourcroydes* var. Yaax ki. Una secuencia de *A. fourcroydes* var. Yaax ki fue considerada como el grupo V. Por último, el grupo VI está constituido por 7 secuencias, 3 de *A. angustifolia* var. Marginata, 2 de *A. angustifolia* var. Chelem ki y 2 de *A. fourcroydes* var. Kitam ki.

En el caso del cebador ATsat-109, el grupo I resultó de secuencias específicas para *A. angustifolia* y se observó una mayor divergencia en la composición de las secuencias; esto tal vez podría traducirse en un mayor número de subfamilias de ADN satélite que se distribuyen y algunas de estas se comparten entre las diferentes especies de *Agave* analizadas, sugiriendo una divergencia estrecha en las secuencias de ADN satélite de las especies de la sección *Rigidae*.

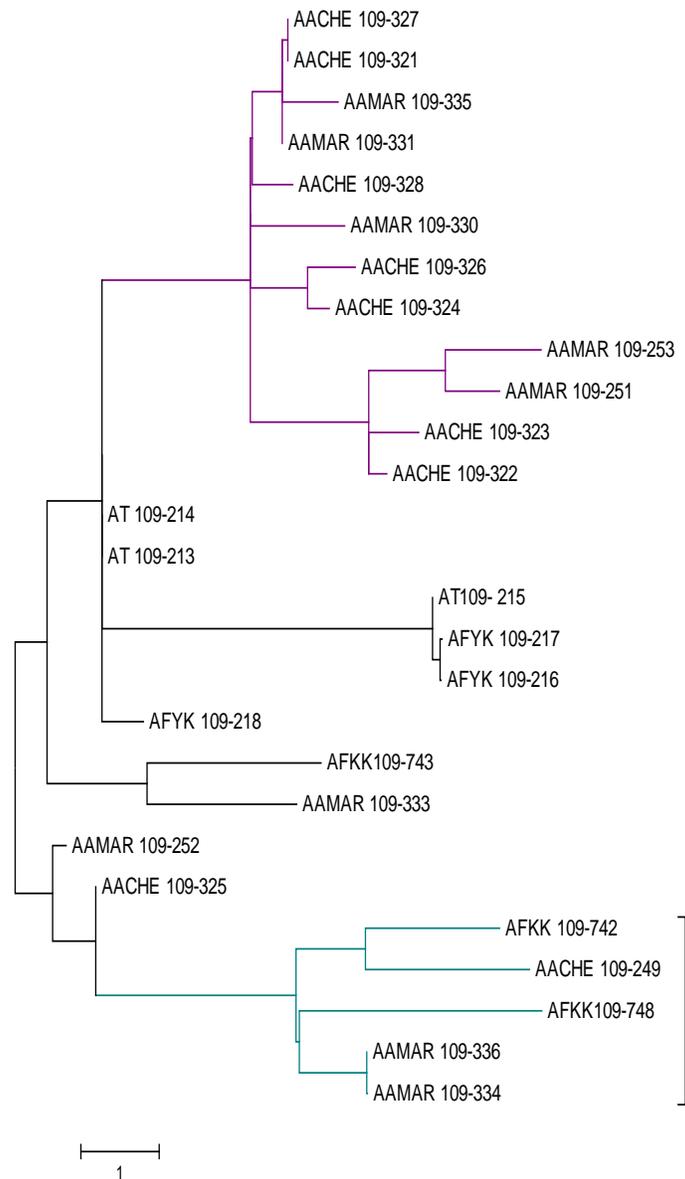


Figura 2.3. Fenograma de secuencias amplificadas con el cebador ATsat-109 en las especies *A. tequilana*, *A. fourcroydes*, *A. angustifolia*. El análisis fue realizado en secuencias con el Método Máxima verosimilitud con un muestreo de 1000 réplicas. El grupo I contiene secuencias de *A. angustifolia* (en color morado). El grupo V contiene secuencias de *A. angustifolia* y *A. fourcroydes*.

Hasta el momento no conocemos ningún reporte de ADN satélites en especie de *Agave* con las que podamos comparar nuestros resultados por lo que se decidió compararlos con las secuencias que están depositadas en el PlantSat (Macas *et al.*, 2002).

Para esto, se determinó en cada grupo de secuencias (del análisis anterior) los motivos consenso de cada secuencia en la base de datos del PlantSat, las secuencias del

PlantSat fueron descargadas y usadas para un alineamiento por grupo, es decir, las secuencias del grupo y los motivos consenso de las mismas. La información obtenida de estos alineamientos y búsqueda de motivos se presentan en los Tablas 2.2 y 2.3 y AteqSat109, respectivamente.

Este análisis permitió conocer los nucleótidos de los motivos consenso que comparten las secuencias obtenidas para *Agave* y secuencias de ADN satélite de otras especies pero no se logró observar las mismas regiones consenso entre grupos, ni entre secuencias. Los motivos de las secuencias del ADN satélite de *Agave* aislados con los cebadores AteqSat102 y AteqSat109 se han descrito para las familias Polygonaceae, Malvaceae, Poaceae, Cucurbitaceae, Solanaceae y Amaryllidaceae, siendo esta última la más cercanamente relacionada con la familia Agavaceae (Tablas 2.2 y 2.3).

	AAMAR102-571 AACHE102-565	GGTTTTG GGTTTTG			
Grupo III	AFKK102-534	ACTCCCAACTA	44-54	Rumex RAYSII (Steflova <i>et al.</i> , 2013)	Polygonaceae
	AFKK102-539 AFYK102-542 AT102-524 AT102-526 AT102-547 AFYK102-545 AFCHE102-555	GGATTTCG GATTTCG GATTTCG GACCTCG	29-35	Secale_cereale_JNK (Nagaki <i>et al.</i> , 1999)	Poaceae

Tabla 2.3. Motivos consenso generados a partir del cebador ATsat-109 que se encuentran presentes en regiones de ADN satélites reportados en el PlantSat (Macas *et al.*, 2002).

Grupo	Secuencia de Agave	Motivo consenso	Posición	ADN satélite	Familia
Grupo I	AACHE109-327 AACHE109-326 AAMAR 109-324 AACHE109-321 AACHE109-328 AAMAR 109-251 AACHE 109-323 AAMAR 109-253 AAMAR109-330	AAAGAAAAGA AAAAGAAAAA AGGATGAAGA AAAGAAAAGA AAAAGAAAAA TTGGAATGGA TTTTTATTGT ATTAAC TAACA TATAGAAAAAT	88-97	Secale_cereale E3900 Cucumis_metuliferus_350	Poaceae Cucurbitaceae
Grupo II	AT 109-214 AT 109-213	AAAGAAAAGA AAAGAAAAGA		Secale_cereale E3900	Poaceae
Grupo III	AT109-215 AFYK 109-216 AFYK 109-217	TAGTGGAATTCT TAGTGGAATTCT TAGTGGAATTCT	171-183	Rumex_RAYS1	Poaceae
Grupo IV	AFYK 109-218	ATAACATATGAA TATAACATATGAAA	175-186 478-491	Solanum_circaefolium_170	Solanaceae
Grupo V	AAMAR 109-333 AFKK109-743	AAAGTAGTGGA ACAATTTATC	114-124 128-138	Zea_mays_CentC Zea_mays_MBSc216	Poaceae
Grupo VI	AACHE109-325 AFKK109-742 AACHE109-249 AFKK109-748 AAMAR109-336 AAMAR109-334	AAATATAAGCATT TATTTTGATTA AGTAGAAATAT AATACATTGAGA CTCCCAAATTA CTCCCAAATTA	349-361	Rumex_RAYS1 Rumex_RAE180 Rumex_RAYSII Rumex_RAE730	Poaceae

2.3.3. Análisis e identificación de motivos repetidos en secuencias de ADN satélite de *Agave*

Los motivos de ADN satélite pueden estar compuestos de motivos más cortos dentro de su secuencia, esta característica es necesaria para la descripción y reporte de ADN satélites nuevos (Garrido-Ramos, 2015; Plohl *et al.*, 2008). El análisis de las secuencias obtenidas se realizó a través de la búsqueda de motivos repetidos a lo largo de todas estas y los criterios empleados fueron que pertenezcan a la misma especie, es decir, a las secuencias pertenecientes a las especies *A. angustifolia*, *A. fourcroydes* y *A. tequilana*.

Para la especie *A. angustifolia* se observaron 20 motivos repetidos y compartidos entre las secuencias de ésta. Estos motivos tienen de 5 a 7 nucleótidos y los que se encontraron en la mayoría de las secuencias son: CCCGC (30%), GGCCG (30%), CGACT (20%) y GAAAG (20%). En la Figura 2.4 se muestran los motivos repetidos y aquellas secuencias que los comparten.

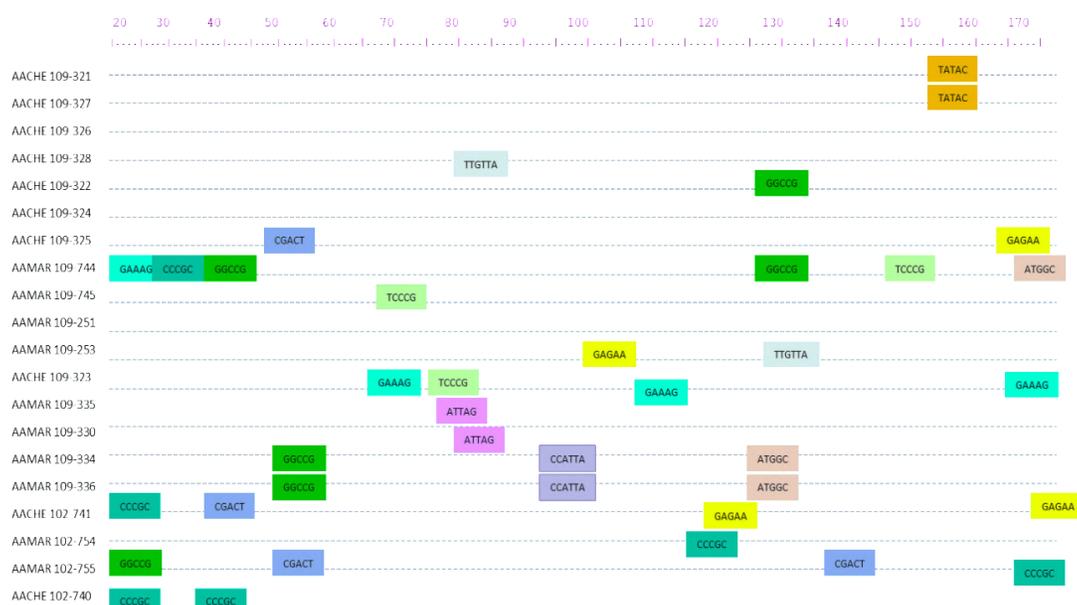


Figura 2.4. Representación de los motivos repetidos internos en la especie *A. angustifolia*. Cada color representa un motivo interno diferente, los más abundantes para esta especie son: CCCGC, GGCCG, CGACT y GAAAG.

En el caso de *A. fourcroydes* se observaron 22 motivos repetidos y compartidos entre las secuencias de esta especie. Los motivos repetidos se encontraron entre 5 y 6 nucleótidos y los que se encontraron en la mayoría de estas secuencias son: GAAAG (22.7%),

En *A. tequilana* se observaron 18 motivos repetidos y compartidos entre las secuencias de esta especie y los tamaños de estos se encontraron entre 5 y 27 nucleótidos; los que se encontraron en la mayoría de las secuencias son: GAACA (38.8%), GTATAA (27%), GCGTT (22%) y TGTTTTG (22%). En la Figura 2.6, se encuentra el resumen de los motivos repetidos y las secuencias de esta especie que los comparten.

Se encontraron motivos repetidos compartidos entre las secuencias de distintas especies; los motivos GAGAA, GGCCG, GAAAG y ATAAG están presente en las secuencias de *A. fourcroydes* y *A. angustifolia*. También se observó que el motivo TGTTTTG es compartido entre *A. angustifolia* y *A. tequilana*.

Es interesante resaltar que el análisis de motivos compartidos entre secuencias de la misma especie indicó que *A. tequilana* tiene el menor número de motivos conservados o compartidos, lo que sugiere que la región de ADN satélite de esta especie presenta una mayor divergencia con respecto a las especies *A. fourcroydes* y *A. angustifolia*, esto probablemente se deba a la selección de plantas elite en *A. tequilana* y por lo tanto podría estar sucediendo una especialización del genoma.

Los motivos compartidos en secuencias de ADN satélite son comunes en especies que se encuentran en el mismo género, esto se observó en un trabajo realizado por Navajas-Pérez *et al.*, (2006) en el que identificaron los motivos repetidos en especies *Oryza sativa*, *Vitis vinifera*, *Sorghum bicolor*, *Carica papaya* y *Arabidopsis thaliana*; estos autores identificaron que las regiones repetidas pertenecientes a estas especies presentan motivos repetidos de entre 9 y 30 pb. Esta características en las secuencias repetidas indican que la presencia de motivos repetidos en la secuencia es propio de los cambios que ocurren durante la evolución de estas secuencias, propiciando la acumulación y fijación de mutaciones de una secuencia o incluso en las familias de ADN satélites. De igual manera, algunas secuencias generadas por ambos cebadores, para todas las especies, presentaron el motivo CAAAA; este motivo ha sido descrito en la caracterización de ADN satélites y se cree que está involucrado en el plegamiento del ADN (Apple *et al.*, 1989; Macas *et al.*, 2002).

2.3.4. Obtención de secuencias consenso de ADN satélite de las especies de *Agave*

Las secuencias obtenidas en este trabajo fueron alineadas empleando el programa BioEdit (Hall, 1999). Posteriormente, con el programa MEGA 6 (Kumar *et al.*, 2013) se obtuvo un fenograma usando el método de Máxima verisimilitud (Figura 2.7). En este

análisis se observaron dos grupos (en color verde y rosa, respectivamente) y una dispersión generalizada de otras secuencias representadas en el fenograma.

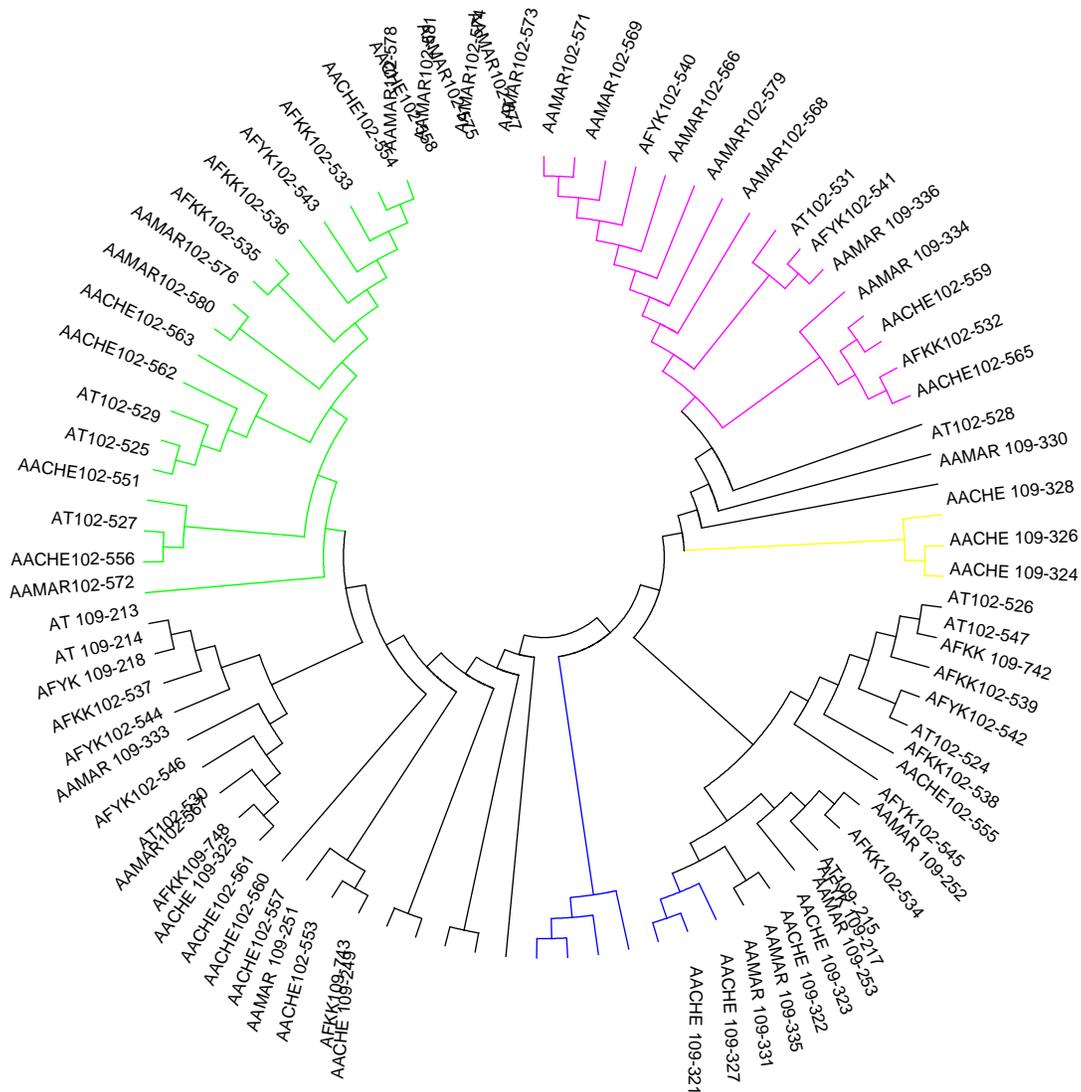


Figura 2.7. Distancias genéticas entre las secuencias ADN satélite de las especies de *Agave* de este estudio. Las distancias fueron inferidas por el método de Máxima verisimilitud. El análisis tuvo un muestreo de 1000 réplicas, utilizando el programa MEGA6 (Kumar *et al.*, 2013).

También fue notable el agrupamiento de secuencias pertenecientes a ATsat-109 de *A. angustifolia* (en azul), siendo un pequeño grupo específico para *A. angustifolia* var. Chelem ki (en amarillo). Las diferencias observadas en el análisis entre las secuencias de ATsat-102-113 y ATsat-109, pueden sugerir que ATsat-102-113 amplifica principalmente secuencias más homogéneas para *A. angustifolia*, *A. fourcroydes* y *A. tequilana* de ADN

satélite, al menos eso resulta con las secuencias obtenidas. La distancia genética entre los dos grupos fue menor al 0.05%, mientras que esta fue mayor en cualquiera de las secuencias analizadas en la dispersión.

En los análisis previos para las secuencias obtenidas con el cebador ATsat102-113 se observaron dos grupos por lo que se realizó un análisis de secuencias en estos grupos, que fueron denominados grupo A y B. En los grupos A y B se encontraron 19 y 15 secuencias, respectivamente y fueron generadas a partir del cebador ATsat-102-113.

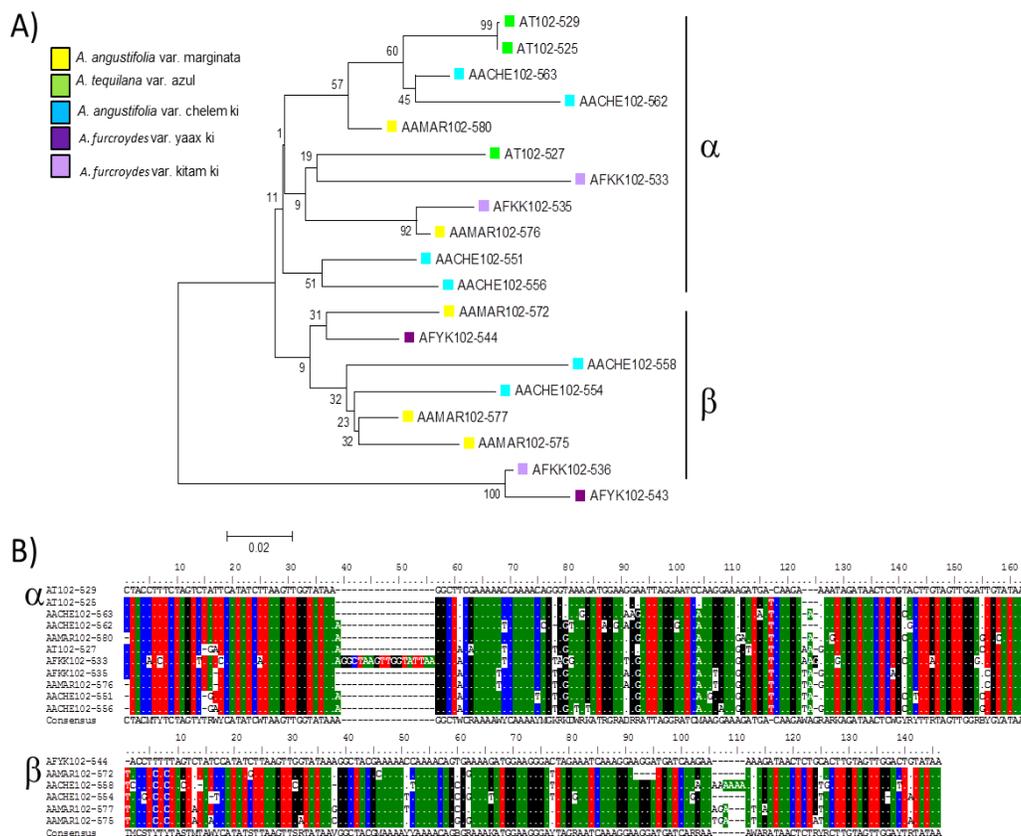


Figura 2.8. Diversidad y generación de secuencias consenso de ADN satélite para el grupo A en especies de *Agave*. **A)** dendrograma inferido usando el método de evolución mínima con un muestreo de 1000 réplicas. El código de color indica las diferentes accesiones de *Agave*. **B)** alineamiento de secuencias mostrando las sustituciones de nucleótidos y las secuencias consenso obtenidas en los subgrupos α y β con el programa BioEdit (Hall, 1999).

Posteriormente, a partir de los grupos A y B se generaron secuencias consenso para facilitar el análisis. La Figura 2.8 muestra un análisis filogenético empleando el método de Evolución mínima para el grupo A. Se observó una distribución homogénea entre las secuencias de las diferentes accesiones de *Agave* y se generaron dos subgrupos α y β , del cual se identificaron dos secuencias consenso, en el que se incluyó la información de

los nucleótidos y que para este análisis se denominaron secuencias consenso del grupo 1: 1 α y 2 β .

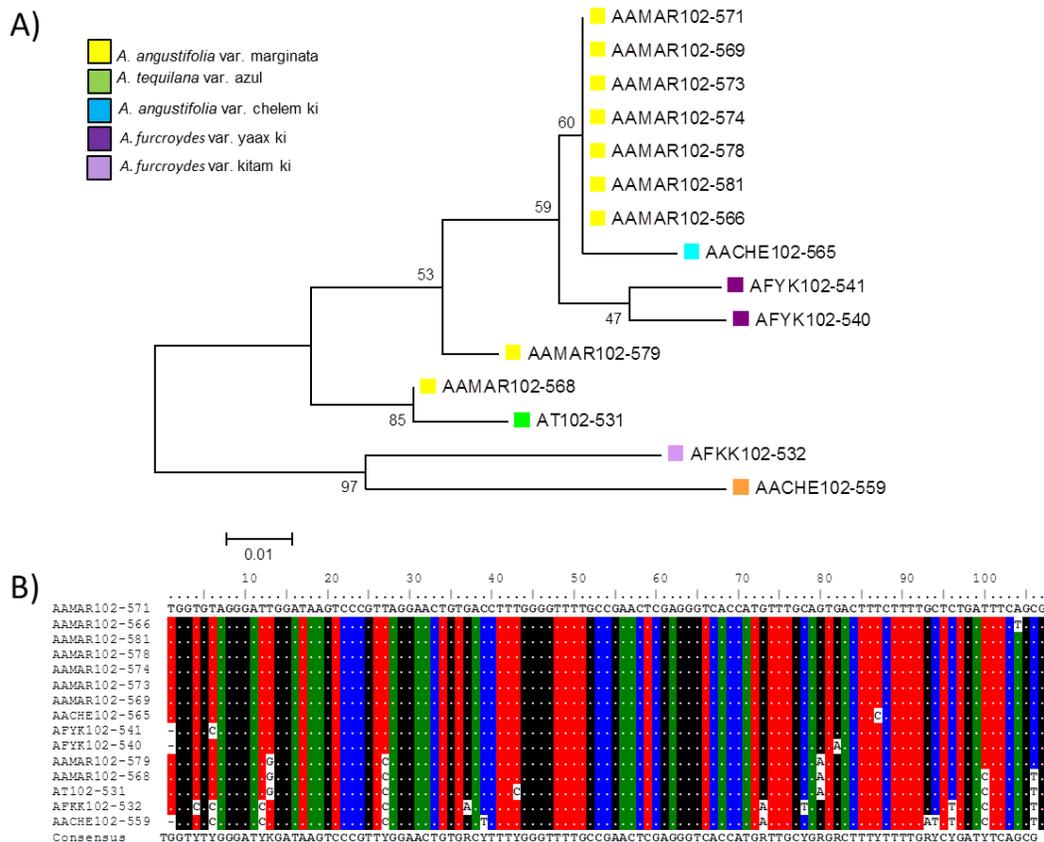


Figura 2.9. Diversidad y generación de secuencias consenso de ADN satélite para el grupo B en especies de *Agave*. **A)** dendrograma inferido usando el método de evolución mínima con un muestreo de 1000 réplicas. El código de color indica las diferentes accesiones de *Agave*. **B)** alineamiento de secuencias mostrando las sustituciones de nucleótidos y la secuencias consenso obtenidas con el programa BioEdit (Hall, 1999).

En el grupo B también se generó una secuencia consenso (Figura 2.9) a partir del fenograma general, en el que se observó una distribución homogénea en las diferentes accesiones de *Agave*, aunque 7 de las secuencias de *A. angustifolia* var. Marginata (AAMAR) son idénticas. La distancia genética fue menor a 0.05 por lo que estas secuencias se utilizaron para generar una secuencia consenso de ADN satélite para el grupo 2.

Las secuencias consenso obtenidas fueron empleadas para realizar un análisis con BLAST en la base de datos de transcriptomas (TSA), con las especies del género *Agave* (taxid: 39509), con las secuencias de una clona de una biblioteca BIBAC de *A. tequilana*

(Tamayo-Ordoñez *et al.*, 2012), en la base de datos de ADN satélites de plantas PlantSat (Macas *et al.*, 2002) y la biblioteca general del Genbank (Tabla 2.4).

Tabla 2.4. Identificación de secuencias consenso con el PlantSat, el Genbank, con secuencias de TSA (Transcriptome Shotgun Assembly) y con secuencias de una clona BIBAC de *A. tequilana*.

Secuencia	TSA	BIBAC	PlantSat	NCBI**
Consenso				
Grupo I α	GAHT01082464.1 (2-40)*	contig01749 (2-36)	N.I.	N.I.
Grupo I β ***	GAHT01104603.1 (2-39)	contig01749 (3-33)	N.I.	N.I.
Grupo II	GAHU01183764.1 (3-22)	contig07395 (2-15)	N.I.	N.I.
Representativa				
Grupo I α	GAHT01007348.1 (2-51)	contig01749 (5-44)	AF07127/3(0.068)	AP014962.1 (1.7)
Grupo I β	GAHT01104603.1 (4-48)	contig01749 (4-39)	AF137608/3(2.8)	HG975444.1(0 .47)
Grupo II	GAHU01183764.1 (1-40)	contig07395 (3-26)	AF025353 (2.1)	LT594793.1 (0.34)

TSA: Transcriptome shotgun assembly en Agave (taxid: 395909)

BIBAC de *A. tequilana* (Tamayo-Ordoñez *et al.*, 2012)

*Números de accesión seguidos entre paréntesis los e-value

**Fue seleccionada para indicar la identidad más cercana, la primera secuencia de plantas correspondiente.

***Tres nucleótidos degenerados fueron substituidos por los más abundantes para analizar esta secuencia.

N.I. Sin identidad.

A nivel de secuencia consenso se pudo identificar su transcripción *in silico* en *A. tequilana*, en varios fragmentos de secuencias (contigs) de la clona BIBAC de esta especie y no se identificaron en el PlantSat y el Genbank. Cabe mencionar, que los valores de e-value obtenidos para estas secuencias fueron bajos, por lo que su identidad podría ser alta en el género *Agave* y la información obtenida del transcriptoma puede sugerir que estos están transcribiéndose; sin embargo, para este último punto es necesarios realizar otros estudios.

Para determinar la existencia de identidad de estas secuencias consenso en el PlantSat o GenBank, se eligieron secuencias representativas para cada grupo, considerando la primera secuencia de cada grupo independientemente de la especie de *Agave*, y se repitió el análisis. Los resultados de este último demostró que las secuencias representativas de ADN satélite de cada grupo tienen pequeñas secuencias (o firmas) en común con las secuencias depositadas en el PlantSat y en el GenBank, sugiriendo que

las degeneraciones introducidas en las secuencias consenso pueden afectar su análisis con ambas bases de datos.

2.3.5. Southern Blot y posible organización de dos subfamilias satélites en *Agave*

Para observar la posible organización en tándem de estas secuencias se realizó a partir de un Southern Blot con las sondas del grupo 1 y 2, que a partir de ahora se denominaron subfamilia 1 y 2, respectivamente. Estas sondas son AACHE102-554 de la subfamilia 1 y AT102-531 de la subfamilia 2. El Southern se realizó con ambas sondas sobre el producto de PCR generado con el cebador ATsat-102-113 (Figura 2.10).

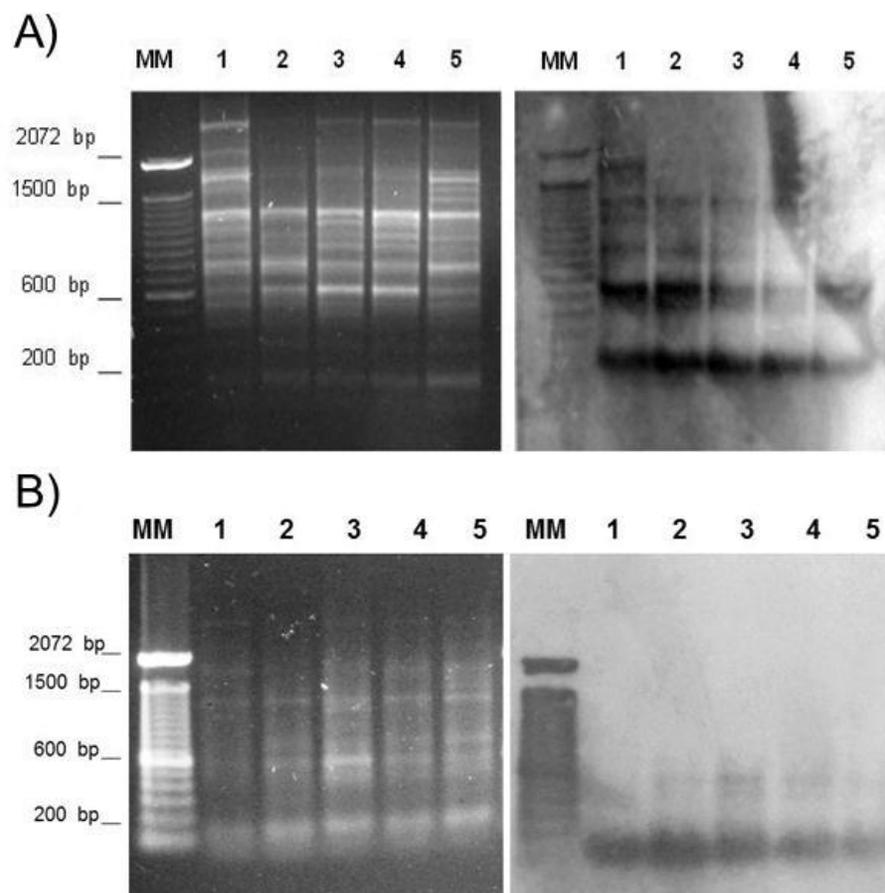


Figura 2.10. Southern Blot para las subfamilias 1 y 2 de satélites. A) Múltiples bandas observadas para la subfamilia 1 en todas las especies evaluadas. B) Dos bandas observadas para la subfamilia 2, una de ~100pb mostró una mayor intensidad en todas las especies evaluadas. 1) *A. tequilana*, 2) *A. fourcroydes* var. Kitam ki, 3) *A. fourcroydes* var. Yaax ki, 4) *A. angustifolia* var. Chelem ki, 5) *A. angustifolia* var. Marginata. MM) Marcador de peso molecular 100pb.

El Southern Blot de la subfamilia 1 (AACHE102-554) y subfamilia 2 (AT102-531) se realizó para el producto de PCR generado por el cebador ATsat-102-113.

El patrón de hibridación generado para la subfamilia 1 fue de múltiples bandas, observándose una mayor intensidad de estas en ~150pb y 550pb; el patrón de multímeros generado por esta subfamilia muestra que se trata de una región de ADN repetida en tándem. Cabe mencionar que se encontraron presentes en todas las especies evaluadas, siendo *A. tequilana* la especie en la que más bandas se observaron. Para la subfamilia 2 se observaron dos bandas en todas las especies evaluadas, siendo la banda de 100pb muy intensa comparada con la de ~500 pb.

Los patrones observados para cada una de las subfamilias fueron distintos. Estas diferencias sugieren que se tratan de dos grupos de secuencias que están organizados de manera diferente en el genoma de estas especies de *Agave*.

2.3.6. Diseño de cebadores para las subfamilias A y B de ADN satélite en *A. tequilana*, *A. fourcroydes* y *A. angustifolia*

Las secuencias obtenidas fueron analizadas y filtradas por tamaños para identificar las aquellas que comparten entre especies, a estos grupos de secuencias similares los denominamos tipos de ADN satélite. Los tamaños de los tipos para el cebador ATsat-102-113 se encuentran en un rango de 46 y 441 pb. A partir de estas secuencias se seleccionaron dos grupos: el tipo I y II que está presente en las especies evaluadas, y tienen un tamaño de 140 y 107 pb, respectivamente.

Los tipos I y II de ATsat-102-113 fueron analizados con el objetivo de determinar los cambios en la composición de la secuencia y porque éstos se encuentran presentes en las 3 especies estudiadas; motivo por el cual se diseñaron dos pares de cebadores, denominados cebadores tipo I y tipo II (Tabla 2.5).

Tabla 2.5. Cebadores diseñados a partir de las secuencias del Tipo I y II de ATsat-102-113.

Cebador	Secuencia
FW tipo I (140)	5' -AGTTGGTATAAGGCTTCGAA- 3'
RV tipo I (140)	5' -CAATCCAACACTACAAGTACAGAGT- 3'
FW tipo II (107)	5' -TGTAGGGATGGGATAAGTCC- 3'
RV tipo II (107)	5' -TCAGAGCAAAAGAAAGTCATT- 3'

Los cebadores tipo I y II fueron usados para la obtención de fragmentos por PCR de los ADN genómicos de las tres especies de este estudio: *A. tequilana*, *A. angustifolia* y *A. fourcroydes*. Los resultados para la amplificación del cebador tipo I fueron tres fragmentos de 140pb, 500pb y 800pb, y para el cebador tipo II solo se observó una banda de aproximadamente 120pb (Figura 2.11). Los patrones de electroforesis de cada cebador, para todas las especies, fue el mismo.

El patrón generado por el cebador tipo I, con al menos tres bandas, muestra que existen dos fragmentos mayores al primero, la cual es una característica de las regiones repetidas en tándem, Sin embargo, para concluir esta información es necesario realizar la secuenciación del probable re-arreglo de estos fragmentos (la banda de ~800 pb).

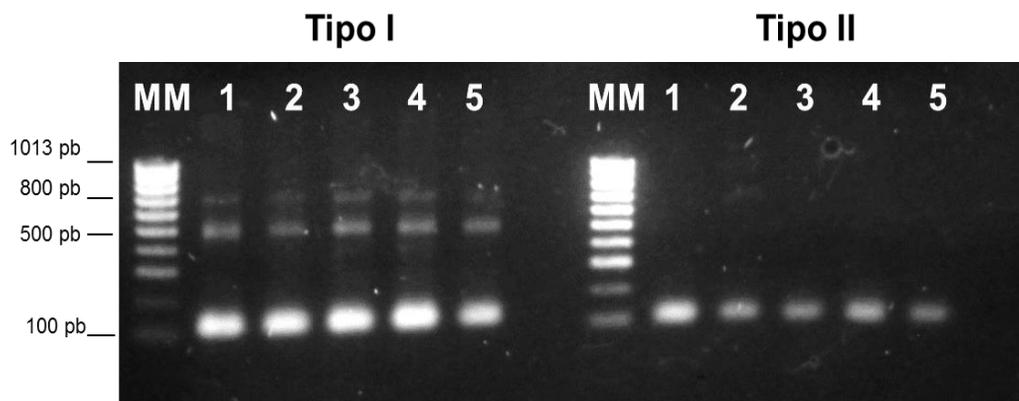


Figura 2.11. Electroforesis de los fragmentos generados por los cebadores Tipo I y Tipo II en las especies de *Agave*. 1) *A. tequilana*, 2) *A. fourcroydes* var. Kitam ki, 3) *A. angustifolia* var. Yaax ki, 4) *A. angustifolia* var. Chelem ki y 5) *A. angustifolia* var. Yaax ki. MM: Marcador de peso molecular 100 pb lines bioline.

2.4. DISCUSIÓN

El ADN satélite es una región repetida en tándem con un tamaño específico en su motivo y los rangos de este motivo en pares de bases (pb) puede ser variable para cada especie de planta en estudio (Neumann, 2001; Ribeiro *et al.*, 2016; Tek *et al.*, 2005). Además del motivo repetido, pueden contener distintos motivos cortos dentro de su secuencia y pueden estar relacionado con otras regiones repetidas como los elementos móviles (retroelementos o transposones) (Biscotti, Olmo y Heslop-Harrison, 2015) y la localización en los cromosomas, es preferencialmente, en centrómeros y regiones subteloméricas (Torres *et al.*, 2011), aunque en menor frecuencia pueden encontrarse en regiones intersticiales de los cromosomas (Ramos *et al.*, 2012).

2.4.1. Aislamiento e identificación de ADN satélite en *Agave*

Los resultados obtenidos en la electroforesis de los fragmentos amplificados con los cebadores ATsat-102-113 y ATsat-109 mostraron un perfil de bandeo en escalera; el patrón generado por los fragmentos amplificados con ATsat102-113 fue homogéneo para las especies evaluadas, pero distintas a los patrones generados por ATsat-109. Estas diferencias también se observaron al analizar las secuencias obtenidas de ambos cebadores en las tres especies de agave.

La variación encontrada para estas secuencias podría sugerir que los fragmentos generados para ATsat-102-113 y ATsat-109 son debido a que son dos miembros distintos de la misma familia de ADN satélite y según la teoría de “La librería” diversas familias de repeticiones pueden existir en el genoma de los eucariotas (Fry y Salser, 1977).

La teoría de “La librería” establece que especies relacionadas tienen un conjunto de diversas familias de ADN satélites y cada una puede ser amplificada o reemplazada de manera diferente; la librería persiste a largos periodos de tiempo y reduce la acción de mecanismos moleculares (Fry y Salser, 1977; Plohl *et al.*, 2012). Un ejemplo de esta teoría es el caso de la familia satélite HinfI, que esta conservada en el genoma de las especies de la tribu Cardueae (familia Compositae), está presente en el ancestro común de Cardueae y en esta familia se encontraron 9 subfamilias (I-IX) distribuidas en los diferentes géneros de las subtribus (Garrido-Ramos *et al.*, 2014).

La electroforesis de los fragmentos generados por ambos cebadores mostraron un perfil de bandeo diferencial notable (Figura 2.2), sugiriendo una baja homogenización y una mayor divergencia del ADN satélite en estas especies. En ocasiones estas diferencias

pueden estar dadas por el reemplazo de nucleótidos a través de mutaciones. Esto se ha observado en las especies de *Nicotiana nesophila*, *Nicotiana stocktonii* y *Nicotiana repanda*, un nuevo y abundante ADN satélite, NNE10, ha evolucionado reemplazando a las repeticiones ancestrales encontradas en las especies *N. sylvestris* y *N. obtusifolia* que son parentales de dichas especies (Koukalova *et al.*, 2010).

Las secuencias obtenidas en de este trabajo fueron clasificadas en tipos, siendo 7 y 5 tipos para ATsat-102-113 y ATsat-109, respectivamente. El rango de tamaño para los tipos fue variable, ATsat-102-113 presentó un rango de tamaño en las secuencias de 46 y 441 pb, y para ATsat-109 fue de 63 pb hasta 562 pb. Los ADN satélites reportados son consistentes con un rango de tamaño del motivo, que van desde 135-195 pb o entre 315-375 pb (Macas *et al.*, 2002), estos tamaños de las unidades monoméricas básicas del ADN repetido en plantas corresponden al largo de este ADN que se encuentra alrededor del nucleosoma, aunque aún no se encuentra evidencia contundente para esta suposición (Macas *et al.*, 2002; Schmidt y Heslop-Harrison, 1998).

A pesar de que hay un tamaño esperado para los motivos de ADN satélite, se han reportado que estos pueden tener diversos tamaños. En el PlantSat (Macas *et al.*, 2002) se han reportado tamaños que van desde 81 pb del satélite familia81 en *Olea europaea* ssp. *sativa* (Katsiotis *et al.*, 1998), 92 pb como el satélite SB92 en *Glycine max* (Vahedian *et al.*, 1995), NTRS9 con 218 pb de *Nicotiana tomentosiformis* (Matyasek *et al.*, 1997), 223pb de LPD220 en el género *Larix* (Hizume *et al.*, 2002), COM2 con 476 pb en *Helictotrichon compressum* (Greibenstein *et al.*, 1996), el satélite Tail con 570 pb en *Leymus racemus* (Kishii *et al.*, 1999), los satélites en *Oryza sativa* Os48, OSG3-498 y OSG5-756, con tamaños de 355, 498 y 756 pb, respectivamente (Wu *et al.*, 1991), entre otros.

Los motivos identificados en las secuencias de las diferentes especies de *Agave* comprendieron un tamaño de 7 a 13 nucleótidos. Entre las diferentes especies en las que se han identificado ADN satélites se encuentra la monocotiledónea *Prospero autumnale* con PaB6, de 249 pb, que mostró motivos repetidos de 6 nucleótidos a lo largo de la secuencia (Emadzade *et al.*, 2014); también en dos monocotiledóneas *Oryza sativa* y *Sorghum bicolor* se identificaron motivos repetidos de entre 3 y 9 nucleótidos, en sus secuencias repetidas de 150 y 135 pb, respectivamente; estas especies pertenecen a la clase Liliopsida, misma a la que pertenece el género *Agave* (Navajas-Pérez y Paterson, 2009). Así como, el ADN satélite Spelt-1, presente en las especies de la tribu Triticeae, que ha evolucionado a partir de un bajo número de copias hasta uno mayor, alcanzando

con esto su propagación en esta tribu (Salina *et al.*, 1998). Se ha documentado que el ADN satélite contiene motivos que son altamente homogenizados debido a la evolución concertada (John, Elder y Turner, 1995).

2.4.2. Identificación de motivos que comprenden las secuencias ADN satélite en *Agave*

La identificación de motivos internos en las secuencias obtenidas se logró con ayuda de la base de datos del PlantSat (Macas *et al.*, 2002) y se observaron diversos motivos cortos, de entre 7 pb y 13 pb; también se realizó la búsqueda de los motivos cortos en las secuencias pertenecientes a la misma especie y entre las especies evaluadas. Se ha documentado que la variabilidad en la secuencia de los miembros de las familias de ADN satélite está dada por los distintos motivos que puede contener en su composición, estos motivos van desde dinucleótidos, trinucleótidos, pentanucleótidos, etcétera; además, estos motivos pueden ser cortos, directos, inversos y palíndromos (Plohl *et al.*, 2008).

El motivo CAAAA se ha encontrado en diversos ADN satélites (Appels y Moran, 1986; Katsiotis *et al.*, 1998; Samoluk *et al.*, 2017) y se sugiere que está involucrado en el plegamiento de la molécula de ADN, que a su vez proporciona las características adecuadas para la amplificación o mantenimiento del ADN satélite en el genoma; además, se cree que este motivo puede estar implicado en procesos de recombinación entre las repeticiones (Appels y Moran, 1986; Macas *et al.*, 2002).

Los distintos motivos que se encuentran presentes dentro de las secuencias de ADN satélite pueden participar en la recombinación homóloga como regiones de mayor similitud y también pueden ser determinantes en modificaciones epigenéticas (Plohl *et al.*, 2012). Una característica del ADN satélite es que aunque los monómeros pueden estar presentes en miles de copias por genoma, la divergencia en la secuencia entre repeticiones de la misma familia a menudo es baja, usualmente hasta un 15% (Plohl *et al.*, 2008).

Las mutaciones en la secuencia de ADN satélite, dentro y entre arreglos de la repetición en tándem y su unidad variante, también provocan cambios en los perfiles del ADN satélite; debido a la abundancia de cambios dentro de estas secuencias se cree que pueden afectar al cariotipo y por lo tanto juegan un papel en la especiación (Biscotti *et al.*, 2015). La acumulación de mutaciones en familias de ADN satélite no es la única manera de alterar los perfiles de repeticiones en el satélite en periodos cortos. En adición a los cambios en la secuencia, los ADN satélites permanecen alterados en el número de copias

por la amplificación o pérdida de arreglos de monómeros en el ADN satélite (Rojo *et al.*, 2015).

2.4.3. Uso de información generada por ADN satélite en plantas

En la actualidad, siguen reportándose el aislamiento de ADN satélites en plantas. Por ejemplo, se ha reportado por primera vez tres ADN satélites en rábano (*Raphanus sativus* L.) que están asociados al centrómero (CL1), a la región subteloamérica (CL25) y al telómero (CL43) de los cromosomas para esta especie; la proporción de regiones ADN satélites fueron más abundantes con un 12% (He *et al.*, 2015). En *Phaseolus vulgaris* se ha encontrado que el genoma está compuesto del 50% de regiones repetidas (Fonseca *et al.*, 2010). Específicamente en *P. vulgaris*, se han identificado los ADN satélites CC4 (Almeida *et al.*, 2012) y khipu (David *et al.*, 2009; Richard *et al.*, 2013) y el retrotransposon Tpv2 (Garber *et al.*, 1999). Es interesante resaltar, que en la especie *Phaseolus microcarpus* además del satélite khipu se encuentra un nuevo satélite adyacente a éste, denominado *junper* que resultó ser específico para esta especie y para una especie relacionada *Phaseolus oligospermus*; la presencia de *junper* en estas dos especies sugiere que pudo originarse y evolucionar de manera independiente de khipu, lo cual ejemplifica la posible amplificación y pérdida de los ADN satélites (Ribeiro *et al.*, 2017).

Entre la información que los ADN satélites pueden brindar se encuentran el análisis comparativo de dos ADN satélites, pCtKpnI-I y pCtKpnI-II, realizado en 20 taxa de la familia Asteraceae y 43 taxa que representan a 36 familias de las angiospermas. Este estudio reveló que pCtKpnI-I está presente en 18 taxa de las evaluadas, mientras que el segundo satélite pCtKpnI-II es específico para el género *Carthamus*. Este resultado puede estar dado por la variación en los motivos internos que presentan ambos ADN satélites (Mehrotra *et al.*, 2014).

El satélite ATR-2 de la sección *Arachis* del género *Arachis*, mostró una distribución diferencial en los cromosomas entre siete especies pertenecientes a esta sección. ATR-2 exhibe un alto porcentaje en A/T (75%, el más alto reportado en el PlantSat), su abundancia en los distintos genomas evaluados fue distinta (apoyando la teoría de la librería), y por último, se demostró que la participación de ATR-2 en las especies de *Arachis* contribuye a la diversificación de la heterocromatina para estas especies (Samoluk *et al.*, 2017).

En el género *Nicotiana* se han identificados una gran variedad de repeticiones de ADN satélite como NTRS (Matyasek *et al.*, 1997) y NTS9 (Jakowitsch *et al.*, 1998) e incluso se

ha reportado una familia, denominada HRS60 (Koukalová *et al.*, 1989). En las últimas dos décadas se han realizado aislamientos, caracterizaciones moleculares y citogenéticas de ADN satélites en especies de diferentes niveles de ploidía e híbridos (Clarkson *et al.* 2005; Lim *et al.*, 2004; Lim *et al.*, 2000; Lim *et al.*, 2006). La información que los ADN satélites han brindado al género *Nicotiana* es la comprensión de la divergencia en el genoma y en los cromosomas que provoca la presencia de secuencias repetidas en especies que tienen diferentes niveles de ploidía, especies con diferentes edades, híbridos y en diferentes taxa de este género (Jo *et al.*, 2011; Jo *et al.*, 2009; Koukalova *et al.*, 2010; Matyasek *et al.*, 2011).

CAPÍTULO III MAPEO FÍSICO DE ADN SATÉLITE EN LAS ESPECIES *A. tequilana*, *A. fourcroydes* Y *A. angustifolia*

3.1. INTRODUCCIÓN

La hibridación *in situ* fluorescente (FISH) es un método efectivo para el mapeo físico de genes y regiones repetidas en núcleos y cromosomas (Kinoshita *et al.*, 2010). La hibridación *in situ* contribuye al entendimiento en la estructura, función, organización y evolución en el genoma (Schwarzacher y Heslop-Harrison, 2000). El mapeo físico de secuencias de ADN en cromosomas, la identificación y caracterización de cromosomas o segmentos de estos provee un indicador de su arreglo en el genoma. La información que proporciona FISH es variada y tiene relación con el número de señales en núcleo o cromosoma, la intensidad de la señal de hibridación, la distancia entre señales de hibridación y morfología de la señal (Ma *et al.*, 1997; Ohmido *et al.*, 2010).

En el estudio del ADN satélite la técnica de FISH ayuda a conocer su distribución y localización; esta información es usada para observar los cambios de organización en el genoma y los posibles eventos por los cuales surge una reorganización. Un ejemplo de este estudio es el realizado en patata, en el cual se determinó la localización de un ADN satélite es específico del centrómero de un solo cromosoma, característica que permite la discriminación entre otros cromosomas, haciéndolo un marcador cromosómico (Gong *et al.*, 2012).

A través del estudio de los cariotipos se ha determinado que las especies del género *Agave* presenta un cariotipo bimodal, es decir, consta de cromosomas pequeños y grandes (Castorena-Sánchez *et al.*, 1991). En 2008, Robert *et al.*, emplearon la técnica de FISH para la localización y el estudio de los genes ribosomales 5S y 35S en distintas especies poliploides de *Agave* (Robert *et al.*, 2008). Posteriormente, esta información fue complementada con la localización del gen ribosomal 5S en los cromosomas pequeños y al 18S en cromosomas grandes (Gómez-Rodríguez *et al.*, 2013). La información generada puede utilizarse para sugerir y explorar los mecanismos que rigen esos cambios en los arreglos y evolución de los genes ribosomales, lo cual es posible con la determinación de su posición en el cromosoma.

En este estudio se llevó a cabo la localización de dos satélites en tres especies de *Agave* a través de la técnica FISH. Este es la primera exploración del mapeo de estas secuencias repetidas en el genoma de agave como parte complementaria a la

caracterización de las mismas. La comparación de la cantidad, posición e intensidad de las señales de ambos satélites mostraron un patrón diferente para las especies evaluadas. En este capítulo se discuten estas diferencias observadas con base en la organización que los ADN satélites presentan en el genoma de plantas.

3.2. MATERIALES Y MÉTODOS

3.2.1. Obtención de preparaciones con células en metafase de *Agave*

La colecta de raíces se realizó aproximadamente 15 días después de podar los tallos de plantas de *A. angustifolia* var. Marginata ($2n=2x=60$), *A. angustifolia* var. Chelem ki ($2n=6x=180$), *A. fourcroydes* var. Kitam ki ($2n=3x=90$), *A. fourcroydes* var. Yaax ki ($2n=5x=150$), y *A. tequilana* ($2n=2x=60$), que se encuentran en maceta. Los cromosomas en metafase se obtienen a partir de ápices de la raíz que se colectan a las 7:00 horas. Las raíces son lavadas con agua y se colocan en un frasco que contiene una solución saturada de Para-diclorobenceno (PDB) durante 3 horas a temperatura ambiente. Transcurrido el tiempo las raíces se lavan para eliminar residuos del agente mitostático y se fijan en una solución de alcohol absoluto y ácido acético glacial (3:1) durante 24 horas. La pared celular de las células de las raíces es degradada con Viscozyme (SIGMA cat. V2010), un coctel enzimático que se incuba a 37 °C a baño maría, durante 25 minutos. Posteriormente, se realizan dos lavados con amortiguador de citratos o agua destilada estéril. La raíz se coloca en un portaobjeto que contiene una gota de ácido acético al 45% y se corta con una navaja de afeitar hasta que el tejido esté cortado. El análisis de preparaciones se realizó en un microscopio de contraste de fases para identificar las células en metafase. Al término del análisis de las preparaciones se realizó el “squash” presionando el cubreobjeto sobre el tejido.

Una vez obtenidas preparaciones con material de buena calidad (células en metafase) se retira el cubreobjetos de las preparaciones con nitrógeno líquido y se dejan secando en la placa de calentamiento a 42 °C durante 24 horas. Las preparaciones se pueden almacenar a -20 °C hasta su uso.

3.2.2. Marcaje de la sonda con Digoxigenina (dig-11-dUTP)

El marcaje de sondas se realizó por PCR utilizando el programa para ATsat-102-113 y digoxigenina (DIG) con 'DIG DNA Labeling Kit', (Roche™). La verificación de la sonda

marcada se realiza por electroforesis en un gel al 0.8% de agarosa, a 70 V durante 40 min.

3.2.3. Hibridación *in situ* fluorescente de ADN satélite en especies de *Agave*

Las preparaciones, previamente expuestas a 42 °C durante toda la noche, fueron tratadas con 100 µg/ml de RNasa A (79U/ml), cubiertas con un cuadro a medida de Parafilm y se incubaron durante 1 hora a 37 °C en una cámara húmeda. Después de la incubación se el parafilm fue removido y las preparaciones lavadas dos veces con 2X SSC, durante 5 min. Transcurrido el tiempo, las preparaciones se incuban en 10 mM de HCL por 5 min y se les adicionó 10 µl/ml de pepsina (SIGMA cat. P6887). Otro cuadro a medida de parafilm fue empleado para cubrir las preparaciones; estas se incubaron a 37 °C por 15 min y fueron lavadas dos veces en 2X SSC por 5 min.

Las preparaciones se colocan en una caja Coplin con Paraformalehido 4% (p/v) (SIGMA cat. P6148), durante 10 min. Transcurrido el tiempo las preparaciones se lavaron dos veces en 2X SSC por 5 min y se deshidratan con etanol a diferentes concentraciones 70%, 90% y 100% v/v, por 3 min en cada uno. Por último, se dejaron secar a temperatura ambiente. Una vez secas, se procede a la desnaturalización para lo cual las preparaciones se colocan en una caja Coplin con 70% de formamida (v/v) (SIGMA cat. F5786), previamente atemperada a 75 °C a baño maría. Inmediatamente a la desnaturalización, fueron deshidratadas en hielo durante 5 min en una serie de etanol al 70%, 90% y 100%, respectivamente.

Al término de la deshidratación se realizó la hibridación de la sonda interés. A cada preparación se le aplicó 40 µl, de la mezcla de hibridación compuesta de 50% de formamida (v/v) (SIGMA cat. F5786), 2X SSC, 10% de Sulfato de Dextrano (p/v) (SIGMA cat. D8906), 0.17% de SDS y 100 ng de sonda (previamente desnaturalizada a 75 °C durante 10 min y mantenida en hielo hasta su aplicación). Se cubrió el área con las células en metafase con Parafilm y se incubó a 42 °C a baño maría durante toda la noche. Posterior a la hibridación, se procedió a la detección de la sonda en donde las preparaciones fueron lavadas dos veces en 2X SSC por 5 min. Después de los lavados fueron colocadas por 5 min (una preparación a la vez) en una serie de 20% de formamida (v/v) y 2X SSC atemperadas a 42 °C a baño maría. Enseguida, las preparaciones se lavaron dos veces en 4X SSC y 0.2% (v/v) de Tween 20 (SIGMA cat. P1379) durante 5 min.

Transcurrido el tiempo de lavados astringentes, las preparaciones se escurren y se aplica 5% BSA (p/v) (SIGMA cat. A2153). Posteriormente, se cubrieron las preparaciones con Parafilm y se dejaron incubar a 37 °C durante 20min. Se retiró la cubierta de Parafilm y se aplicó 5 µl (0.75 U/µl) de Anti-digoxigenina conjugada con Fluoresceína (ROCHE cat. 11207741), y se incubó a 37 °C durante una hora. Pasado el tiempo de detección, las preparaciones se lavaron dos veces en 4x SSC y 0.2% de tween 20 (v/v) (SIGMA cat. P1379). Las preparaciones se escurrieron y se aplicaron 30 µl (1.5 µg/ml) del medio de montaje Vectashield con DAPI (VECTOR cat. H-1200), un cubreobjetos fue colocado sobre la preparación y por último se mantuvo en oscuridad por 10 min.

3.2.4. Fotografía de células en metafase y análisis de las señales

Las preparaciones se analizaron con el juego de filtros 09 (BP 450-490 nm excitación, FT 510 filtro divisor de haz, y LP 515 nm emisión) para detectar la señal de la fluorescencia en verde, y el juego de filtros 01 (BP 365/12; FT 395 y LP 397 nm) para la emisión en azul del complejo DAPI-ADN, en un microscopio de Carl Zeiss Axioplan, equipado con un sistema de Epifluorescencia para su análisis. Las imágenes se capturan mediante una cámara AxioCam MRm y el software AxioVision de Carl Zeiss.

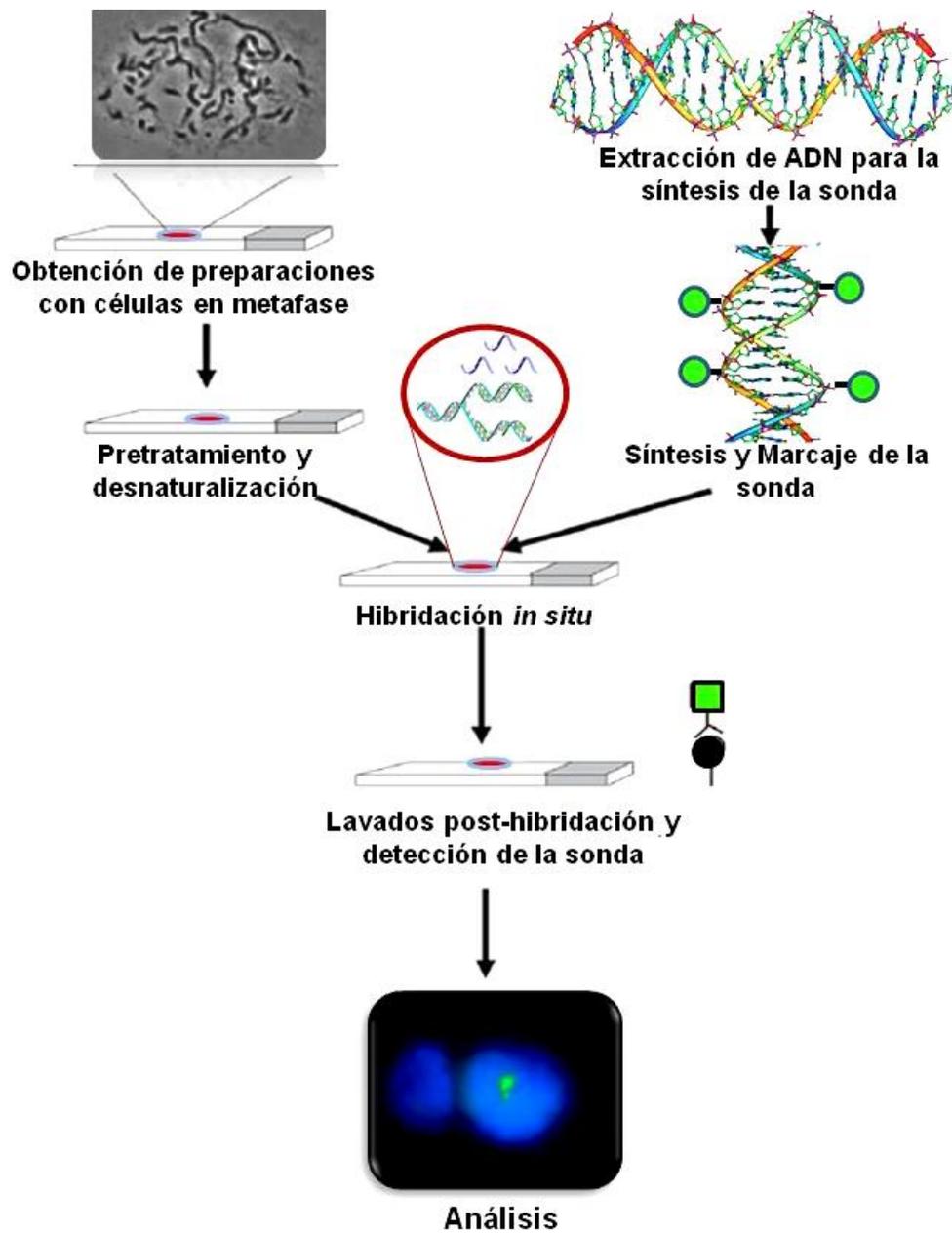


Figura 3.1. Estrategia experimental para la hibridación *in situ* de ADN satélite en especies de *Agave*. (Rescalvo-Morales Tesis de maestría, 2011).

3.3. RESULTADOS

3.3.1. Mapeo físico de ADN satélite en tres especies de la sección *Rigidae* del género *Agave*

A partir del análisis de las secuencias de ADN satélite y la agrupación de las mismas se utilizaron 2 sondas para el mapeo físico, una para el grupo A y una para el grupo B. Para el grupo A y el grupo B se seleccionaron clonas pertenecientes a la especie *A. angustifolia* (AACHE102-554) y *A. tequilana* (AT102-531), respectivamente. En la figura 3.2 se observa las metafases mitóticas de dos especies diploides ($2n=2x=60$), *A. tequilana* y *A. angustifolia* var. Marginata, dos citotipos de *A. fourcroydes*, *A. f.* var. Kitam Ki ($2n=3x=90$), y el otro pentaploide *A. fourcroydes* var. Yaax ki ($2n=5x=150$) y un citotipo hexaploide *A. angustifolia* var. Chelem ki ($2n=6x=180$).

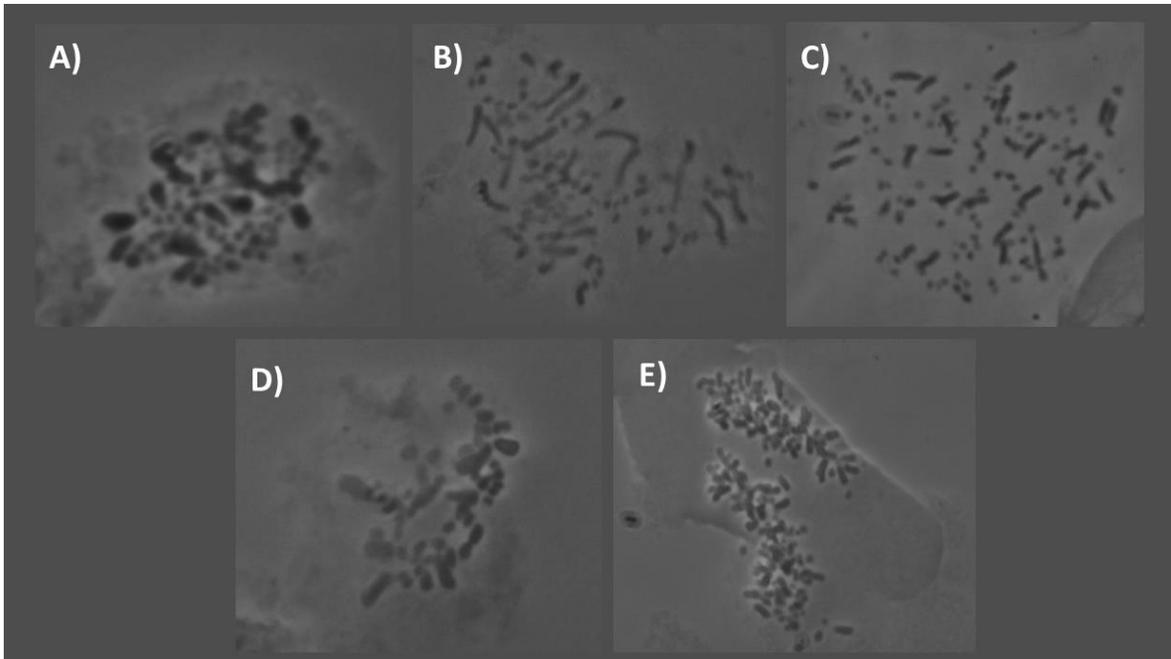


Figura 3.2. Fotografías de preparaciones de células en metafase de *A. tequilana*, *A. fourcroydes* y *A. angustifolia*. A) *A. tequilana*, B) *A. fourcroydes* var. Kitam ki, C) *A. fourcroydes* var. Yaax ki, D) *A. angustifolia* var. Marginata y E) *A. angustifolia* var. Chelem ki.

Los resultados de la hibridación *in situ* con la sonda A (AACHE102-554; Fig. 3.3) fueron los siguientes: para *A. tequilana* se detectaron 4 señales en cromosomas pequeños en la posición subtelomérica; en *A. fourcroydes* para ambas variedades se encontraron señales en cromosomas grandes y pequeños pero en un número distinto de señales, siendo más

de 20 señales en *A. fourcroydes* var. Kitam ki y 10 en *A. fourcroydes* var. Yaax ki, la posición de estas señales fue en regiones centroméricas. En *A. angustifolia* var. Marginata se encontraron 16 señales en regiones centroméricas, subteloméricas e intersticiales de los cromosomas pequeños y grandes; en *A. angustifolia* var. Chelem ki se encontraron 12 señales en cromosomas grandes y pequeños en regiones subteloméricas e intersticiales (Tabla 3.1).

La hibridación para la sonda B (AT102-531; Fig. 3.4) dio como resultado para *A. tequilana* 17 señales en cromosomas pequeños y grandes ubicados en regiones centroméricas, subteloméricas e intersticiales de los cromosomas; en *A. fourcroydes* var. Kitam ki se observaron 10 señales en regiones subteloméricas de cromosomas pequeños y para *A. fourcroydes* var. Yaax ki se observaron 10 señales en regiones subteloméricas e intersticiales de cromosomas grandes y pequeños; para *A. angustifolia* var. Marginata las señales observadas fueron 13 en regiones centroméricas, subteloméricas e intersticiales en cromosomas grandes y pequeños, por último para *A. angustifolia* var. Chelem ki se encontraron 12 señales en regiones subteloméricas e intersticiales de cromosomas grandes y pequeños (Tabla 3.2).

Entre las especies diploides hibridadas, *A. tequilana* y *A. angustifolia* var. Marginata, encontramos menos señales de la sonda A (AACHE102-554) en *A. tequilana* y estas señales, 4 comparadas con 16, fueron ubicadas en regiones subteloméricas de cromosomas pequeños. Comparando las señales encontradas para la sonda B (AT102-531) en *A. tequilana* y *A. angustifolia* var. Marginata sucedió algo distinto pues la cantidad de señales para cada diploide fueron 17 y 13, respectivamente; además las regiones en las que se encontraron fueron centroméricas, subteloméricas e intersticiales de cromosomas pequeños y grandes. Siendo la sonda B de *A. tequilana* la que mostró una alta cantidad de señales en el genoma de estas especies diploides (Tablas 3.1 y 3.2).

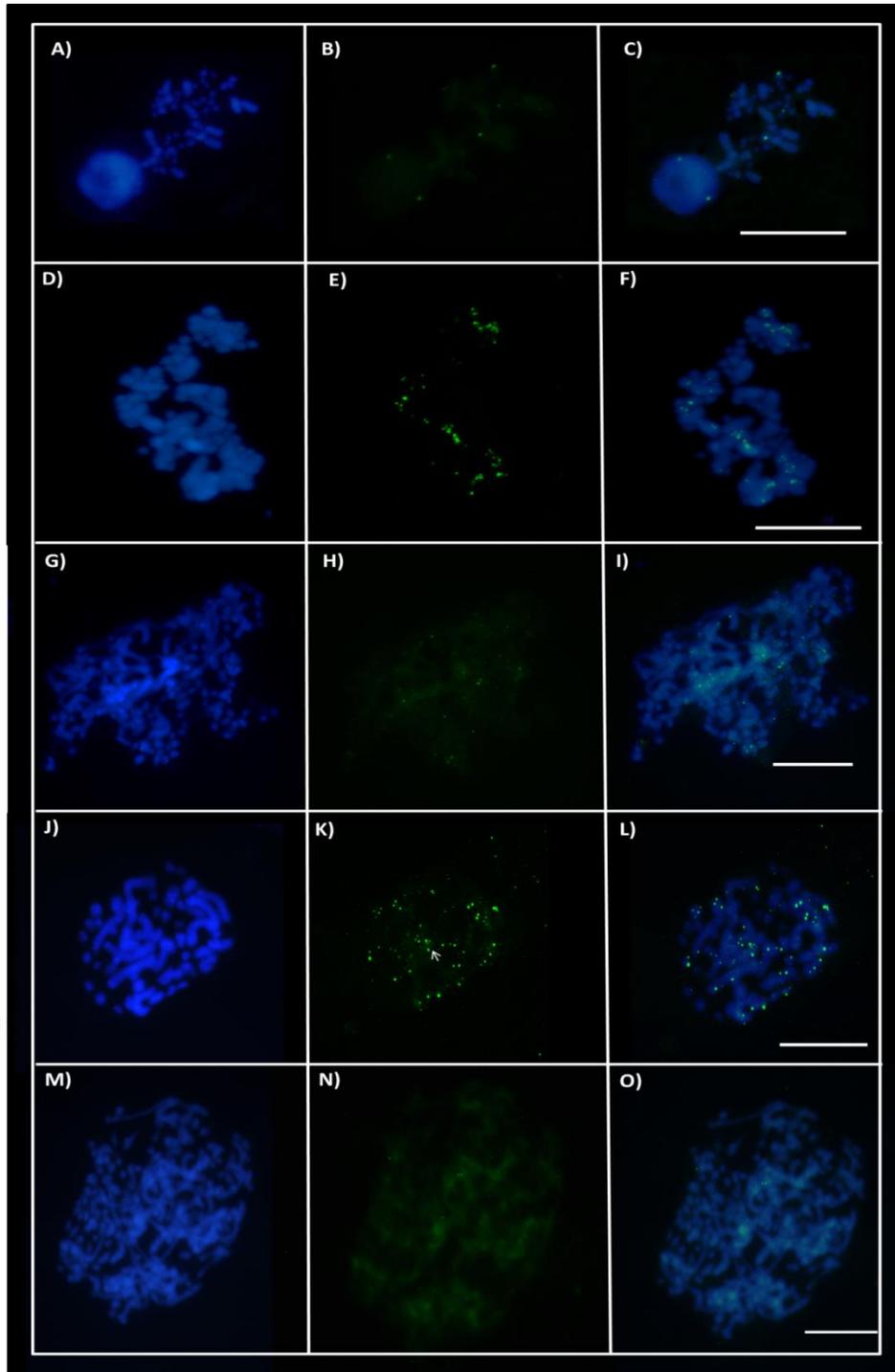


Figura 3.3. Hibridación *in situ* fluorescente en especies de *Agave* con la sonda I (AACHE102-554). Se realizó la localización de AACHE102-554 en cromosomas de *A. tequilana* (A-C), *A. fourcroydes* var. *kitam ki* (D-F), *A. fourcroydes* var. *Yaax ki* (G-I), *A. angustifolia* var. *Marginata* (J-L) y *A. angustifolia* var. *Chelem ki* (M-O). Los cromosomas fueron teñidos con DAPI y la sonda I se identificó con FICT. Escala 10um.

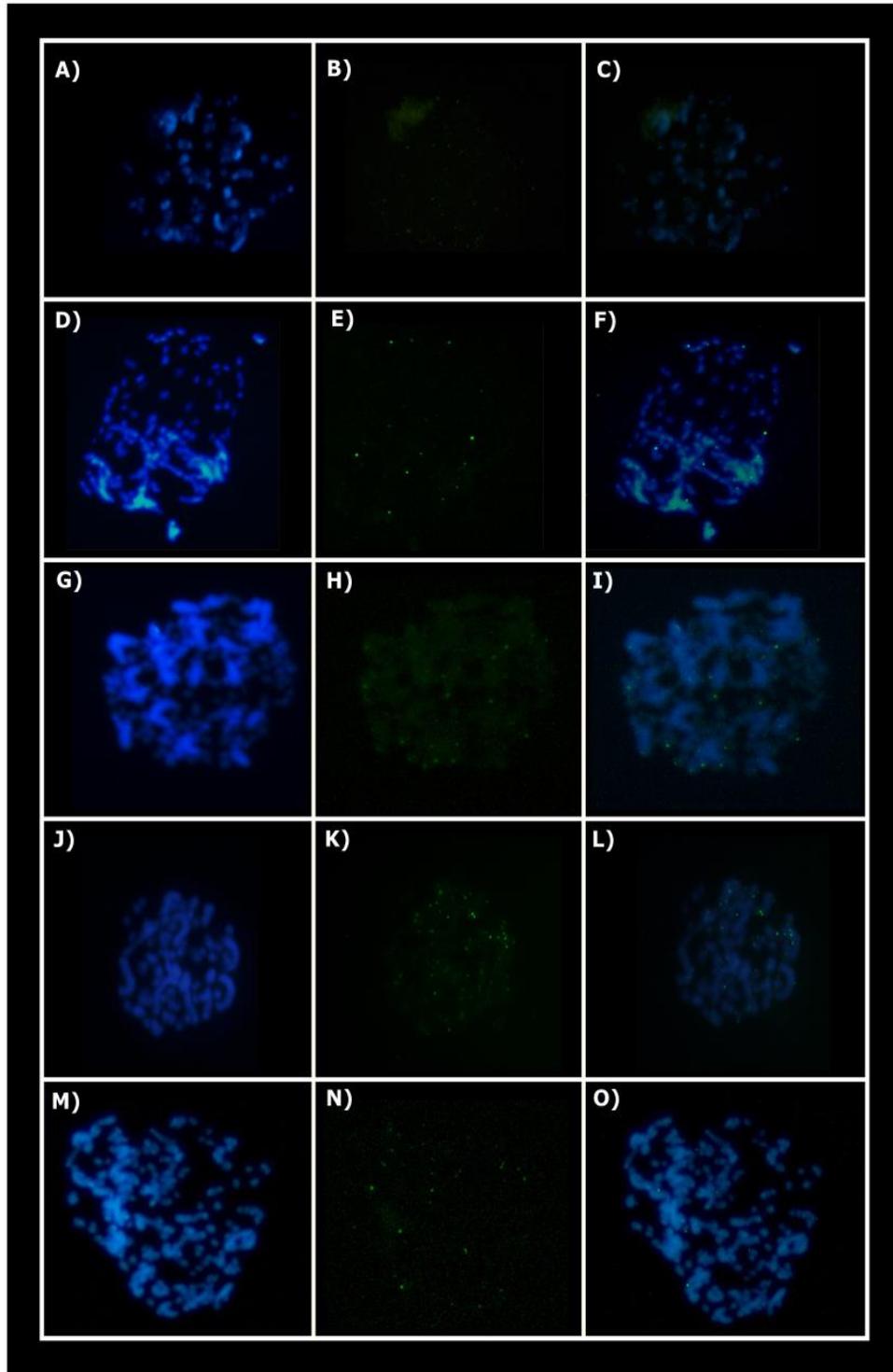


Figura 3.4. Hibridación *in situ* fluorescente en especies de *Agave* con la sonda II (AT102-531). Se realizó la localización de AT102-531 en cromosomas de *A. tequilana* (A-C), *A. fourcroydes* var. Kitam ki (D-F), *A. fourcroydes* var. Yaax ki (G-I), *A. angustifolia* var. Marginata (J-L) y *A. angustifolia* var. Chelem ki (M-O). Los cromosomas fueron teñidos con DAPI y la sonda II se identificó con FICT. Escala 10um.

Tabla 3.1. Identificación de señales en cromosomas de 3 especies de *Agave*. Sonda I AACHE102-554. El cariotipo bimodal de agave consta de cromosomas pequeños y grandes. Las posiciones en las que se encuentran las señales fueron centroméricas (C), subtroméricas (ST) e intersticiales (I).

Especie	Señales	Cromosomas	Posición		
			C	ST	I
<i>A. tequilana</i>	4	Pequeños		X	
<i>A. fourcroydes</i> var. Kitam ki	>20	Pequeños Grandes		X	X
<i>A. fourcroydes</i> var. Yaax ki	10	Pequeños Grandes	X		
<i>A. angustifolia</i> var. Marginata	16	Pequeños Grandes	X	X	X
<i>A. angustifolia</i> var. Chelem ki	12	Pequeños Grandes		X	X

La sonda A (AACHE102-554) proviene de la especie *A. angustifolia* var. Chelem ki y mostró una gran diversidad de señales en distintas regiones de los cromosomas tanto pequeños como grandes de las 3 especies de agave evaluadas. Cabe mencionar que *A. fourcroydes* var. Kitam ki es la que mostró mayor cantidad de señales incluso que en *A. angustifolia* var. Chelem ki que es la especie a la que pertenece la sonda. La región más común en la que se encontró fue en regiones subtroméricas de cromosomas pequeños y grandes.

Tabla 3.2. Identificación de señales en cromosomas de 3 especies de *Agave*. Sonda II AT102-531. El cariotipo bimodal de agave consta de cromosomas pequeños y grandes. Las posiciones en las que se encuentran las señales fueron centroméricas (C), subtroméricas (ST) e intersticiales (I).

Especie	Señales	Cromosomas	Posición		
			C	ST	I
<i>A. tequilana</i>	17	Pequeños Grandes	X	X	X
<i>A. fourcroydes</i> var. Kitam ki	10	Pequeños		X	
<i>A. fourcroydes</i> var. Yaax ki	10	Pequeños Grandes		X	X
<i>A. angustifolia</i> var. Marginata	13	Pequeños grandes	X	X	X
<i>A. angustifolia</i> var. Chelem ki	12	Pequeños Grandes		X	X

La sonda B (AT102-531) proviene de la especie *A. tequilana*, para las 3 especies evaluadas se encontraron al menos 10 señales y la región más común en la que se encontró fue en subteloméricas de cromosomas pequeños y grandes. También es importante mencionar que se encontró en regiones intersticiales a excepción de la especie *A. fourcroydes* var. *Kitam ki*.

Estos resultados tal vez sugieren que se han identificado al menos dos subfamilias de ADN satélite encontradas en especies de la sección *Rigidae* del género *Agave*. Esta información concuerda con la identificación de más de una secuencia de ADN satélite en un solo genoma de planta, pues como se ha reportado en la base de datos del PlantSat (Macas *et al.*, 2002) en el que se enlistan secuencias de ADN satélite en el mismo genoma de distintas plantas con diferencias en los motivos, en el tamaño y en la localización de estas.

3.4. DISCUSIÓN

El ADN satélite está asociado al centrómero y regiones subteloméricas de los cromosomas, en la heterocromatina (Heslop-Harrison, Brandes y Schwarzacher, 2003), y en menor frecuencia ha sido reportado en regiones intersticiales de los cromosomas (Zhao *et al.*, 2011). La posición regiones repetidas en los cromosomas puede sugerir los mecanismos por los cuales se están generando cambios en la organización y las posibles funciones que estas tienen en el genoma. Por ejemplo, la dispersión de señales en los cromosomas puede estar relacionada con la presencia de elementos móviles como los retroelementos. En el caso de *Pisum sativum* se encontró un patrón disperso del ADN satélite PisTR-A en los cromosomas por medio de FISH y los autores observaron que se debe a que está asociado a la región variable UTR del retrotransposon Ogre (Neumann *et al.*, 2001).

3.4.1. Estudios de ADN satélites implicados en los centrómeros

Los resultados obtenidos en este estudio mostraron una baja presencia de ADN satélites en el centrómero de los cromosomas de *Agave*, siendo AT102-531 el satélite en centrómeros de los cromosomas de la especie *A. fourcroydes* var. Kitam ki. En plantas, los centrómeros están compuestos de regiones repetidas como los ADN satélites y los retroelementos (Ma y Jackson, 2006), por esta razón es común encontrar variación en las secuencias que lo componen.

Debido a la variabilidad que puede existir entre los miembros de una familia de ADN satélite se ha evidenciado que pueden ser variables en el centrómero, como en el caso de *Cucumis melo* en el cual se determinó que el satélite centromérico CEM es específico para el cromosoma de esta especie, estando ausente en miembros del mismo subgénero *Melo* (Zhang *et al.*, 2015). A través de un estudio computacional de las secuencias satélites del centrómero para plantas y animales se determinó que estas secuencias son distintas para las especies que divergen en 50 millones de años (Melters *et al.*, 2013). Por el contrario, también se ha observado que las regiones de ADN satélite que se encuentran en el centrómero pueden ser conservadas entre especies cercanamente relacionadas (Dong *et al.*, 1998; Gindullis *et al.*, 2001; Lee *et al.*, 2005).

Los cambios que existen entre los satélites del centrómero son un reflejo del cambio constante a las que están sometidas estas secuencias. En papa (*Solanum tuberosum*), a través del mapeo físico de distintos ADN satélites (st24, st57, st49, st18) se determinó que

estos pueden estar presentes en los centrómeros de cromosomas específicos, como el ADN satélite St24 que está solo en el cromosoma 1 de esta especie; las diversas posiciones que estos ocupan en los cromosomas sugiere que eventos de transposición pueden estar implicados en el origen de ADN satélites nuevos (Gong *et al.*, 2012).

3.4.2. Estudios de ADN satélites implicados en regiones subtelo méricas y teloméricas

Los resultados de FISH para las tres especies de agave estudiadas en este trabajo demuestran que ambas sondas están presentes preferentemente en las regiones subtelo méricas de los cromosomas. Es interesante mencionar que en la especie *A. tequilana* hibridada con la sonda perteneciente a esta misma especie, AT102-531, se encontró una alta presencia de la misma en regiones subtelo méricas de una gran cantidad de cromosomas pequeños y grandes; una gran diferencia de este resultado fue observado al realizar la hibridación in situ de AACHE102-554 en *A. tequilana* pues solo fue posible observarla en cuatro regiones subtelo méricas de cromosomas pequeños.

Al igual que en los centrómeros, en las regiones subtelo méricas pueden estar una gran variedad de familias de ADN satélites (Torres *et al.*, 2011). En un trabajo realizado por Richard *et al.* (2013) describen al ADN satélite khipu, este se encuentra en la región subtelo mérica de los cromosomas y es específico para el género *Phaseolus* (Richard *et al.*, 2013). Posteriormente, esta investigación fue complementada con la caracterización de khipu en otras especies del género y se caracterizó un satélite adyacente a khipu que se denominó jumper que es específico para *P. microcarpus*. Debido a que se encontró una baja representatividad y su ausencia en otros cromosomas de jumper en *P. microcarpus* los autores sugieren que su origen pueda estar dado por un evento de transposición (Ribeiro *et al.*, 2017).

Las sondas AACHE102-531 y AACHE102-554 pertenecen a las especies *A. angustifolia* y *A. tequilana*, respectivamente. Es interesante destacar que a pesar de que ninguna sonda es específicamente de *A. fourcroydes* se observó la presencia de ambas sondas en regiones subtelo méricas e intersticiales de cromosomas grandes y pequeños pero no en todos los cromosomas de las variedades *A. fourcroydes* var. Kitam ki y *A. fourcroydes* var. Yaax ki. Por otro lado, *A. angustifolia* var. Marginata mostró una alta presencia de ambas sondas en regiones subtelo méricas, incluso la AT102-531 perteneciente a la especie *A. tequilana*.

La región subtelomérica de los cromosomas puede cambiar por los rearrreglos que surgen en las regiones terminales de estos, los cambios en estas regiones en los cromosomas no homólogos puede ser un mecanismo que está generando la aparición o pérdida de distintos miembros de una familia de ADN satélite o incluso nuevas familias (Kishii *et al.*, 1999; Koukalová *et al.*, 2010; Macas *et al.*, 2006; Trofimova *et al.*, 2015). Se han descrito ADN satélites en regiones subteloméricas de las especies papa (*Solanum tuberosum*) (Gong *et al.*, 2012; Torres *et al.*, 2011) cebolla (*Allium fistulosum*) (Kirov *et al.*, 2017), centeno (*Secale cereale*) (Salina *et al.*, 2009), plátano (*Musa spp.*) (Čížková *et al.*, 2013), frijol (*Phaseolus vulgaris*) (Richard *et al.*, 2013). En familias de ADN satélite subteloméricas incluso se han identificado satélites teloméricos que están adyacentes a los telómeros y se refieren a ellos como secuencias asociadas al telómero (TASs telomere-associated sequences) (Louis y Vershinin, 2005).

3.4.3 Análisis citogenético comparativo de ADN satélites en plantas cercanamente relacionadas

En este estudio se logró identificar la presencia de al menos dos ADN satélites en tres especies de *Agave* cercanamente relacionadas. *A. angustifolia*, *A. fourcroydes* y *A. tequilana* contienen a las sondas AACHE102-554 y AT102-531 en regiones subteloméricas de cromosomas pequeños y grandes. Las distintas posiciones en las que se ubicaron sugieren la conservación de estas secuencias a pesar de que se encuentran en distinta cantidad de señales o posición en el cromosoma para estas especies. Sin embargo, para poder concluir este análisis es necesaria la ubicación de marcadores cromosómicos en agaves pues hasta el momento solo se conoce la posición de los genes ribosomales 45S y 18S (Gomes-Rodríguez *et al.*, 2013; Robert *et al.*, 2008).

Algunos ADN satélites en plantas han sido empleados como marcadores cromosómicos, tal es el caso de Tail, que se encuentra en una sola posición en un solo cromosoma de la especie *Leymus racemosus* (Neumann *et al.*, 2001). En el género *Cucumis* se ha logrado inferir la relación filogenética de tres variedades de la especie *Cucumis sativus* a través de la localización de ADN satélites (Zhang *et al.*, 2011).

Las repeticiones en tándem HAT58 y CAT36 fueron probadas en el híbrido *Allium wakegi* (*A. fistulosum* X *A. cepa*) y se determinó que ambas son específicas de *A. fistulosum* pues estuvieron ausente en la proporción genética de *A. cepa* de este híbrido (Kirov *et al.*, 2017).

El género *Nicotiana* es uno de los mejores modelos que existe para el estudio de regiones repetidas, pues este cuenta con especies híbridas, con distintas secciones taxonómicas, un extenso estudio en sus cariotipos, también cuentan con la caracterización molecular y citogenética de regiones repetidas con el objetivo de encontrar los posibles mecanismos que dan origen al ADN satélite en especies poliploides (Clarkson *et al.*, 2004; Lim *et al.*, 2004; Lim *et al.*, 2000).



Identification and characterization of a new satellite-like DNA family in three *Agave* species



Espinosa-Barrera L.A.^a, Sánchez-Teyer L.F.^a, Quiroz-Moreno A.^a, Narváez-Zapata J.A.^{b,*}

^a Centro de Investigación Científica de Yucatán, A.C. Calle 43 No. 130, Col. Chuburná de Hidalgo, C.P. 97200, Mérida, Yucatán, Mexico

^b Instituto Politécnico Nacional. Centro de Biotecnología Genómica. Boulevard del Maestro s/n esq. Elías Piña, Col. Narciso Mendoza, C.P. 88710 Cd., Reynosa, Tamaulipas, Mexico

ARTICLE INFO

Keywords:

DNA satellite
Agave
Short motifs
FISH
Transposable elements

ABSTRACT

Repeated DNA analysis was conducted in three *Agave* species (*A. angustifolia* Haworth, *A. tequilana* Weber and *A. fourcroydes* Lemaire). Self-priming PCR was initially used to generate repeated PCR products and their sequence information was then used to build different primers sets. The ATsat102-113 primer set was able to amplify repeated PCR products, including probably multimers without becoming a typical ladder profile, in all *Agave* accessions analyzed. Sequence analysis of these amplifications identified two subfamilies of repeated DNA which share six conserved short motifs but exhibit a high genetic distance and different sequence traits between them. Consensus monomers obtained have a length higher than 100 bp and are rich in Adenine/Thymine percentage, as many satellites of plants. A representative probe of these consensus monomers was able to hybridize in some PCR amplified products and the more frequent short motif 1 (CTYRAGTTSG) also exhibits probably multimers without becoming a typical ladder profile. These *Agave* satellite-like subfamilies have different homogenization rates and the subfamily 2 is more conserved. These satellite-like monomers are transcribed and exhibited a strong relation with transposable element signatures of the LTR/Gypsy families and with *A. tequilana* TSA sequences, some which exhibits transposon-related domains in different locations to the satellite-like sequences identified. FISH analysis of these satellites-like subfamilies only may be reliable determined in the diploid (*A. angustifolia* Haworth var. Marginata and *A. tequilana* Weber 'Azul') accessions. In these generally showed a subtelomeric hybridization with less interstitial signals.

1. Introduction

Sequence motifs that are repeated hundreds or thousands of times in the nuclear genome are an abundant part of all eukaryotic genomes. These repeated sequences may be organized as satellite DNA regions (Biscotti et al., 2015b). These are non-coding sequences with highly repetitive monomers (150–180 bp or 300–350 bp) mainly arrayed in tandem (Plohl et al., 2008; Garrido-Ramos, 2015). Usually, long arrays of satellite DNA are composed of tens to thousands of tandem copies that occur in centromeric or subtelomeric regions, and less commonly in interstitial regions (Torres et al., 2011; Plohl et al., 2012). Plant satellite sequences also showed the presence of short motifs as direct and inverted repeats, and palindromes (Mehrotra et al., 2014). These sequences have a high Adenine/Thymine percentage (AT.%) and are located in heterochromatin regions, although in some cases may also be transcribed (Ugarković and Plohl, 2002; Ugarkovic, 2005; Biscotti et al., 2015a). Furthermore, several studies support the relationship

between these DNA regions and the Transposable Elements (TEs), particularly for the presence of short TE signatures, although they differed in structure, genome organization, spreading mechanism and evolutionary dynamics (Macas et al., 2009; Meštrović et al., 2015). The origin of satellite DNA can be explained using different proposed mechanisms such as unequal crossover, genic conversion, circle rolling replication and reinsertion, and transposition (Tek et al., 2005; Gong et al., 2012; Biscotti et al., 2015b). These mechanisms can generate changes in sequence composition and in nucleotide rearrangement (Hemleben et al., 2007). These changes may increase the number of motifs giving rise to different satellite DNA families (Yoong Lim et al., 2006; Plohl and Chaves, 2014). Families of satellite DNA could be conserved in different taxonomic groups (Quesada Del Bosque et al., 2014). They also exhibited variability in sequences (Lim et al., 2004; Matyasek et al., 2011; Melters et al., 2013), unit length ranges (Plohl et al., 2008), abundances (Garrido-Ramos, 2015) and could occur in different chromosome locations (Gong et al., 2012). Sequence

* Corresponding author.

E-mail address: jnarvaez@ipn.mx (J.A. Narváez-Zapata).

<https://doi.org/10.1016/j.plgene.2018.08.002>

Received 12 January 2018; Received in revised form 7 August 2018; Accepted 10 August 2018

Available online 14 August 2018

2352-4073/ © 2018 Elsevier B.V. All rights reserved.

variability of these different plant satellite families has been used to conduct evolutionary and speciation studies in *Nicotiana* (Pires et al., 2004; Koukalova et al., 2010), *Arabidopsis* (Martinez-Zapater et al., 1986; Heslop-Harrison et al., 2003), *Musa* (Čížková et al., 2013), *Triticum* (Salina et al., 1998; Kishii and Tsujimoto, 2002) and *Solanum* (Tek et al., 2005).

The genus *Agave* contains > 200 species with different ploidy levels (Castorena-Sánchez et al., 1991; Good-Avila et al., 2006; Moreno-Salazar et al., 2007; Palomino et al., 2008) and high genomic variability (Gil-Vega et al., 2006; Sánchez-Teyer et al., 2009). This variability may be explained as a consequence of allopolyploidy (McKain et al., 2012), and presence of TEs (Khaliq et al., 2012). These last are highly abundant in the genome and transcriptome of these plants (Khaliq et al., 2012; Gross et al., 2013). Therefore, in this study, we carried out for the first time the isolation and characterization of new satellite-like DNA family in *Agave* plants, which could provide valuable information for evolutionary and speciation studies in this important plant genus.

2. Materials and methods

2.1. Plant materials

Five commercial *Agave* L. accessions were selected; *A. tequilana* Weber, 1902 'Azul' (AT), two varieties of *A. fourcroydes* Lemaire, 1864 classified as 'Kitam ki' (AFKK) and 'Yaax ki' (AFYK), and two varieties of *A. angustifolia* Haworth, 1915 classified as 'Marginata' (AAMAR) and 'Chelem ki' (AACHE). *A. tequilana* Weber 'Azul' has been propagated in vitro and adapted to greenhouse conditions. All the accessions form part of the botanical garden "Roger Orellana" of the Research Center of Yucatan (CICY), México.

2.2. Isolation of repeated sequences

A self-priming PCR was initially conducted using digested DNA of *A. tequilana*. Genomic DNA was isolated from leaves tissues by the CTAB protocol (Doyle, 1991). DNA integrity was verified by electrophoresis in agarose gel at 1% and the DNA purity was measured for spectrophotometry with a ratio of 260/280 nm using the NANODROP 1000 equipment (Thermo Scientific, USA). Different restriction enzymes (*Mse*I, *Hind*III and *Taq*I) were assayed to digest the genomic DNA. PCR products were run on acrylamide gel (6%) for eight hours. Finally, it was selected the self-priming PCR generated by using the *Taq*I since this enzyme produce discrete and abundant DNA fragments (200, 450 and 700 bp) on acrylamide gels. Then, a direct PCR cloning was conducted by using the pGEM-T Easy Vector Systems (PROMEGA, USA) and six positive clones were selected to further sequencing using the MACROGENE (<http://www.macrogen.com/>) services. Sequences obtained (codes; 101, 102, 105, 109, 110 and 113) were analyzed using the BioEdit program (Hall, 1999). Specific primer sets for each sequence were designed by using the OligoAnalyzer Tool 3.1 (<http://www.idtdna.com/calc/analyzer>) program. These primer sets were; ATsat-101 Forward 5'-GCAAATACAGATCATATCATA-3' and ATsat-101 Reverse 5'-CTTATACAGCGAAAGTCAT-3'; ATsat-102-113 Forward 5'-ACTAGTTGATGAGTCTGA-3' and ATsat-102-113 Reverse 5'-CGAGCAGCCAGCCG-3'; ATsat-109 Forward 5'-AGACATGACTCGCTTCCC-3' and ATsat-109 Reverse 5'-TCAGACAAACAAGTTACCC-3'; and ATsat-110 Forward 5'-TAGATACAAACATAATGCGT-3' and ATsat-110 Reverse 5'-CTCTCACAGGAGCGA-3'.

Genomic DNA was isolated from all *Agave* accessions as was previously described. Then, to amplify repeated DNA was done a reaction mixture that contained 50 ng of genomic DNA, 5 pmol of each primer set, 3 mM of MgCl₂, 0.2 mM of dNTP's, 1 × reaction buffer, and 0.5 U of DNA Taq polymerase, and adjusted to 25 µl with ultrapure water. PCR conditions were firstly a denaturation step of 5 min at 95 °C follows of 25 cycles of 1 min to 95 °C, 30 s at 49 °C and 1 min at 72 °C elongation, and then a last elongation step of 5 min at 72 °C. PCR products were

visualized in agarose gel (1.5%), cut from the gel and purified by using the Wizard® Plus SV Minipreps DNA Purification System (PROMEGA, USA). Purified fragments were cloned in the pGEM-T Easy Vector Systems (PROMEGA, USA) and send to MACROGENE (<http://www.macrogen.com/>) for sequencing. Ten amplicons of each different *Agave* accession were sequenced. Most of these sequences ranged from 100 to 200 bp.

2.3. Sequence classification

A database containing all cloned sequences for each primer sets and for each *Agave* accession was built. A code was designed to each sequence, with the abbreviation of the *Agave* accession followed by the number of the primer set, and lastly the clone number. All the sequences were aligned and a distance radial tree was generated by using the MEGA7 program (Kumar et al., 2016). This tree was inferred by using the Minimum Evolution method. The evolutionary distances were computed using the p-distance method. The Neighbor-joining algorithm was used to generate the initial tree. A bootstrap test of 1000 replicates also was implemented. Contig determination of these sequences was realized by the Codon code aligner program (<http://www.codoncode.com/aligner/>). This determination includes the sequence length, the AT.% and the restriction enzyme map for each contig. Finally, the identification of the consensus sequences was conducted from more abundant contigs by using the CLS sequence viewer program ver 7.7.1 (<https://www.qiagenbioinformatics.com/products/clc-sequence-viewer>). Homogenization and mutation analysis of the sequences belonging to these selected clades was conducted by using the haplotype networks analysis. For this, specific sequences within of each clade, identified in the previous radial tree, were trimmed at 110 bp using the Bioedit program (Hall, 1999). Then, the Network program ver. 5 was used (<https://www.fluxus-engineering.com>; Bandelt et al. 1999) assuming that epsilon is 0 and rate of transversion/transitions of 1:2. The network was used to visualize the specific mutations sites for each clade.

2.4. Repeated regions analysis

Two analysis were conducted to detect the repeated behavior of these sequences. First one; repetitive motifs identification was conducted by using the MEME program, version 4.2.1 with a maximum of 6 motifs, and a range of 6 to 10 bp (Bailey and Elkan, 1994). Weblogo motif signatures were identified and a motif map was built on aligned sequences of each main contig. More abundant motif, considering its occurrence in all sequences grouped within of a specific contig, was selected to design a new specific PCR primer set to test if these motifs exhibited a repeated pattern. This motif-1 specific primer was: 5'-ATGCCAATTCAGACCT-3' which was amplified using the reverse ATsat-102-113 primer and the same PCR conditions previously described. Second one; Southern blot hybridization was performed on ATsat102-113 amplification products of the different *Agave* accessions. PCR products were electrophoresed on a 1% agarose gel and blotted on a membrane Hybond-N+ (Amersham, GE Healthcare) by capillary transfer. Then, membrane was hybridized with the AACHE102-554 (subfamily 1) and AT102-531 (subfamily 2) probes. Previously, this was labeled with Digoxigenine-11-dUTP (Roche, USA) and incubated at 42 °C overnight. Hybridization and post-hybridization washes were conducted as recommend the manufacture (DIG-easy-hyb; Roche USA). Finally, membrane was incubated with CSPD substrate (Roche, USA) and exposed to a Hyperfilm ECL (Amersham, GE Healthcare) for 20 min and then revealed.

2.5. Expression analysis

Two analysis were conducted to detect the expression of these repeated sequences. First one; representative sequences of *A. tequilana* of

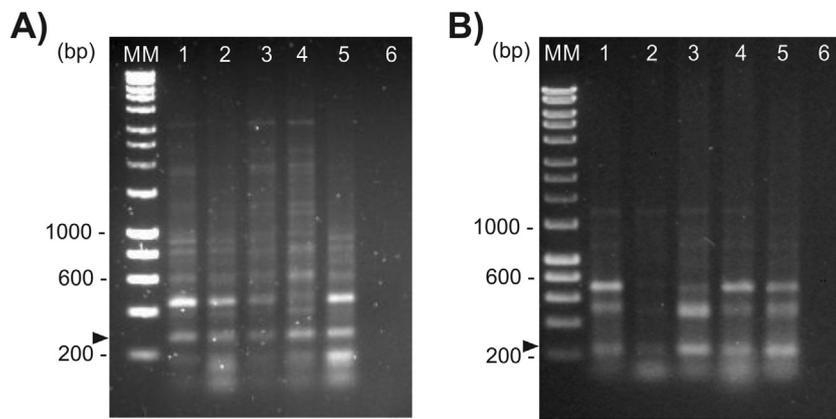


Fig. 1. DNA amplification by using the ATsat-102-113 (A) and ATsat-109 (B) primers sets Amplification from genomic DNA of *A. tequilana* var. Azul (1), *A. fourcroydes* var. kitam ki (2), *A. fourcroydes* var. yaax ki (3), *A. angustifolia* var. chelem ki (4), *A. angustifolia* var. Marginata (5) and negative control without DNA (6). Arrow shows the cut bands and the molecular marker (MM) of 1000 bp used.

the more abundant contigs were selected to be blasted in GenBank database, specifically using Transcriptomic Shotgun Assembly (TSA) of *A. tequilana* (taxid: 386106) with the aim to identify if these are probably transcribed. In addition, the presence of conserved domains on the TSA sequences that exhibit some class of identity with these representative sequences was analyzed in order to identify some relation with coding regions. Conserved domains were analyzed according to (Marchler-Bauer et al. (2017) by using the tool of search in the Conserved domain database (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Second one; PCR amplifications by using the previous described conditions with the At-102-113 and motif-1 primer sets were conducted on cDNA obtained from leaves of *A. tequilana*. Total RNA was extracted by using the Trizol® (Thermo Fisher) and the cDNA synthesis by using the Reverse transcription System (Promega, USA). 18S rDNA amplification by using the primers and conditions describes by Tamayo-Ordóñez et al. (2015) was used as positive control. Finally, PCR products were visualized by electrophoresis in agarose gel at 1%.

2.6. Relation with transposable elements

The relationship of the representative *A. tequilana* sequences with plant TEs was analyzed, beginning with the specific relation of these repeated regions with *A. tequilana* TEs. Firstly, the TSA dataset of *A. tequilana* (GenBank: GAHU00000000) was retrieved and TEs were identified by using the RepeatMasker (ver. 3.2.9) with the *Zea mays* repeat models present in Repbase Update v 17.08 as described (Gross et al., 2013). Then, identified TEs were compared with the experimental repeated sequences obtained in this study to identify the possible match between both analyses. It was used the gene structure display server (<http://gsds.cbi.pku.edu.cn/>) to establish the organization of the satellite-like domain with respect to the other transposon-related domains on these *A. tequilana* TSA sequences. Second one; representative *A. tequilana* DNA sequences were also blasted with the Gypsy database (http://gydb.org/index.php/Main_Page), specifically using the LTR database, in order to identify TEs signatures (Llorens et al., 2011).

2.7. Hybridization in situ fluorescent

Root tips from *A. tequilana*, *A. fourcroydes* and *A. angustifolia* were treated with a saturated paraformaldehyde solution for 3 h. After that, root tips were fixed in ethanol:acetic acid (3:1) for 24 h and then, digested with Visozyme® (SIGMA) for 30 min at 37 °C. Squash preparations were made in a drop of 45% acetic acid and frozen in liquid nitrogen. The cover slips were removed with a razor blade and the slides were incubated for 15 h to 42 °C. For localization analysis we selected two different probes; AACHE102-554 and AT102-531. Hybridization mixture consisting of 50% formamide, 10% dextran sulfate in 1 × SSC

and 1 µg/ml labeled probe was added onto slides and denatured at 75 °C for 2 min. The hybridization was carried out at 42 °C overnight. The sites of probe hybridization were detected using anti-digoxigenin-FITC (Roche Applied Science) and the chromosomes were counterstained with DAPI. The slides were examined on a microscope Axioplan I.

3. Results

3.1. Isolation of satellite sequences

Genomic DNA obtained from *A. tequilana* was digested with different restriction enzymes, and a self-priming was conducted. It was selected the *TaqI* enzyme to carried out this analysis since *TaqI* produces several discrete bands (~250, 450, and 700 bp) on acrylamide gel (gel not shown). These amplicons were cloned, sequenced and finally, six different PCR products with different length size (198, 312, 263, 206 and 180 bp), corresponding codes as 101, 102, 105, 109, 110 and 113 respectively, were obtained. A sequence analysis was conducted on these six PCR products. Then, specific PCR primers, trying to cover the total sequence length, were designed from these PCR products, and only the ATsat-102-113 and ATsat-109 primer set were able to generate discrete bands in all *Agave* accessions analyzed (Fig. 1). PCR products display discrete bands, although become a typical ladder profile. Specifically, ATsat-102-113 primer set generates discrete bands (~100, 200, 500, 600, 800 and 1000 bp) in all *Agave* accessions. The more abundant band of ~200 bp was cut in all *Agave* accessions for further analysis.

3.2. Sequences analysis

These PCR fragments (~200 bp) were cloned, and at least ten clones were selected to sequence by each *Agave* accession and primer set. Sequences obtained have a wide range of sizes (44 to 412 bp). After edition process only the sequences higher of 90 bp were included to further analysis. Information of these sequences was used to build a distance radial tree by using the minimal evolution method (Fig. 2). This analysis allows discriminating two specific clades (1 and 2), particularly, for some ATsat-102-113 primer related sequences although most of them ungroup. Sequences of these clades belonged to all *Agave* accessions analyzed (GenBank accessions; MG029581 to MG029613). A more detailed analysis of these sequences by using the CodonCode Aligner program allows grouping many (45.5%) of them in 7 contigs according to its size, alignment, TA % content and its restriction map (Table 1). Particularly, the more abundant contig 1 (19 sequences) and 2 (15 sequences) correspond to clade 1 and 2 of the radial distance tree, respectively. All the ATsat-109-primer related sequences were ungrouped and, therefore, they were left for further analysis. Clade 1 and 2 exhibits a size of 165 and 107 bp, respectively. In addition, sequences in both clades have a high percentage (> 50) of TA.

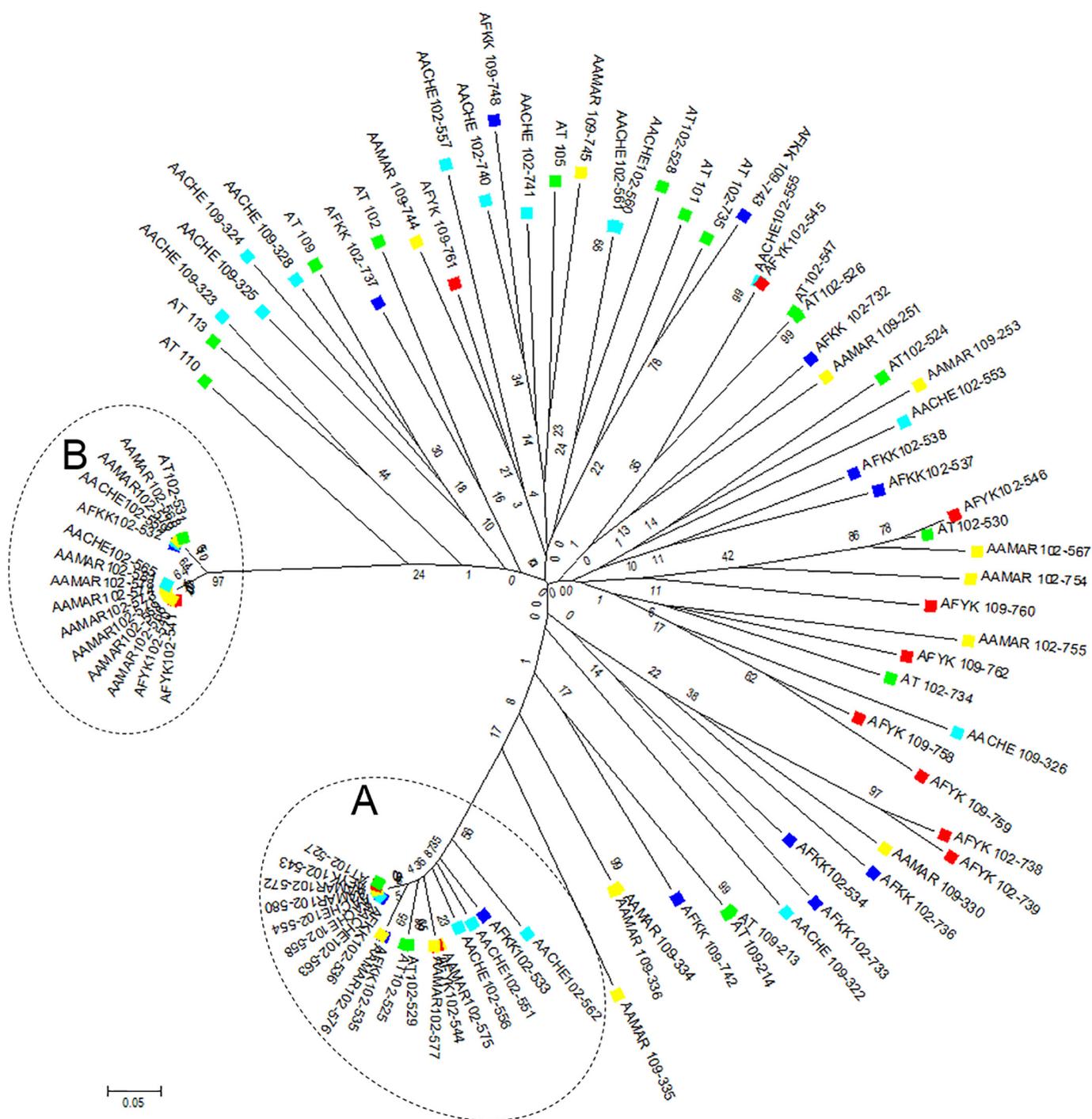


Fig. 2. Evolutionary relationships of amplified sequences. Tree was inferred using the Minimum Evolution method. The evolutionary distances were computed using the p-distance method. The Neighbor-joining algorithm was used to generate the initial tree. Circles A and B showed the sequences within of the clades 1 and 2 used to further analysis. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016). Agave color code is green: AT, blue: AACHE, yellow: AAMAR, red: AFYK, violet: AFKK indicates to the different accessions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Identification of repeated regions

The clade 1 and 2 are conserved and their sequences exhibit traits associates with satellite DNA of plants, such as its sequence length and AT%. We obtain consensus sequences even though its sequences are originated from different *Agave* accessions. CLS program was chosen to make the alignment for this consensus sequences (Fig. 3A and B). Homogenization of both clades was determined by using the haplotype network analysis and it was observed only 10 mutation in the clade 2

regarding found in the clade 1 (data not shown). Therefore, the clade 2 exhibits a higher homogenization tendency in these *Agave* accessions. Repetitive motif analysis was conducted by using MEME program. This program was selected since it allows searching repeated and ungrouped sequence patterns in a sequence database (Bailey et al., 2006). Sequences of the more abundant clades exhibited a length of 165 or 107 bp. Therefore, a search of maximum 6 motifs with a range of 6 to 10 bp was conducted (Table S1; Supplementary material 2). This search discovered 6 motifs distributed in 6 contigs (Contig 1, 2, 3, 4, 6, and 7)

Table 1
Contig assembling and MEME motif analysis of the sequences obtained.

Contig ^a	Accession occurrence	Evolutionary analysis ^b	Number	% ^c	Size (pb)	TA %	MEME motif analysis ^d						Restriction map ^e
							-1	-2	-3	-4	-5	-7	
1	AACHE, AAMAR, AFYK, AFKK, AT	Clade I	19	41.3	165	62	100	89	100	89	79	95	<i>BstXI</i> (77)
2	AACHE, AAMAR, AFYK, AFKK, AT	Clade II	15	32.6	107	51.4	100	100	93	100	100	93	<i>XhoI</i> (59)
3	AACHE, AAMAR, AFYK, AT	Ungroup	4	8.7	441	57.6	0	100	0	0	25	0	<i>Acc65I</i> (70), <i>EcoRV</i> (75), <i>KpnI</i> (189)
4	AACHE, AFYK	Ungroup	2	4.35	199	40.8	100	0	0	0	0	0	<i>HincII</i> (144, 168), <i>Sall</i> (142, 166), <i>XmiI</i> (143,167)
5	AT	Ungroup	2	4.35	188	42	0	0	0	0	0	0	<i>HindIII</i> (178)
6	AACHE	Ungroup	2	4.35	94	52.1	0	0	100	0	0	100	<i>XbaI</i> (67)
7	AFYK, AFKK	Ungroup	2	4.35	46	45.7	0	0	0	100	0	0	<i>HincII</i> (32)

^a According to CodonCode Aligner program (www.codoncode.com/index.htm).

^b According to radial tree (Fig. 2).

^c According to total sequence number.

^d MEME Motif (Meme-suite.org) (Motif length 10 pb).

^e Analyzed enzymes (*Acc65I*, *ApaI*, *BamHI*, *Bsp68I*, *EcoRI*, *EcoRV*, *HincII*, *HindIII*, *KpnI*, *NotI*, *PaeI*, *PstI*, *SacI*, *SacII*, *Sall*, *SmaI*, *XbaI*, *XhoI*, *XmiI*).

previously obtained by the CodonCode Aligner program, but only the contigs 1 and 2 (clade 1 and 2) have all the discovered motifs and exhibit low e-values (range 5×10^{-5} and 9.3×10^{-13}) although in some cases these were found in the reverse complementary sequence (Fig. 3). Contig 3, 4, 6 and 7 also showed only one or two motifs with high e-value (0.016 to 7.7) and therefore were discarded for further analysis.

Most of the sequences belonging to clades 1 and 2 have the 6 motifs, being the more abundant the motif 1 (100%) (Table 1). Distribution of these motifs on sequences of the clades 1 and 2 is shown in Fig. 3. In addition, the weblogo signatures corresponding to these motifs are also shown along of the sequences. Sense weblogo sequence of these 6 motifs are as follows: Motif 1; CTYRAGTTSG, Motif 2; TRGATAASTC, Motif 3; GRAATCARAG, Motif 4; AAAACCMMAA, Motif 5; GSAAASATGG and Motif 6; CAAMWRMAAG. Motif 1 appears in all sequences of these two repeated DNA groups both in sense as in reverse complementary direction even though not exhibit the same localization (Fig. 3). Taken together these results, specifically the presence of repeated multimers without becoming a typical ladder profile and the presence in these consensus sequences of multiple shorts motifs with different localization, it is possibly suggested that the clades 1 and 2 represent related sequences or subfamilies of the same repeated DNA region.

To explore with more detailed the abundance of these repeated regions on *Agave* DNA, it was used the sequence information of the motif 1 (CTYRAGTTSG) to build a new specific PCR primer set to detect its presence along the genomic DNA in the different *Agave* accessions (Fig. SA; Supplementary material 1). PCR products obtained by using this forward primer and the original ATsat102-113 reverse primer generated a ladder pattern on an agarose gel. This repeated pattern exhibits conserved bands in all *Agave* accessions that started in approximately 110 bp, although without becoming a typical ladder profile. In a more detailed analysis, it was conducted a Southern blot hybridization by using the representative AT102-531 probe (subfamily 2) on original ATsat-102-113 primer set amplification in all the *Agave* accessions. This showed a clear signal in the most abundant band (~200 bp) and weakly signals in others PCR products. This signal on this single band suggests a high conservation of this satellite-like region (subfamily 2) and more variation on the other PCR products amplified. However, when the AACHE102-554 probe (subfamily 1) was used, clear signals were detected in the more abundant PCR products (~200, ~600, ~800, ~1200, ~1800) suggested the presence of multimers highly variables of these repeated DNA regions.

3.4. Expression analysis and transposable element analysis

Specific primer sets (ATsat-102-113 and Motif 1) were used to amplify cDNA from *A. tequilana* leaves that allowed obtaining a discrete band of 150 bp, a similar length to consensus sequence characterized. However, these primers sets on cDNA were not able to amplify other longer discrete bands or a ladder PCR products profile (Fig. 4A). PCR amplification on cDNA support the transcriptional presence of these repeated DNA regions although only of a single copy. To obtain information on this transcriptional behavior, the representative sequences of *A. tequilana* (AT102-529 and AT102-531) belonging to subfamilies 1 and 2 were chosen to run a local BLAST search against the *A. tequilana* TSA database. A total of 11 and 17 TSA accessions exhibited some degree of identity (86–97%) with the representative AT102-529 and AT102-531 sequences, respectively (Table 2). Only, the TSA accessions associated with the AT102-531 sequence (subfamily 2) showed some conserved domains, particularly for the retrotransposon gag protein and RNase_H like super family, which suggests a probable relation of this repeated DNA region with some TEs. In fact, some of these related TSA were also identified as a probably TEs by using the RepeatMasker program on all *A. tequilana* TSA database. Identification of repeat class by using the *Zea mays* repeat models showed identity in these related TSA accessions with the LTR/Gypsy (Gypsy65-ZM_I-int and Gypsy-71-I_ZM families). However, this analysis is limited since only considered the *Zea mays* repeat models and no other Liliopsida plants. Therefore, an additional analysis considering all LTR/Gypsy sequences was conducted by using the Gypsy database (Llorens et al., 2011). The result showed a high amount of TEs in these short (AT102-529 and AT102-531) and representative sequences. These motifs are Cereba, GypsySL, CopiaSL and Ogre, among others (Table S2). The presence of a single band in the RT-PCR analysis was confirmed by identify only a single match on closest *A. tequilana* TSA (Fig. 4B). The localization of these TEs motifs varies between *A. tequilana* sequences, similarly to the MEME motifs previously identified although in any case are overlapping with other transposable domains identified. Identification of repeated and transposable element motifs associated to these sequences suggests an evolutionary relationship of these repeated DNA sequences with these TEs.

3.5. Physical mapping of satellite sequences

Physical mapping of AACHE102-554 (subfamily 1) and AT102-531 (subfamily 2) probes were built for FISH analysis on these *Agave* accessions. FISH results showed a different number and signal behavior

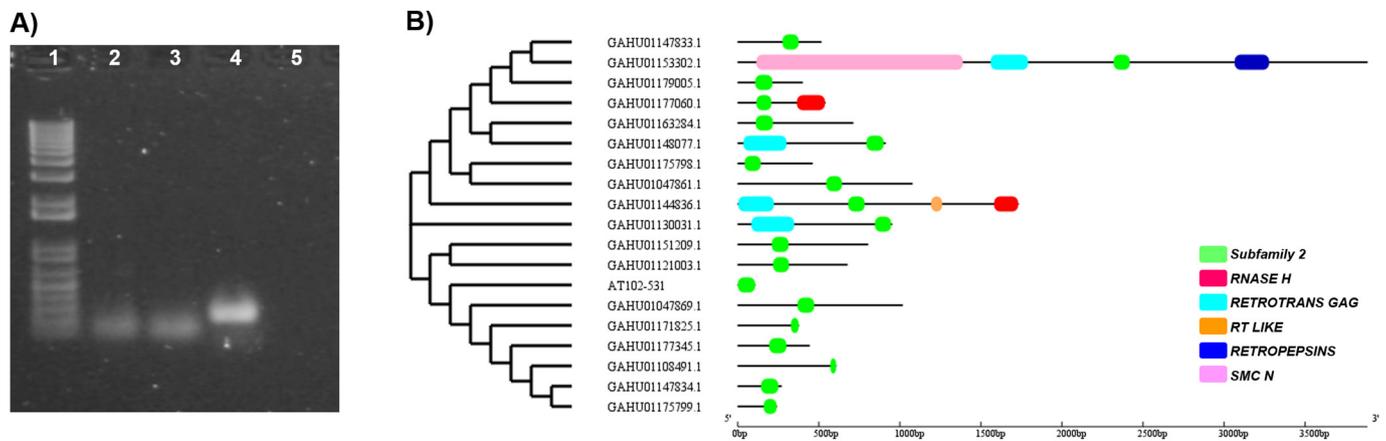


Fig. 4. Expression analysis of satellite-like regions by direct expression on *A. tequilana* cDNA and by Blast analysis on *A. tequilana* TSA sequences. **A)** PCR amplification by using the ATsat-102-113 and Motif 1 primers set on *A. tequilana* cDNA. Lane 1 (WM 100 bp), lane 2 (ATsat-102-113), lane 3 (Motif 1), lane 4 (18S rDNA), lane 5 (negative control). **B)** Satellite-like domain (subfamily 2) and transposon-related domain organization on closest *A. tequilana* TSA sequences. Tree inferred by using the Minimum Evolution method. Identified transposon and satellite-like domains are showed with different color codes.

characterization of a repeated DNA region in three *Agave* species. DNA amplification showed the presence of similar fragments (~100, 200, 500, 600, 800 and 1000 bp) without becoming a ladder profile in different species of *Agave*, of which the 200 bp fragment is the most abundant. These results are consistent with the satellite DNA monomer abundance, previously detected as thousands of copies per genome in

plants (Plohl et al., 2008). A further characterization (in at least 10 clones of each *Agave* accession) of the 200 pb fragment yields two related repeated DNA regions (grouped in the clades 1 and 2) in all *Agave* accessions (Fig. 2; Table 1). The consensus sequences of these DNA regions are amplified with the same primer set but exhibit a high genetic distance and different sequence traits, and therefore in this study

Table 2
Expression analysis and TEs identification of the representative *A. tequilana* sequences of subfamilies 1 and 2.

Probes	Max score	e-value	Identity (%)	Accession ^a	Size (bp)	Alignment length	Match	Start	End	CDS search ^b	
AT102-529	195	3.9E-53	95.2	GAHU01132883.1	613	124	1	1	124	N.d.	
(Subfamily 1)	195	3.9E-53	95.2	GAHU01132881.1	535	124	1	1	124	N.d.	
	195	3.9E-53	91.5	GAHU01109566.1 ^c	3149	142	1	2251	2392	N.d.	
	193	1.4E-52	91.6	GAHU01004324.1 ^d	1904	143	1	249	389	N.d.	
	187	6.6E-51	90.8	GAHU01134683.1	475	142	1	141	280	N.d.	
	187	6.6E-51	90.8	GAHU01134682.1	665	142	1	169	308	N.d.	
	187	6.6E-51	90.8	GAHU01134681.1	580	142	1	131	270	N.d.	
	180	1.1E-48	89.6	GAHU01136532.1	260	144	1	110	253	N.d.	
	163	1.1E-43	87.5	GAHU01111957.1	2279	144	1	888	1031	N.d.	
	161	4.0E-43	88.3	GAHU01135813.1 ^f	4397	137	1	494	630	N.d.	
	156	1.9E-41	88.7	GAHU01145357.1	461	133	1	1	130	N.d.	
	145	4E-38	91.6	GAHU01081317.1	659	107	1	15	121	N.d.	
	(Subfamily 2)	161	2.9E-43	94.3	GAHU01151209.1	803	105	1	210	314	N.d.
		148	2.2E-39	91.6	GAHU01163284.1	712	107	1	111	217	N.d.
		143	1.0E-37	90.7	GAHU01179005.1	400	107	1	108	214	N.d.
143		1.0E-37	90.7	GAHU01148077.1	912	107	1	794	900	Retrotransposon gag protein	
143		1.0E-37	92.1	GAHU01130031.1	952	101	1	846	946	Retrotransposon gag protein	
141		3.8E-37	93.6	GAHU01177060.1	540	94	1	115	208	RNase_H_like super family	
139		1.4E-36	91.2	GAHU01144836.1 ^e	1732	102	1	682	783	RNase_H_like, Retrotrans_gag and RT_like super families	
139		1.4E-36	92	GAHU01121003.1	676	100	1	218	316	N.d.	
137		4.9E-36	89.7	GAHU01147834.1	270	107	1	144	250	N.d.	
137		4.9E-36	91.1	GAHU01147833.1	515	101	1	276	376	N.d.	
132		2.3E-34	90.1	GAHU01153302.1	3885	101	1	2319	2419	Retropepsin_like, Retrotrans_gag and SMC_N super families	
126		1.1E-32	89.1	GAHU01175798.1	461	101	1	42	142	N.d.	
121		4.9E-31	86.9	GAHU01177345.1	443	107	1	194	300	N.d.	
115		2.3E-29	86.5	GAHU01047869.1	1017	104	1	369	472	N.d.	
115	2.3E-29	87.8	GAHU01047861.1	1077	98	1	546	643	N.d.		
111	3.0E-28	91.4	GAHU01175799.1	240	81	1	160	240	N.d.		
75	3.9E-17	93.9	GAHU01171825.1	376	49	1	328	376	N.d.		
63.9	8E-14	97.3	GAHU01108491.1	607	37	1	571	607	N.d.		

N.d. No detected.

^a TSA dataset of *A. tequilana* (GenBank:GAHU00000000).

^b Database (CDD 3.16–50,369 PSSMs).

^c Related with the TEs

^d Related with the Gypsy65-ZM_I-int

^e Related with the A-rich and Gypsy-71-I_ZM, respectively.

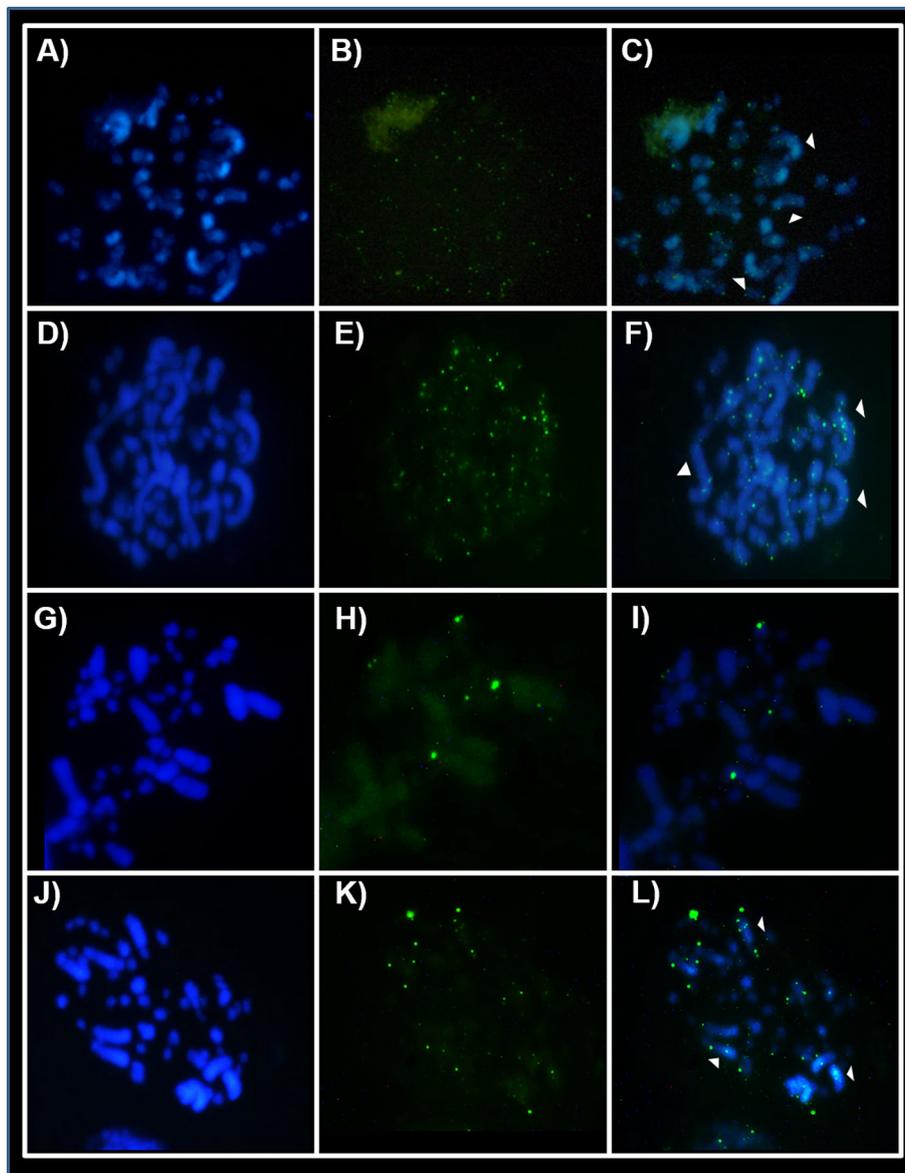


Fig. 5. FISH analysis of the AT-531 and AACHE-554 probes on the different diploid *Agave* accessions. Left column (DAPI detection), middle column (fluorescein detection), and right (merge). Arrow heads shows two or more signals on a single chromosome. A. *tequilana* (A–C) and *A. angustifolia* var. *Marginata* (D–F) hybridized with AT-531 (Subfamily 2) probe. A. *tequilana* (G–H) and *A. angustifolia* var. *Marginata* (J–L) hybridized with AACHE-554 (Subfamily 1) probe.

are considered as subfamilies (Fig. 2). Satellite DNA repeat monomers are originated as consequence of different amplification cycles and sequence divergence of shorter basic monomers (Stupar et al., 2002; Navajas-Pérez et al., 2005; Neumann et al., 2006). Consensus sequences of these repeated DNA regions have a length higher than 100 bp, and are AT.% rich, particularly, the sequences belonging to the subfamily 1 with a length of 165 bp and 62% of AT (Table 1). The preferential monomer length of 150–180 bp and 300–360 bp detected in many satellites of plants is often considered to mirror requirements of DNA length wrapped around one or two nucleosomes (Biscotti et al., 2015b; Schmidt and Heslop-Harrison, 1998). In addition, satellite DNA are generally AT.% rich, especially in the case of centromeric satellite DNAs (He et al., 2015) where this richness and AT.% periodical distribution are probably involved in the DNA and proteins packing of the heterochromatin (Pezer et al., 2012). According to the anterior, the consensus sequences identified in these two sequence subfamilies shows several traits associated with the satellite DNAs. On this way, it has also been reported the existence of several satellite DNA subfamilies as in some other plant species (Kazama et al., 2006; Navajas-Pérez et al., 2006).

When several of these satellite DNA subfamilies or families coexists constituted a library (Fry and Salser, 1977). This represents a permanent source of sequences that can be independently amplified in each specific genome into a dominant satellite DNA, a phenomenon known as the ‘library’ hypothesis (Fry and Salser, 1977; Plohl et al., 2012). Repeated DNA regions characterized in this study exhibited a wide variety of sequences composition. In this sense, it has been reported in other plant species, that satellite DNA subfamilies shared mutations that are unambiguously observed in repeats of the same subfamily in exactly the same positions, while at the same sites, all the sequences of the other families had different nucleotides (Navajas-Pérez et al., 2006; Quesada Del Bosque et al., 2013, 2014). A comparative analysis of tandem repeats from hundreds of species revealed that repeat monomers of different satellite DNA families are highly variable in its sequence composition and showed little evidence of sequence conservation (Melters et al., 2013). In fact, according to experimental evidence, any random sequence can lead to a family of tandem repeats (Navajas-Pérez et al., 2005). PCR amplifications by using specific primers in the DNA of these *Agave* accessions did not show a ladder profile. Although

all sequences grouped in both subfamilies if shows 6 conserved motifs (memes analysis). These short motifs were found in different positions probably by sequence rearrangements or in the reverse complementary sequences (Table 2). Monomer-repeat sequences give some clues about the mechanisms involved in the origin of a family of tandem repeats. Thus, repeat sequences are often composed of direct sub repeats or shorter motifs that remain as remnants of past events of sequence duplications or divergence events of an initial shorter repeat (Neumann et al., 2006). Particularly, in this study, a short motif occurs in all the monomers of both subfamilies (Tables 1 and 2). This motif 1 (CTYRA-GTTSG) exhibited discrete bands produced by its specific amplification, with an initial length of 110 bp (Fig. SA). Repetitive satellite DNA sequences may present short motifs as dinucleotides (i.e. AA/TT), pentanucleotides (i.e. CAAAA) and 9–30 bp repeats, as in this current study (Katsiotis et al., 1998; Macas et al., 2002; Navajas-Pérez and Paterson, 2009; Samoluk et al., 2017). In addition, southern blot hybridization on the amplification obtained by using specific primer sets on these *Agave* accessions showed a differential signal pattern according to subfamily probe used. The AT102–531 probe (subfamily 2) showed a relatively strong signal in the most abundant band (~200 bp) and weakly signals in the others bands (~500 bp) which suggested that this repeated region might not have a typical tandem organization (Fig. SA; Supplementary material 1). However, when the AACHE102–554 probe (subfamily 1) was used several signals were detected in the main PCR bands (~200, 600, 800, 1200, 1800). These multiple hybridizations suggested the presence of multimers and probably a tandem organization, particularly for subfamily 1, with different sequence variation on these PCR products. The clear signal intensity of some bands may also be related with the presence of a different copy number (Čížková et al., 2013). Considering together these results, it is possible to suggest that these repeated DNA regions share a common origin and a strong relation with plant satellite sequences.

Sequence divergence is low in the specific monomers within of each subfamily (Fig. 3). Particularly, 10 mutations were only detected in the case of the subfamily 2 (data not shown). Usually, sequence divergence between monomers of the same subfamily or family is often very low, usually up to 15% (King and Cummings, 1997; Plohl et al., 2008). This DNA satellite family conservation has also been reported in other plants (Mehrotra et al., 2014; Quesada Del Bosque et al., 2014). This homogeneity may be achieved by a non-independent evolution of monomers, a process in which mutations are homogenized throughout members of a repetitive family, and concomitantly fixed within of a group of species phylogenetically related (Dover, 1982; Plohl et al., 2008). *Agave* species are classified within the Asparagaceae family (The Angiosperm Phylogeny Group III, 2009) but the phylogenetic relationship among the species of this study (*A. angustifolia*, *A. fourcroydes* and *A. tequilana*) is not clearly defined. However, the high sequence identity observed among monomers of each *Agave* subfamily regardless of the *Agave* accession origin indicates that these accessions are phylogenetically closely related. Genetic homogenization and closest phylogenetic relationships have also been observed in these *Agave* accessions by using ribosomal gene analysis, specifically using 5S rDNA genes (Tamayo-Ordóñez et al., 2015).

Specific primers set designed from these repeated DNA regions allowed amplify a single PCR products on cDNA from *A. tequilana* (Fig. 4A). Transcripts of satellite DNA have been detected in rice and maize indicating that transcription of these may be possible in plants (Topp et al., 2004; Lee et al., 2006). In this study, some representative *A. tequilana* Weber (AT102-529 and AT102-531) monomers of each *Agave* consensus region were blasted against *A. tequilana* TSA dataset to explore its transcriptional potential. Results suggested that 28 TSA accessions have a high identity (< 86%) with these monomers (Table 2). Satellite DNA repeats are also transcriptionally active in some cases. For example, the CentO transcripts in rice are processed into siRNA, suggesting a potential role of this satellite repeat family in epigenetic chromatin modification (Lee et al., 2006). Gross et al. (2013) identified

TE transcripts in *A. tequilana* and *A. deserti*. Most of these TEs in *Agave* were derived from retrotransposons. In this study, the sequences representatives of the different satellite DNA families exhibited identities with some of these TSA related to TEs previously identified (Gross et al., 2013). Results shows that *A. tequilana* TSA might belong to the LTR/Gypsy (Gypsy65-ZM_I-int and Gypsy-71-I_ZM) family, and considering all the LTR/Gypsy database (Gypsy database) was also possible identity TEs signatures corresponding to the Cereba, GypsySL, CopiaSL and Ogre families (Table S2; Supplementary material 3). Here it is important to note that although these repeated DNA sequences may contain short TEs signatures, these are not part of the transposable domains of the closest TSA sequences (Fig. 4B). Direct and inverse repeated sequence elements in satellite monomers could be related to transposition as a mechanism in the process of concerted evolution (Dover, 2002). Different classes of transposable elements were recognized as sources of centromeric satellites in various organisms, for example in *Drosophila* (Heikkinen et al., 1995), rice (Ma and Jackson, 2006) and potato (Tek et al., 2005). Several studies have supported the idea that tandem repetitive families with homology to parts of retrotransposons derive from this class of transposable elements by random recombination and amplification (Macas et al., 2009; Gong et al., 2012). More studies must be conducted to further explore the importance of the TEs in *Agave* plants. The repeat DNA regions identified in the present study share many characteristics with the satellite DNA but fail in shows a tandem organization. Therefore, in this study, these only may be considered as satellite-like sequences.

4.2. Physical mapping of satellite sequences

Hybridization with the probes of each satellite-like DNA subfamily showed a differential signal pattern in diploid *Agave* accessions. These results are consistent with the identification of several satellite DNA sequences on different locations of the genomes of plants that have been reported in the PlantSat database (Macas et al., 2002; Čížková et al., 2013; Ribeiro et al., 2017; Samoluk et al., 2017). In this study, satellite-like DNA from subfamily 1 was more variable than those belonging to subfamily 2. Sequences homogenization can also be supported for the FISH signal number variation found between both satellite-like subfamilies. Subfamily 1 ranged between 4 and > 20 signals whereas than subfamily 2 ranged between 12 and 27 clear signals in these diploid *Agave* accessions. Sequences homogenization phenomenon in satellite DNA can fix different members of a family in a group of species phylogenetically related (Dover, 1982; Plohl et al., 2008). In this study, considering all the *Agave* accessions (Fig. SB and SC; Supplementary material 1), the average of hybridization signals in the subfamily 2 was 12.4 regardless of *Agave* accession analyzed. Similarly to this study, in *Musa* has been documented the existence of two subtelomeric satellite DNA families, the CL18 and the CL33, which showed a high level of sequence conservation and a high identity among different *Musa* species (Čížková et al., 2013). In *Solanum*, the subtelomeric CL14 and CL34 families are organized as independent long arrays. CL34 family is younger and more diverse, including the presence of short direct subrepeats and various subfamilies sequences than the CL14 family (Torres et al., 2011). FISH signal variations may be generated for different genetic mechanisms in these repetitive sequences thus affecting its chromosomal localization and distribution (He et al., 2015). In *Agave*, few studies about hybridization of repetitive sequence are reported, for example, the hybridization of 5S and 18S ribosomal DNA (rDNA) (Robert et al., 2008; Gomez-Rodriguez et al., 2013) where is suggested that these 5S loci rDNA are in the small chromosomes of the diploid species as results of an unequal recombination and transposition events (Gomez-Rodriguez et al., 2013).

Differently, in this study, it was not possible to identify the location of most of the signals detected in all *Agave* accessions. The above, probably due to the high signal number and the high ploidy level of some of these accessions (Castorena-Sánchez et al., 1991; Good-Avila

et al., 2006; Moreno-Salazar et al., 2007; Palomino et al., 2008). In addition, in *Agave* species, not all chromosomes are metacentric and, until now, there is no reliable centromeric markers to support this location (Gomez-Rodriguez et al., 2013). In general, FISH signal localization was subtelomeric for the probes of each satellite-like DNA subfamily and for the diploid *Agave* accessions analyzed (Table S3; Supplementary material 4). Many satellite DNA sequences are subtelomeric, and it is common to find several subtelomeric satellite DNA families within the same specie (Cuadrado and Jouve, 1994). In this study, few FISH signals were detected on interstitial chromosomal regions. In general, satellite DNA is less frequent in interstitial regions of the chromosomes (Zhao et al., 2011). Specifically, the *A. angustifolia* Haworth var. *Marginata* accession exhibits many hybridization signals for all chromosomal regions and for both satellite DNA subfamilies. In fact, only the last accession and *A. tequilana* Weber exhibit signals in all chromosomal regions, specifically, for the more homogenized subfamily 2. Presence of FISH signals in multiple positions suggests a high genomic diversity in these diploid *Agave* accessions. Previously, it had been reported an evolutionary differential pressure on the diploid *A. tequilana* Weber accession by using ribosomal genes (Tamayo-Ordóñez et al., 2015). In this last accession exist a high presence of Ty-1 copia retrotransposons which are considered important for organization and evolution of its genome (Bousios et al., 2007; Khaliq et al., 2012). Similarly at this study, TEs and satellite DNA differ in structure, origin mechanisms and evolutionary dynamics, but share sequence similarity and organization patterns, which indicates a mutual relationships between them (Neumann et al., 2001; Macas et al., 2009; Meštrović et al., 2015). More studies must be conducted to explore with more detailed the relations between these two class of repeat regions and its contribution in the genomic variation of this important plant genus.

5. Conclusion

This study reports for the first time an analysis of repeated DNA in *Agave* species. We found two subfamilies of satellite-like DNA with different length, composition and AT.% presents in all *Agave* accessions analyzed. The monomers of these subfamilies share short motifs although in different sequence position which suggests a genetic relationship between them. The more frequent motif 1 (CTYRAGTTSG) also generate discrete bands without becoming a ladder pattern when was amplified from genomic DNA, and the southern blot hybridization allows to detect signals in different discrete amplification bands. Particularly, sequence monomers of the subfamily 2 are more conserved suggesting a homogenization phenomenon in these satellite-like DNA regions. Experimental and bioinformatics analysis of these satellite monomers showed that are transcribed and exhibit a strong relation with TEs signatures, particularly with LTR/Gypsy family. Although not any of these TEs domains is overlapping with these satellite-like regions. FISH analysis with representative probes of these satellite-like subfamilies showed, in general, a subtelomeric hybridization with less interstitial signals. Specifically, for the subfamily 2, the diploid accessions *A. angustifolia* Haworth var. *Marginata* and *A. tequilana* Webber showed multiple chromosome localization signals, suggesting a high genomic diversity in these diploid *Agave* accessions. Finally, the presence of this satellite-like DNA in all the analyzed *Agave* accessions allows suggesting its usefulness in evolutionary and speciation studies within the genus.

Acknowledgements

This work was supported by CONACYT projects 50268 and 180757 of the Center of Scientific Research of Yucatán. Narváez-Zapata JA thanks to CONACYT for the national sabbatical support (application 265546) and to the IPN Sabbatical Program (application 10230). Thanks to Antonio Rescalvo-Morales and to Jorge Humberto Ramirez-Prado for their FISH and bioinformatics assistance, respectively.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.plgene.2018.08.002>.

References

- Bailey, T.L., Elkan, C., 1994. Fitting a mixture model by expectation maximization to discover motifs in bipolymers. In: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28–36. [citeulike-article-id:878292](https://doi.org/10.1093/nar/gkl198).
- Bailey, T.L., Williams, N., Misleh, C., Li, W.W., 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34. <https://doi.org/10.1093/nar/gkl198>.
- Bandelt, H.J., Forster, P., Rohlf, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>.
- Biscotti, M.A., Canapa, A., Forconi, M., Olmo, E., Barucca, M., 2015a. Transcription of tandemly repetitive DNA: functional roles. *Chromosom. Res.* 23, 463–477. <https://doi.org/10.1007/s10577-015-9494-4>.
- Biscotti, M.A., Olmo, E., Heslop-Harrison, J.S.P., 2015b. Repetitive DNA in eukaryotic genomes. *Chromosom. Res.* 23, 415–420. <https://doi.org/10.1007/s10577-015-9499-z>.
- Bousios, A., Saldana-Oyazabal, I., Valenzuela-Zapata, A.G., Wood, C., Pearce, S.R., 2007. Isolation and characterization of Ty1-copia retrotransposon sequences in the blue agave (*Agave tequilana* Weber var. azul) and their development as SSAP markers for phylogenetic analysis. *Plant Sci.* 172, 291–298. <https://doi.org/10.1016/j.plantsci.2006.09.002>.
- Castorena-Sánchez, I., Escobedo, M.R., Quiroz, A., 1991. New cytotoxicological determinants recognized in six taxa of agave in sections rigidae and sisalanae. *Can. J. Bot.* 69, 1257–1264.
- Čížková, J., Hřibová, E., Humplíková, L., Christelová, P., Suchánková, P., Doležal, J., 2013. Molecular analysis and genomic organization of Major DNA satellites in banana (*Musa* spp.). *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0054808>.
- Cuadrado, A., Jouve, N., 1994. Mapping and organization of highly-repeated DNA sequences by means of simultaneous and sequential FISH and C-banding in 6 × -triticale. *Chromosom. Res.* 2, 331–338. <https://doi.org/10.1007/BF01552727>.
- Dover, G., 1982. Molecular drive: a cohesive mode of species evolution. *Nature* 299, 111–117. <https://doi.org/10.1038/299111a0>.
- Dover, G., 2002. Molecular drive. *Trends Genet.* [https://doi.org/10.1016/S0168-9525\(02\)02789-0](https://doi.org/10.1016/S0168-9525(02)02789-0).
- Doyle, J., 1991. CTAB total DNA isolation. *Mol. Tech. Taxon.* 283–293. https://doi.org/10.1007/978-3-642-83962-7_18.
- Fry, K., Salsler, W., 1977. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12, 1069–1084. [https://doi.org/10.1016/0092-8674\(77\)90170-2](https://doi.org/10.1016/0092-8674(77)90170-2).
- Garrido-Ramos, M.A., 2015. Satellite DNA in plants: more than just rubbish. *Cytogenet. Genome Res.* 146, 153–170. <https://doi.org/10.1159/000437008>.
- Gil-Vega, K., Díaz, C., Nava-Cedillo, A., Simpson, J., 2006. AFLP analysis of *Agave tequilana* varieties. *Plant Sci.* 170, 904–909. <https://doi.org/10.1016/j.plantsci.2005.12.014>.
- Gomez-Rodriguez, V.M., Rodriguez-Garay, B., Palomino, G., Martínez, J., Barba-Gonzalez, R., 2013. Physical mapping of 5S and 18S ribosomal DNA in three species of *Agave* (Asparagales, Asparagaceae). *Comp. Cytogenet.* 7, 191–203. <https://doi.org/10.3897/CompCytogen.v7i3.5337>.
- Gong, Z., Wu, Y., Koblížková, A., Torres, G.A., Wang, K., Iovene, M., Neumann, P., Zhang, W., Novák, P., Buell, C.R., Macas, J., Jiang, J., 2012. Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell* 24, 3559–3574. <https://doi.org/10.1105/tpc.112.100511>.
- Good-Avila, S.V., Souza, V., Gaut, B.S., Eguiarte, L.E., 2006. Timing and rate of speciation in *Agave* (Agavaceae). *Proc. Natl. Acad. Sci. U. S. A.* 103, 9124–9129. <https://doi.org/10.1073/pnas.0603312103>.
- Gross, S.M., Martin, J.A., Simpson, J., Abraham-Juarez, M.J., Wang, Z., Visel, A., 2013. De novo transcriptome assembly of drought tolerant CAM plants, *Agave deserti* and *Agave tequilana*. *BMC Genomics* 14, 563. <https://doi.org/10.1186/1471-2164-14-563>.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98. [citeulike-article-id:691774](https://doi.org/10.1093/nar/gkl198).
- He, Q., Cai, Z., Hu, T., Liu, H., Bao, C., Mao, W., Jin, W., 2015. Repetitive sequence analysis and karyotyping reveals centromere-associated DNA sequences in radish (*Raphanus sativus* L.). *BMC Plant Biol.* 15, 105. <https://doi.org/10.1186/s12870-015-0480-y>.
- Heikkinen, E., Launonen, V., Müller, E., Bachmann, L., 1995. The pB370 BamHI satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic pDv elements. *J. Mol. Evol.* 41, 604–614. <https://doi.org/10.1007/BF00175819>.
- Hemleben, V., Kovarik, A., Torres-Ruiz, R.A., Volkov, R.A., Beridze, T., 2007. Plant highly repeated satellite DNA: molecular evolution, distribution and use for identification of hybrids. *Syst. Biodivers.* 5, 277–289. <https://doi.org/10.1017/S147220000700240X>.
- Heslop-Harrison, J.S., Brandes, A., Schwarbacher, T., 2003. Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. *Chromosom. Res.* 11, 241–253. <https://doi.org/10.1023/A:1022998709969>.

- Katsiotis, A., Hagidimitriou, M., Douka, A., Hatzopoulos, P., 1998. Genomic organization, sequence interrelationship, and physical localization using *in situ* hybridization of two tandemly repeated DNA sequences in the genus *Olea*. *Genome* 41, 527–534.
- Kazama, Y., Sugiyama, R., Suto, Y., Uchida, W., Kawano, S., 2006. The clustering of four subfamilies of satellite DNA at individual chromosome ends in *Silene latifolia*. *Genome* 49, 520–530. <https://doi.org/10.1139/G05-130>.
- Khalil, I., Khan, M.A., Pearce, S., 2012. Ty1-Copia retrotransposons are heterogeneous, extremely high copy number and are major players in the genome organization and evolution of *Agave tequilana*. *Genet. Resour. Crop. Evol.* 59, 575–587. <https://doi.org/10.1007/s10722-011-9705-6>.
- King, L.M., Cummings, M.P., 1997. Satellite DNA repeat sequence variation is low in three species of burying beetles in the genus *Nicrophorus* (Coleoptera: Silphidae). *Mol. Biol. Evol.* 14, 1088–1095.
- Kishii, M., Tsujimoto, H., 2002. Genus-specific localization of the tai I family of tandem-repetitive sequences in either the centromeric or subtelomeric regions in *Triticeae* species (Poaceae) and its evolution in wheat. *Genome* 45, 946–955. <https://doi.org/10.1139/g02-059>.
- Koukalova, B., Moraes, A.P., Renny-Byfield, S., Matyasek, R., Leitch, A.R., Kovarik, A., 2010. Fall and rise of satellite repeats in allopolyploids of *Nicotiana* over c. 5 million years. *New Phytol.* 186, 148–160. <https://doi.org/10.1111/j.1469-8137.2009.03101.x>.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. <https://doi.org/10.1093/molbev/msw054>.
- Lee, H.R., Neumann, P., Macas, J., Jiang, J., 2006. Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. *Mol. Biol. Evol.* 23, 2505–2520. <https://doi.org/10.1093/molbev/msl127>.
- Lim, K.Y., Skalicka, K., Koukalova, B., Volkov, R.A., Matyasek, R., Hemleben, V., Leitch, A.R., Kovarik, A., 2004. Dynamic changes in the distribution of a satellite homologous to intergenic 26-18S rDNA spacer in the evolution of *Nicotiana*. *Genetics* 166, 1935–1946. <https://doi.org/10.1534/genetics.166.4.1935>.
- Llorens, C., Futami, R., Covelli, L., Domínguez-Escribá, L., Viu, J.M., Tamarit, D., Aguilar-Rodríguez, J., Vicente-Ripolles, M., Fuster, G., Bernet, G.P., Maumus, F., Munoz-Pomer, A., Sempere, J.M., Latorre, A., Moya, A., 2011. The gypsy database (GyDB) of mobile genetic elements: release 2.0. *Nucleic Acids Res.* 39, 70–74. <https://doi.org/10.1093/nar/gkq1061>.
- Ma, J., Jackson, S.A., 2006. Retrotransposon accumulation and satellite amplification mediated by segmental duplication facilitate centromere expansion in rice. *Genome Res.* 16, 251–259. <https://doi.org/10.1101/gr.4583106>.
- Macas, J., Meszaros, T., Nouzova, M., 2002. PlantSat: a specialized database for plant satellite repeats. *Bioinformatics* 18, 28–35. <https://doi.org/10.1093/bioinformatics/18.1.28>.
- Macas, J., Koblížková, A., Navrátilová, A., Neumann, P., 2009. Hypervariable 3' UTR region of plant LTR-retrotransposons as a source of novel satellite repeats. *Gene* 448, 198–206. <https://doi.org/10.1016/j.gene.2009.06.014>.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.L., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., Bryant, S.H., 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45, D200–D203. <https://doi.org/10.1093/nar/gkw1129>.
- Martínez-Zapater, J.M., Estelle, M.A., Somerville, C.R., 1986. A highly repeated DNA sequence in *Arabidopsis thaliana*. *MGG Mol. Gen. Genet.* 204, 417–423. <https://doi.org/10.1007/BF00331018>.
- Matyasek, R., Fulneck, J., Leitch, A.R., Kovarik, A., 2011. Analysis of two abundant, highly related satellites in the allotetraploid *Nicotiana glauca* using double-strand conformation polymorphism analysis and sequencing. *New Phytol.* 192, 747–759. <https://doi.org/10.1111/j.1469-8137.2011.03827.x>.
- McKain, M.R., Wickert, N., Zhang, Y., Ayyampalayam, S., Richard McCombie, W., Chase, M.W., Chris Pires, J., Depamphilis, C.W., Leebens-Mack, J., 2012. Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *Am. J. Bot.* 99, 397–406. <https://doi.org/10.3732/ajb.1100537>.
- Mehrotra, S., Goel, S., Raina, S.N., Rajpal, V.R., 2014. Significance of satellite DNA revealed by conservation of a widespread repeat DNA sequence among angiosperms. *Appl. Biochem. Biotechnol.* 173, 1790–1801. <https://doi.org/10.1007/s12010-014-0966-3>.
- Melters, D.P., Bradnam, K.R., Young, H.A., Telis, N., May, M.R., Ruby, J.G., Sebra, R., Peluso, P., Eid, J., Rank, D., Garcia, J.F., Derisi, J.L., Smith, T., Tobias, C., Ross-Ibarra, J., Korf, I., Chan, S.W.L., 2013. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* 14. <https://doi.org/10.1186/gb-2013-14-1-r10>.
- Meštrović, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Šatović, E., Plohl, M., 2015. Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosom. Res.* 23, 583–596. <https://doi.org/10.1007/s10577-015-9483-7>.
- Moreno-Salazar, S.F., Esqueda, M., Martínez, J., Palomino, G., 2007. Nuclear genome size and karyotype of *Agave angustifolia* and a rhodantha from Sonora, México. *Rev. Fitotec. Mex.* 30, 13–23.
- Navajas-Pérez, R., Paterson, A.H., 2009. Patterns of tandem repetition in plant whole genome assemblies. *Mol. Gen. Genomics* 281, 579–590. <https://doi.org/10.1007/s00438-009-0433-y>.
- Navajas-Pérez, R., De La Herrán, R., Jamilena, M., Lozano, R., Rejón, C.R., Rejón, M.R., Garrido-Ramos, M.A., 2005. Reduced rates of sequence evolution of Y-linked satellite DNA in *Rumex* (Polygonaceae). *J. Mol. Evol.* 60, 391–399. <https://doi.org/10.1007/s00239-004-0199-0>.
- Navajas-Pérez, R., Schwarzacher, T., De La Herrán, R., Ruiz Rejón, C., Ruiz Rejón, M., Garrido-Ramos, M.A., 2006. The origin and evolution of the variability in a Y-specific satellite-DNA of *Rumex acetosa* and its relatives. *Gene* 368, 61–71. <https://doi.org/10.1016/j.gene.2005.10.013>.
- Neumann, P., Nouzová, M., Macas, J., 2001. Molecular and cytogenetic analysis of repetitive DNA in pea (*Pisum sativum* L.). *Genome* 44, 716–728. <https://doi.org/10.1139/gen-44-4-716>.
- Neumann, P., Koblížková, A., Navrátilová, A., Macas, J., 2006. Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics* 173, 1047–1056. <https://doi.org/10.1534/genetics.106.056259>.
- Palomino, G., Martínez, J., Méndez, I., 2008. Karyotype studies in cultivars of *Agave tequilana* Weber. *Caryologia* 61, 144–153. <https://doi.org/10.1080/00087114.2008.10589622>.
- Pezer, Brajković, J., Felicello, I., Ugarković, D., 2012. Satellite DNA-mediated effects on genome regulation. In: *Repetitive DNA*, pp. 153–169. <https://doi.org/10.1159/000337116>.
- Pires, J.C., Lim, K.Y., Kovarik, A., Matyasek, R., Boyd, A., Leitch, A.R., Leitch, I.J., Bennett, M.D., Soltis, P.S., Soltis, D.E., 2004. Molecular cytogenetic analysis of recently revealed a karyotype that is additive of the. *Am. J. Bot.* 91, 1022–1035. <https://doi.org/10.3732/ajb.91.7.1022>.
- Plohl, M., Chaves, R., 2014. Evolutionary story of a Satellite DNA from phodopus 6, 2944–2955. <https://doi.org/10.1093/gbe/evu233>.
- Plohl, M., Luchetti, A., Meštrović, N., Mantovani, B., 2008. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 409, 72–82. <https://doi.org/10.1016/j.gene.2007.11.013>.
- Plohl, M., Meštrović, N., Mravinac, B., 2012. Satellite DNA evolution. *Genome Dyn.* 7, 126–152. <https://doi.org/10.1159/000337122>.
- Quesada Del Bosque, M.E., López-Flores, I., Suárez-Santiago, V.N., Garrido-Ramos, M.A., 2013. Differential spreading of HinfI satellite DNA variants during radiation in Centaureinae. *Ann. Bot.* 112, 1793–1802. <https://doi.org/10.1093/aob/mct233>.
- Quesada Del Bosque, M.E., López-Flores, I., Suárez-Santiago, V., Garrido-Ramos, M.A., 2014. Satellite-DNA diversification and the evolution of major lineages in *Cardueae* (Carduoideae Asteraceae). *J. Plant Res.* 127, 575–583. <https://doi.org/10.1007/s10265-014-0648-9>.
- Ribeiro, T., dos Santos, K.G.B., Richard, M.M.S., Sévignac, M., Thureau, V., Geffroy, V., Pedrosa-Harand, A., 2017. Evolutionary dynamics of satellite DNA repeats from Phaseolus beans. *Protoplasma* 254, 791–801. <https://doi.org/10.1007/s00709-016-0993-8>.
- Robert, M.L., Lim, K.Y., Hanson, L., Sanchez-Teyer, F., Bennett, M.D., Leitch, A.R., Leitch, I.J., 2008. Wild and agronomically important *Agave* species (Asparagaceae) show proportional increases in chromosome number, genome size, and genetic markers with increasing ploidy. *Bot. J. Linn. Soc.* 158, 215–222. <https://doi.org/10.1111/j.1095-8339.2008.00831.x>.
- Salina, E.A., Pestsova, E.G., Adonina, I.G., Vershinin, A.V., 1998. Identification of a new family of tandem repeats in *Triticeae* genomes. *Euphytica* 100, 231–237.
- Samoluk, S.S., Robledo, G., Bertoli, D., Seijo, J.G., 2017. Evolutionary dynamics of an at-rich satellite DNA and its contribution to karyotype differentiation in wild diploid *Arachis* species. *Mol. Gen. Genomics* 292, 283–296. <https://doi.org/10.1007/s00438-016-1271-3>.
- Sánchez-Teyer, F., Moreno-Salazar, S., Esqueda, M., Barraza, A., Robert, M.L., 2009. Genetic variability of wild *Agave angustifolia* populations based on AFLP: a basic study for conservation. *J. Arid Environ.* 73, 611–616. <https://doi.org/10.1016/j.jaridenv.2009.01.008>.
- Schmidt, T., Heslop-Harrison, J.S., 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci.* 3, 195–199. [https://doi.org/10.1016/S1360-1385\(98\)01223-0](https://doi.org/10.1016/S1360-1385(98)01223-0).
- Stupar, R.M., Song, J., Tek, A.L., Cheng, Z., Dong, F., Jiang, J., 2002. Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. *Genetics* 162, 1435–1444.
- Tamayo-Ordóñez, Y.J., Narváez-Zapata, J.A., Sánchez-Teyer, L.F., 2015. Comparative characterization of ribosomal DNA regions in different *Agave* accessions with economical importance. *Plant Mol. Biol. Report.* 33, 2014–2029. <https://doi.org/10.1007/s11105-015-0895-5>.
- Tek, A.L., Song, J., Macas, J., Jiang, J., 2005. Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. *Genetics* 1238, 1231–1238. <https://doi.org/10.1534/genetics.105.041087>.
- The Angiosperm Phylogeny Group III, 2009. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161, 105–121. <https://doi.org/10.1111/j.1095-8339.2009.00996.x>.
- Topp, C.N., Zhong, C.X., Dawe, R.K., 2004. Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc. Natl. Acad. Sci. USA* 101, 15986–15991. <https://doi.org/10.1073/pnas.0407154101>.
- Torres, G.A., Gong, Z., Iovene, M., Hirsch, C.D., Buell, C.R., Bryan, G.J., Novak, P., Macas, J., Jiang, J., Birchler, J.A., 2011. Organization and evolution of subtelomeric satellite repeats in the potato genome. *Genes|Genomes|Genetics* 1, 85–92. <https://doi.org/10.1534/g3.111.000125>.
- Ugarkovic, D., 2005. Functional elements residing within satellite DNAs. *EMBO Rep.* 6, 1035–1039. <https://doi.org/10.1038/sj.embor.7400558>.
- Ugarković, D., Plohl, M., 2002. Variation in satellite DNA profiles—causes and effects. *EMBO J.* 21, 5955–5959. <https://doi.org/10.1093/emboj/cdf612>.
- Yoong Lim, K., Kovářik, A., Matyasek, R., Chase, M.W., Knapp, S., McCarthy, E., Clarkson, J.J., Leitch, A.R., 2006. Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section *Alatae*. *Plant J.* 48, 907–919. <https://doi.org/10.1111/j.1365-3113.2006.02930.x>.
- Zhao, X., Lu, J., Zhang, Z., Hu, J., Huang, S., Jin, W., 2011. Comparison of the distribution of the repetitive DNA sequences in three variants of *Cucumis sativus* reveals their phylogenetic relationships. *J. Genet. Genomics* 38, 39–45. <https://doi.org/10.1016/j.jcg.2010.12.005>.

DISCUSIÓN GENERAL

En este estudio se describió por primera vez una caracterización comparativa de una región de ADN satélite en 3 especies de Agave. Los cebadores ATsat-102-113 y ATsat-109 dan como resultado la amplificación de fragmentos que generan un perfil de múltiples bandas. A partir de la clonación y secuenciación de fragmentos de ~200pb, uno de los fragmentos que mostraron más abundancia, para cada uno de los cebadores, se obtuvieron 101 secuencias pertenecientes a las 3 especies de agave evaluadas, con sus respectivas variedades. Entre las características de estas secuencias destacan: la longitud promedio de estas secuencias que es de 119 pb, la presencia de motivos internos cortos dentro de las mismas, también se observó que comparten ciertos motivos cortos con otras regiones repetidas de plantas del PlantSat (Macas *et al.*, 2002). Los resultados obtenidos son similares y consistentes con lo observado en plantas como *Musa* (Čížková *et al.*, 2013), *Arachis* (Samorak *et al.*, 2017), *Paphiopedilum* (Lee *et al.*, 2018), entre otros.

Las secuencias obtenidas con el cebador ATsat-102-113 y ATsat-109 y su posterior análisis arrojaron que hay una gran variabilidad entre ellas. Sin embargo, para dos grupos de secuencias generados con el cebador ATsat-102-113 mostraron características similares en la secuencias, el grupo de secuencias A tienen un tamaño 140 pb y el grupo B tiene 107 pb. También se observó que estas secuencias están presentes en todas las especies evaluadas. Se determinó las secuencias consenso para cada grupo y el grupo A mostró una variación en la composición de la misma, mientras el grupo B es más conservado. La presencia de más de una secuencia de ADN satélite en una especie o especies es común; especies como *Cucumis sativus* (Zhang *et al.*, 2011), el género *Nicotiana* (Lim *et al.*, 2006; Kovarick *et al.*, 2009; Jo *et al.*, 2009) o la tribu Cardue (Quesada del bosque *et al.*, 2014) son algunos ejemplos de esta característica. Los monómeros (o motivos) de ADN satélite pueden estar ordenados en arreglos o matrices en tándem dentro del genoma en eucariotas (Garrido-Ramos *et al.*, 2015). Estos ADN satélites pueden estar representados en el genoma por miles de copias e incluso pueden ser diferentes entre ellos o formar parte de una “familia de ADN satélites” (Plohl *et al.*, 2008). La librería del ADN satélite es una teoría que establece que más de una familia de estas secuencias coexiste, lo cual puede representar una fuente importante para la generación de nuevas familias de repeticiones en el genoma (Fry y Salsler, 1977).

La evidencia en el Southern Blot para las sondas AACHE102-554 y AT102-531, ambas generadas por el cebador ATsat-102-113, mostró que existe una probable organización diferencial para ambas. El perfil de múltiples bandas y el tamaño observado en la electroforesis y en el Southern Blot sugieren que estos fragmentos se encuentran en tándem característico del ADN satélite y es probable que la presencia de fragmentos del mismo tamaño en algunas de las especies evaluadas se deba a que están relacionados, es decir, que pertenecen al mismo género siendo especies cercanas relacionadas (Macas *et al.*, 2000).

La información obtenida por hibridación fluorescente *in situ* (FISH) para los cromosomas de las 3 especies de agave evaluadas mostró una acumulación diferencial de las sondas AACHE102-554 y AT102-531. Las diferencias encontradas en el mapeo de ambas sondas fueron notables pues no fue posible determinar con exactitud la posición de estas. Otro problema es la falta de un marcador cromosómico en el centrómero en estas especies, pues no todos los cromosomas en agaves son metacéntricos (Castorena *et al.*, 1991; Robert *et al.*, 2008, Gomes-Rodríguez *et al.*, 2013). Los rDNA 45S y 5S han sido mapeados en cromosomas de *A. tequilana*, *A. fourcroydes* y *A. angustifolia* con diferentes niveles de ploidía, encontrando una relación entre la cantidad de señales encontradas y el nivel de ploidía (Robert *et al.*, 2008). En 2013, Gómez-Rodríguez *et al.*, determinaron que el gen ribosomal 18S se encuentra en cromosomas grandes acrocéntricos y el gen ribosomal 5S en cromosomas pequeños metacéntricos o submetacéntricos de la especie *A. tequilana*. La especie *A. angustifolia* var. *Marginata* mostró una amplia distribución de ambas sondas

Se ha documentado que los ADN satélites pueden tener una estrecha relación con otras regiones repetidas como los elementos móviles o retroelementos (Ma y Jackson, 2006, fritillaria 2009 o 10; Mestrova *et al.*, 2015). En *A. tequilana* se ha reportado una alta presencia de retroelementos, aproximadamente de un 80% de su genoma (Khaliq *et al.*, 2012) y en especies como *A. amaniensis* y el Híbrido H11648 se ha observado una amplia distribución de Ty1-copya (Tesis de maestría Huijara-Vasconcelos, 2010; tesis de maestría Rescalvo-Morales, 2011). Un análisis *in silico* realizado con AT102-531 en secuencias de los Ty3-gypsy demostró que puede estar contenido en la secuencia de este (datos no mostrados) para complementar este análisis se realizó la amplificación de un solo fragmento de ~100pb en cDNA, este tamaño concuerda con el tamaño de esta secuencia (107 pb). Sin embargo, es importante realizar más estudios para corroborar esta información y complementarla de ser necesario.

El género *Agave* está clasificado dentro de la familia Asparagaceae (APG III, 2009) pero la relación filogenética entre las especies evaluadas en este estudio no es del todo clara. Sin embargo, el alto índice de identidad de secuencia observada entre los monómeros de cada subfamilia *Agave* independientemente del origen de la accesión de *Agave* indica que estas accesiones son filogenéticamente estrechamente relacionados. Homogenización genética y más cercanas relaciones filogenéticas también se han observado en estas accesiones de *Agave* mediante el uso de análisis de genes ribosomales, específicamente utilizando genes 5S rDNA (Tamayo-Ordoñez *et al.*, 2015). Diferentes estudios mostraron una amplia variabilidad genómica en los *Agaves*, por ejemplo en genes rDNA (Tamayo-Ordoñez *et al.*, 2015), en poliploides (Tamayo-Ordoñez *et al.*, 2016) y AFLP (Infante *et al.*, 2002; Gil-Vega *et al.*, 2006; Sánchez-Teyer *et al.*, 2009), que puede ser consecuencia de un proceso de aloploidía (McKain *et al.*, 2012), eventos epigenéticos (Díaz-Martínez *et al.*, 2012; Us-Camas *et al.*, 2017) y elementos repetidos (Macas *et al.*, 2009; Macas *et al.*, 2015). Estos últimos están muy representados en el genoma (Khaliq *et al.*, 2012) y en el transcriptoma (Gross *et al.*, 2013) de estas plantas.

A pesar de que hasta la fecha no se tenga clara la probable función que los ADN satélites tienen en los genomas de las plantas su alta presencia es intrigante pues se cree que son más importantes de lo que se sabe. La identificación de DNA satélites abre una ventana de oportunidad para conocer la importancia de estas regiones dentro del genoma en diversos organismos, las plantas son un buen ejemplo de esto pues pueden contener más de un complemento cromosómico. Los *agaves* son un modelo interesante y un reto para el estudio de regiones repetidas.

CONCLUSIONES

1. Se identificaron dos subfamilias de ADN satélites en tres especies de *Agave* cercanamente relacionadas. Estas subfamilias presentan características de ADN satélites como lo son la longitud del motivo, motivos cortos y repetidos internos y organización en tándem.
2. La composición y organización de los ADN satélites son similares entre ellas, siendo *A. tequilana* la que mostró una probable diferencia en su organización.
3. El mapeo en los cromosomas indica una distribución variada entre los cromosomas de cada especie, con preferencia en las regiones subteloméricas de estos.
4. El análisis realizado en la caracterización de las secuencias obtenidas indicó que los ADN satélites de las subfamilias A y B son variables entre ellas pero son compartidos en *A. tequilana*, *A. fourcroydes* var. Kitam ki, *A. fourcroydes* var. Yaax ki, *A. angustifolia* var. Marginata y *A. angustifolia* var. Chelem ki.

PERSPECTIVAS

1. Caracterización y localización de las secuencias con el cebador ATsat-109, incluso el aislamiento de secuencias nuevas.
2. Análisis de secuencias ADN satélites en especies de la sección *Rigidae* e incluso *Littae* (presencia/ausencia).
3. Mapeo físico de regiones repetidas en *Agave* para estudiar la evolución citogenética (ADN ribosomal, ADN satélite, ADN telomérico y retroelementos).
4. Explorar la relación de las subfamilias de ADN satélite con retroelementos.

BIBLIOGRAFÍA

- Ananiev, E. V., Phillips, R. L., y Rines, H. W. (1998). Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. *Proceedings of the National Academy of Sciences of the United States of America*, 95(22), 13073–13078.
- Appels, R., & Moran, L. B. (1986). Rye heterochromatin . I . Studies on clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. *Canadian Journal of Genetics and Cytology*, 28(Singer 1982), 645–657.
- Banerjee, S., & Sharma, A. K. (1987). Cytologia Cytophotometric Estimation of Nuclear DNA in Different Species and Varieties of Agave. *Cytologia*, 52, 85-90.
- Benjamin Lewin. (2008). *Genes IX*. (McGraw-Hill Interamericana Editores, Ed.) (novena). México.
- Bennett, M. D., y Leitch, I. J. (2005). Plant genome size research: A field in focus. In *Annals of Botany* (Vol. 95, pp. 1-6).
- Biscotti, M. A., Olmo, E., y Heslop-Harrison, J. S. P. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 23(3), 415-20.
- Buntjer, J. B., & Lenstra, J. A. (1998). Self-amplification of satellite DNA in vitro. *Genome . June*, 41(3), 429–434.
- Čížková, J., Hřibová, E., Humplíková, L., Christelová, P., Suchánková, P., & Doležel, J. (2013). Molecular Analysis and Genomic Organization of Major DNA Satellites in Banana (*Musa* spp.). *Public Library of Science ONE*, 8(1).
- Clarkson, J. J., Knapp, S., Garcia, V. F., Olmstead, R. G., Leitch, A. R., & Chase, M. W. (2004). Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. *Molecular Biology*, 10-18.
- Clarkson, J. J., Lim, K. Y., Kovařík, A., Chase, M. W., Knapp, S., & Leitch, A. R. (2005). Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae). *New Phytologist*, 168(1), 241-252.
- Dong, F., Miller, J. T., Jackson, S. a, Wang, G. L., Ronald, P. C., & Jiang, J. (1998). Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 95(14), 8135–8140.
- Eguiarte, L. E., Souza, V., & Silva-Montellano, A. (2000). Evolución de la Familia Agavaceae: Filogenia, Biología Reproductiva y genética de Poblaciones. *Bol. Soc. Bot. México*, (66), 131-150.

- Emadzade, K., Jang, T. S., Macas, J., Kovařík, A., Novák, P., Parker, J., & Weiss-Schneeweiss, H. (2014). Differential amplification of satellite PaB6 in chromosomally hypervariable *Prospero autumnale* complex (Hyacinthaceae). *Annals of Botany*, 114(8).
- Ferree, P. M., & Prasad, S. (2012). How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. *Genetics Research International*, 2012(Figure 1), 430136.
- Flavell, R. B., Bennett, M. D., Smith, J. B., & Smith, D. B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochemical Genetics*, 12(4), 257–269.
- Fry, K., & Salser, W. (1977). Nucleotide sequences of HS- α satellite DNA from kangaroo rat *dipodomys ordii* and characterization of similar sequences in other rodents. *Cell*, 12(4), 1069–1084.
- Garcia-Mendoza, A. (2007). Los Agaves de México. *Ciencias*, 14–23.
- Garrido-Ramos, M. A. (2015). Satellite DNA in plants: More than just rubbish. *Cytogenetic and Genome Research*, 146(2), 153–170.
- Gindullis, F., Desel, C., Galasso, I., & Schmidt, T. (2001). The large-scale organization of the centromeric region in Beta species. *Genome Research*, 11(2), 253–265.
- Gomez-Rodriguez, V. M., Rodriguez-Garay, B., Palomino, G., Martínez, J., & Barba-Gonzalez, R. (2013). Physical mapping of 5S and 18S ribosomal DNA in three species of Agave (Asparagales, Asparagaceae). *Comparative Cytogenetics*, 7(3), 191–203.
- Gong, Z., Wu, Y., Koblížková, A., Torres, G. a, Wang, K., Iovene, M., ... Jiang, J. (2012). Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell*, 24(9), 3559–74.
- Good-Avila, S. V., Souza, V., Gaut, B. S., & Eguiarte, L. E. (2006). Timing and rate of speciation in Agave (Agavaceae). *Proceedings of the National Academy of Sciences of the United States of America*, 103(24), 9124–9129.
- Grebenstein, B., Grebenstein, O., Sauer, W., & Hemleben, V. (1996). Distribution and complex organization of satellite DNA sequences in Aveneae species. *Genome*, 39(6), 1045–1050.
- Guadalupe, P., Martínez, J., Méndez, (2008). Karyotype studies in cultivars of *Agave tequilana* Weber. *Caryologia*, 61(2), 144–153.
- Hall Thomas A. (1999). BioEdit: a user-friendly biological sequence alignment editor and

- analysis program for Windows 95/98/NT.
- He, Q., Cai, Z., Hu, T., Liu, H., Bao, C., Mao, W., & Jin, W. (2015). Repetitive sequence analysis and karyotyping reveals centromere-associated DNA sequences in radish (*Raphanus sativus* L.). *BMC Plant Biology*, *15*(1), 105.
- Hemleben, V., Kovarik, A., Torres-Ruiz, R. A., Volkov, R. A., & Beridze, T. (2007). Plant highly repeated satellite DNA: molecular evolution, distribution and use for identification of hybrids. *Systematics and Biodiversity*, *5*(3), 277–289.
- Herrera, J. C. (2007). La citogenética molecular y su aplicación en el estudio de los genomas vegetales. *Red de Revistas Científicas de América Latina, El Caribe, España Y Portugal*, *25*, 26–35.
- Heslop-Harrison, J. S., Brandes, A., & Schwarzacher, T. (2003). Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. *Chromosome Research*, *11*(3), 241–253.
- Hizume, M., Shibata, F., Matsumoto, A., Maruyama, Y., Hayashi, E., Kondo, T., ... Hong, D. (2002). Tandem repeat DNA localizing on the proximal DAPI bands of chromosomes in *Larix*, Pinaceae, *783*, 777–783.
- Hourcade, D., Dressler, D., & Wolfson, J. (1973). The amplification of ribosomal RNA genes involves a rolling circle intermediate. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(10), 2926–30.
- Jo, S., Koo, D., Kim, J. F., Hur, C., Lee, S., Yang, T., Choi, D. (2009). Evolution of ribosomal DNA-derived satellite repeat in tomato genome. *BMC Plant Biology*, *9*, 42.
- Jo, S., Park, H., Kim, S.-M., Kim, H. H., Hur, C., & Choi, D. (2011). Unraveling the sequence dynamics of the formation of genus-specific satellite DNAs in the family solanaceae. *Heredity*, *106*(5), 876–885.
- John F. Elder, J., & Turner, B. J. (1995). Concerted Evolution of Repetitive DNA Sequences in Eukaryotes. *The Quarterly Review of Biology*, *70*(3), 297–320.
- Katsiotis, A., Hagidimitriou, M., Hagidimitriou, M., & Douka, A. (1998). Genomic organization, sequence interrelationship, and physical localization using in situ hybridization of two tandemly repeated DNA sequences in the genus *Olea*. *Genome / National Research Council Canada = Génome / Conseil National de Recherches Canada*, *41*(February 2014), 527–534.
- Kirov, I. V, Kiseleva, A. V, Laere, K. Van, Roy, N. Van, & Khrustaleva, L. I. (2017). Tandem repeats of *Allium fistulosum* associated with major chromosomal landmarks. *Molecular Genetics and Genomics*, *292*(2), 453–464.

- Kishii, M., Nagaki, K., & Tsujimoto, H. (2001). A tandem repetitive sequence located in the centromeric region of common wheat (*Triticum aestivum*) chromosomes. *Chromosome Research*, 9(5), 417–428.
- Kishii, M., Nagaki, K., Tsujimoto, H., & Sasakuma, T. (1999). Exclusive localization of tandem repetitive sequences in subtelomeric heterochromatin regions of *Leymus racemosus* (Poaceae, Triticeae). *Chromosome Research*, 7(7), 519–529.
- Koukalova, B., Moraes, A. P., Renny-Byfield, S., Matyasek, R., Leitch, A. R., & Kovarik, A. (2010). Fall and rise of satellite repeats in allopolyploids of *Nicotiana* over c. 5 million years. *New Phytologist*, 186(1), 148–160.
- Lee, H., Neumann, P., Macas, J., & Jiang, J. (2005). Transcription and Evolutionary Dynamics of the Centromeric Satellite Repeat CentO in Rice.
- Lim, K. Y., Kovařík, A., Matyasek, R., Bezdek, M., Lichtenstein, C. P., & Leitch, A. R. (2000). Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units. *Chromosoma*, 109(3), 161–172.
- Lim, K. Y., Matyásek, R., Lichtenstein, C. P., & Leitch, a R. (2000). Molecular cytogenetic analyses and phylogenetic studies in the *Nicotiana* section *Tomentosae*. *Chromosoma*, 109(4), 245–258.
- Lim, K. Y., Skalicka, K., Koukalova, B., Volkov, R. A., Matyasek, R., Hemleben, V., ... Kovarik, A. (2004). Dynamic Changes in the Distribution of a Satellite Homologous to Intergenic 26-18S rDNA Spacer in the Evolution of *Nicotiana*. *Genetics*, 166(4), 1935–1946.
- Ma., G. I., D., T. M., M.L., N., A., D., J., S. M., J.M., G. B., & A., O. (1997). Hibridacion in situ. *Química Clínica*, 49(6), 255–8.
- Ma, J., & Jackson, S. A. (2006). Retrotransposon accumulation and satellite amplification mediated by segmental duplication facilitate centromere expansion in rice. *Genome Research*, 16(2), 251–259.
- Macas, J., Meszaros, T., & Nouzová, M. (2002). PlantSat: a specialized database for plant satellite repeats. *Bioinformatics*, 18(1), 28–35.
- Macas, J., Navrátilová, A., & Koblížková, A. (2006). Sequence homogenization and chromosomal localization of VicTR-B satellites differ between closely related *Vicia* species. *Chromosoma*, 115(6), 437–447.
- Matyasek, R., Fulnecek, J., Leitch, A. R., & Kovarik, A. (2011). Analysis of two abundant, highly related satellites in the allotetraploid *Nicotiana arentsii* using double-strand

- conformation polymorphism analysis and sequencing. *New Phytologist*, 192(3), 747–759.
- Matyšek, R., Koukalová, B., & Reich, J. (1989). Isolation and sequencing of HRS60dim1, a dimeric member of the HRS60-family of a tobacco DNA repeat. *Nucleic Acids Research*, 17(11), 4377.
- Mehrotra, S., Goel, S., Raina, S. N., & Rajpal, V. R. (2014). Significance of satellite DNA revealed by conservation of a widespread repeat DNA sequence among angiosperms. *Applied Biochemistry and Biotechnology*, 173(7), 1790–1801.
- Melters, D. P., Bradnam, K. R., Young, H. A., Telis, N., May, M. R., Ruby, J. G., Chan, S. W. L. (2013). Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biology*, 14(1).
- Meštrović, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Šatović, E., & Plohl, M. (2015). Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosome Research*, 23(3), 583–596.
- Narváez-Zapata, J. a, & Sánchez-Teyer, L. F. (2009). Agaves as a raw material: recent technologies and applications. *Recent Patents on Biotechnology*, 3(3), 185–191. <https://doi.org/10.2174/187220809789389144>
- Navajas-Pérez, R., & Paterson, A. H. (2009). Patterns of tandem repetition in plant whole genome assemblies. *Molecular Genetics and Genomics*, 281(6), 579–590.
- Neumann, P., Koblížková, A., Navrátilová, A., & Macas, J. (2006). Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics*, 173(2), 1047–1056.
- Neumann, P., Nouzová, M., & Macas, J. (2001). Molecular and cytogenetic analysis of repetitive DNA in pea (*Pisum sativum* L.). *Genome*, 44(4), 716–728.
- Nouzová, M., Kubaláková, M., Želová, M. D., Koblížková, A., Neumann, P., Doležel, J., & Macas, J. (1999). Cloning and characterization of new repetitive sequences in field bean (*Vicia faba* L.). *Annals of Botany*, 83(5), 535–541.
- Ohmido, N., Fukut, K., & Kinoshita Toshiro. (2010). Recent advances in rice genome and chromosome structure research by uorescence in situ hybridization (FISH). *Proceeding of the Japan Academy*, 86(2), 103–116.
- Pires, J. C., Lim, K. Y., Kovarik, A., Matyasek, R., Boyd, A., Leitch, A. R., ... Soltis, D. E. (2004). Molecular Cytogenetic Analysis of Recently Reveal a Karyotype That Is Additive of the. *Amer. J. Bot.*, 91(7), 1022–1035.
- Plohl, M., Luchetti, A., Meštrović, N., & Mantovani, B. (2008). Satellite DNAs between

- selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene*, 409(1-2), 72–82.
- Plohl, M., Meštrović, N., & Mravinac, B. (2012). Satellite DNA evolution. *Genome Dynamics*, 7, 126–152.
- Quesada del Bosque, M. E., López-Flores, I., Suárez-Santiago, V., & Garrido-Ramos, M. A. (2014). Satellite-DNA diversification and the evolution of major lineages in Cardueae (Carduoideae Asteraceae). *Journal of Plant Research*, 127, 575–583.
- Ramos, I. L.-F. M. A. G.-. (2012). The Repetitive DNA Content of Eukaryotic Genomes. *Genome Dyn Basel, Kargen*, 7, 1–28.
- Ribeiro, T., dos Santos, K. G. B., Richard, M. M. S., Sévignac, M., Thareau, V., Geffroy, V., & Pedrosa-Harand, A. (2017). Evolutionary dynamics of satellite DNA repeats from Phaseolus beans. *Protoplasma*, 254(2), 791–801.
- Ribeiro, T., Marques, A., Novák, P., Schubert, V., Vanzela, A. L. L., Macas, J., Pedrosa-Harand, A. (2016). Centromeric and non-centromeric satellite DNA organisation differs in holocentric Rhynchospora species. *Chromosoma*.
- Richard, M. M., Chen, N. W., Thareau, V., Pflieger, S., Blanchet, S., Pedrosa-Harand, A., Geffroy, V. (2013). The Subtelomeric khipu Satellite Repeat from Phaseolus vulgaris: Lessons Learned from the Genome Analysis of the Andean Genotype G19833. *Front Plant Sci*, 4, 109-119.
- Robert, M. L., Lim, K. Y., Hanson, L., Sanchez-Teyer, F., Bennett, M. D., Leitch, A. R., & Leitch, I. J. (2008). Wild and agronomically important Agave species (Asparagaceae) show proportional increases in chromosome number, genome size, and genetic markers with increasing ploidy. *Botanical Journal of the Linnean Society*, 158(2), 215–222.
- Rojo, V., Martínez-Lage, A., Giovannotti, M., González-Tizón, A. M., Cerioni, P. N., Barucchi, V. C., ... Naveira, H. (2015). Evolutionary dynamics of two satellite DNA families in rock lizards of the genus Iberolacerta (Squamata, Lacertidae): different histories but common traits. *Chromosome Research*, 23(3), 441–461.
- Salina, E. A., Pestsova, E. G., Adonina, I. G., & Vershinin, A. V. (1998). Identification of a new family of tandem repeats in Triticeae genomes. *Euphytica*, 100, 231–237.
- Samoluk, S. S., Robledo, G., Bertoli, D., & Seijo, J. G. (2017). Evolutionary dynamics of an at-rich satellite DNA and its contribution to karyotype differentiation in wild diploid Arachis species. *Molecular Genetics and Genomics*, 292(2), 283–296.
- Schmidt, T., & Heslop-Harrison, J. S. (1998). Genomes, genes and junk: The large-scale

- organization of plant chromosomes. *Trends in Plant Science*, 3(5), 195–199.
- Soltis, P. S., & Soltis, D. E. (2012). *Polyploidy and Genome Evolution*. (P. S. Soltis & D. E. Soltis, Eds.). Gainesville, FL.
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729.
- Tek, A. L., Song, J., Macas, J., & Jiang, J. (2005). Sobo , a Recently Amplified Satellite Repeat of Potato , and Its Implications for the Origin of Tandemly Repeated Sequences. *Genetics*, 1238(July), 1231–1238.
- Torres, G. a., Gong, Z., Iovene, M., Hirsch, C. D., Buell, C. R., Bryan, G. J., ... Birchler, J. a. (2011). Organization and Evolution of Subtelomeric Satellite Repeats in the Potato Genome. *Genes/Genomes/Genetics*, 1(2), 85–92.
- Trofimova, I., Chervyakova, D., & Krasikova, A. (2015). Transcription of subtelomere tandemly repetitive DNA in chicken embryogenesis. *Chromosome Research*, 23(3), 495–503.
- Ugarkovic, D. (2005). Functional elements residing within satellite DNAs. *European Molecular Biology Organization*, 6(11), 5–9. <https://doi.org/10.1038/sj.embor.7400558>
- Vahedian, M., Shi, L., Zhu, T., Okimoto, R., Danna, K., & Keim, P. (1995). Genomic organization and evolution of the soybean SB92 satellite sequence, 11026, 857–862.
- Wolfgruber, T. K., Sharma, A., Schneider, K. L., Albert, P. S., Koo, D. H., Shi, J., Presting, G. G. (2009). Maize centromere structure and evolution: Sequence analysis of centromeres 2 and 5 reveals dynamic loci shaped primarily by retrotransposons. *PLoS Genetics*, 5(11), 13–16.
- Yoong Lim, K., Kovařík, A., Matyasek, R., Chase, M. W., Knapp, S., McCarthy, E., ... Leitch, A. R. (2006). Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section *Alatae*. *Plant Journal*, 48(6), 907–919.
- Zhang, Y., Cheng, C., Li, J., Yang, S., Wang, Y., Li, Z., Lou, Q. (2015). Chromosomal structures and repetitive sequences divergence in *Cucumis* species revealed by comparative cytogenetic mapping. *BMC Genomics*, 16(1), 730-742.
- Zhao, X., Lu, J., Zhang, Z., Hu, J., Huang, S., & Jin, W. (2011). Comparison of the distribution of the repetitive DNA sequences in three variants of *Cucumis sativus* reveals their phylogenetic relationships. *Journal of Genetics and Genomics*, 38(1), 39–45.