



Centro de Investigación Científica de Yucatán A.C.



Posgrado en Ciencias Biológicas

Identificación de familias de retrotransposones LTR en  
tres especies de *Agave*.

Tesis que presenta:

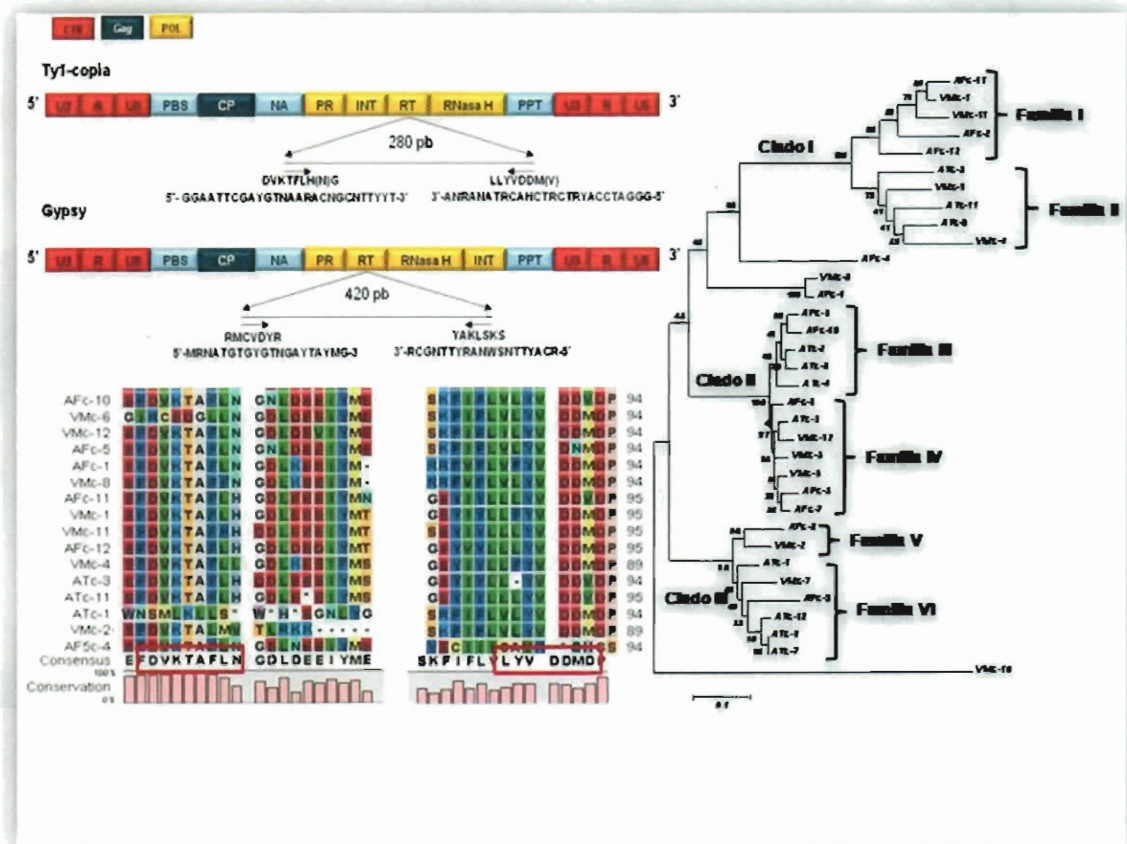
I.B.Q José Javier Huijara Vasconcelos

En opción al título de  
**MAESTRO EN CIENCIAS BIOLÓGICAS**  
Opción Biotecnología



Posgrado en Ciencias Biológicas

Identificación de familias de retrotransposones LTR en tres especies de *Agave*.



Tesis que presenta:

I.B.Q José Javier Huijara Vasconcelos

En opción al título de

MAESTRO EN CIENCIAS BIOLÓGICAS

Opción Biotecnología



## RECONOCIMIENTO



Por medio de la presente, hago constar que el trabajo de tesis titulado identificación de familias de retrotransposones LTR en tres especies de Agave fue realizado en los laboratorios de la Unidad de Biotecnología del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del (los) Dr. Felipe Sánchez Teyer, dentro de la Opción Biotecnología, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Para los efectos que sean necesarios,

Dr. Oscar Moreno Valenzuela

Director Académico

Centro de Investigación Científica de Yucatán A.C.

## DECLARACIÓN DE PROPIEDAD

Mérida, Yucatán, 22 de Marzo 2010

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 dicha información le pertenece en términos de la Ley de la Propiedad Industrial, por lo que no me reservo ningún derecho sobre ello.



I.B.Q José Javier Huijara Vasconcelos.

## **AGRADECIMIENTOS.**

Al Centro de Investigación Científica de Yucatán en especial a la unidad de Biotecnología por proporcionar los equipos necesarios para el desarrollo de este trabajo.

Al CONACyT por el proyecto de ciencia básica No. 24574 y por la beca otorgada No. 213382.

A mi asesor el Dr. Felipe Sánchez Teyer por la excelente formación académica brindada. Sin su asesoría e ideas este trabajo no se hubiese realizado.

A la técnica M.C. Adriana Quiroz Moreno por el apoyo y asesoramiento en las técnicas de clonación, extracción de plásmido, preparación y transformación de células competentes.

Al técnico M.C. Bartolomé Chi Manzanero por el apoyo técnico y asesoramiento en las técnicas de cultivo y transformación de células competentes.

A la técnica M.C. Miriam del Socorro Monforte González por el apoyo técnico y proporcionarnos el protocolo para la preparación de células competentes.

Al Dr. Jorge Alberto Rubio Piña por el apoyo técnico en la preparación de células competentes y clonación de los fragmentos.

A mi comité Tutoral el Dr. Manuel Martínez Estévez, el Dr. Víctor Suarez Solís, el Dr., Luis Carlos Rodríguez Zapata y el Dr. Felipe Sánchez Teyer por las importantes observaciones a este trabajo durante los 4 exámenes tutorales.

A mi comité evaluador el Dr. Manuel Martínez Estévez, el Dr. Víctor Suarez Solís, el Dr., Luis Carlos Rodríguez Zapata, el Dr. Luis Sáenz Carbonell y el Dr. Felipe Sánchez Teyer, por las valiosas observaciones a la redacción del trabajo.

## ÍNDICE GENERAL

ÍNDICE GENERAL.....	i
ÍNDICE DE FIGURAS.....	iv
ÍNDICE DE CUADROS.....	iv
RESUMEN.....	1
ABSTRACT.....	2
CAPÍTULO I ANTECEDENTES.....	3
1.1 INTRODUCCIÓN.....	3
1.2 CLASIFICACIÓN TAXONÓMICA Y DESCRIPCIÓN DEL GÉNERO AGAVE.....	4
1.3 IMPORTANCIA.....	5
1.4 NIVEL DE PLOÍDIA Y TAMAÑO DEL GENOMA.....	6
1.5 ELEMENTO MÓVILES.....	7
1.5.1 ELEMENTOS GENÉTICOS TIPO I (RETROTRANSPOSONES).....	7
1.5.2 ESTRUCTRA DE LOS RETROTRANSPOSONES.....	7
1.5.3 RETROTRANSPOSONES DEL TIPO LTR.....	10
1.5.4 RETROTRANSPOSONES SIN LTR.....	10
1.5.5 MODO DE TRANSPOSICIÓN.....	11
1.6 DISTRIBUCIÓN Y CONTRIBUCIÓN AL TAMAÑO DEL GENOMA.....	13
1.7 LOCALIZACIÓN EN EL CROMOSOMA.....	15
1.8 VARIACIÓN Y POLIMORFISMO DE LAS SECUENCIAS DE LOS RETROTRANSPOSONES.....	17
1.9 JUSTIFICACIÓN.....	19
1.10 HIPÓTESIS.....	20

1.11 OBJETIVO GENERAL .....	20
1.12 OBJETIVOS PARTICULARES.....	20
<b>CAPÍTULO II. AISLAMIENTO DE LOS DOMINIOS PARCIALES DE TRANSCRIPTASA REVERSA.....</b>	<b>27</b>
2.1 INTRODUCCIÓN.....	27
2.2 MATERIALES Y MÉTODOS .....	30
2.2.1 EXTRACCIÓN DE ADN.....	30
2.2.2 AISLAMIENTO DE TR.....	30
2.2.3 CLONACIÓN Y SECUENCIACIÓN.....	31
2.3 RESULTADOS .....	33
2.3.1 EXTRACCIÓN DE ADN.....	33
2.3.2 AISLAMIENTO DE TR.....	33
2.3.3 EXTRACCIÓN DEL PLÁSMIDO .....	34
2.3.4 CARACTERIZACIÓN DE LAS SECUENCIAS .....	36
2.4 DISCUSIÓN.....	38
2.5 BIBLIOGRAFÍA .....	39
<b>CAPÍTULO III. ANÁLISIS DE LAS SECUENCIAS PARCIALES DE RT DE RETROTRANSPONES LTR.....</b>	<b>41</b>
3.1 INTRODUCCIÓN.....	41
3.2 MATERIALES Y MÉTODOS.....	41
3.2.1 COMPARACIÓN GENBANK.....	42
3.2.2 ANÁLISIS FILOGENÉTICO.....	42
3.2.3 IDENTIFICACIÓN DE FAMILIAS .....	42
3.3 RESULTADOS.....	42
3.3.1 TY1-COPIA.....	42

3.3.1.1 COMPARACIÓN DE LAS SECUENCIAS CON LAS REPORTADAS EN EL GENBANK .....	42
3.3.1.2 ANÁLISIS FILOGENÉTICO.....	44
3.3.2 GYPSY .....	48
3.3.2.1 COMPARACIÓN DE LAS SECUENCIAS EN EL GENBANK.....	48
3.3.2.2 ANÁLISIS FILOGENÉTICO.....	49
3.4 DISCUSIÓN.....	52
<b>CAPÍTULO IV. DISCUSIÓN GENERAL.....</b>	<b>59</b>
4.1 CONCLUSIONES.....	64
4.2 PERSPECTIVAS.....	65
4.2 BIBLIOGRAFÍA .....	66



## ÍNDICE DE FIGURAS.

<i>Figura 1.1 Plantas de A. americana en estado de reproducción.</i>	5
<i>Figura 1.2 Estructura general de los retrotransposones LTR.</i>	7
<i>Figura 1.3 Estructura de los retrotransposones tipo LTR.</i>	8
<i>Figura 1.4 Estructura de los retrotransposones sin LTR.</i>	10
<i>Figura 1.5 Transposición de elementos LTR.</i>	12
<i>Figura 1.6 Retrotranscripción del elemento.</i>	13
<i>Figura 1.7 Modelo de distribución de los diferentes tipos de retrotransposones dentro del cromosoma.</i>	16
<i>Figura 2.1 Estructura típica de un retroelemento.</i>	27
<i>Figura 2.2 Extracción de ADN.</i>	33
<i>Figura 2.3 Amplificación por PCR de TR Gypsy y ty1-copia.</i>	34
<i>Figura 2.4. Clonación y extracción del plásmido.</i>	35
<i>Figura 2.5 Traducción in silico.</i>	36
<i>Figura 3.1 Alineamiento de las 34 secuencias Ty1 copia.</i>	45
<i>Figura 3.2 Árbol filogenético Ty1-copia.</i>	46
<i>Figura 3.3 Alineamiento de las 30 secuencias de Gypsy.</i>	50
<i>Figura 3.4 Árbol filogenético Gypsy.</i>	51

## ÍNDICE DE CUADROS

<i>Cuadro 2.1.</i>	29
<i>Cuadro 2.2.</i>	36
<i>Cuadro 3.1.</i>	43
<i>Cuadro 3.2.</i>	48

## RESUMEN

Existen dos tipos de elementos móviles los transposones (tipo II) y retrotransposones (tipo I) que difieren en su mecanismo de transposición. Los primeros cambian de un lugar a otro a través de un mecanismo de corta y pega, mientras que los retrotransposones lo hacen a través de un intermediario de ARN. Se ha demostrado que los retrotransposones están ampliamente distribuidos en el genoma de las plantas y en diversas especies representan una porción considerable de su genoma además de participar en diversos mecanismos relacionados con evolución.

Por otro lado, el género *Agave* ha acompañado al hombre americano por varios siglos, y sus especies son utilizadas como plantas ornamentales y para la producción de fibra, esteroides, pulque, tequila, por mencionar algunos usos. De acuerdo a las características florales, las especies de *Agave* se clasifican en dos subgéneros: *Littae* y *Agave*.

Mencionado lo anterior como marco conceptual, el objetivo de este trabajo fue identificar y caracterizar diferentes familias de retrotransposones Ty1-copia y Gypsy en tres especies de *Agave* (*A. tequilana*, *A. fourcroydes* y *A. angustifolia*).

Para identificar las diferentes familias, se aislaron fragmentos de retrotransposones, por PCR con cebadores degenerados que reconocen un dominio de la región transcriptasa reversa de los retrotransposones Ty1-copia y Gypsy presente en éstos.

Se realizaron alineamientos múltiples y se encontró que los retrotransposones Gypsy son más heterogéneos que los Ty1-copia con un 11% y 22 % de identidad respectivamente.

Con base en el agrupamiento de las secuencias, en este trabajo, se logró identificar 6 familias de retrotransposones Ty1-copia y 2 de Gypsy. Su posible implicación en evolución de poliploides de *Agave* es discutida.

## ABSTRACT

There are two types of mobile elements that differ in their mechanism of transposition; the transposons (type II) and the retrotransposons (type I). Transposons change from one site to another through a cut and paste mechanism, whereas retrotransposons insert a new copy through an intermediary RNA codified by Reverse transcriptase (TR). It has been showed that retrotransposons are widely distributed in all analyzed plant genomes, being a significant portion of the whole genome for several species and could be participating in several mechanisms related with evolution.

*Agave* genus has accompanied the American's man for many centuries, and its species are used as ornamentals and fiber production, steroids, and alcoholic beverages such as pulque, bacanora or tequila, to name a few. According to the floral characteristics the species *Agave* are classified into two subgenus: *Littae* and *Agave*.

Taking in account the above mentioned as a framework, the goal of this study was to identify and characterize different families of retroelements Ty1-copia and Gypsy in three *Agave* species (*A. tequilana*, *A. fourcroydes* and *A. angustifolia*).

To identify the different families of retrotransposons, a group of fragments were isolated by using PCR and degenerated primers that recognize a region of the conserved domain from reverse transcriptase present in both types of retroelementos, Ty1-copia and Gypsy. Based on the isolated sequences a multiple alignments were performed showing that Gypsy retrotransposons are more heterogeneous than Ty1-copia with 11% and 22% of identity respectively. Based on the grouping of the sequences, in this work we identify six families of retrotransposons Ty1-copia and two of Gypsy. Its possible involvement in *Agave* polyploid evolution is discussed.

## Capítulo I Antecedentes

### 1.1 INTRODUCCIÓN

Con base en las características florales, Gentry (1982)<sup>1</sup> agrupó a las especies del género *Agave* en dos subgéneros: *Littae* y *Agave*. Muchas especies de este género son utilizadas como plantas de ornato, en la producción de bebidas alcohólicas (tequila, bacanora, pulque) o la extracción de diferentes compuestos como los esteroides y saponinas<sup>1</sup>. Es un género relativamente reciente con una edad evolutiva entre 7 y 11 millones de años<sup>2</sup>, en comparación con las Angiospermas que tienen una edad evolutiva alrededor de 130 millones de años<sup>3</sup>. Las especies del género *Agave* presentan diferentes niveles de ploidia (desde diploides hasta octoploides) con un cariotipo bimodal de 5 cromosomas grandes y 25 pequeños<sup>4-8</sup>, y su contenido de ADN por genoma haploide (1C) se mantiene constante<sup>8</sup>. Sin embargo, se sabe poco acerca del origen de la ploidia y divergencia de la gran diversidad de especies del género.

En diversas especies vegetales, se han reportado dos tipos de elementos móviles: los transposones (tipo II) y retrotransposones (tipo I) que se diferencian por su secuencia y por el mecanismo de transposición; los transposones cambian de un lugar a otro del genoma a través de un mecanismo de corta y pega, mientras que los retrotransposones lo hacen a través de un intermediario de ARN.<sup>9-11</sup> Los retrotransposones se dividen en dos grupos: los retrotransposones que contienen regiones terminales largas, por sus siglas en inglés (LTR) y aquellos que no presentan este tipo de repeticiones (non LTR)<sup>10-11</sup>. El mecanismo de transposición de los retrotransposones incrementa el número de copias y el tamaño del genoma huésped<sup>12</sup>. Se ha reportado que son componentes ancestrales y que están ampliamente distribuidos en el genoma de las plantas<sup>13-14</sup>, así mismo son fuente de variación genética y en algunas especies son parte fundamental en la organización de los cromosomas, pudiendo participar en el mecanismo de evolución de los genomas de las plantas<sup>15-17</sup>.

Dicho lo anterior, el objetivo de este trabajo fue identificar diferentes familias de retrotransposones presentes en el genoma de tres especies de la sección *Rigidae*, como parte de un estudio integral encaminado a determinar el tipo y origen de los poliploides en *Agave*.

## 1.2 CLASIFICACIÓN TAXONÓMICA Y DESCRIPCIÓN DEL GÉNERO AGAVE.

Reino: Plantae.

División: Magnoliophyta.

Clase: Liliopsida.

Orden: Asparagales.

Familia: Agavaceae

Género: *Agave*.

Sección: Rigidae.

El género *Agave* se ubica dentro de la familia Agavaceae. De acuerdo a su descripción taxonómica, basada en la biología floral y el tamaño y forma de las hojas, se clasifica en dos subgéneros *Littaea* y *Agave*<sup>1</sup>. Según Gentry (1982) el subgénero *Littaeae* incluye 54 especies, 4 subespecies, 6 variedades y 7 formas, distribuidas en 8 secciones, mientras que el subgénero *Agave* está integrado por 82 especies, 21 subespecies y 23 variedades distribuidas en 12 secciones.

Las especies del género son monocarpicas es decir, mueren después de la floración, por lo que su reproducción es predominantemente asexual. La reproducción sexual se lleva a cabo cuando alcanzan la edad adulta (entre 8 y 10 años en la mayoría de los agaves), emergiendo del centro de la roseta un tallo largo que puede alcanzar hasta los 10 m de altura, a la mitad de este crecen ramificaciones con numerosos grupos de flores tubulares<sup>18</sup> (**Figura 1.1**). La reproducción asexual se da a través de bulbillos y rizomas siendo éstos últimos extensiones de la raíz que posteriormente se re-diferencian y se especializan.



**Figura 1.1 Plantas de *A. americana* en estado de reproducción.**

*Muestra el tallo largo que sale de entre la roseta y las flores en la parte final de este.*

### 1.3 IMPORTANCIA

El cultivo de las especies del género *Agave*, durante siglos ha acompañado al ser humano proporcionando una gran variedad de productos como bebidas (pulque, mezcal y tequila), alimento y fibras naturales, entre otros<sup>1</sup>. Una muestra de lo anterior se observa en *Agave fourcroydes* que en el siglo XIX se convirtió en la principal fuente de ingresos para el Estado de Yucatán, y ahora *Agave tequilana*, que por denominación de origen, es la única especie utilizada para la producción de Tequila.

Los *agaves* son plantas suculentas que resisten prolongados periodos de sequia, lo que les da una ventaja adaptativa sobre otras especies bajo estas condiciones. Se ha sugerido que la edad del género *Agave* se ubica entre los 7 y 11 millones de años<sup>2</sup>, si se compara con el origen de las angiospermas hace aproximadamente 132 millones de años<sup>3</sup>, se puede decir que el género *Agave* tiene un origen reciente.

Un género de origen reciente y con una gran diversidad de especies debe poseer una elevada tasa de especiación. Se ha reportado que la tasa de diversificación de las especies del género es de 0.32 especies por millón de año<sup>2</sup> y son las más altas

reportadas para angiospermas cuyo valor varía entre (0.077 y 089)<sup>3</sup> especies por millón de años. Sin embargo; hasta el momento no se conocen con exactitud los eventos que ocasionan esta diversificación, pero se sugiere que el gran número de especies presentes en el género es debido a una radiación adaptativa<sup>19</sup>, iniciada con el proceso de aridificación en Norteamérica que coincide con la fecha de origen del género. Teóricamente, el proceso de aridificación incrementó la temperatura de esta región y disminuyó la disponibilidad de agua, por lo que se puede inferir que el primer motor de la radiación adaptativa fue la adaptación a los nuevos hábitats áridos.

Por otro lado, la gran diversidad topogeográfica y climática de México, genera una gran variedad de ambientes ecológicos (microambientes) propicios para acelerar los procesos de diferenciación genética y de aislamiento poblacional (adaptativos y por deriva génica), dirigiendo así la especiación, así mismo la influencia del hombre también ha contribuido con dicha especiación por medio de la selección y traslado de algunas especies donde antes no existían<sup>1</sup>.

Se ha planteado que el proceso de especiación en los *Agaves* es debido, entre otros, a los rearrreglos cromosómicos producidos por mutaciones heterocigóticas espontaneas como: perdida, translocaciones y fusiones en los cromosomas y a mutaciones puntuales, que han sido observados como variaciones de citotipos numéricos y estructurales<sup>20</sup>.

La alta tasa de diversificación, el origen reciente, el tipo de reproducción y el nivel de ploidía de *Agave*, hacen a este género un modelo de estudio interesante en evolución de poliploides.

#### **1.4 NIVEL DE PLOÍDIA Y TAMAÑO DEL GENOMA**

Las especies del género *Agave* presentan niveles de ploidía que van desde diploides (*A. tequilana* var. Azul) hasta octoploides (*A. decipiens*) y se ha sugerido que esto pudo haberse originado tanto por aloploidía como autopoliploidía<sup>4-7</sup>. Por otro lado, su genoma haploide es casi 7 veces más grande que el de *Arabidopsis* y se ha reportado que aumenta proporcionalmente con el nivel de ploidía, sin embargo su valor 1C permanece constante<sup>8</sup>. Es importante destacar que independientemente del nivel de ploidía, las especies del género *Agave* presentan un cariotipo bimodal con 5 cromosomas grandes y 25 de cromosomas chicos<sup>5, 7, 8, 21</sup>.

## 1.5 ELEMENTO MÓVILES

Los elementos genéticos móviles son secuencias de ADN que se mueven dentro del genoma. De acuerdo a la forma en que se mueven, se conocen dos tipos: tipo I (Retrotransposones) y II (transposones). Los elementos del tipo II cambian de un lugar a otro por un mecanismo de corte y pega, mientras que los del tipo I se mueven a través de un intermediario de ARN (copia-pega)<sup>9</sup>.

### 1.5.1 ELEMENTOS GENÉTICOS TIPO I (RETROTRANSPONES)

Los elementos móviles del tipo I, a diferencia del tipo II se transponen a partir de un intermediario de ARN y están clasificados en dos grandes grupos los LTR retrotransposones y los Non-LTR retrotransposones<sup>11</sup>. Están ampliamente distribuidos en el genoma de las plantas<sup>13-14</sup>. Los retrotransposones tienen características muy parecidas a los retrovirus, lo que hace suponer que pudieron haber evolucionado a partir de éstos, integrándose como parte constituyente del genoma del hospedero, sin causar daños letales y desde entonces han estado presentes en el genoma de las plantas<sup>10,22</sup>.

### 1.5.2 ESTRUCTRA DE LOS RETROTRANSPONES

Los retrotransposones con la región LTR son elementos móviles entre 2-15 Kb de tamaño, se mueven a través de un intermediario de RNA que contiene los componentes necesarios que le permiten copiarse e insertarse en otro sitio del genoma pero manteniendo la copia parental<sup>11</sup>. La estructura general de los retrotransposones se muestra en la (Figura 1.2) y a continuación se describen cada una de las regiones.



**Figura 1.2 Estructura general de los retrotransposones LTR.**

*Los retrotransposones están flanqueados por una región LTR en sus extremos que contiene secuencias reguladoras que le permiten transcribirse, copiarse a ADN e insertarse en otro sitio del genoma. La región Gag codifica para una proteína que funciona como la cápside de los virus. La región Pol codifica para una proteasa, transcriptasa reversa, integrasa y ribonucleasa (RNasa H).*



### Región terminal larga repetida (por sus siglas en inglés LTR).

La secuencia LTR es una secuencia entre 100 pb y 5 kb situadas en ambos extremos del retrotransposon y hasta el momento se sabe que no codifica para algún gen conocido, sin embargo, contiene las secuencias promotoras y terminadoras asociadas con la transcripción del retrotransposon LTR<sup>11</sup>. La transcripción inicia en 5'LTR y termina en la región 3'LTR.

Como se muestra en la (Figura 1.3) la región LTR contiene:



**Figura 1.3 Estructura de los retrotransposones tipo LTR.**

Los retrotransposones contienen en sus extremos una región LTR (rojo), una región Gag (Azul marino) y una región pol (Amarillo).

### Región LTR

- Una región U3 que está situada desde el lugar de inicio de la región LTR hasta el inicio de la transcripción y contiene una secuencia promotora (caja TATA) para la ARN pol II<sup>11, 23</sup>.
- Una región R que es la zona comprendida entre el inicio y el final de la transcripción, se encuentra en ambos extremos del transcrito y por ser repeticiones invertidas permiten a la cadena corta 3'—5' de cADN circularizarse por complementariedad de la 5'-LTR a la 3'-LTR durante la transcripción inversa.
- Una región U5 que se encuentra desde el sitio de polia-adenilación (final de la transcripción) y el final del LTR.
- PBS (Primer Binding Site) es una secuencia complementaria a una parte del Met<sup>t</sup>RNA. El Met<sup>t</sup>RNA funciona como un cebador que reconoce la transcriptasa reversa (TR) para el inicio de la síntesis de la cadena de cADN.

- PPT (Poly Purine Track) actúa como lugar de inicio de la síntesis de la segunda cadena de cADN durante la retrotranscripción. Esta región no es detectada por la RNasa H y por lo tanto no es degradada, esta región es reconocida por la TR para la síntesis de la segunda cadena.

### La región Gag

La región **Gag** codifica un tipo de proteína que sirve para formar una estructura similar a la capsida viral. Esta región a nivel de secuencia es poco conservada<sup>11</sup>.

### Región Pol

La región **pol** esta situada siempre a continuación de la región *gag*, se traduce por uno o diferentes marcos de lectura como un solo polipéptido, que posteriormente se fragmentan en diferentes proteínas con diferentes actividades catalíticas<sup>11,23</sup>. Los constituyentes de la región Pol se muestran a continuación.

- La proteasa (PT) fragmenta al polipéptido dando lugar a las diferentes proteínas funcionales<sup>23</sup>.
- La transcriptasa reversa (TR) tiene actividad ADN polimerasa y puede copiar el ARN genómico a cADN, la TR es la región más conservada a nivel aminoácidos entre los LTR y retrovirus<sup>9</sup>. En la secuencia de la región que codifica para la TR existen una serie de secuencias altamente conservadas (dominio) tales como YVDDLL y TAFLHG<sup>14</sup>, para el cual se han diseñado cebadores degenerados para amplificar por PCR la región comprendida entre este dominio de los LTR-retrotransposon tipo *Ty1-copia*.
- Una ribonucleasa (RNasa H) con actividad 3' y 5' exonucleasa que degrada el ARN del híbrido ARN-ADN en la retrotranscripción, este dominio también es conservado dentro de los retrotransposones para los cuales se han definido dos motivos altamente conservados, en los distintos retrotransposones: el motivo1 RTKHI(D/E) y el motivo2 AD(I/M/F/L)LTK<sup>24</sup>,
- Una región *int* que codifica para una integrasa o endonucleasa (IN) que es necesaria para la inserción del cADN del retrotransposon en el genoma.
- Una *EN* que codifica para glicoproteínas de la cubierta que reconocen proteínas receptoras de la superficie de la célula huésped permitiendo su fusión con la membrana

plasmática y posteriormente la entrada a la célula facilitando la movilidad celular y su transposición en otras células.

### 1.5.3 RETROTRANSPOSONES DEL TIPO LTR.

De acuerdo a la similitud de sus secuencias y al orden en que se codifican los polipéptidos de la región *pol* los retrotransposones LTR se dividen en dos grupos los Ty1-*copia* y Ty3 *gypsy*. Los Ty1-*copia* transcriben la región *pol* en el siguiente orden INT-TR-RNasa H, mientras que, los Ty3 *gypsy* primero transcriben la TR luego la RNasa H y por último la INT<sup>11</sup> (Figura 1.3).

### 1.5.4 RETROTRANSPOSONES SIN LTR.

Los retrotransposon sin LTR (Non-LTR) no están flanqueados por la región LTR, sin embargo sus extremos presentan una secuencia repetida simple (UTR) que va seguida, generalmente, de una cola de poli(A). La ARN polimerasa II reconoce un promotor interno dentro del retrotransposon sin LTR para iniciar la transcripción.

Dependiendo del tamaño y la capacidad con que se transponen estos elementos pueden dividirse en elementos LINE (long interspersed repetitive elements) y SINE (short interspersed repetitive elements)<sup>11</sup>.



Figura 1.4 Estructura de los retrotransposones sin LTR.

### Elementos LINE (Elementos Nucleares Intercalados Largos).

Los elementos LINE (Long Interspersed Nuclear Element) son de un tamaño entre 1- 8 kb conteniendo uno o dos marcos de lectura y su extremo 3' presenta una cola de poli(A) o una secuencia repetida. El primer marco de lectura parece ser equivalente a la secuencia *gag* y el segundo a la región *pol*, sin embargo no contiene las secuencias que codifican para la actividad proteasa e integrasa. La expresión involucra la presencia de promotores

para ARN pol II y III, su transcripción reversa ocurre durante la integración de este directamente en el lugar de inserción<sup>11,23</sup>.

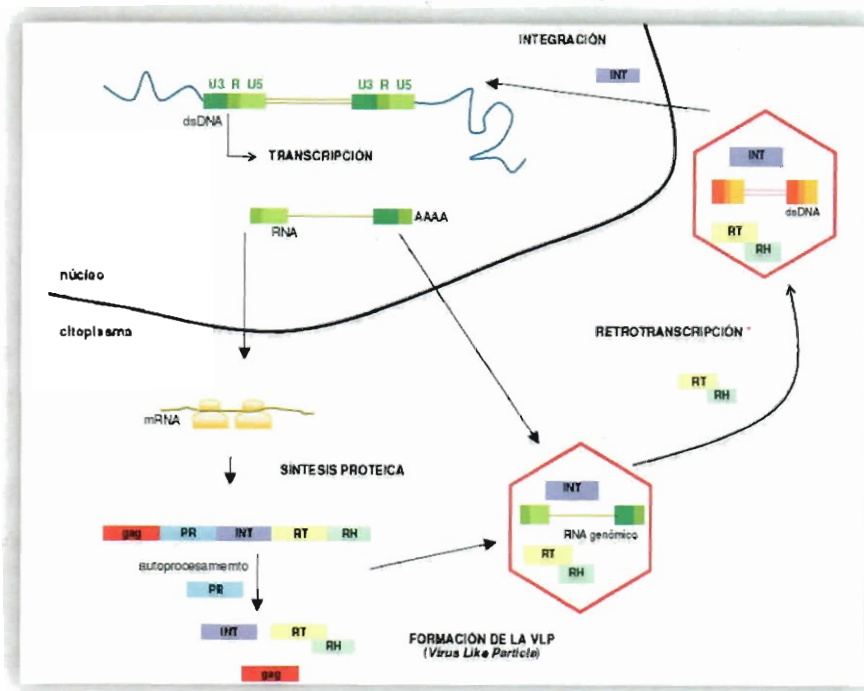
#### **Elementos SINEs (Elementos Nucleares Intercalados Cortos).**

Los pequeños retrotransposones (100 a 300 pb) son diferentes de las otras clases de retrotransposones, no codifican para ninguna proteína funcional de la transposición, sin embargo tienen las regiones promotoras para la Pol III y las secuencias característica involucradas en integración por proteínas codificadas por los elementos LINE ó LTR-retrotransposones actuando como elementos<sup>11, 23</sup>.

#### **1.5.5 MODO DE TRANSPOSICIÓN.**

La transposición inicia con la transcripción del mRNA sintetizado por la Pol II celular en el extremo 5'-LTR y termina en el extremo 3'-LTR formándose una molécula de mRNA. Posteriormente el mRNA es traducido en un complejo proteico (**Figura 1.5**) y la actividad de la proteasa de la región Pol fragmenta y libera las diferentes proteínas funcionales descritas en la sección 1.5.2. Posteriormente, la proteína codificada por la región Gag se une al ARN encerrándolo en partículas parecidas a los virus<sup>25</sup>.

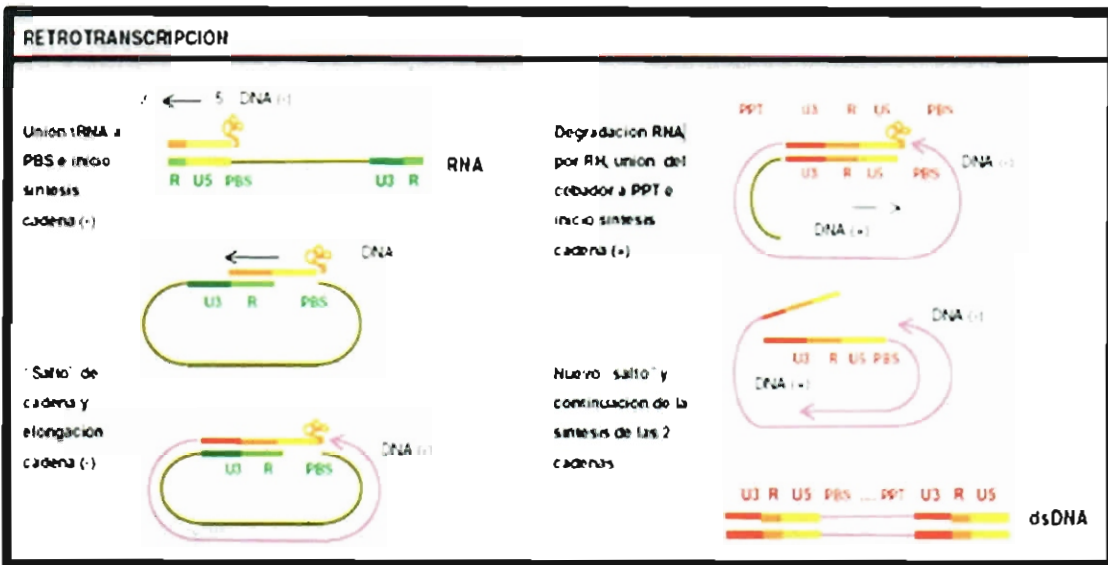
El mecanismo de transcripción inversa inicia con la unión del <sup>Met</sup>tARN celular a la región PBS en el extremo 5'LTR, esta región sirve como templado para que la TR inicie la síntesis de una región corta del extremo LTR en dirección 5'—3', esta deja de copiar cuando encuentra una secuencia de terminación en la región R de la cadena, la RNasa H degrada la región R y U5 del ARN y a continuación la cadena corta de ADN sintetizada se une por complementariedad al extremo 3'LTR en las regiones R de ambos extremos. Esta región sirve como templado para continuar la síntesis de la cadena hasta llegar a la señal de parada de la región R del extremo 3', conforme se avanza en la síntesis del ADN, la RNasa H va degradando el ARN del híbrido ARN-ADN. La síntesis de la segunda cadena de ADN se inicia cuando la TR se une a la región PPT utilizándolo como iniciador para iniciar la síntesis de las regiones U3-R-U5-PBS.



**Figura 1.5 Transposición de elementos LTR**

Esquema general de la transposición de los elementos LTR<sup>26</sup>. Este tipo de elementos se transponen a partir de un intermediario de mRNA, que codifica para las proteínas involucradas en la transposición. Las proteínas se ensamblan en una estructura parecida a la viral donde se lleva a cabo la retrotranscripción por la acción de la TR y RNasa H y posteriormente la integrasa (IN) es la encargada de abrir las cadenas del ADN blanco y por último la nueva copia del retrotransposon es integrada al ADN.

La nueva cadena se circulariza por complementariedad al extremos 5', ahora al PBS, se continúa con la síntesis de la segunda cadena de ADN<sup>25</sup> (Figura 1.6). Cuando el nuevo retroelemento está listo, la INT unida al LTR realiza un corte en el ADN blanco en donde el nuevo LTR es insertado. Los extremos del ADN que fueron cortados son nuevamente reparados y ahora hay una nueva copia del retrotransposon en el ADN<sup>25</sup>



**Figura 1.6 Retrotranscripción de los elementos LTR<sup>26</sup>.**

En la etapa inicial de la retrotranscripción una región del tARN que es complementaria al PBS funciona como cebador para iniciar la retrotranscripción, y sintetizar la región U5 y R del extremo 3'. Posteriormente la cadena forma un bucle debido a que los extremos se unen por complementariedad en la región R posteriormente se sintetiza la cadena complementaria al ARN y la cadena del mRNA es degradado por la RNasa H. Por último ocurre otro cambio de lugar de la cadena complementaria en la región U3-R-U5-PBS y por último se termina de sintetizar la segunda cadena.

### 1.6 DISTRIBUCIÓN Y CONTRIBUCIÓN AL TAMAÑO DEL GENOMA.

Los retrotransposones están ampliamente distribuidos en los organismos eucariotas, desde organismos unicelulares hasta organismos superiores<sup>14, 27</sup>. La distribución de los retrotransposones entre las plantas, puede explicarse, asumiendo que ya estaban presentes desde las algas verdes<sup>28</sup>. Estas regiones han divergido en grupos heterogéneos en las plantas actuales y pueden haber sido transmitidos de manera vertical u horizontal a las especies descendientes, como lo han hecho otros componentes del genoma<sup>14, 29-31</sup>.

Kumekawa y colaboradores (1999)<sup>11</sup> aislaron y secuenciaron una región parcial de TR en especies de angiospermas (monocotiledóneas y dicotiledóneas), coniferófitos, ginkgofias, pteridofitas, briofitas y algas, y se observó que estas secuencias quedaron agrupadas en dos clados principales y dentro de cada clado se agrupan en diferentes subfamilias de retrotransposones, que interesantemente, las secuencias de mono y dicotiledóneas siempre quedan agrupadas. Lo anterior sugiere que los retrotransposones Ty1 Copia divergieron después del origen de mono y dicotiledóneas y que la transmisión vertical (que transmite de generación en generación) ha tenido un rol importante en la evolución de los retrotransposones en mono y dicotiledóneas.

Debido a que los retrotransposones generan una copia de sí mismos que se inserta en algún lugar del genoma, conservando la copia original, puede contribuir significativamente a incrementar el tamaño del genoma. Por ejemplo, el genoma de Maíz está representado por el 80 % de este tipo de elementos móviles<sup>12</sup>, mientras que en arroz representan el 17%<sup>32</sup>, 4% en *Arabidopsis thaliana*, 3.1% en *Saccharomyces servisiae*<sup>32</sup>, entre otros.

En cinco especies de trigo se identificaron siete familias de transcriptasa reversa que son comunes a todas las especies analizadas, lo cual sugiere que estos retrotransposones se originaron antes de la especiación<sup>34</sup>.

La comparación de los elementos presentes en Maíz (2290 Mpb) y Sorgo (735 Mpb), que divergieron de un ancestro común aproximadamente hace 16 millones de años, ha sugerido que las diferencias en el tamaño de estas dos especies fue debida a la actividad de los retrotransposones que estuvieron activos durante su divergencia<sup>12</sup>.

Por otro lado, Flavell y colaboradores (1992)<sup>14</sup> utilizaron la PCR para aislar fragmentos de Ty1 copia en 6 especies de la división Briofitas, 6 pteridofitas, 9 monocotiledóneas y 29 dicotiledóneas, sugiriendo que los retrotransposonesTy1-copia están ampliamente distribuidos en las plantas. El análisis filogenético de sus secuencias mostro que el grado de divergencia, en algunas especies es proporcional a la distancia evolutiva entre ellas.

Hirochika & Hirochika (1993)<sup>35</sup>, amplificando una región conservada de TR con cebadores degenerados diseñados para reconocer el dominio conservado de aminoácidos QMDVKT y YVDDM demostraron que los retrotransposones están ampliamente distribuidos en 29 especies de angiosperma, tres gimnosperma, dos briofitas y una pteridofitas.

Los estudios realizados por Flavell y colaboradores (1992)<sup>14</sup> y Hirochika & Hirochika (1993)<sup>34</sup> incluían plantas que se reproducen sexualmente y se había sugerido que los retrotransposones podrían persistir solamente en estas plantas<sup>36-37</sup>. Sin embargo, amplificando dominios de TR de 4 especies de plantas que se reproducen asexualmente se demostró que los retrotransposones pueden estar presentes en cualquier planta independientemente del mecanismo de reproducción<sup>38</sup>. A la fecha no existe evidencia que sugiera que estos jueguen un papel predominante relacionado con mecanismos adaptativos de acuerdo al tipo de reproducción.

### **1.7 LOCALIZACIÓN EN EL CROMOSOMA.**

La localización de los retrotransposones en regiones específicas de los cromosomas puede sugerir alguna función relacionada con el sitio donde se encuentra.

Se han identificado secuencias de retrotransposones en regiones centroméricas y se ha sugerido que estos participan en la división celular. Las histonas se unen a estas secuencias permitiendo la migración de los cromosomas en la división celular<sup>15-17</sup>.

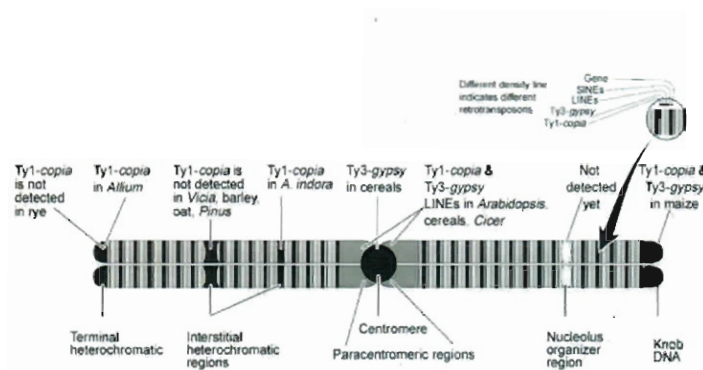
Sondas similares a la secuencia de Ty1-copia fueron localizadas solo en regiones centroméricas de *Sorghum bicolor*, *Oriza sativa*, *Bambusa vulgaris*, *Saccharum officinaru*, *Hordeum vulgare*, *Agropyrun intermedium*, *Brachypodium syvaticum*, *Avena sativa* entre otras<sup>16</sup>. La presencia de Ty3-copia en las regiones centroméricas de todas las especies analizadas sugiere que estos retrotransposones son inserciones ancestrales que se originaron antes de la divergencia de las gramíneas y se han desarrollado independientemente insertándose preferencialmente en regiones centroméricas. Además el grado de conservación de las secuencias de este retrotransposon en regiones centroméricas, entre la familia de las gramíneas, indican una posible asociación de este tipo de elemento móvil con las funciones centroméricas<sup>16</sup>

Por otro lado Belyayev y colaboradores (2001)<sup>39</sup> utilizaron las secuencias de dominios de TR de tres tipos de retrotransposones Ty1-copia, Ty3-gypsy and LINE para realizar una hibridación *in situ para ver* su distribución dentro del cromosoma en dos especies de *Triticum*, Revelando que estos están presentes en regiones génicas (heterocromatinas, eucromatina y en regiones pericentroméricas cerca de las heterocromatinas). Adicionalmente proponen que este tipo de elementos móviles pudieron causar mutaciones



al insertarse en las regiones cercanas a los genes (modificando su expresión) o entre genes (cambiando la funcionalidad del gen). En otro estudio recientemente con los mismos dominios para TR de *Ty1 copia* se encontró que éstos están distribuidos a lo largo del cromosoma excepto en regiones centroméricas y sitios de rADN en especies de la familia Asteraceae<sup>40</sup>.

Los LINES y SINE están distribuidos por todo el cromosoma en tres especies del género *Beta* (*B. vulgaris*, *B. iomatogona*, *B. nana*), en tres especies monocotiledóneas (*Allium cepa*, *O. sativa* y *Secale cereale*) y dos dicotiledóneas (*Nicotiana tabacum* y *Antirrhinum majus*)<sup>41</sup>. Así mismo, Kubis y colaboradores (1998)<sup>41</sup> observaron por hibridación *in situ* de secuencias de TR de los elementos LINE de la remolacha (*B. vulgaris*) que los cromosomas en metafase hibridan solo en regiones subtelerómicas de los brazos de los cromosoma.



**Figura 1.7 Modelo de distribución de los diferentes tipos de retrotransposones dentro del cromosoma.**

Los retrotransposones están distribuidos a lo largo de todo el cromosoma. Sin embargo cada grupo de retrotransposones también tiene familias que están presentes o ausentes en regiones específicas del cromosoma, por ejemplo, en los cereales algunos *Ty3-gypsy* están agrupados en regiones centroméricas. Mientras que, *Ty1-copia* está distribuido en regiones terminales de heterocromatinas del cromosoma de cebolla (*Allium cepa*). También muchos retrotransposones se encuentran preferencialmente en regiones eucromáticas en donde se encuentran localizados regiones génicas<sup>11</sup>.

## **1.8 VARIACIÓN Y POLIMORFISMO DE LAS SECUENCIAS DE LOS RETROTRANSPOSONES.**

Al comparar los retrotransposones de una misma familia se observa que no son idénticos y muestran polimorfismo entre especies y familias de plantas. Hay varias causas relacionadas con el polimorfismo de los retrotransposones. Se ha sugerido que las principales causas son por la relación del elemento mismo con el genoma hospedero y por factores externos<sup>23</sup>.

A continuación se describen las principales causas reportadas de mutación en retrotransposones:

### **Variación ocasionada por el elemento mismo.**

La principal variación es la que se provoca por la transcriptasa reversa que muestra una tasa de mutación de  $2.5 \times 10^{-05}$  errores/ciclo, 100 a 1000 veces más que la ADN polimerasa del genoma huésped<sup>42</sup>. Diferentes errores de la transcripción han sido identificados: baja discriminación de dNTPs, modificación de los marcos de lectura, cambio de bases por no tener la actividad exonucleasa 3' o 5' (no hay un mecanismo de reparación) y la adición de bases al extremo final de la secuencia sintetizada modifica la región LTR<sup>43-44</sup>. Todos estos errores pueden ocasionar la acumulación de copias deficientes o pseudoretrotransposones perdiendo la capacidad de transponerse independientemente.

### **Variación en el genoma de la especie.**

Los procesos de duplicación de ADN y de recombinación son una fuente de variación genética que ocasiona polimorfismos en el genoma de las células. Aunque las polimerasas de la célula tienen la capacidad de corregir los errores en la duplicación, la tasa de mutación de la polimerasa es alrededor de 1 por cada  $10^6$  nucleótidos<sup>45</sup>. Las secuencias repetidas en el ADN generan puntos calientes para que ocurra la recombinación, los retrotransposones pueden generar estos puntos de recombinación<sup>11</sup>.

Los retrotransposones pueden ser una fuente de variación, cambiando los niveles de expresión génica llevando a la formación de nuevos fenotipos, además de que pueden ser puntos donde ocurra la recombinación.

Por su mecanismo de transposición quizá participen en mecanismos de especiación. Por ejemplo, si un retroelemento se insertara en genes involucrados en la síntesis de atrayente de polinizadores, pudiera propiciar un cambio en el modo reproductivo y si la selección natural permite la preservación de este cambio se originaría una nueva especie. Sin embargo, se necesitan más estudios para evaluar la implicación de los posibles cambios mediados por retrotransposones en la evolución.

En *Agave* los retrotransposones pueden estar asociados con el tamaño del genoma y pudieran participar en el origen de los diferentes citotipos así como los procesos de poliploidización, relacionados con la diversificación de las especies del género.

## 1.9 JUSTIFICACIÓN

El género *Agave* es relativamente joven en términos evolutivos y con una elevada tasa de diversificación; sin embargo, no se conocen los mecanismos que han favorecido tal diversificación.

Por otro lado, hasta la fecha se desconoce el papel que juegan los retrotransposones en el mecanismo de evolución, de su distribución y localización en el genoma y si estos pudieran generar variabilidad genética dentro de las especies en este género. El análisis de los retrotransposones en diversas especies de *Agave* permitirá un acercamiento para generar conocimiento acerca del papel que estos elementos juegan en el genoma, y se obtendrá información acerca del número de familias de elementos y las relaciones que guardan en cada una de ellas a un nivel intraespecífico, como base para estudios evolutivos posteriores.

### **1.10 HIPÓTESIS.**

Si los retrotransposones están ampliamente distribuidos en el genoma de las plantas, y las familias de estos se originan a través de la diversificación de las especies. Entonces es posible encontrar diferentes familias de retrotransposones que se han originado a través de la diversificación de *A. tequilana*, *A. fourcroydes* y *A. angustifolia*.

### **1.11 OBJETIVO GENERAL**

Identificar diferentes familias de retrotransposones LTR con base en dos dominios conservados de transcriptasa reversa en *A. tequilana*, *A. fourcroydes* y *A. angustifolia*.

### **1.12 OBJETIVOS PARTICULARES.**

- Aislar las diferentes familias de TR en *A. tequilana*, *A. fourcroydes* y *A. angustifolia*
- Caracterizar las diferentes familias de TR con base a la identidad de sus secuencias de aminoácidos con las reportadas.
- Identificar diferentes familias de retrotransposones encontradas en estas tres especies.

### 1.13 BIBLIOGRAFÍA.

1. H. S. Gentry *Agaves of continental north America*, (The university of Arizona press), 1982.
2. S. V. Good-Avila; V. Souza; S. Gaut Brandon; L. E. Eguiarte, Timing and rate of speciation in *Agave* (Agavaceae). *PNAS*. 103:24, 9124–9129, 2006.
3. S. Magañón; M. J. Sanderson, Absolute diversification rates in angiosperm clades. *Evolution*, 55:9, 1762–1780, 2001,
4. L. R. Doughty, Chromosome behaviour in relation to genetics of *Agave*. I. seven species of fibre *Agave*. *J. Genetics*, 33, 197-205, 1936.
5. G. Backman, A Karyosystematic study of the genus *Agave*. *Am. J. Bot.* 32:5, 283-298, 1944.
6. M. Bennett; I. Leitch, Nuclear DNA amounts in Angiosperms. *Annals of Botany*. 76, 113-176, 1995.
- 7 G. Palomino; J. Martínez; I. Méndez, Citotipos en *Agave angustifolia* Haw. determinados por citometria de flujo y análisis de sus cariotipos. *Rev. Int. Contam. Ambient*, 21:1, 49-54, 2005.
8. M. L. Robert; K. Y. Lim, L. Hanson; F. Sanchez-Teyer; M. D. Bennett; A. R. Leitch; I. J. Leitch, 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, 215–222, 2008.
9. Y. Xion; T. H. Eickbush, Similarity of Reverse Transcriptase-like Sequences of Viruses, Transposable Elements, and Mitochondrial Introns. *Mol. Biol. Evol.* 5:6, 675-690, 1988.
10. Y. Xion; T. H. Eickbush, Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO J.* 9:10, 3353-3362, 1990.
11. A. Kumar; J. L. Bennetzen, Plant retrotransposons. *Annu Rev. Genet.* 33, 429-532, 1999.

12. P. San Miguel; J. I. Bennetzen, Evidence that a Recent Increase in Maize genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. *Annals of Botany* 82: A pp 37-44 (1998).
13. D. F. Voytas; M. P. Cummings; A. Konieczny; F. M. Ausubel; S. R. Rodermel, Copia-like retrotransposons are ubiquitous among plants. *PNAS*, 89, 7124–7128, 1992.
14. Flavell A. J., Smith D. B. and Kumar A. 1992. Extreme heterogeneity of Ty1-copia group retrotransposons in plants. *Mol. Gen. Genet.* 231: pp 233-242 (1992)
15. E. V. Ananiev; R. L. Phillips; H. W. Rines Complex Structure of Knob DNA on Maize Chromosome 9: Retrotransposon Invasion into Heterochromatin. *Genetics.* 149, 2025–2037, 1998.
16. J. T. Miller; F. Dong; S. A. Jackson; J. Song; J. Jiang, Retrotransposon related DNA sequences in the centromeres of grass chromosomes. *Genetics*, 150, 1615–1623, 1998.
17. G. G. Presting; L. Malysheva; J. Fuchs; I. Schubert, A TY3/GYPSY retrotransposon-like sequence localises to the centromeric regions of cereal chromosomes. *Plant J.* 16, 721–728, 1998.
18. J. Cervantes-Martínez; R. Flores-Hernández; B. Rodríguez-Garay, F. Santacruz-Ruvalcaba, Detection of bacterial infection of Agave plants by laser-induced fluorescence. *Appl. Opt.* 41, 2541–2545, 2002.
19. P. Colunga-García Marín; A. Larqué Saavedra; L. E. Eguiarte & D. Zizumbo-Villarreal, En lo ancestral hay futuro: del tequila, los mezcales y otros agaves. CICY-CONACYT-CONABIO-INE. Mérida, Yucatán. 2007, 402
20. L. Castorena-Sanchez; R. M. Escobedo; A. Quiroz, New cytological determinants recognized in six taxa of *Agave* in the section *rigidae* and *sisalanae*. *Can. J. Bot.* 69: 1257-1264, 1991
21. S. F. Moreno Salazar; M. Esqueda; J. Martínez; G. Palomino, Tamaño del genoma y cariotipo en *Agave angustifolia* y *A. rhodacantha* de Sonora, México. *Rev. Fitotecnia Mexicana*, 30:1, 13-23, 2007.

22. M. S. Springer; R. J. Britten, Phylogenetic Relationships of Reverse Transcriptase and RNase H Sequences and Aspects of Genome Structure in the Gypsy Group of Retrotransposons. *Mol. Biol. Ed.* 10:6, 1370-1379, 1993.
23. F. Sabot; D. Simon; M. Bernard, Plant transposable elements, with an emphasis on grass species. *Euphytica* 139, 227–247, 2004.
24. S. R. Pearce; C. Stuart-Rogers; M. R. Knox; A. Kumar; T. H. N. Ellis; A. J. Flavell, Rapid isolation of plant Ty1-copia group retrotransposon LTR sequences for molecular marker studies. *The plant J.* 19:6, 711-717, 1999.
25. F. Sabot; A. H. Schulman, Parasitism and the retrotransposon life cycle in plants: a hitchhiker's guide to the genome. *Heredity* 97, 381–388, 2006.
26. L. Rico Cabanas, Nuevos retrotransposones de naranjo. Desarrollo de marcadores basados en retrotransposones para la evaluación de la biodiversidad de los cítricos. Tesis Doctoral en el Departamento de Bioquímica y Biología Molecular de la universidad de Barcelona, España. 2005.
27. N. Kumekawa; E. Ohtsubo; H. Ohtsubo, Identification and phylogenetic analysis of gypsy-type retrotransposons in the plant kingdom. *Genes Genet. Syst.* 74, 299-307 1999.
28. Flavell A. J., Pearce S. R., (J.S.) Heslop-Harrison P. and Kumar A. 1997. The evolution of Ty1-copia group retrotransposons in eukaryote genomes. *Genetica* 100: 185–195 1997
29. A. Belyayev; O. Raskina; A. Korol; E. Nevo, Coevolution of A and B genomes in allotetraploid *Triticum dicoccoides*. *Genome* 43, 1021–1026, 2000.
30. S. M. Tam; M. Caussem; C. Garchery; H. Burck; C. Mhiri; M. A. Grandbastien, The distribution of copia-type retrotransposons and the evolutionary history of tomato and related wild species. *Journal compilation. European society for evolutionary biology.* 20, 1056–1072, 2007.
31. A. Dixit; J. Ho Ma; J. Woong Yu; E. Gi Cho; J. Jin Park, Reverse transcriptase domain sequences from Mungbean (*Vigna radiata*) LTR retrotransposons: Sequence characterization and phylogenetic analysis. *Heredity*, 97, 381–388, 2006.



32. S. P. Wang; N. Liu; K. M. Peng; Q. F. Zhang, The distribution and copy number of copia-like retrotransposons in rice (*Oryza sativa*) and their implications in the organization and evolution of the rice genome, *PNAS*, 96, 6824–6828, 1999.
33. J. M. Kim; S. Vanguri; J. D. Boeke; A. Gabriel; D. F. Voytas, Transposable elements and genome organisation: A comparative survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.* 8, 464–478, 1998.
35. H. Hirochika; R. Hirochika, Ty1-copia group retrotransposons as ubiquitous components of plant genome. *Jpn. J. Genet.* 68, 35–46, 1993.
36. D. A. Hickey, Selfish DNA a sexually-transmitted nuclear parasite. *Genetics*, 101, 519–531, 1982.
37. S. Wright; D. Finnegan, Genome evolution sex and the transposable element. *Curr. Biol.* 11, 296–299, 2001.
38. T. R. Docking; F. E. Saadé; C. M. Elliot; D. J. Schoen; Retrotransposon sequence variation in four asexual plant species. *J. Mol. Evol.* 62, 375–387, 2006.
39. A. Belyayev; O. Raskina; E. Nevo, Chromosomal distribution of reverse transcriptase containing retroelements in two Triticeae species. *Chromosome Res.* 9, 129–136, 2001.
40. C. F. Ruas; H. Weiss-Schneeweiss; T. F. Stuessy; M. R. Samuel; A. Pedrosa-Harand, K. Tremetsberger; P. M. Ruas; P. M. Schlüter; M. A. Ortiz Herrera; C. König; N. I. Matzenbacher, Characterization, genomic organization and chromosomal distribution of Ty1-copia retrotransposons in species of *Hypochaeris* (Asteraceae). *Gene*, 15:412 (1-2), 39–49, 2008.
41. S. E. Kubis; J. S. Heslop-Harrison; C. Desel; T. Schmidt, The genomic organization of nonLTR retrotransposons (LINEs) from three Beta species and five other angiosperms *Plant Molecular Biology*, 36, 821–831, 1998.
42. A. Gabriel; M. Willems; E. H. Mules; J. D. Boeke, Replication infidelity during a single cycle of Ty1 retrotransposition. *Proc. Natl. Acad. Sci. USA* 93, 7767–7771 1996

43. B. Preston, Error-prone retrotransposition: Rime of the ancient mutators Commentary, PNAS, 93, 7427-7431, 1996.

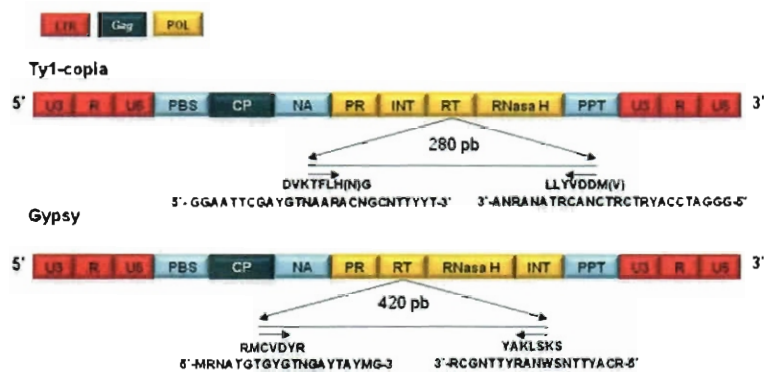
44. M. Boutabout; M. Wilhelm; F.-X. Wilhelm, DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1. Nucleic Acids Research. 29:11, 2217-2222, 2001.

45. B. Lewin 2000. Genes VII. Oxford University Press 2000.

## Capítulo II. Aislamiento de los dominios parciales de transcriptasa reversa

### 2.1 INTRODUCCIÓN.

Los retrotransposones LTR son elementos móviles que a diferencia de los transposones, se insertan en el genoma por el mecanismo de copia y pega, a través de un intermediario de ARN. Constan de tres regiones principales: la región LTR contiene secuencias regulatorias de la transcripción y transcripción reversa, la región Gag que es similar a la estructura que forma la cápside de los retrovirus y una región POL encargada del mecanismo de la transcripción reversa (**Figura 2.1**). Se dividen en dos grandes grupos los retrotransposones Ty1-copia y los Gypsy. Se diferencian principalmente por el orden en que codifican sus genes (**Figura 2.1**) y por las diferencias en sus secuencias<sup>1</sup>.



**Figura 2.1 Estructura típica de un retroelemento**

Representación esquemática de los retrotransposones LTR. La figura representa las regiones estructurales y las diferencias de los retrotransposones Ty1-copia y Gypsy. La RT está localizada en la región POL en el extremo 3' de la integrasa de los retrotransposones Ty1-copia, mientras que en los Gypsy se encuentra en el extremo 3' de la proteasa. Las líneas sin flechas representan el fragmento el tamaño de fragmento amplificado con los cebadores degenerados diseñados para reconocer los dominios DVKTFH(N)G-LLYVDDM(V) para Ty1-copia y RMCVDYR-YAKLSKS para Gypsy.

Las regiones de los retrotransposones más conservadas son la RT y RNasa H (por sus siglas en inglés)<sup>1</sup>. Sin embargo, para estudios filogenéticos Springer & Britten (1993)<sup>2</sup> reportan que la región RT es más sencilla de analizar que la región RNasa H, esta última no alcanza a diferenciar entre los diversos tipos de retrotransposones, ni entre familias dentro de la misma especie y cuando el análisis alcanza a separar las diferentes familias, el soporte de los clados es muy bajo. La transcriptasa reversa RT está localizada en el extremo 3' de la integrasa de los retrotransposones Ty1-copia, mientras que en los Gypsy se encuentra en el extremo 3' de la proteasa (**Figura 2.1**).

Esta región ha sido utilizada para determinar el posible origen de los retrotransposones<sup>3-4</sup> así como para identificar diferentes familias en maíz, trigo, frijol, chícharo, cebolla, plantas de ornato etc<sup>1, 5-7</sup>. Esta región ha permitido aislar retrotransposones completos del genoma de manzana con la técnica avance en el cromosoma conocida como "Chromosomal Walking"<sup>8</sup>.

Las variaciones en las secuencias de los diferentes dominios **Cuadro 2.1** han sido utilizadas en diversos géneros, para identificar las familias de retrotransposones en: *Triticum*, *Vigna*, *Hypochoeris*, *Gossypium*, *Oriza*, entre otros<sup>5, 6, 9-10</sup>. Además ha sido posible determinar que familias estuvieron presentes antes y las que se originaron después de la divergencia.

En un estudio filogenético de secuencias parciales de RT de retrotransposones Gypsy, Kumekawa y colaboradores (1999)<sup>11</sup> sugieren que las familias de Gypsy se originaron después de la divergencia entre mono y dicotiledóneas.

En la bibliografía revisada (**Cuadro 2.1**), existen 6 juegos de cebadores degenerados reportados, diseñados para amplificar los dominios conservados de RT característicos de los diferentes tipos de retrotransposones.

**Cuadro 2.1 Secuencias de los cebadores degenerados Ty1-copia y Gypsy**

CEBADOR DEGENERADO	DOMINIO	TIPO LTR	TAMAÑO pb	ESPECIE	REFERENCIAS
GGAATTCGAYGTNAAACACNGCNTTYT GGGATCCAYRTCRCTCNACRTANARNA	(1) DVKTFLH(N)G LLYVDDM(V)	Ty1-copia	280	<i>Gossypium barbadense</i> (algodón), <i>Lycopodas</i> , <i>Coníferas</i> , 38 <i>Angiospermas</i> . Entre otras	12-13
TNAA(C/T)ACNATGATGGA(C/T)GA TCTTG(T/C)GCATG(A/T)TCCCA	(2) N/D N/D	Ty1-copia	319	10 especies de <i>Echinoid</i>	14
CARATGGARGTNAARAC CATRTRCTCNACRTA	(3) QMDVKT VDDM	Ty1-copia	280	Especies de <i>Ingo</i>	5, 9
ACNGCNTT(C/T)(C/T)TNCA(C/T)GG AT(G/A)CANCT(G/A)CT(G/A)TAC(G/A)A	(4) TAFLHG YVDDML	Ty1-copia	260	<i>Sorgo</i> , <i>bríofitas</i> , <i>teridofitas</i> , <i>gimnospermas</i> (mono y <i>dicotiledóneas</i> )	15, 6
MRNATGTGYGTNGAYTAYMG RCAYTNSWARYTTNGCR	(5) RMCVDYR YAKLSKS	Gypsy	420	<i>Vigna radiata</i> .	10
AGMGRATGTGYGTS GAYTAT AARTGTGWRTTYKGGKTG	(6) N/D N/D	Gypsy	431	Monocotiledóneas (arroz, Ingo, maiz, cebada.) Dicotiledóneas (espinaca, papa, tabaco <i>arabidopsis</i> )	11

N/D= no disponible. Código de las secuencias V=A+C+G, N=A+C+G+T, D=A+T+G, B=T+C+G, H=A+T+C, W=A+T, S=C+G, K=T+G, M=A+C, Y=C+T, R=A+G

Los dominios (1), (3) y (4) han sido los más utilizados en plantas, para amplificar por PCR la región parcial RT de los retrotransposones Ty1-copia. Mientras que para los retrotransposones Gypsy solo dos dominios (5) y (6) han sido reportados para amplificar la región parcial RT **Cuadro 2.1**.

El objetivo de este capítulo fue aislar diferentes secuencias parciales de RT de cada tipo de retrotransposon LTR en tres especies de *Agave* incluyendo a *A. tequilana*, *A. fourcroydes* y *A. angustifolia* var. *Marginata*, a partir de iniciadores degenerados reportados que reconocen los dominios conservados.

## 2.2 MATERIALES Y MÉTODOS

### 2.2.1 EXTRACCIÓN DE ADN

Se tomaron muestras de hojas de tres especies de *Agave*: *A. angustifolia* var. *marginata*, *A. fourcroydes* Sac Ki y *A. tequilana* del jardín botánico del Centro de Investigación Científica de Yucatán (CICY).

Para la extracción de ADN se utilizó un protocolo de Roman en el que se adicionó 1gr de tejido fresco previamente macerado a 5 ml de buffer de extracción (Tris-Base 100mM, EDTA 50mM y NaCl 500mM) y 3.5 ml de  $\beta$ -Mercaptoetanol (100mM). Posteriormente se adicionó 500  $\mu$ l de SDS (20%).

Se incubó en hielo por 30 min y enseguida en baño maría a 65 °C por 30-40 min. Al terminar la incubación se adicionaron 1.6 ml de la solución alta en sales (6M de acetato de potasio, pH 5.5 con Ac. Acético) y se incubó en hielo 30 min. Las muestras se centrifugaron a 10000 rpm por 15 min. El sobrenadante fue separado y se le adicionaron 0.8 volúmenes de isopropanol previamente enfriado, se agitó suavemente hasta que el ADN precipitó y se pudo separar. El ADN precipitado se tomó con una punta de pipeta y se pusieron en 1 ml de etanol al 70%, repitiendo este paso dos veces más con 500 $\mu$ l de etanol al 70% y posteriormente se dejó secar la pastilla y se resuspendió en agua.

### 2.2.2 AISLAMIENTO DE TR

Diferentes fragmentos de ADN fueron amplificados por PCR con iniciadores degenerados que reconocen los dominios conservados de transcriptasa reversa DVKTFLH(N)G-LLYVDDM(V) y RMCVDYR-YAKLSKS de Ty1-copia y Gypsy respectivamente (**Figura 2.1 y Cuadro 2.1**) en *A. fourcroydes*, *A. angustifolia* y *Agave tequilana*, en dos mezclas de reacción independientes que contenían 500 ng de ADN, 3.5 mM de MgCl<sub>2</sub>, 0.4mM de dNTPs, 2U Taq ADN polimerasa (Invitrogen), 1x PCR buffer y 1 pmol de su respectivo juego de cebador, en 50  $\mu$ l de volumen final.

Las condiciones de PCR para Ty1-copia iniciaron con una temperatura de desnaturalización a 94°C por 4 minutos seguido de 30 ciclos de una desnaturalización a 94°C por 1 min, un alineamiento a 47°C por un min y una extensión a 72°C por 2 min.

Posteriormente una extensión final a 72°C por 7 min<sup>11-12</sup>. Para Gypsy consistió de una temperatura inicial de desnaturalización a 94°C por 3 min, seguido de 40 ciclos con una temperatura de desnaturalización 94°C por un minuto, un alineamiento a 50°C por 50 s una extensión a 72°C por 40 s y finalmente una extensión a 72°C por 10 min<sup>10</sup>. Los fragmentos amplificados se separaron mediante electroforesis en gel de agarosa al 1%.

### **2.2.3 CLONACIÓN Y SECUENCIACIÓN**

#### **Ligación**

Antes de la clonación los productos de PCR se purificaron del gel de agarosa con el kit de extracción Quiaex<sup>®</sup> II siguiendo las instrucciones del proveedor. (QUIAGENE cat. 20021).

La clonación inicio con la ligación del fragmento al vector de clonación pGEM-T Easy siguiendo las instrucciones del proveedor (Promega cat. A1360).

#### **Transformación**

Para la transformación se utilizaron bacterias *E. coli* JM109 (Promega) químicamente competentes. Se adicionó 1µl de la mezcla de ligación a 100 µl de células competentes e incubadas en hielo por 20 min. La transformación de las bacterias se realizo por choque térmico a 42°C por 1 minuto. Las células transformadas fueron crecidas en 500 µl de medio líquido LB (Invitrogen Cat. 22780-052) por 2 hr a 37°C con agitación constante 200 rpm. Antes del cultivo en medio semisólido LB agar (Invitrogen Cat. 22700-025), en la caja de petri se disperso 100 µl de isopropil-β-D-thiogalactopiranosido (IPTG) 100mM y 20 µl de X-Gal (50mg/ml) como activador y como marcador de selección respectivamente. Posteriormente, 100 µl de cada cultivo fue crecido en medio solido en cajas de petri.

#### **Extracción del plásmido.**

Las colonias blancas fueron seleccionadas, para comprobar la presencia del fragmento de interés mediante PCR con los cebadores específicos (M13 forward y M13 reverse), que amplifican el fragmento de interés más 200 pb del vector.

Las colonias que contenían el inserto de interés fueron crecidas 12 h en 5 ml de medio líquido LB a 37°C con agitación constante 200 rpm

Para extraer el plásmido se utilizó el protocolo STET, 3 ml de cultivo fueron centrifugados a 8000 rpm por 1 min, el sobrenadante fue desechado y la pastilla se resuspendió en 300  $\mu$ l de buffer STET (8% glucosa, 5% Triton x 100, 50mM EDTA y 50 mM Tris HCL pH 8.0). Se adiciono 10  $\mu$ l de lisosima (10mg/ml) y se incubo en agua hirviendo por 40 s. posteriormente se centrifugo a 12500 rpm y el la pastilla formada fue desechada con un palillo estéril. Para precipitar el plásmido, el sobrenadante se incubo 10 min en frio con 300  $\mu$ l de isopropanol previamente enfriado, centrifugo a 12500 rpm por 10 min. El sobrenadante fue retirado y la pastilla se lavo 2 veces con 500  $\mu$ l de etanol al 70%. Por último, la pastilla sin residuos de alcohol se diluyo en 20  $\mu$ l de agua estéril.

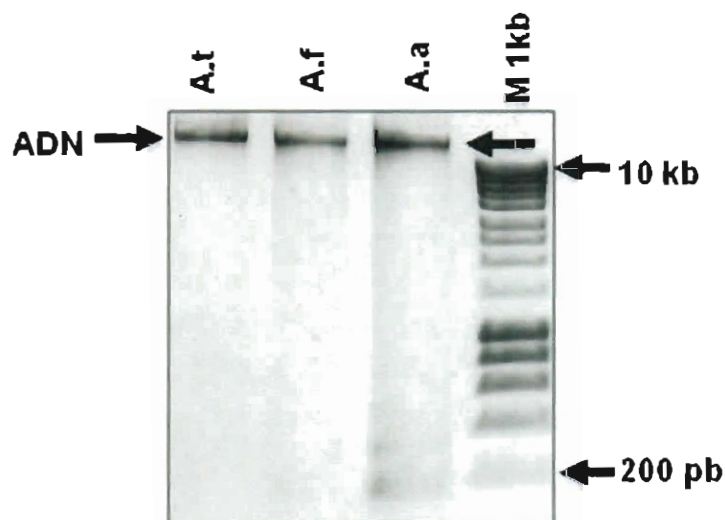
Para secuenciar los fragmentos de interés, 20  $\mu$ l de plásmido (100ng/ $\mu$ l) se mandaron a secuenciar con la empresa coreana (MACROGEN).



## 2.3 RESULTADOS

### 2.3.1 EXTRACCIÓN DE ADN

El ADN aislado de plantas de *A. fourcroydes* (A.f), *A. angustifolia* (A.a) y *A. tequilana* (A.t) fue tratado con 1  $\mu$ l de RNasa (0.034 mg/ $\mu$ l), se obtuvo un rendimiento promedio de 24.7  $\mu$ g de ADN por gramo de tejido. La (Figura 2.2) muestra las bandas características del ADN integro (limpio y sin degradar).



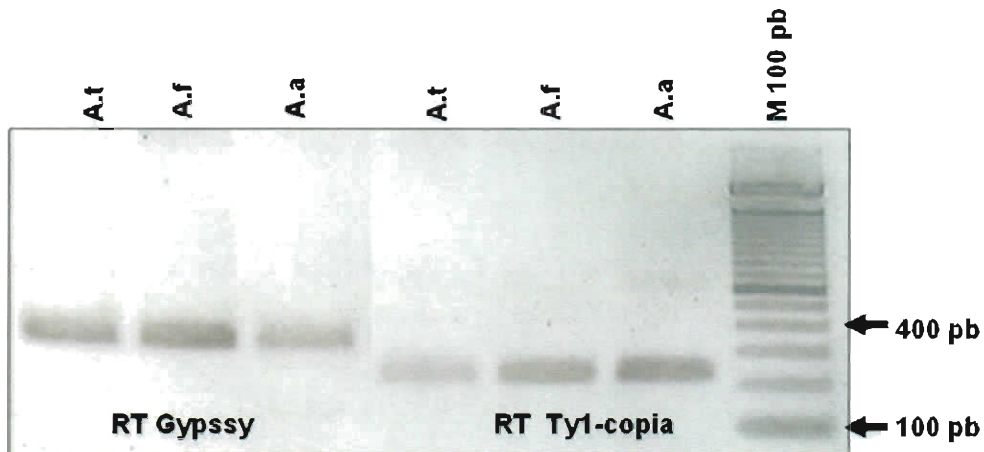
**Figura 2.2 Extracción de ADN.**

Gel de agarosa al 1% con 0.5  $\mu$ l de cada muestra de ADN, teñido con Bromuro de Etidio. M=Marcador de tamaño 1kb. A. *fourcroydes* (A.f, A. *angustifolia* (A.a) y *Agave tequilana* (A.t)

### 2.3.2 AISLAMIENTO DE TR

Del ADN aislado se amplificaron por PCR los fragmentos de RT, los productos de PCR fueron desplegados en gel de agarosa al 1% a 80 Volts (V) para comprobar los tamaños esperados 280 y 420 pb para Ty1-copia y Gypsy respectivamente (Voytas *et al.* 1992, Ditix *et al.* 2006) (Figura 2.3).

Las bandas desplegadas fueron cortadas, purificadas y ligadas al vector pGEM-T Easy (Promega cat. A1360). Posteriormente se realizó la transformación de *E. coli* con el vector.



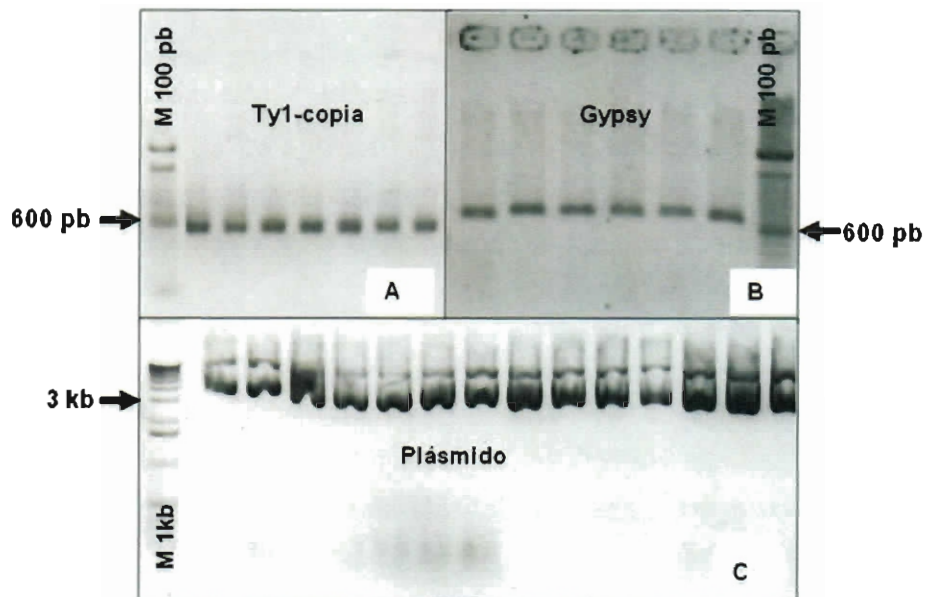
**Figura 2.3 Amplificación por PCR de TR Gypsy y ty1-copia.**

El panel A-B corresponde a fragmentos amplificados con iniciadores degenerados que reconocen los dominios conservados en la región TR en *A. tequilana*, *A. fourcroydes* y *A. angustifolia*. A) Dominio RMCVDYR-YAKLSKS (Gypsy). B) Dominio DVKTFLH(N)G-LLYVDDM(V) (Ty1-copia). M = marcador de peso 100 pb.

Veinte colonias blancas se seleccionaron al azar, para comprobar si estas contenían el inserto de interés por PCR con los cebadores específicos M13F y M13R. Para comprobar el tamaño de los fragmentos 480 pb para ty1-copia y 680 pb para gypsy, el producto fue desplegado en una electroforesis en gel de agarosa (Figura 2.4 A y B).

### 2.3.3 EXTRACCIÓN DEL PLÁSMIDO

Las colonias que contenían inserto se cultivaron para realizar la extracción del plásmido (Figura 2.4)



**Figura 2.4. Clonación y extracción del plásmido.**

El panel A y B gel de agarosa 1% muestra la PCR de las colonias blancas seleccionadas con fragmentos de Ty1-copia y gypsy respectivamente. C) gel de agarosa 1% con 1  $\mu$ l de la extracción del plásmido. M= marcador 100pb y 1kb.

En la (Figura 2.4 C) se aprecia más de una banda por arriba del tamaño del plásmido (3kb), esto es debido a la concatenación del plásmido. El rendimiento promedio de la extracción fue de 329 ng ADN/ml de cultivo.

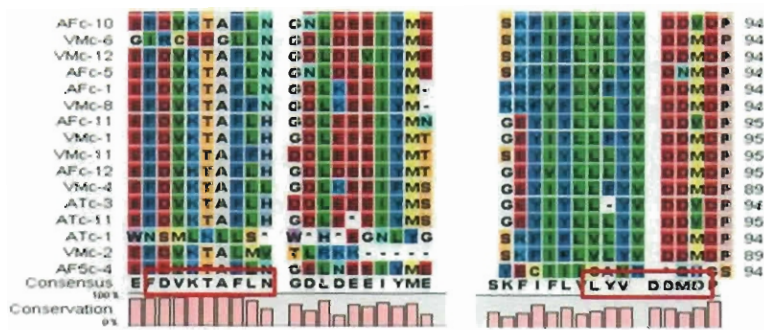
Se mandaron a secuenciar 20  $\mu$ l de plásmido (100ng/ $\mu$ l), de 34 fragmentos para Ty1-Copia: 11 de *A. tequilana*, 12 de *A. fourcroydes* y 11 de *A. angustifolia*. Mientras que para Gypsy fueron 41 secuencias de las cuales 12 son de *A. tequilana*, 18 de *A. fourcroydes* y 11 de *A. angustifolia*.

### 2.3.4 CARACTERIZACIÓN DE LAS SECUENCIAS

#### Secuencias Ty1 copia.

Las secuencias se editaron en el programa BioEdit para encontrar el extremo 5' y 3' de las secuencias que corresponde a los cebadores degenerados. El tamaño de los fragmentos en *A. tequilana* vario entre 263 y 300 pb, en *A. fourcroydes* de 283 a 286 y para *A. angustifolia* de 272 a 286.

Posteriormente, con el programa CLC Sequence Viewer se realizó la traducción *in silico* de las secuencias (Figura 2.5), para encontrar el marco de lectura en el que se encontraran los dominios DVKTFLH(N)G-LLYVDDM(V) y de terminar de esta forma que la región analizada corresponde a la región parcial de TR de Ty1-copia. Sin embargo, en algunas secuencias se necesitaba más de un marco hipotético para encontrar los dominios Cuadro 2.2.



**Figura 2.5 Traducción *in silico*.**

Muestra solamente la región de los dominios DVKTFLH(N)G-LLYVDDM(V) (cuadro rojo) para los que fueron diseñados los cebadores degenerados.

**Cuadro 2.2 diferentes marcos de lectura.**

	Marcos hipotéticos	Total de secuencias
<i>A. tequilana</i>	(+1), (+2), (+3)	11
<i>A. fourcroydes</i>	(+1), (+2), (+3)	12
<i>A. angustifolia</i>	(+2), (+3)	11

### **Secuencias Gypsy.**

Las secuencias en el caso de Gypsy se editaron en el programa BioEdit, se encontraron 2 secuencias de *A. agustifolia*, 4 de *A. fourcroydes* y 2 de *A. tequilana* idénticas, estas secuencias se descartaron del análisis. Las secuencias de Gypsy fueron traducidas *in silico* en los diferentes marcos de lectura, para encontrar los dominios RMCVDYR-YAKLSKS y comprobar que la secuencia amplificada es de retrotransposones Gypsy. Al igual que en las secuencias de Ty1-copia, diferentes marcos de lectura (+1 +2, -2 y -3) fueron utilizados para encontrar el dominio. Las secuencias que no presentaban el dominio fueron eliminadas de los análisis posteriores que serán discutidos en el capítulo III.

Solo 30 secuencias de 41 en total fueron diferentes, 9 de *A. angustifolia*, 11 de *A. fourcroydes* y 10 de *A. tequilana*.

## 2.4 DISCUSIÓN.

Para poder hacer estudios en Biología molecular es necesario contar con un método de extracción de ADN, eficiente y que permita obtenerlo libre de contaminantes como carbohidratos, proteínas y/o sales que pueden interferir con las reacciones necesarias para su desarrollo, principalmente que pueda interferir con la PCR. El rendimiento y calidad del ADN fue suficiente para realizar la amplificación por PCR de los diferentes fragmentos de TR de Ty1- copia y Gypsy (**Figura 2.1 y 2.2**).

Los dominios seleccionados en este trabajo, para los cuales se han diseñado los cebadores degenerados, han sido utilizados para amplificar la región parcial de RT en diferentes especies de plantas <sup>10, 12, 13</sup>.

Los fragmento amplificados en este trabajo corresponden al tamaño esperado (280 pb Ty1-copia y 420 Gypsy) y la traducción *in silico* de las secuencias mostró que los fragmentos contienen los dominios DVKTFLH(N)G-LLYVDDM(V) y RMCVDYR-YAKLSKS característicos de los retrotransposones Ty1-copia y Gypsy respectivamente (**Figura 2.5**). Confirmando así que se pudieron aislar fragmentos que correspondes a retrotransposones en *Agave*.

Los retrotransposones LTR están ampliamente distribuidos en el genoma de las plantas<sup>15, 11, 9</sup>. Bousios *et al.* (2007)<sup>16</sup>, aisló regiones LTR de los retrotransposones Ty1-copia para emplear en SSAP para evaluar diversas especies de *Agaves* y *Yucas* y mostro evidencia de que estaban presentes tanto en el genoma de especies y variedades de *Agaves* como de *Yuca*, sugiriendo un componente ancestral.

Los resultados obtenidos en este trabajo confirman la existencia de retrotransposones Ty1-copia en el genoma de *Agave*. Adicionalmente se demostró que también los retrotransposones Gypsy están presentes en los genomas de *A. tequilana*, *A. fourcroydes* y *A. angustifolia*.

## 2.5 BIBLIOGRAFÍA

1. A. Kumar; J. L. Bennetzen. Plant retrotransposons. *Annu Rev. Genet.* 33: pp 429-532 (1999).
2. M. S. Springer; R. J. Britten, Phylogenetic Relationships of Reverse Transcriptase and RNase H Sequences and Aspects of Genome Structure in the Gypsy Group of Retrotransposons. *Mol. Biol. Evol.* 10:6, 1370-1379, 1993.
3. Y. Xian; T. H. Eickbush, Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. *Mol. Biol. Evol.* 5, 675-690, 1988.
4. Y. Xian; T. H. Eickbush, Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO J.* 9:10, 3353-3362, 1990.
5. Y. Matsuoka; K. Tsunewaki, Wheat retrotransposon families identified by reverse transcriptase domain analysis. *Mol. Biol. Evol.* 13:10, 1384-1392, 1996.
6. Galosso; G. E. Harrison; D. Pignone; A. Brandes; J. S. Heslop-Harrison, The distribution and organization of Ty1-copia-like retrotransposons elements in the genome of *Vigna unguiculata* (L.) Walp. (Cowpea) and its relatives. *Annals of Botany*, 80, 327-333. 1997.
7. C. F. Ruas; H. Weiss-Schneeweiss; T. F. Stuessy; M. R. Samuel; A. Pedrosa-Harand, K. Tremetsberger; P. M. Ruas; P. M. Schlüter; M. A. Ortiz Herrera; C. König; N. I. Matzenbacher, Characterization, genomic organization and chromosomal distribution of Ty1-copia retrotransposons in species of *Hypochaeris* (Asteraceae). *Gene*, 15:412 (1-2), 39-49, 2008.
8. G., Zhao; Z. Zhang; H. Sun; H. Li; H. Dai, Isolation of Ty1-copia-like Retrotransposon Sequences from the Apple Genome by Chromosome Walking Based on Modified Site Finding-polymerase Chain Reaction. *ABBS* 39:9, 675-683, 2007.
9. H. Hirochika; R. Hirochika, Ty1-copia group retrotransposons as ubiquitous components of plant genome. *Jpn. J. Genet.* 68, 35-46, 1993.
10. A. Dixit; J. Ho Ma; J. Woong Yu; E. Gi Cho; J. Jin Park, Reverse transcriptase domain sequences from Mungbean (*Vigna radiata*) LTR retrotransposons: Sequence characterization and phylogenetic analysis. *Heredity*, 97, 381-388, 2006.

11. N. Kumekawa; E. Ohtsubo; H. Ohtsubo, Identification and phylogenetic analysis of gypsy-type retrotransposons in the plant kingdom. *Genes Genet. Syst.* 74, 299-307, 1999.
12. A. Ghany; Saki, Isolation characterization and analysis of copia-like retrotransposons in the Egyptian cotton *Gossypium barbadense* and its progenitors. *African Journal of biotechnology.* 2:6, 165-168. 2003.
13. D. F. Voytas; M. P. Cummings; A. Konieczny; F. M. Ausubel; S. R. Rodermel, Copia-like retrotransposons are ubiquitous among plants. *PNAS.* 89, 7124–7128, 1992.
14. M. S. Springer; N. A. Tusneen; E. H. Davidson; R. J. Britten, Phylogeny, rates of evolution, and patterns of codon usage among sea urchin retroviral-like elements, with implications for the recognition of horizontal transfer. *Mol. Biol. Evol.* 12:2, 219-230. 1995.
15. A. J. Flavell; D. B. Smith; A. Kumar, Extreme heterogeneity of Ty1-copia group retrotransposons in plants. *Mol. Gen. Genet.* 231, 233-242, 1992
16. A. Bousios; I. Saldana-Oyarzabala; A. G. Valenzuela-Zapata; C. Wooda; S. R. Pearce, 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 Science*, 172:2, 291-298, 2007,



## Capítulo III. Análisis de las secuencias parciales de RT de retrotransposones LTR.

### 3.1 INTRODUCCIÓN.

Los retrotransposones son un tipo de elementos móviles clase 1, se dividen en dos grupos: los retrotransposones LTR y los retrotransposones sin LTR<sup>1</sup>. Los primeros presentan una región larga repetida en sus extremos LTR por sus siglas en inglés (Long terminal repeats), estos se dividen a su vez, en dos grandes grupos los retrotransposones Ty1-copia y los Gypsy que pertenecen a la familia *pseudoviridae* y *metaviridae* respectivamente<sup>2</sup>. Su transposición en el genoma, se lleva a cabo, por un mecanismo de copia y pega, a través de un intermediario de RNA<sup>3</sup>. Estos elementos están ampliamente distribuidos en el genoma de las plantas<sup>4-6</sup>. Hay evidencia de que los retrotransposones contribuyen en la estructura de los cromosomas principalmente en centrómeros y telómeros, además de que pueden ser puntos de recombinación<sup>7</sup>.

Por otro lado existe evidencia, de que éstos pueden ser activados en condiciones de estrés<sup>8</sup> y recientemente se demostró que la metilación está involucrada en el proceso de regulación de la transposición en el caso de *Arabidopsis thaliana*<sup>9-10</sup>.

En muchas especies de plantas representan más del 50% de su genoma total y están representados por diferentes familias<sup>1</sup>, estas familias son reconocidas con base en la comparación de sus secuencias por lo que es posible la reconstrucción filogenética e inferir la evolución de una familia de un retrotransposon en particular<sup>11</sup>. Por ejemplo, en especies del género *Oryza* han sido identificadas 32 familias<sup>12</sup>, 7 en *Triticum*<sup>13</sup>, al menos 7 *Hypochoeris*<sup>14</sup>, 13 en *Arabidopsis thaliana*<sup>15</sup>, 11 en *Vigna radiata*<sup>16</sup>.

El objetivo de esta parte del trabajo fue identificar, con base en las secuencias, diferentes familias de retrotransposones Ty1-copia y Gypsy en tres especies de *Agave*.

### 3.2 MATERIALES Y MÉTODOS.

Se analizaron 34 secuencias de Ty1 copia: 11 de *A. tequilana*, 12 *A. fourcroydes* y 11 *A. angustifolia*. Para Gypsy 33 secuencias: 14 de *A. fourcroydes*, 10 de *A. tequilana*, 9 de *A. angustifolia*.

Las claves para las secuencias de Ty1-copia fueron ATc<sub>1...12</sub> (*A. tequilana*), AFc<sub>1...12</sub> (*A. fourcroydes*) y VMc<sub>1...12</sub> (*A. angustifolia*). Mientras que para Gypsy fue ATg<sub>1...10</sub>, AFg<sub>1...13</sub>, y VMg<sub>1...16</sub>.

### 3.2.1 COMPARACIÓN GENBANK.

La comparación de las secuencias con las reportadas en el GenBank (solo las bases de datos EMBL+DDBJ+PDB de secuencias de nucleótidos) se realizó con el programa (BLASTN 2.2.19)<sup>17</sup>.

### 3.2.2 ANÁLISIS FILOGENÉTICO

El análisis filogenético se realizó usando el programa MEGA versión (4)<sup>18</sup>. Los parámetros para el alineamiento fueron los marcados por default (una penalización por abrir un gap de 15 y 6.66 por la extensión del mismo), utilizando la matriz IUB.

Para la construcción del árbol filogenético se utilizó el método Neighbor-Joining<sup>19</sup>. Para evaluar el soporte de los grupos formados se utilizó un análisis de remuestreo (bootstrap) con 1000 replicas<sup>20</sup>. Las distancias evolutivas fueron computadas con el método Maximum Composite Likelihood<sup>21</sup>.

### 3.2.3 IDENTIFICACIÓN DE FAMILIAS

Para la identificación de familias se considero los agrupamientos con valores de bootstrap mayores al 90% y un porcentaje de identidad del grupo mayor al 60%.

## 3.3 RESULTADOS.

### 3.3.1 TY1-COPIA

#### 3.3.1.1 COMPARACIÓN DE LAS SECUENCIAS CON LAS REPORTADAS EN EL GENBANK

Las secuencias fueron comparadas con las reportadas en el GenBank encontrando homología con los retrotransposones Ty1-copia de diversas especies. Las especies en las que se encontró la mayor homología (más del 83% de identidad y el valor más alto de  $E_{value} = 0.09$ ) son: *Spirantes honkongensis*, *Spirantes cinensis*, *Cicer arietinum*, *Panax quinquefolius*, *Eucalyptus globulus*, *Ipomea batatas*, *Gopssypium hirsutum*, *Gopssypium herbaceum*, *Solanum tuberosum*, *Oriza sativa*, *Amaranthus cruentus*, *Vigna radiata*, *Diospiros kacki*, y *Aegiceras parmiculata*. En algunas secuencias se encontró una mayor

homología consecuencias de cromosomas completos no caracterizadas de *Vigna radiata*, *Arabidopsis thaliana* y *Oryza sativa*.

**Cuadro 3.1**

Analogía de las secuencias Ty1-copia con las reportadas en el GenBank usando blastn.

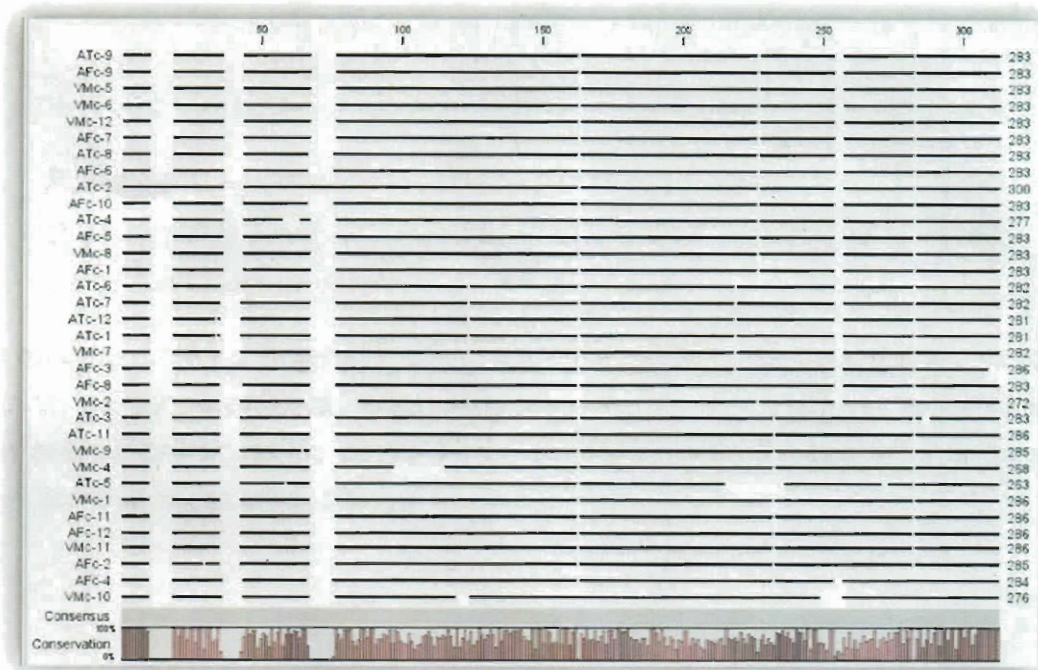
Secuencia	Especie con mayor homología	Score	E <sub>value</sub>	Identidad
AFc1	<i>Spirantes honkongensis</i>	68	2x10 <sup>-08</sup>	90%
AFc2	<i>Cicer arietinum</i>	78	3x10 <sup>-11</sup>	91%
AFc3	<i>Cicer arietinum</i>	56	3x10 <sup>-04</sup>	100%
AFc4	<i>Panax quinquefolius</i>	72	2x10 <sup>-09</sup>	92%
AFc5	<i>Eucalyptus globulus</i>	111	2x10 <sup>-21</sup>	85%
AFc6	<i>Ipomoea batatas</i>	133	5x10 <sup>-28</sup>	80%
AFc7	<i>Eucalyptus globulus</i>	121	2x10 <sup>-24</sup>	84%
AFc8	<i>Cicer arietinum</i>	56	1x10 <sup>-4</sup>	97%
AFc9	<i>Eucalyptus globulus</i>	149	9x10 <sup>-33</sup>	85%
AFc10	<i>Ipomoea batatas</i>	127	3x10 <sup>-26</sup>	81%
AFc11	<i>Gopssypium hirsutum</i>	78	3x10 <sup>-11</sup>	84%
AFc12	<i>Solanum tuberosum</i>	68	3x10 <sup>-08</sup>	93%
ATc1	<i>Ipomoea batatas</i>	65	1x10 <sup>-07</sup>	100%
ATc2	<i>Eucalyptus globulus</i>	145	1x10 <sup>-31</sup>	87%
ATc3	<i>Gopssypium hirsutum</i>	71	2x10 <sup>-09</sup>	92%
ATc4	<i>Eucalyptus globulus</i>	121	2x10 <sup>-24</sup>	84%
ATc5	<i>Spiranthes sinensis</i>	68	2x10 <sup>-08</sup>	95%
ATc6	<i>Spirantes honkongensis</i>	58	2x10 <sup>-05</sup>	96%
ATc7	<i>Spirantes honkongensis</i>	58	2x10 <sup>-05</sup>	96%
ATc8	<i>Eucalyptus globulus</i>	149	9x10 <sup>-33</sup>	86%
ATc9	<i>Eucalyptus globulus</i>	145	1x10 <sup>-31</sup>	87%
ATc11	<i>Gopssypium hirsutum</i>	73	4x10 <sup>-10</sup>	83%
ATc12	<i>Oryza sativa</i>	46	0.09	93%
VMc1	<i>Amaranthus cruentus</i>	82	2x10 <sup>-12</sup>	88%
VMc2	<i>Vigna radiata</i>	56	9x10 <sup>-05</sup>	87%
VMc4	<i>Cicer arietinum</i>	64	4x10 <sup>-07</sup>	95%

VMc5	<i>Eucalyptus globulus</i>	157	$4 \times 10^{-35}$	86%
VMc6	<i>Eucalyptus globulus</i>	165	$1 \times 10^{-37}$	87%
VMc7	<i>Diospyros kaki</i>	50	0.006	86%
VMc8	<i>Spirantes honkongensis</i>	67	$2 \times 10^{-08}$	90%
VMc9	<i>Spiranthes sinensis</i>	60	$6 \times 10^{-06}$	92%
VMc10	<i>Aegiceras paniculata</i>	58	$2 \times 10^{-5}$	96%
VMc11	<i>Gopssypium herbaceum</i>	80	$7 \times 10^{-12}$	89%
Vm12	<i>Ipomoea batatas</i>	137	$3 \times 10^{-29}$	85%

### 3.3.1.2 ANÁLISIS FILOGENÉTICO

Como se discutió en el capítulo 2, el análisis filogenético se realizó con las secuencias de nucleótidos debido principalmente a que en la traducción *in silico* las secuencias fueron idénticas, además de que se pierde información por la degeneración del código genético.

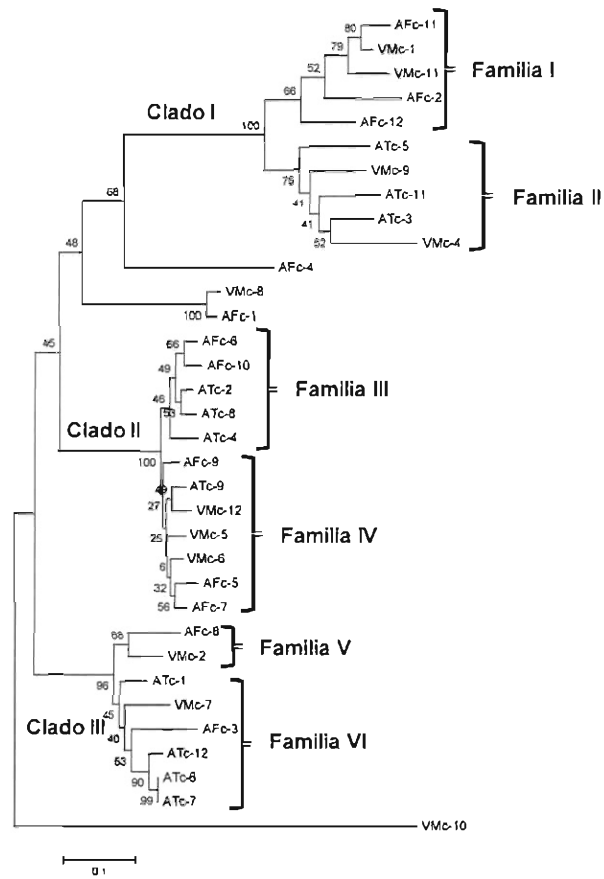
Se realizó un alineamiento múltiple con el programa MEGA 4.1 y CLC Sequence Viewer (**Figura 3.1**) con los parámetros por default que consideran gaps de 15 nucleótidos máximos y con la matriz IUB y se encontró que entre todas las secuencias hay mucha heterogeneidad, con solo un 22 % de identidad.



**Figura 3.1 Alineamiento de las 34 secuencias Ty1 copia.**

*El alineamiento cuenta con solo el 22 % de identidad. Sin embargo la mayoría de los nucleótidos se conservan por arriba del 50%.*

Se construyó un árbol filogenético con el programa MEGA utilizando el método Neighbour Joining, la distancia evolutiva fue calculada con el método compuesto de máxima similitud con análisis de remuestreo de 1000 replicas (Figura 3.2.) En este árbol se pueden identificar tres clados principales.



**Figura 3.2** *Árbol filogenético Ty1-copia.*

La historia evolutiva fue inferida utilizando el método Neigbor-Joinig<sup>19</sup> con un análisis de remuestreo con 1000 replicas<sup>20</sup>. Las secuencias quedaron agrupadas en tres principales clados y en cada clado se pueden identificar 2 familias.

La constitución de los agrupamientos se describe a continuación:

El clado I se encuentra representado por el 30% de las secuencias, 3 secuencias de *A. tequilana* (ATc1, ATc3 y ATc11), 4 de *A. angustifolia* Var. *marginata* (VMc1, VMC11, VMc9 y VMc4) y 3 de *A. fourcroydes* (AFc11, AFc2 y AFc12). La identidad entre todas las secuencias de este clado es del 50%. En este clado se pueden identificar dos familias (Familia I y II) tomando como parámetro de identificación el análisis de remuestreo (mayor del 90%) y el porcentaje de identidad (mayor del 60%). La familia I solamente incluye secuencias de *A. fourcroydes* y *A. angustifolia* (AFc11, VMc1, VMC11, AFc2 y AFc12) con un 70% de identidad. La familia II presenta un 62% de identidad, las secuencias presentes en esta familia son de *A. tequilana* y *A. angustifolia* var. *marginata* (ATc1, ATc3, ATc11 VMc9 y VMc4) (**Figura 3.2**).

Se puede sugerir que estas familias pueden ser específicas para estas especies en las que *A. angustifolia* var. *marginata* es común. Sin embargo es necesario confirmar experimentalmente esta suposición a través del diseño de iniciadores específicos para amplificar en las especies correspondientes, esperando que amplifiquen selectivamente.

En el clado II se agruparon 12 secuencias que corresponde a un 35 % del total de las secuencias analizadas. Se pudieron identificar dos familias (familia III y IV). La familia III (AFc6, AFc10, AFc, ATc2, ATc8 y ATc4) incluye solamente especies de *A. fourcroydes* y *A. tequilana* con un 81% de identidad. Mientras que la familia IV incluye secuencias de las tres especies (AFc9, AFc5, AFc7, ATc9, VMc12, VMc5 y VMc6) con un 87% de identidad.

Por último, en el clado III quedaron agrupadas 8 secuencias dos de *A. fourcroydes* (AFc8, AFc3), 2 de *A. angustifolia* (VMc2 y VMc7) y 4 de *A. tequilana* (ATc1, ATc12, ATc6 y ATc7). Se pueden identificar dos familias (familia V y VI) aunque en la familia V solo se encuentran dos secuencias. El porcentaje de identidad de la familia V es del 85% mientras que para la familia VI es de 75%. En la familia VI quedaron agrupadas secuencias de las tres especies.

### 3.3.2 GYPSY

#### 3.3.2.1 COMPARACIÓN DE LAS SECUENCIAS EN EL GENBANK

Las secuencias fueron comparadas con las reportadas en el GenBank, se observó que éstas se alinearon exitosamente con retrotransposones del tipo Gypsy con un porcentaje de identidad mayor al 71% y valor E mayor que 0.002. Nueve de las secuencias alinearon con las reportadas para *Gopssypium herbaceum*, 4 de *Vigna radiata*, 2 de *Orobancha crenata*, 2 de *Arabidopsis lyrata*, 1 de *Vicia melanops*, 1 de *Helianthus annuus*, 1 de *Helianthus deserticola*, 1 de *Heliantus anomalus*, 1 de *Malus x domestica*, 1 de *Medicago truncatula*, 1 de *Orobancha cernua*, 1 de *Phelipanche bungean*, 1 de *Phelipanche tunetana*, 1 de *Spinacia oleracea*, 3 alinean con secuencias no caracterizadas y 3 no alinean con ninguna reportada,

#### Cuadro 3.2

Analogía de las secuencias Gypsy con las reportadas en el GenBank.

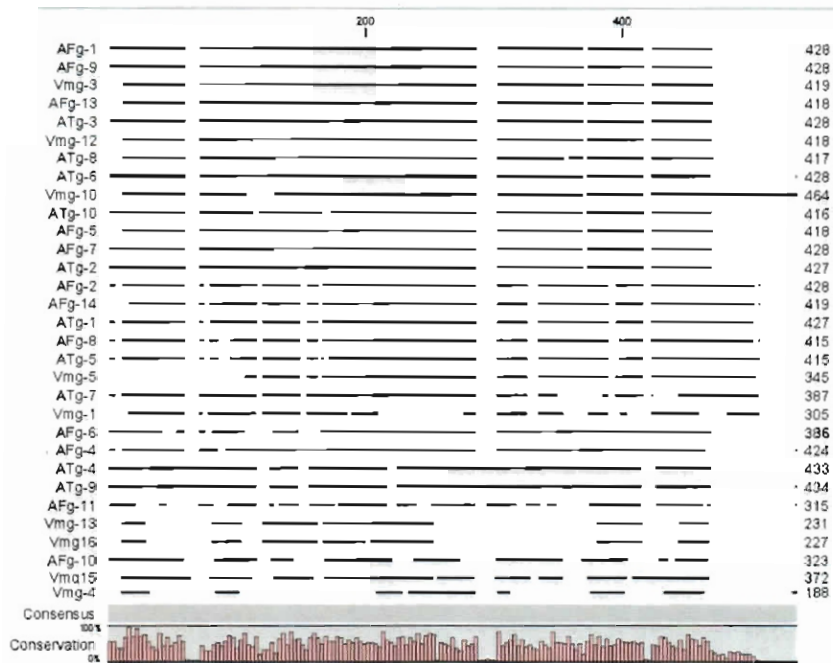
Secuencia	Especie con mayor homología	Score	E <sub>value</sub>	Identidad
VMg-6	<i>Orobancha cernua</i>	109	$1 \times 10^{-20}$	83%
ATg-1	<i>Vicia melanops</i>	62	$4 \times 10^{-06}$	93%
ATg-2	<i>Gopssypium herbaceum</i>	100	$1 \times 10^{-17}$	88%
AFg-2	<i>Gopssypium herbaceum</i>	92	$3 \times 10^{-15}$	86%
AFg-4	<i>Gopssypium herbaceum</i>	68	$4 \times 10^{-08}$	79%
AFg-9	<i>Vigna radiata</i>	56	$4 \times 10^{-05}$	85%
AFg-11	<i>Vigna radiata</i>	52	0.002	96%
AFg-13	<i>Gopssypium herbaceum</i>	100	$1 \times 10^{-17}$	95%
AFg-16	<i>Gopssypium herbaceum</i>	80	$1 \times 10^{-11}$	79%
VMg-3	<i>Gopssypium herbaceum</i>	90	$1 \times 10^{-14}$	86%
VMg-5	<i>Gopssypium herbaceum</i>	66	$1 \times 10^{-07}$	87%
VMg-10	<i>Orobancha crenata</i>	80	$1 \times 10^{-11}$	84%
VMg-12	<i>Gopssypium herbaceum</i>	80	$1 \times 10^{-11}$	84%
AFg-1	<i>Vigna radiata</i>	262	$9 \times 10^{-67}$	75%
ATg-3	<i>Phelipanche tunetana</i>	253	$5 \times 10^{-64}$	75%



ATg-8	<i>Heliantus anomalus</i>	226	$6 \times 10^{-58}$	73%
VMg-12	<i>Orobanche crenata</i>	239	$1 \times 10^{-59}$	76%
AFg-7	<i>Phelipanche bungeana</i>	269	$6 \times 10^{-69}$	74%
AFg-5	<i>Gossypium herbaceum</i>	235	$1 \times 10^{-50}$	74%
ATg-6	<i>Vigna radiata</i>	210	$5 \times 10^{-61}$	71%
ATg-10	<i>Helianthus deserticola</i>	176	$1 \times 10^{-40\%}$	75%
ATg-7	No alineo	-	-	-
VMg-15	No alineo	-	-	-
AFg-8	<i>Malus x domestica</i>	138	$2 \times 10^{-29}$	70%
ATg-5	<i>Spinacia oleracea</i>	113	$1 \times 10^{-21}$	73%
VMg-4	Secuencias no caracterizadas	-	-	-
ATg-4	<i>Arabidopsis lyrata</i>	347	$3 \times 10^{-92}$	100%
ATg-9	<i>Arabidopsis lyrata</i>	347	$3 \times 10^{-92}$	100%
VMg13	<i>Medicago truncatula</i>	75	$1 \times 10^{-10}$	83%
VMg-16	Secuencias no caracterizadas	-	-	-
VMg-1	<i>Helianthus annuus</i>	59	$1 \times 10^{-05}$	79%
AFg-6	Secuencias no caracterizadas	-	-	-
AFg-10	No alineo	-	-	-

### 3.3.2.2 ANÁLISIS FILOGENÉTICO

Se realizó un alineamiento múltiple en el programa MEGA 4.1 y CLC Sequence (Figura 3.3) con los parámetros por default con gaps de 15 nucleótidos máximos y con la matriz IUB. y se encontró que entre todas las secuencias hay mucha heterogeneidad, con solo un 11 % de identidad.

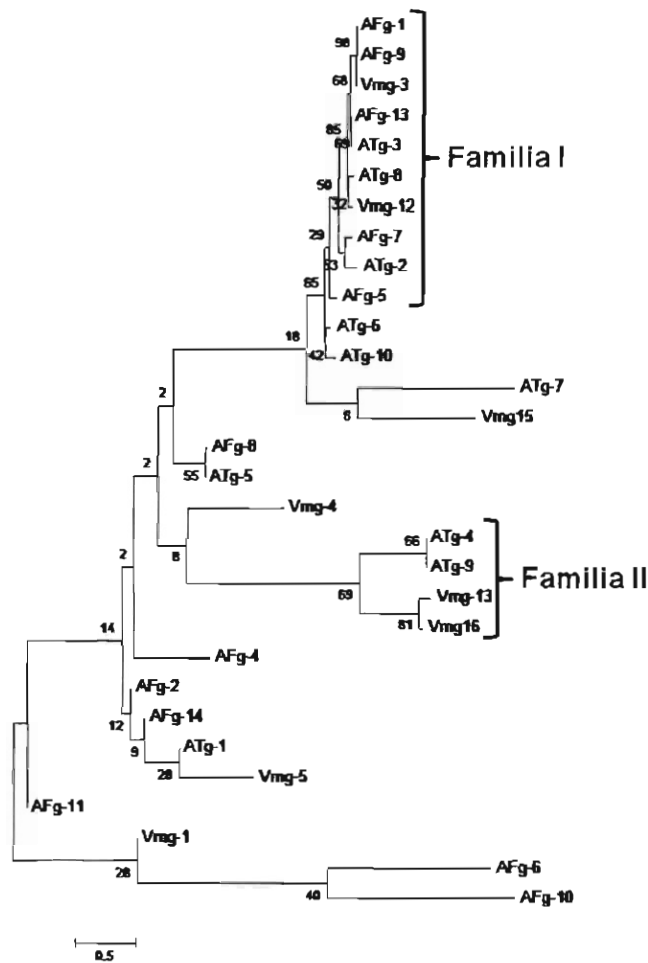


**Figura 3.3 Alineamiento de las 30 secuencias de Gypsy.**

El alineamiento cuenta con solo el 11 % de identidad. La conservación de los nucleótidos apenas alcanza el 60%.

Se construyo un árbol filogenético con el programa MEGA utilizando el método Neighbour Joining, la distancia evolutiva fue calculada con el método compuesto de máxima similitud con análisis de remuestreo de 1000 replicas (**Figura 3.4**).

De acuerdo a los parámetros de identificación de familias, solo se pueden identificar II familias. La Familia I incluye secuencias de las tres especies de *Agave* con un porcentaje de identidad del 60%. Mientras que la Familia II solamente incluye secuencias de *A. tequilana* y *A. fourcroydes* con un porcentaje de identidad del 66%. El 32% y 13% de las secuencias está representado en la familia I y II respectivamente.



**Figura 3.4** Árbol filogenético Gypsy.

La historia evolutiva fue inferida utilizando el método Neigbor-Joinig<sup>19</sup> con un análisis de remuestreo con 1000 replicas<sup>20</sup>. Con base en los parámetros de identificación de familias, se pudieron diferenciar dos familias.

### 3.4 DISCUSIÓN.

El género *Agave* es un género relativamente joven 7-11 millones de años<sup>22</sup> en comparación con las angiospermas<sup>23</sup> presenta un cariotipo bimodal con diferentes niveles de ploidía<sup>24-27</sup> y el tamaño del genoma aumenta proporcionalmente al nivel de ploidía<sup>27</sup>. Así mismo, las especies de este género se reproducen predominantemente asexualmente<sup>28</sup>. Este lo hace un género interesantes para realizar estudio de evolución de los genomas y una estrategia de análisis es a través del estudio de las regiones repetidas como satélites de ADN y retrotransposones.

Por otro lado, el genoma de algunas especies de otros géneros está representado por más del 50% de retrotransposones<sup>1, 7</sup>. Se ha discutido que estos elementos son una fuente de diversidad genética, actuando como agentes mutagénicos debido a su movilidad en el genoma<sup>29</sup>. Estos elementos contribuyen en la estructura de los cromosomas y pueden ser regiones en donde ocurre la recombinación<sup>1</sup>.

Este es el primer trabajo que trata de identificar diferentes familias de retrotransposones LTR en especies de *Agave* con base en el dominio conservado de transcriptasa reversa, a diferencia del trabajo previo, enfocado en estudios de variabilidad genética, para discriminar entre diferentes especies de *Agave* y variedades de *A. tequilana*<sup>30</sup>.

Para la identificación de familias se utilizaron las secuencias descritas en el capítulo II y se encontró que las secuencias de Ty1-copia y Gypsy presentan homología con secuencias reportadas en el GenBank pertenecientes a diversas especies de plantas. Sin embargo, presentan una elevada heterogeneidad, con un 22 y 11% de identidad entre todas las secuencias analizadas de Ty1-copia y Gypsy respectivamente. Esto ya ha sido reportado en estudios previos con valores que van desde 10 hasta 60% de identidad entre especies de diversos géneros<sup>31, 13-14</sup>. En algunas especies del género *Vigna*, por ejemplo, se ha observado que los retrotransposones del tipo Gypsy son más heterogéneos que los Ty1-copia<sup>16</sup>. La heterogeneidad de las secuencias de los retrotransposones ha sido atribuida a la elevada tasa de mutación ( $2.5 \times 10^{-5}$  nucleótidos por ciclo de transposición) asociada a la transcripción reversa<sup>32</sup>. Así mismo, cada ciclo de transposición puede originar diferentes secuencias pero estrechamente relacionadas y originar diferentes familias<sup>33, 11</sup>.

En el presente trabajo, con base en el agrupamiento y el porcentaje de identidad ( $\geq 60\%$ ), fue posible identificar al menos 6 familias de Ty1-copia y 2 de Gypsy. Por lo que en *Agave*, como en otras especies éste tipo de elementos está presente en el genoma con un elevado nivel de diversidad debido a mutaciones.

Los resultados obtenidos en el presente trabajo son similares a los reportados para otros géneros como *Oryza* donde han sido identificadas 32 familias<sup>12</sup>, 7 en *Triticum*<sup>13</sup>, al menos 7 *Hypochoeris*<sup>14</sup>. 13 en *Arabidopsis thaliana*<sup>15</sup>, siendo 6 y 5 familias de Ty1-copia y Gypsy respectivamente. Pareciera que los retrotransposones más abundantes son los Ty1-copia, esto es porque han sido los más estudiados.

En resumen en el genoma de *Agave*, como en otros genomas vegetales, éste tipo de elementos está presente y pudiera estar asociado a procesos que dirigen la evolución.

La información disponible hasta el momento sugiere que estos elementos son abundantes en el genoma de *Agave*. Pueden tener un papel importante en evolución, en diversos procesos celulares como respuesta a estrés ambiental o el papel que estos puedan jugar durante rearrreglos genómicos, como por ejemplo durante la ploidización del genoma o durante los procesos de recombinación en la reproducción sexual.

Es necesario entender con mayor profundidad como se regulan, si existen sitios específicos de inserción o cual es la contribución de cada una de las familias propias de una especie o en qué momento se dispara la transposición.

### 3.5 BIBLIOGRAFÍA

1. A. Kumar; J. L. Bennetzen 1999, Plant retrotransposons. *Annu Rev. Genet.* 33, 429-532, 1999.
2. Y. Xion; T. H. Eickbush, Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO J.* 9:10, 3353-3362, 1990.
3. B. Lewin, *Genes VII.* Oxford University Press 2000.
4. A. J. Flavell; D. B. Smith; A. Kumar, Extreme heterogeneity of Ty1-copia group retrotransposons in plants. *Mol. Gen. Genet.* 231, 233-242, 1992.
5. H. Hirochika; R. Hirochika, Ty1-copia group retrotransposons as ubiquitous components of plant genome. *Jpn. J. Genet.* 68, 35-46, 1993.
6. N. Kumekawa; E. Ohtsubo; H. Ohtsubo, Identification and phylogenetic analysis of gypsy-type retrotransposons in the plant kingdom. *Genes Genet. Syst.* 74, 299-307. 1999.
7. P. Sanmiguel; J. I. Bennetzen, Evidence that a recent Increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* 82: A, 37-44, 1998.
8. R. Wu; W. L. Guo; X. R. Wang; X. L. Wang; T. T. Zhuang; L. C. Jihong; B. Liu, Unintended consequence of plant transformation: biolistic transformation caused transpositional activation of an endogenous retrotransposon Tos17 in rice ssp. japonica cv. Matsumae. *Plant Cell Rep.* 28:7, 1043-51, 2009.
9. S. Tsukahara; A. Kobayashi; A. Kawabe; O. Mathieu; A. Miura; T. Kakutani, Bursts of retrotransposition reproduced in Arabidopsis. *Nature* 461. pp 423-427 2009.
10. M. Mirouze; J. Reinders; E. Bucher; T. Nishimura; K. Schneeberger; S. Ossowski; J. Cao; D. Weigel; J. Paszkowski; O. Mathieu, Selective epigenetic control of retrotransposition in Arabidopsis. *Nature*, 461, 427-43, 2009.
11. K. Alix; J. S. Heslop-Harrison, The diversity of retroelement in diploid and allotetraploid Brassica species *Plant Mol. Biol.* 54, 895-909, 2004.

12. H. Hirochika, Retrotransposons of rice: their regulation and use for genome analysis. *Plant Mol. Biol.* 35, 231–240, 1997.
13. Y. Matsuoka; K. Tsunewaki, Wheat retrotransposon families identified by reverse transcriptase domain analysis. *Mol. Biol. Evol.* 13:10, 1384-1392, 1996.
14. C. F. Ruas; H. Weiss-Schneeweiss; T. F. Stuessy; M. R. Samuel; A. Pedrosa-Harand; K. Tremetsberger; P. M. Ruas; P. M. Schlüter; M. A. Ortiz Herrera; C. König; N. I. Matzenbacher, Characterization, genomic organization and chromosomal distribution of Ty1-copia retrotransposons in species of *Hypochaeris* (Asteraceae). *Gene*, 15:412 (1-2), 39-49, 2008.
15. A. Konieczny; D. F. Voytas; M. P. Cummings; F. M. Ausubel, A superfamily of *Arabidopsis thaliana* retrotransposon. *Genetics*, 127, 801-809, 1991.
16. A. Dixit; J. Ho Ma; J. Woong Yu; E. Gi Cho; J. Jin Park, Reverse transcriptase domain sequences from Mungbean (*Vigna radiata*) LTR retrotransposons: Sequence characterization and phylogenetic analysis. *Heredity*, 97, 381–388, 2006.
17. S. F. Altschul; T. L. Madden; A. A. Schäffer; J. Zhang; Z. Zhang; W. Miller; D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389-3402, 1997.
18. K. Tamura; J. Dudley; M. Nei; S. Kumar, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596-1599, 2007.
19. N. Saitou; M. Nei, The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406-425, 1987.
20. J. Felsenstein, Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783-791.
21. K. Tamura; M. Nei; S. Kumar, Prospects for inferring very large phylogenies by using the neighbor-joining method. *PNAS*, 101, 11030-11035, 2004.
22. S. V. Good-Avila; V. Souza; S. Gaut Brandon; L. E. Eguiarte, Timing and rate of speciation in *Agave* (Agavaceae). *PNAS*. 103:24, 9124–9129, 2006.

23. S. Magañon; M. J. Sanderson, Absolute diversification rates in angiosperm clades. *Evolution*, 55:9, 1762–1780, 2001,
24. G. Backman, A Karyosystematic study of the genus *Agave*. *Am. J. Bot.* 32:5, 283-298 1944.
25. S. F. Moreno Salazar; M. Esqueda; J. Martínez; G. Palomino, Tamaño del genoma y cariotip en *Agave angustifolia* y *A. rhodacantha* de Sonora, México. *Rev. Fitotecnia Mexicana* 30:1, 13-23, 2007.
26. G. Palomino; J. Martínez; I. Méndez, Citotipos en *Agave angustifolia* Haw. determinados por citometria de flujo y análisis de sus cariotipos. *Rev. Int. Contam. Ambient.* 21:1 pp 49-54, 2005.
27. M. L. Robert; K. Y. Lim; L. Hanson; F. Sanchez-Teyer; M. D. Bennett; A. R. Leitch; I. J. Leitch, 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, 215–222 (2008).
28. H. S. Gentry, *Agaves of continental north America*. (The university of Arizona press). 1982
29. J. M. Deragon; P. Cappy, Impact of transposable elements on the human genome. *Ann. Med.* 32, 264-273, 2000.
30. A. Bousios; I. Saldana-Oyarzabala; A. G. Valenzuela-Zapatab; C. Wooda; S. R. Pearce, 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 Science*, 172:2, 291-298, 2007,
31. Z. Price; F. Dumortier; D. W. MacDonald; S. Mayes, Characterization of copia-like retrotransposons in oil palm (*Elaeis guineensis* Jacq.). *Theor. Appl. Genet.* 104, 860–867, 2002.
32. B. Preston, Error-prone retrotransposition: Rime of the ancient mutators, *Commentary PNAS*, 93, 7427-7431.1996.



33. J. M. Casacuberta; S. Vernhettes; C. Audeon; M. A. Grandbastien. Quasispecies in retrotransposons: a role for sequence variability in Tnt1 evolution. *Genetica*, 100, 109–117 1997.

#### Capítulo IV. Discusión general.

El género *Agave* es un género con una edad de 7 a 11 millones de años, siendo evolutivamente joven dentro de las angiospermas<sup>1</sup> y muestra una gran capacidad de adaptación a diferentes ambientes, con una tasa de especiación de 0.32 especies por millón de año en comparación con la tasa de diversificación de las angiospermas (0.077-.089) poniendo de manifiesto su plasticidad genética<sup>2</sup>.

Los agaves como modelo para estudios evolutivos son interesantes ya que presentan diversas características entre las que se pueden mencionar un cariotipo bimodal con diferentes niveles de ploidía desde diplóides hasta octoploides<sup>3-6</sup> y un tamaño de genoma que aumenta proporcionalmente con el nivel de ploidía<sup>6</sup>. Así mismo, las especies de este género se reproducen de manera predominantemente asexualmente a través de la formación de rizomas o bulbillos<sup>7</sup> sin embargo, muchas de las especies del género presentan reproducción sexual igualmente eficiente, sobre todo en aquellas especies diploides ó tetraploides, no así en triploides y pentaploides. Todas estas características en su conjunto permiten plantear preguntas desde un punto de vista de la evolución de su genoma.

Los retrotransposones LTR, que son elementos móviles que se mueven dentro del genoma a través de un intermediario de ARN<sup>8</sup> están ampliamente distribuidos en el genoma de las plantas<sup>9-10</sup> de manera importante y en algunas especies constituye más del 50% de su genoma<sup>11</sup>.<sup>8</sup> Este tipo de elementos están involucrados en otras especies vegetales en cambios genéticos y muchas veces morfológicos, como el caso de la uva roja<sup>12</sup>, incrementando la variabilidad los genomas necesaria para su adaptación, a través de la generación de mutaciones permanentes y una contribución importante en el incremento del tamaño del genoma vegetal<sup>13</sup>.

Se han reportado diversos tipos de retrotransposones entre los que se encuentran los Ty1-copia y los Gypsy<sup>14</sup> y se propone que tanto los transposones como los retrotransposones son una fuente generadora de cambios genéticos que favorecen la plasticidad genómica, debido a su movilidad dentro del genoma<sup>15</sup>, y en particular se ha propuesto que los retrotransposones pueden actuar como "sitios calientes" en donde

ocurre la recombinación, por su alto nivel de repetición y similitud en secuencia como parte de la estructura de los cromosomas<sup>8</sup>.

En un gran número de especies vegetales, tanto monocotiledóneas como dicotiledóneas, se han aislado y caracterizado retrotransposones a través de análisis de secuencias del dominio conservado de transcriptasa reversa, pudiendo diferenciar entre TY1-Copia y Gypsy<sup>16-20, 9-10</sup>.

En este trabajo se aislaron dos diferentes dominios de RT empleando iniciadores degenerados que se han utilizado de manera exitosa en diferentes especies<sup>9, 17-18</sup>. Las secuencias aisladas de diversas especies de *Agave* mostraron una elevada heterogeneidad y valores de identidad de entre 11% y 22% para Gypsy y Ty1-copia respectivamente. Esto ya ha sido reportado en otros estudios y los valores van desde 10-60% de identidad para especies de los géneros *Vigna*, *Triticum*, *Hypochaeris*<sup>16, 19-20</sup>, entre muchos otros.

Estos valores de identidad observados en todos los reportes sugiere que estos elementos de origen viral, presentes en aún en especies ancestrales<sup>14</sup> presentan una elevada tasa de mutación, posiblemente debido a la incapacidad de la transcriptasa reversa de editar la secuencia de novo<sup>21-22</sup>.

Hasta nuestro conocimiento, este es el primer trabajo que trata de identificar diferentes familias de retrotransposones LTR en especies de *Agave* con base en el dominio conservado de transcriptasa reversa y nuestros resultados son comparables con lo reportado para especies del género *Vigna*, donde los retrotransposones del tipo Gypsy son más heterogéneos que los Ty1-copia<sup>17</sup>.

Si bien estos elementos contienen regiones conservadas en regiones génicas como lo son los dominios de RNAsaH y transcriptasa reversa, tiene una elevada heterogeneidad del resto de la secuencia misma que ha sido atribuida a la tasa de mutación propia asociada a la transcripción reversa<sup>22-24</sup> y mecanismos de reinserción dentro otros retrotransposones como se ha demostrado para tabaco<sup>25</sup>.

En el presente trabajo, con base en el agrupamiento y el porcentaje de identidad ( $\geq 60\%$ ), fue posible identificar al menos 6 familias de Ty1-copia y 2 de Gypsy en las tres especies

de *Agave* analizadas, demostrando que, como en otras especies, éste tipo de elementos está presente en el genoma con un elevado nivel de diversidad debido a mutaciones.

Los resultados obtenidos en el presente trabajo son comparables a los reportados para géneros como *Oryza*<sup>26</sup>, *Triticum*<sup>20</sup> *Hypochaeris*<sup>16</sup> ó *Arabidopsis*<sup>27</sup>, en los cuales hay evidencia de la existencia de retrotransposones del tipo Ty1-copia y Gypsy en diferentes proporciones y representados en un gran número de familias que difieren entre sí de manera considerable. La evidencia en la literatura sugiere que los retrotransposones Ty1-copia son más abundantes que el resto de los retrotransposones.

Con los resultados obtenidos del agrupamiento para los retrotransposones Ty1-copia se sugieren que las familias I y II, III y IV, V y VI comparten el mismo ancestro común pero sin embargo han seguido dos rutas evolutivas diferentes independientemente de la especie. Así mismo, se puede sugerir que las familias VI y IV se originaron antes de que divergieran estas tres especies por que están representadas en las tres especies en estudio. Mientras que las familias I y V pueden estar mayormente representadas en *A. fourcroydes* y *A. angustifolia*, mientras que la familias II está mayormente representada en *A. tequilana* y de *A. angustifolia*, la III de *A. tequilana* y de *A. fourcroydes*. Sin embargo, los resultados obtenidos hasta el momento no son concluyentes y es necesario profundizar en el conocimiento del número de copias y su distribución en los diversos genomas a través de Hibridación Fluorescente in Situ (FISH) y mediante el diseño iniciadores específicos para las familias I, II, III y V para dar lugar a la amplificación por PCR específicamente en las especies mencionadas y determinar el número de copias en cada genoma y su nivel de expresión para esclarecer si continúan activos. De comprobarse la especificidad de cada familia en las diversas especies del género, se podría sugerir que las familias I, II y V se originaron inicialmente en *A. angustifolia* pero siguieron dos rutas evolutivas diferentes en la divergencia entre *A. tequilana* y *A. fourcroydes*, mientras que la familia III se origino antes de la divergencia entre *A. tequilana* y *A. fourcroydes*. Esto se puede sustentar con la hipótesis de que *A. angustifolia* es el ancestro común de *A. tequilana* y *A. fourcroydes*<sup>28</sup> y las características genómicas de poliploidía de los agaves, como ha sido reportado para tabaco donde se han reportado "comportamientos específicos" de diversas poblaciones de retrotransposones dependiendo de la historia evolutiva del hospedero con períodos rápidos de recambio de secuencias estimulado por procesos de aloploidización<sup>25</sup>.

En un reporte previo, Bouosius y colaboradores (2008)<sup>29</sup> se demostró la existencia de retrotransposones TY1-Copia en diversas especies de agaves y se concluye que existen retrotransposones ancestrales presentes en diversas especies de la familia Agavaceae, sugiriendo la existencia de retrotransposones en puntos iniciales de la divergencia de las especies del género.

El agrupamiento de los retrotransposones Gypsy sugiere por un lado que la familia I es ancestral en el genoma de las tres especies analizadas, mientras que la familia II se origina después de la divergencia de *A. angustifolia*, ó que ambas familias estuvieron presentes en la especie ancestral pero las secuencias de la familia I dieron origen a la familia dos mediante mutaciones adquiridas durante el movimiento de dichas familias. Esto último sugeriría que existen familias con mayor actividad insercional que origina diversos niveles de variación.

Los retrotransposones pudieran ayudar a esclarecer las relaciones filogenéticas en algunos géneros, por ejemplo en *Solanum* los resultados filogenéticos con secuencias parciales de TR de Ty1-copia concuerda con la historia evolutiva de tomate (*Solanum lycopersicum*) y las especies relacionadas filogeneticamente<sup>30</sup>.

De acuerdo a la hipótesis planteada la familia III pudo originarse después de que *A. tequilana* y *A. fourcroydes* divergieron de *A. angustifolia*. Sin embargo nuestro alcance no es suficiente para concluir con esta hipótesis. Pero al menos fue posible identificar 8 familias de retrotransposones.

La información generada hasta el momento demuestra que diversas familias de estos elementos están presentes en el genoma de *Agave* y éstos muestran divergencia en secuencia, de manera similar a muchas otras especies vegetales tanto monocotiledóneas como dicotiledóneas, pudiendo tener un papel importante en evolución, participando en diversos procesos celulares ó como respuesta a estrés ambiental o jugando un papel importante como agentes que generan variación genética durante rearrreglos genómicos, como por ejemplo durante la ploidización del genoma o durante los procesos de recombinación en la reproducción sexual.

Aunque se ha reportado que la transposición de éstos elementos en arroz se dispara en condiciones de estrés<sup>31</sup> así mismo en *Arabidopsis thaliana* la transposición está regulada

por metilación<sup>32-33</sup> o de que el ambiente de cultivo *in vitro* induce su activación, es necesario entender con mayor profundidad como se regula su activación y si existen sitios específicos de inserción dentro del genoma. Adicionalmente es indispensable determinar cuál es la contribución de cada una de las familias propias de una especie al tamaño del genoma o en qué momento y bajo que circunstancia se dispara su transposición.

En *Agave* se desconoce el papel que puedan jugar los retrotransposones: en los procesos de poliploidización, de recombinación, de especiación o la forma en que se regulan. Se desconoce la distribución y contribución al tamaño del genoma de estos elementos. Toda esta información puede ser valiosa para poder determinar el origen de los poliploides.

#### 4.1 CONCLUSIONES.

- ❖ Los retrotransposones LTR (Gypsy y Ty1copia) están presentes en el genoma de las tres especies analizadas.
- ❖ Los retrotransposones Gypsy son más heterogéneos que los Ty1-copia con un porcentaje de identidad de 11% y 22% respectivamente.
- ❖ En este trabajo se identificaron 6 familias para Ty1-copia y 2 en Gypsy.

## 4.2 PERSPECTIVAS.

Para identificar el número de familias de retrotransposones en *Agave* y poder hacer inferencias evolutivas, es necesario incluir las especies representativas del género e identificar las familias de retrotransposones ancestrales y las que se originaron después de la divergencia de las especies. La localización física en los cromosomas por hibridación *in situ* fluorescente (FISH) puede ayudar a esclarecer el origen de los poliploides y a sugerir alguna función relacionada con su localización

Para tener un acercamiento de la contribución al tamaño del genoma de cada una de las familias, se puede determinar el número de copias por PCR en tiempo real y considerando el tamaño promedio de los retrotransposones.

Para entender si la transposición se dispara en condiciones de estrés. Se puede medir los niveles de expresión por PCR en tiempo real en diferentes estreses.

Se ha propuesto que los retrotransposones generan variabilidad genética en plantas que se reproducen asexualmente, esto se puede ser comprobado evaluando la variabilidad entre hija y planta madre por medio de marcadores moleculares SSAP.



## 4.2 BIBLIOGRAFÍA

1. M. Rocha; S. V. Good-Avila; F. Molina-Freaner; H. Arita; A. Castillo; A. García-Mendoza; A. Silva-Montellano; B. Gaut; V. Souza; L. Eguiarte, *Monocots: Comparative Biology and Evolution*. Columbus J. T., Friar E. A., Porter J. M., Prince L. M., Simpson M. G., editors. Clairmont, CA: Rancho Santa Ana Botanical Gardens; 2006. in press
2. S. Magaña; M. J. Sanderson, Absolute diversification rates in angiosperm clades. *Evolution*, 55:9, 1762–1780, 2001,
3. G. Backman, A Karyosystematic study of the genus *Agave*. *Am. J. Bot.* 32:5, 283-298 1944.
4. S. F. Moreno Salazar; M. Esqueda; J. Martínez; G. Palomino, Tamaño del genoma y cariotip en *Agave angustifolia* y *A. rhodacantha* de Sonora, México. *Rev. Fitotecnia Mexicana* 30:1, 13-23, 2007.
5. G. Palomino; J. Martínez; I. Méndez, Citotipos en *Agave angustifolia* Haw. determinados por citometria de flujo y análisis de sus cariotipos. *Rev. Int. Contam. Ambient.* 21:1 pp 49-54, 2005.
6. M. L. Robert; K. Y. Lim; L. Hanson; F. Sanchez-Teyer; M. D. Bennett; A. R. Leitch; I. J. Leitch, 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, 215–222 (2008).
7. H. S. Gentry, *Agaves of continental north America*. (The university of Arizona press). 1982
8. A. Kumar; J. L. Bennetzen 1999, Plant retrotransposons. *Annu Rev. Genet.* 33, 429-532, 1999.
9. D. F. Voytas; M. P. Cummings; A. Konieczny; F. M. Ausubel; S. R. Rodermel, Copia-like retrotransposons are ubiquitous among plants. *PNAS*, 89, 7124–7128, 1992.
10. A. J. Flavell; D. B. Smith; A. Kumar, Extreme heterogeneity of Ty1-copia group retrotransposons in plants. *Mol. Gen. Genet.* 231, 233-242, 1992.

11. P. Sanmiguel; J. I. Bennetzen, Evidence that a recent Increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* 82: A, 37-44, 1998.
12. S. Kobayashi; N. Goto-Yamamoto; H. Hirochika, Retrotransposon-Induced Mutations in Grape Skin Color. *Science*, 304:5673, 982, 2004.
13. A. Belyayev; O. Raskina; A. Korol; E. Nevo, Coevolution of A and B genomes in allotetraploid *Triticum dicoccoides*. *Genome* 43, 1021–1026, 2000.
14. Y. Xion; T. H. Eickbush, Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO J.* 9:10, 3353-3362, 1990.
15. J. M. Deragon; P. Cappy, Impact of transposable elements on the human genome. *Ann. Med.* 32, 264-273, 2000.
16. C. F. Ruas; H. Weiss-Schneeweiss; T. F. Stuessy; M. R. Samuel; A. Pedrosa-Harand; K. Tremetsberger; P. M. Ruas; P. M. Schlüter; M. A. Ortiz Herrera; C. König; N. I. Matzenbacher, Characterization, genomic organization and chromosomal distribution of Ty1-copia retrotransposons in species of *Hypochaeris* (Asteraceae). *Gene*, 15:412 (1-2), 39-49, 2008.
17. A. Dixit; J. Ho Ma; J. Woong Yu; E. Gi Cho; J. Jin Park, Reverse transcriptase domain sequences from Mungbean (*Vigna radiata*) LTR retrotransposons: Sequence characterization and phylogenetic analysis. *Heredity*, 97, 381–388, 2006.
18. A. Ghany; Saki, Isolation characterization and analysis of copia-like retrotransposons in the Egyptian cotton *Gossypium barbadense* and its progenitors. *African Journal of biotechnology*. 2:6, 165-168. 2003.
19. Z. Price; F. Dumortier; D. W. MacDonald; S. Mayes, Characterization of copia-like retrotransposons in oil palm (*Elaeis guineensis* Jacq.). *Theor. Appl. Genet.* 104, 860–867, 2002.
20. Y. Matsuoka; K. Tsunewaki, Wheat retrotransposon families identified by reverse transcriptase domain analysis. *Mol. Biol. Evol.* 13:10, 1384-1392, 1996.

21. M. Boutabout; M. Wilhelm; F.-X. Wilhelm, DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1. *Nucleic Acids Research*. 29:11, 2217-2222, 2001.
22. B. Preston, Error-prone retrotransposition: Rime of the ancient mutators, *Commentary PNAS*, 93, 7427-7431.1996.
23. J. M. Casacuberta; S. Vernhettes; C. Audeon; M. A. Grandbastien. Quasispecies in retrotransposons: a role for sequence variability in Tnt1 evolution. *Genetica*, 100, 109–117 1997.
24. K. Alix; J. S. Heslop-Harrison, The diversity of retroelement in diploid and allotetraploid *Brassica* species *Plant Mol. Biol.* 54, 895-909, 2004.
25. M. Petit; K. Y. Lim; E. Julio; C. Poncet; B. F. Dorlhac; A. Kovarik; A. R. Leitch; M.-A. Grandbastien; M. Corinne, Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*). *Mol Genet Genomics* 278 pp. 1–15 (2007).
26. H. Hirochika, Retrotransposons of rice: their regulation and use for genome analysis. *Plant Mol. Biol.* 35, 231–240, 1997.
27. A. Konieczny; D. F. Voytas; M. P. Cummings; F. M. Ausubel, A superfamily of sradidopsis thaliana retrotransposon. *Genetics*, 127, 801-809, 1991.
28. P. Colunga-García Marín; A. Larqué Saavedra; L. E. Eguiarte & D. Zizumbo-Villarreal, En lo ancestral hay futuro: del tequila, los mezcales y otros agaves. CICY-CONACYT-CONABIO-INE. Mérida, Yucatán. 2007, 402
29. A. Bousios; I. Saldana-Oyarzabala; A. G. Valenzuela-Zapatab; C. Wooda; S. R. Pearce, 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 Science*, 172:2, 291-298, 2007,
30. S. M. Tam; M. Caussem; C. Garchery; H. Burck; C. Mhiri; M. A. Grandbastien, The distribution of copia-type retrotransposons and theevolutionary history of tomato and

related wild species. Journal compilation. European society for evolutionary biology. 20, 1056–1072, 2007.

31. R. Wu; W. L. Guo; X. R. Wang; X. L. Wang; T. T. Zhuang; L. C. Jihong; B. Liu, Unintended consequence of plant transformation: biolistic transformation caused transpositional activation of an endogenous retrotransposon Tos17 in rice ssp. japonica cv. Matsumae. Plant Cell Rep. 28:7, 1043-51, 2009.

32. S. Tsukahara; A. Kobayashi; A. Kawabe; O. Mathieu; A. Miura; T. Kakutani, Bursts of retrotransposition reproduced in Arabidopsis. Nature 461. pp 423-427 2009.

33. M. Mirouze; J. Reinders; E. Bucher; T. Nishimura; K. Schneeberger; S. Ossowski; J. Cao; D. Weigel; J. Paszkowski; O. Mathieu, Selective epigenetic control of retrotransposition in Arabidopsis. Nature, 461, 427-43, 2009.