



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**UN ENFOQUE FILOGENÉTICO SOBRE EL CAMBIO  
EN LA TASA DE MUTACIÓN DEL GENOMA  
MITOCONDRIAL DE LAS GRAMÍNEAS ASOCIADO  
A LA PRESENCIA DE MICROSATÉLITES.**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**BIÓLOGO**

**P R E S E N T A:**

**MARCO ANTONIO ALAN GARDUÑO SÁNCHEZ**



**DIRECTOR DE TESIS:  
Dr. JUAN PABLO JARAMILLO CORREA**

**2018**

CIUDAD UNIVERSITARIA, CD.MX.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## **Dedicatoria.**

Quisiera dedicar el presente trabajo a mis padres, a mis hermanos y mis abuelos, por estar presentes durante mi crecimiento personal y profesional, pues para bien o para mal, gracias a ellos soy la persona en la que me he convertido. También a todos ellos que invirtieron su tiempo y su esfuerzo en mi preparación académica: a Anayansi Sierralta, por sus constantes comentarios académicos; a los profesores con los que tomé clase, como Ricardo García, a mi tutor Juan Pablo y a los doctores Carlos Pedraza y Patricia Ornelas, pues han sido la piedra angular de mi formación profesional.

Por último, este trabajo va dedicado a *Marco del pasado...*

...Lamento no haber tenido las palabras que necesitabas en ese entonces, ni el coraje para que pudieras ser más feliz de lo que eras. Pero si pudiera verte de nuevo, sé exactamente que te diría:

*“Lo que te define no son tus circunstancias del pasado, sino las decisiones que tomarás en el futuro. Tu habilidad para aprender y adaptarte es lo que te define hoy y eso dictará en lo que te convertirás. Tu mayor venganza no es el hecho de sólo estar vivo, sino que ahora eres más fuertes de lo que ellos pensaron jamás...¿Sabes lo increíble que es eso?”.*

Que mi trabajo siempre sirva como testimonio de tu existencia y estés en donde estés, espero que seas feliz.

Marco G.

## **Agradecimientos.**

Quiero agradecer a mi familia: a mis padres, porque me enseñaron la importancia de seguir, aunque no fuera fácil, y de abandonar, aunque fuera difícil. A mis hermanos, por ayudarme a desarrollar las virtudes de la fuerza y la paciencia. A mis abuelos, por enseñarme desde chico el nexo inseparable entre libertad y disciplina.

A mis amigos: A Lulú por ayudarme con los muestreos, a Nancy por ayudarme con el laboratorio, a Tania por ayudarme con el trámite administrativo de la titulación, A Gustavo por ser un excelente estímulo intelectual. A Omar por ayudarme a mejorar en mi trabajo. A Miguel, por las correcciones y su tiempo. A Emma por volver a mi vida y ayudarme a tomar mejores decisiones. Quiero agradecer también a todos aquellos que estuvieron conmigo durante un momento, y luego continuaron con sus vidas, porque las cosas que aprendí a su lado son valiosas, incluso ahora que ya no están.

A Anayansi Sierralta, por darme las herramientas para poder ser yo mismo. Por ser mi colega, mi amiga y mi compañera de juegos, y no permitirme nunca el ser cruel conmigo mismo.

A aquellas personas que formaron parte de mi formación académica: A mis maestros, por mostrarme que aprender es ser autodidacta. A Juan Pablo, por enseñarme a ser crítico con mi trabajo. A Carlos Pedraza, por darme una oportunidad, y mil más. A Patricia Ornelas, por creer en mí, incluso cuando yo no podía hacerlo.

Muchas gracias por todo.

## Tabla de contenidos.

Resumen.....	1
Introducción.....	2
Características principales del genoma mitocondrial de las plantas.....	2
Secuencias repetidas dentro del genoma mitocondrial de las plantas.....	6
Los microsatélites.....	8
Modelo de estudio: las gramíneas.....	13
Evaluación de la diversidad genética.....	16
Justificación.....	21
Hipótesis.....	22
Objetivos.....	23
Objetivo general.....	23
Objetivos específicos:.....	23
Métodos.....	24
Búsqueda de los microsatélites.....	24
Reconstrucción filogenética.....	27
Colecta, extracción y amplificación.....	28
Análisis de la diversidad nucleotídica.....	33
Reconstrucción de caracteres ancestrales.....	34
Resultados.....	345
Búsqueda de microsatélites en los genomas mitocondriales de pastos depositados en el NCBI.....	35
Descripción de las regiones de interés.....	37
Análisis de la diversidad nucleotídica en las regiones de interés.....	48
Reconstrucción filogenética.....	50
Reconstrucción de caracteres ancestrales.....	59
Discusión.....	65
Distribución de elementos repetidos en las distintas regiones de los genomas mitocondriales.....	65
Comparación de la diversidad nucleotídica entre regiones con microsatélites y sin microsatélites.....	68
Evolución y conservación de los microsatélites seleccionados en el genoma mitocondrial de los pastos.....	71
Presencia de alelos nulos en el genoma mitocondrial vegetal.....	74
Durante la amplificación de las regiones portadoras de microsatélites, no fue posible amplificar un producto para algunos organismos. Existen varias razones por las que pudo haberse presentado este fenómeno:.....	74
1) Eventos de recombinación.....	74
2) Altas tasas de evolución en las regiones vecinas.....	75
3) Eventos de transferencia horizontal.....	75
Perspectivas.....	77
Referencias.....	79

## **Resumen.**

En este estudio se sugiere que la presencia de microsatélites puede estar relacionada al aumento en las tasas de diversidad genética del genoma mitocondrial vegetal de los pastos. La distribución de estas secuencias a través del genoma no ocurre de manera aleatoria, pues éstos suelen ser más abundantes en regiones intergénicas, que en regiones codificantes. Estos elementos pueden generarse dentro del genoma a través del deslizamiento de las hebras de DNA -y posteriores errores en el apareamiento de las hebras durante la replicación del material genético-, o ser adquiridos a través de transferencia horizontal, ya sea de otros compartimentos celulares o incluso de especies sexualmente incompatibles. Los microsatélites se encuentran presentes en una gran cantidad de grupos taxonómicos, y pueden ser retenidos durante largos periodos evolutivos. Además, se encontró una correlación positiva entre el número de microsatélites y el tamaño del genoma mitocondrial.

Los resultados obtenidos a partir de la reconstrucción de caracteres ancestrales sugiere que el aumento de las tasas de mutación no ocurre inmediatamente después de ser adquirido el microsatélite, haciendo que el aumento en la diversidad sea un fenómeno posterior a la inserción del mismo.

## **Introducción.**

### **Características principales del genoma mitocondrial de las plantas**

Las mitocondrias son las estructuras celulares encargadas de las funciones de óxido-reducción para la producción de energía de los organismos (Gualberto *et al.*, 2014). La teoría del origen monofilético de las mitocondrias postula un evento endosimbiótico en el cual una  $\alpha$ -proteobacteria y el ancestro común de los eucariontes coevolucionaron (Gray, 2012). La asociación de este complejo se ha mantenido a lo largo de la historia evolutiva, llevando a la divergencia de los genomas mitocondriales de los eucariontes a través del tiempo (Kitazaki & Kubo, 2010; Gray, 2012).

En particular, la evolución del genoma mitocondrial -o mitogenoma- de las plantas ha sido diferente de la de los demás eucariontes, tanto a nivel de secuencia como estructural (Palmer & Herbon, 1998). A lo largo de su historia evolutiva, estos genomas han llevado a cabo modificaciones en su contenido a través de diferentes procesos como la pérdida o ganancia de genes por medio de transferencias del y hacia el núcleo, eventos de duplicación, y re-arreglos genómicos, entre otros (Fauron *et al.*, 1995; Cupp & Nielsen, 2014).

Generalmente, el genoma mitocondrial de las plantas es representado como una sola estructura de forma circular, similar a los genomas bacterianos, llamada cromosoma maestro (Kubo & Newton, 2008; Sloan *et al.*, 2012; Sloan, 2013). Esta estructura presenta características únicas que la distinguen de otros genomas; por ejemplo, el tamaño de los genomas mitocondriales muestra gran variación entre distintos grupos de plantas y presentan una gran cantidad de reacomodos estructurales a través de eventos de recombinación homóloga (Fauron *et al.*, 1995).

Con respecto a la variación del genoma mitocondrial, el rango puede ir desde 66 a 2,000 kilobases (kb; Sloan *et al.*, 2012; Skippington *et al.* 2015; Guo *et al.*, 2016) llegando en algunos casos a ser más grandes que algunos genomas bacterianos. Sin embargo, a pesar del gran tamaño que pueden alcanzar, los genomas mitocondriales vegetales no parecen albergar un número mayor de genes que otros genomas mitocondriales con un tamaño menor (Gualberto, 2014; Skippington *et al.*, 2016). El genoma mitocondrial vegetal está conformado por secuencias codificantes, separadas por regiones intergénicas que pueden llegar a abarcar hasta el 90% del mismo (Alverson *et al.*, 2010, Gualberto *et al.*, 2014). Por ejemplo, el genoma mitocondrial completo de *Arabidopsis thaliana* (376Kb) sólo posee 14 genes más que el genoma mitocondrial humano y es 20 veces más extenso que éste (16.6 kb) (Palmer, 2000).

Por otro lado, el tamaño del genoma mitocondrial vegetal parece variar a lo largo de tiempo, ya que puede llegar a presentar contracciones o expansiones en periodos de tiempo evolutivamente cortos (Guo *et al.*, 2016). Algunas de las expansiones más significativas podemos encontrarlas en la calabaza, *Cucurbita pepo* (Alverson *et al.*, 2010), o en representantes del género *Silene* (Sloan *et al.*, 2012). Estas expansiones se deben principalmente a eventos de duplicación y a la adquisición de DNA de origen exógeno a través de la asimilación de secuencias provenientes del núcleo o del cloroplasto (Palmer *et al.*, 2000, Kubo & Newton, 2008; Alverson *et al.*, 2010), o por acumulación de secuencias repetidas en las regiones intergénicas o incluso en intrones. (Alverson *et al.*, 2010, Jaramillo-Correa *et al.*, 2013). También se tiene evidencia de la exportación de genes del mitogenoma hacia el genoma nuclear, lo cual participaría en eventos de contracción (Skippington *et al.*, 2016).



Otra característica de los genomas mitocondriales vegetales son sus bajas tasas de mutación. Los estimados de éstas para distintas regiones del genoma mitocondrial vegetal suelen ser bajas, llegando a ser de 50 a 100 veces menores que la de genomas mitocondriales de animales, e incluso aún más bajos que en genes nucleares de plantas y animales. (Laroche *et al.*, 1997; Palmer *et al.*, 2000; Galtier, 2011). Como ejemplo de lo anterior, Wolfe y colaboradores (1987) encontraron que las tasas de sustitución sinónimas en los genomas mitocondrial y del cloroplasto de las plantas es de  $0.2 - 0.1 \times 10^{-9}$  y  $0.1 - 0.3 \times 10^{-9}$  respectivamente, mientras que los valores de sustitución en los genomas mitocondrial y nuclear de los primates son  $21.8 - 43.7 \times 10^{-9}$  y  $0.9 - 1.9 \times 10^{-9}$ , respectivamente. Esta característica es compartida entre mitogenomas de varios grupos de plantas, como angiospermas y gimnospermas, suponiéndose entonces como un carácter primitivo en la evolución estos genomas (Richardson *et al.*, 2013) Para explicar las bajas tasas de evolución dentro del genoma mitocondrial se utiliza la hipótesis de la carga mutacional. Ésta se basa en la idea de que los elementos no codificantes imponen un costo selectivo asociado al posible aumento de mutaciones que puedan alterar la función del genoma. Cuando una mutación surge en una población, puede llegar a fijarse por procesos de deriva génica si los tamaños efectivos son pequeños. Esto sugiere que elementos no codificantes pueden imponer un costo selectivo asociado al posible aumento de las mutaciones, por lo cual genomas de gran tamaño suelen presentar bajas tasas de mutación. (Charlesworth & Barton, 2004; Sloan *et al.*, 2012). Sin embargo, existen linajes específicos de plantas que muestran tasas elevadas de sustitución nucleotídica, poniéndolas a la par de genomas mitocondriales de evolución rápida, como los de los animales (Palmer *et al.*, 2000; Sloan *et al.*, 2012; Skippington *et al.*, 2016). Por ejemplo, los géneros *Plantago* y *Pelargonium* presentan tasas de sustitución en su genoma mitocondrial que exceden el

valor de la tasa de mutación del cloroplasto y del genoma nuclear en distintos taxa (Cho *et al.*, 2004; Parkinson *et al.*, 2005). Los mecanismos que llevan a un aumento de estas tasas de mutación en taxa específicos aún son desconocidos.

## **Secuencias repetidas dentro del genoma mitocondrial de las plantas**

Dentro de las regiones intergénicas de los mitogenomas vegetales se pueden encontrar secuencias repetidas (Kubo & Newton, 2008; Galtier, 2011), las cuales pueden llevar a cabo eventos de recombinación entre sí, haciendo que la estructura mitocondrial de las plantas sea multipartita -es decir, que el cromosoma maestro puede transformarse en moléculas subgenómicas luego de estos eventos de recombinación (Kubo & Newton, 2008; Kitaazaki & Kubo, 2010; Sloan *et al.*, 2012; Gualberto, 2014; Guo *et al.*, 2016)-, por lo que el genoma puede ser visto como una mezcla de estructuras alternativas originadas a partir de redes interactivas de secuencias repetidas altamente recombinantes (Maráchel & Brisson, 2010; Sloan *et al.*, 2013; Gualberto, 2014). Estos procesos de recombinación son los responsables del alto grado de polimorfismo estructural que presentan los mitogenomas de las plantas, haciendo de ellos entidades altamente dinámicas (Palmer *et al.*, 2000; Galtier, 2011). Este proceso de recombinación es conservado entre linajes, y hace difícil su representación física (Sloan *et al.*, 2012; Gualberto, 2014).

Las secuencias repetidas involucradas en los eventos de recombinación pueden ser de longitud larga (aproximadamente 1 ó 2 kb), o corta (de 50 a 600 pares de bases). La recombinación a través las primeras ocurre de manera más frecuente (Maréchal & Brisson, 2010), y crea un equilibrio dinámico donde diferentes versiones del genoma mitocondrial existen en mezclas equimolares con estructuras intercambiables entre sí (Dávila *et al.*, 2011). Los eventos de recombinación a través de secuencias cortas ocurren con menor frecuencia, pero ayudan a mantener versiones modificadas del genoma mitocondrial a bajas concentraciones (estas versiones modificadas reciben el nombre de sublimones; Woloszynska, 2009). Los eventos de recombinación asociados a secuencias cortas pueden generar inversiones en las secuencias intermedias, alterando el orden de los genes (Palmer

*et al.*, 2000; Kubo & Newton, 2008), así como duplicaciones o deleciones de segmentos específicos en la secuencia (Kubo & Newton, 2008; Sloan *et al.*, 2012). Estas secuencias tienen una mayor abundancia en el genoma mitocondrial. Por ejemplo, en la calabaza, (*Cucurbita pepo*), las secuencias repetidas más abundantes son aquellas que presentan una longitud de 19 a 40 nucleótidos (Alverson *et al.*, 2010).

## Los microsatélites

Existe un grupo particular de secuencias repetidas conformadas por unidades específicas de nucleótidos que se suceden en tándem (esto es, una después de la otra) un número determinado de veces. A estas unidades conformadoras del arreglo en tándem se les conoce como unidades de repetición, y a este tipo de arreglos se les conoce como Elementos Repetidos en tándem de Longitud Variable (VNTR por sus siglas en inglés). Estos elementos pueden dividirse en dos categorías en función del tamaño de la unidad de repetición: microsatélites, cuyo elemento de repetición va de 1 a 6 nucleótidos y minisatélites, cuyo elemento de repetición puede llegar a tener más de 100 nucleótidos (Buschiazzo & Gemmel, 2006). Ambos se distribuyen por todo el genoma, tanto en regiones codificantes como en regiones no codificantes, llegando algunos a tener implicaciones en desórdenes genéticos y en la regulación de la expresión génica (Buschiazzo y Gemmel, 2006). Con respecto a su historia evolutiva, los microsatélites están presentes tanto en angiospermas como en gimnospermas (Jaramillo-Correa *et al.*, 2013; Gou *et al.*, 2016), y generalmente los mismos loci están compartidos entre organismos filogenéticamente emparentados (Ishii *et al.*, 2006).

Con el advenimiento de la tecnología del DNA replicante, los análisis que utilizan este tipo de marcadores se han popularizado (Wang *et al.*, 2009). Éstos pueden ser amplificados a través del diseño de cebadores (*primers*, en inglés), que se sitúan en las regiones flanqueantes (Schötterer, 2000), y generalmente muestran altos niveles de variabilidad entre individuos, poblaciones o especies (Wang *et al.*, 2009). La gran abundancia y variabilidad de estos marcadores es responsable en parte del incremento del tamaño en los genomas, particularmente el de las mitocondrias vegetales (Lilli & Havey, 2001; Hetch *et al.*, 2011).

Los microsatélites pueden originarse a través de mutaciones puntuales en la secuencia (formación *de novo*) o ser adquiridos a través de elementos móviles (Schötterer, 2000, Buschiazzo & Gemmel, 2006). Una vez formados, estas secuencias pueden llevar a cabo eventos de deslizamiento de las hebras, lo que les permite ganar o perder unidades de repetición, expandiendo o contrayendo la longitud del arreglo a tasas elevadas (Schlötterer, 2000; Buschiazzo & Gemmel, 2006).

Los microsatélites no crecen de manera indefinida, generalmente alcanzan una longitud máxima a través de la introducción de sustituciones nucleotídicas dentro del arreglo de repeticiones (Buschiazzo & Gemmel, 2006). Estas interrupciones son críticas en el proceso evolutivo de los microsatélites. Generalmente se dan a través de sustituciones dentro del arreglo de unidades de repetición, o a través de inserciones y deleciones de corto tamaño (Buschiazzo & Gemmel, 2006). Al ocurrir estas interrupciones, disminuye la tasa de crecimiento por deslizamiento, lo que eventualmente genera la ruptura del patrón de repetición y deja a la secuencia con un número de unidades de repetición insuficientes para llevar a cabo el proceso de deslizamiento. El fragmento se asemeja entonces en estructura a aquellas de las cuales fueron formados los microsatélites originalmente. Es en este punto en el que se completa su “ciclo de vida” y nuevos microsatélites pueden ser originados (Buschiazzo y Gemmel, 2006).

Dado que una secuencia repetida puede formarse en casi cualquier parte del genoma, la posición de los microsatélites puede influenciar su proceso mutacional, contribuyendo a su variabilidad. Por ejemplo, si los microsatélites juegan papeles dentro de la dinámica genómica como reguladores de la expresión, transcripción o función de las proteínas, estos estarían eventualmente sometidos a presiones selectivas. Esto genera que algunos microsatélites que están situados en regiones conservadas presenten bajas tasas de mutación

puntual y sean retenidos después de la divergencia de los linajes (Buschiazzo & Gemmell, 2006).

Por ejemplo, en especies como el maíz se ha sugerido la presencia de microsátélites en relación con la disrupción de la maquinaria celular encargada de los procesos de respiración (Rajendrakumar *et al.*, 2006). Esta disrupción se ha asociado a la formación del fenotipo que da lugar a la esterilidad masculina en plantas (CMS por sus siglas en inglés, Rajendrakumar *et al.*, 2006). Estos no sólo exhiben mayores tasas de sustitución, sino que también mayores tasas de recombinación (Darrecq *et al.*, 2011), las cuales están íntimamente ligadas a la presencia de regiones repetitivas en el genoma.

Por otro lado, en regiones no codificantes del genoma mitocondrial vegetal los microsátélites suelen presentar una mayor variabilidad. En el caso particular de organismos del género *Triticum* y especies emparentadas filogenéticamente, el número de alelos de los microsátélites presentes en las regiones no codificantes es mayor comparado con el número de alelos en regiones codificantes del genoma mitocondrial vegetal (Ishii *et al.*, 2006). A su vez, estos microsátélites han sido retenidos entre clados específicos (Ishii *et al.*, 2006).

La retención de secuencias repetidas a través de largos periodos de tiempo no es exclusiva de las angiospermas. En gimnospermas, diferentes poblaciones de *Pinus banksiana* comparten un fragmento repetido en tándem con una unidad de repetición de 32 nucleótidos en el intrón 1 del gen *nad 7*. Esta región es altamente variable y muestra una elevada estructura geográfica entre poblaciones (Godbout *et al.*, 2005). Este minisatélite está además presente en otras cuatro especies: *P. contorta* var. *Latifolia*, *P. resinosa*, *P. sylvestris* y *P. strobus*.

En resumen, los microsátélites se encuentran asociados a distintos procesos modificadores de la arquitectura del mitogenoma, ya sea a través de eventos de inserción o

delección de nucleótidos (Sloan *et al.*, 2012) o por medio de eventos de recombinación poco frecuentes (Marechál & Brisson, 2010) y éstos pueden ser retenidos por largos intervalos de tiempo (Guo *et al.*, 2016) y compartidos entre linajes emparentados (Ishii *et al.*, 2006; Godbout *et al.*, 2005).

Todas estas características podrían generar alteraciones en las tasas de mutación en aquellas regiones circumvecinas de los lugares donde se encuentran. De hecho, algunos estudios indican que estas tasas tienden a aumentar en regiones que albergan gran cantidad de microsatélites; por ejemplo, se ha demostrado que en el genoma nuclear de los bovinos existe un aumento en las tasas de mutación en las primeras bases de las regiones flanqueantes de los microsatélites, y esta variación fue asociada a posibles eventos de reparación del DNA, así como de recombinación homóloga y replicación (Brodehe & Ellegreen, 1999). Por otro lado, en plantas se demostró que en el mitogenoma de las coníferas hay un aumento de las tasas de mutación en una región intrónica rica en microsatélites, en comparación con regiones codificantes del mismo genoma (Jaramillo-Correa *et al.*, 2013). Sin embargo no es claro si el efecto sobre las tasas de mutación de los microsatélites es homogéneo a lo largo de los genomas. Por ejemplo, no se sabe si esto ocurre de igual manea en intensidad en regiones intergénicas, intrones o exones de genes codificantes, o si el tamaño de las repeticiones tiene un impacto específico en la variación de las tasas en regiones cortas del genoma o si estos efectos se mantienen entre especies emparentadas o son especie-específicos. Dadas estas características, se vuelve de interés la forma y el modo en que los microsatélites han ido evolucionando, y el impacto que éstos han tenido en sus regiones flanqueantes dentro del genoma mitocondrial de las plantas. En esta tesis se busca evaluar el efecto de la presencia de microsatélites en los genomas



mitocondriales de las plantas, sobre las tasas de mutación de las regiones vecinas en un contexto filogenético, utilizando los pastos como modelo de estudio.

### **Modelo de estudio: las gramíneas.**

Las gramíneas -o pastos- son un grupo de monocotiledóneas muy diverso, caracterizado por su estructura floral en forma de espiga (Freeling, 2001). Agrupados en la familia Poaceae, éstos representan un grupo taxonómico que incluye más de 10,000 especies ampliamente distribuidas por todo el globo (Gaut, 2002; Eckardt, 2008; Hochbach *et al.*, 2015). Algunos organismos dentro del grupo de los pastos como el trigo (*Triticum* sp.), el maíz (*Zea* sp.), el centeno (*Secale* sp.), la cebada (*Hordeum* sp.) y la avena (*Avena* sp.) han sido domesticados para el consumo humano y constituyen más del 60% de la producción agrícola mundial total (Goff *et al.*, 2002, Eckardt, 2008; Hochbach *et al.*, 2015). Evolutivamente, los pastos han sido agrupados en dos clados principales, conocidos por sus acrónimos como clado BEP (compuesto por las subfamilias Bambusoideae, Ehrhartoideae y Pooideae) y el clado PACMAD (compuesto por las subfamilias Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae y Danthonioideae). Los grupos más cercanamente emparentados a los pastos son las subfamilias Anomochlooideae, Pharoideae y Puelioideae (Edwards *et al.*, 2011).

La familia de los pastos es fascinante debido a su grado de variación genética, y ecológica, así como por sus niveles de ploidía y números cromosómicos variables. Por ejemplo, *Triticum aestivum* tiene un genoma hexaploide de 17 gigabases, mientras que *Zea mays*, un genoma diploide de 2.5 gigabases (Eckardt, 2008). Se ha propuesto que los pastos han evolucionado a partir de un ancestro común con un número básico de cinco cromosomas nucleares, los cuales han llevado a cabo diversos eventos de fusión, duplicación y translocación para generar estados intermedios, permitiendo los diferentes niveles de ploidia en los cromosomas (Salse *et al.*, 2008). A pesar de presentar diferencias

en el número cromosómico, parece que éste no está relacionado con el orden y número de genes en el genoma (Freeling, 2001).

El genoma mitocondrial de los pastos comparte las características generales del de otros grupos de angiospermas, como una baja tasa de sustitución y un bajo nivel de sintenia (Hisano *et al.*, 2016), producto de constantes rearrreglos genómicos (Oldenburg *et al.*, 2013); sin embargo, bajo ciertas condiciones éste puede evolucionar de manera acelerada. Por ejemplo, bajo condiciones aloplásmicas. Esta condición se da cuando el citoplasma de una especie es reemplazado por la de otra especie, rompiendo así la interacción entre el núcleo y el citoplasma, los cuales han coevolucionado juntos, generando con ello nuevos fenotipos (Noyszewski *et al.*, 2014). La condición aloplásmica de ciertas variedades de trigo y especies emparentadas presentan una mayor aceleración en las tasas de recombinación, formando nuevos genomas mitocondriales con una expresión génica alterada (Noyszewski *et al.*, 2014). En el genoma mitocondrial de los pastos, como en otros mitogenomas vegetales también existen elementos repetidos asociados a eventos de recombinación (Levings & Brown, 1989; Oldenburg *et al.*, 2013; Hisano *et al.*, 2014). En particular, algunos microsatélites han sido retenidos a lo largo de la divergencia de los clados, como en el caso de la avena y especies emparentadas (Ishii *et al.*, 2006). En los mitogenomas de pastos, los microsatélites suelen estar confinados principalmente a regiones intergénicas (Rajendrakumar *et al.*, 2006; Alverson *et al.*, 2010), aunque pueden existir también en regiones intrónicas (Jaramillo-Correa *et al.*, 2013).

Dadas las características de los microsatélites descritas previamente, es de interés conocer si estas regiones se asocian a variaciones en las tasas de evolución. Es de esperarse que si los microsatélites están relacionados con la dinámica mitocondrial, entonces las tasas de mutación serán diferentes entre las regiones que contienen microsatélites y las que están

desprovistas de ellos. La importancia de las gramíneas ha permitido tener una buena representación de genomas mitocondriales secuenciados hasta la fecha en las bases de datos públicas (Eckardt, 2008), lo que permite una oportunidad única para medir la variación de estas tasas de evolución.

## Evaluación de la diversidad genética

Para evaluar el impacto en las tasas de mutación por la presencia de microsatélites, es conveniente hablar sobre el modo en que se estudia la diversidad genética. El estudio de las poblaciones naturales ha mostrado que la cantidad de variación a nivel molecular es sustancial, y una manera de conocer esta variación es analizando por separado segmentos codificantes, intrones, regiones flanqueantes o intergénicas (Nei, 1987). Cuando se dispone de un conjunto de secuencias de DNA de ciertos organismos de interés es normal que éstas sean diferentes unas de otras, por lo que existen varias maneras de medir la cantidad de polimorfismos presentes en el DNA. Una medida es determinando el número de diferencias que presentan entre secuencias, o lo que es lo mismo, el número de sitios segregantes por secuencia (Nei, 1987; Yang 2006). Sin embargo, esta medida depende del tamaño de muestreo, y las especies de donde provengan las secuencias tienen que estar cercanamente emparentadas para representar una medida no sesgada de la variación del DNA (Nei, 1987; Yang, 2006). Otra forma de estimar variación es calculando el número promedio de diferencias nucleotídicas (por ejemplo, sustituciones) entre pares de secuencias. A esta medida se le conoce como diversidad nucleotídica, y es definida por la siguiente ecuación:

$$\pi = \sum_{ij} x_i x_j \pi_{ij}$$

(Nei, 1987, Tajima 1998)

Donde  $X_i$  y  $X_j$  son respectivamente las frecuencias de las secuencias  $i$  y  $j$  en una muestra de secuencias, y  $\pi_{ij}$  es el número de diferencias nucleotídicas por sitio entre dichas secuencias.

En una población panmítica,  $\pi$  es simplemente la heterocigosidad en las secuencias nucleotídicas y puede ser estimada de la siguiente manera:

$$\pi = \sum_{i < j} \frac{\pi_{ij}}{n_c}$$

(Nei, 1987. Sin embargo, se recomienda revisar Tajima, 1998).

Donde  $n_c$  es el número total de comparaciones entre las secuencias.

También es posible conocer la variación dentro de una población, obteniendo la diversidad haplotípica a partir de una colección de secuencias de DNA. El número de haplotipos en un conjunto de secuencias puede ser un indicador de la variación, y se espera que éste sea grande cuando la cantidad de polimorfismos presentes en una población natural es alta. Al igual que con el número de sitios segregantes, este valor es dependiente el tamaño de muestreo, y la comparación entre diferentes poblaciones podría tener sesgos si las muestras no tienen el mismo tamaño (Nei, 1987). Una vez que los diferentes haplotipos han sido identificados, puede calcularse la diversidad haplotípica, que también es una medida equivalente a la heterocigosidad, y señala la diversidad que hay entre y dentro de las poblaciones. La diversidad haplotípica puede calcularse de la siguiente manera:

$$H = \frac{2n(1 - \sum x_i^2)}{2n - 1}$$

(Nei, 1987).

Donde  $n$  es el número de individuos muestreados, y  $X$  es la frecuencia de los distintos haplotipos( $i$ ).

Por otro lado, para estimar la divergencia en un conjunto de secuencias de DNA es necesario tomar en cuenta el número de polimorfismos presentes en dicha colección y calcular la distancia media de todas las secuencias a partir de una secuencia (o grupo de secuencias) que sirva como referencia (Shankarappa *et al.*, 1999; Deng *et al.*, 2010). Suponiendo que existe un número determinado de haplotipos en una región particular del genoma entre el conjunto  $X$  y el conjunto  $Y$  de secuencias de DNA, es posible conocer la divergencia promedio entre los haplotipos de  $X$  y  $Y$  de la siguiente manera:

$$D_{xy} = \sum_{ij} x'_i y'_j d_{ij}$$

(Nei, 1987)

Donde  $x'$  y  $y'$  son las frecuencias de los haplotipos entre las poblaciones y  $d_{ij}$  el número de sustituciones nucleotídicas entre el haplotipo  $i$  de  $X$  y el haplotipo  $j$  de  $Y$ . El significado de esta ecuación es que, en presencia de secuencias polimórficas, la divergencia promedio entre los alelos de las poblaciones está en relación con el tiempo de divergencia entre ellas (Nei, 1987).

El hecho de que las secuencias diverjan unas de otras se debe principalmente a la tasa de mutación a la que ocurren dichos cambios a través del tiempo. Esto hace posible examinar la variación entre muestras de DNA y compararlas con la distribución esperada bajo la teoría neutral (Tajima, 1998). Bajo esta teoría, es posible estimar la tasa de mutación poblacional a partir de la diversidad nucleotídica observada de una población (Watterson, 1975).

Cuando una población se encuentra en equilibrio y las mutaciones son selectivamente neutrales, la proporción de sitios polimórficos está dada por:

$$\theta = 4Ne\mu$$

(Watterson *et al.*, 1975)

Donde  $Ne$  es el tamaño efectivo poblacional y  $\mu$  es la tasa de mutación de la población por generación (Watterson, 1975). Cuando se presume que no hay recombinación es posible estimar esta tasa de cambio utilizando la siguiente fórmula:

$$\theta = Pn/A1$$

(Nei, 1987; Tajima, 1998)

Donde  $Pn$  es el número de sitios polimórficos por nucleótido y  $A = 1 + 1/2 + 1/3 + \dots + (n-1)^{-1}$ , lo que corresponde al número armónico del número total de las secuencias menos uno (Nei, 1987; Tajima, 1998).

El estudio de la diversidad a nivel molecular permite conocer el modo en el que los organismos han divergido. Para ello, es posible emplear herramientas para inferir su historia evolutiva. La reconstrucción filogenética es la herramienta que permite la representación de las relaciones genealógicas entre las especies, genes, poblaciones o individuos a través de la representación gráfica de un conjunto de vértices (nodos) y aristas (ramas) interconectados, (Yang *et al.*, 2006). La inferencia de las relaciones puede llevarse a cabo a través de diferentes métodos como la reconstrucción por parsimonia, máxima verosimilitud o probabilidad bayesiana (Yang & Rannala, 2012). Utilizando estas herramientas para conocer el modo y la forma en la que los organismos han divergido y/o



están relacionados unos con otros, es posible inferir el momento en el que ocurrieron los cambios en la secuencia de DNA y el impacto de dichos cambios.

## **Justificación**

Esta tesis busca ampliar el conocimiento actual sobre el modo de evolución del genoma mitocondrial de las plantas a través de la evaluación, bajo un contexto filogenético, de la posible relación entre microsatélites sobre la variación de las tasas de mutación en sus regiones vecinas. La presencia de dicha variación asociada a los microsatélites dentro del genoma mitocondrial podría sugerir una revisión a la hipótesis de la carga mutacional, la cual esperar que exista una correlación negativa entre el tamaño del genoma mitocondrial y la variación nucleotídica. (Lynch *et al.*, 2006; Skippington *et al.*, 2015).

## **Hipótesis**

Ho. Si las secuencias de microsatélites están relacionadas con la dinámica evolutiva del genoma mitocondrial vegetal de los pastos en sus regiones adyacentes, a través de un aumento en la variación nucleotídica, entonces la diversidad entre regiones con microsatélites serán diferentes de las secuencias desprovistas de microsatélites.

Ha. Si las secuencias de microsatélites no mantienen ningún tipo de relación con la dinámica evolutiva del genoma mitocondrial, entonces las tasas de diversidad en regiones vecinas a microsatélites no presentarán diferencias significativas, al ser comparadas con aquellas regiones desprovistas de ellos.

## **Objetivos**

### **Objetivo general.**

Comparar la diversidad nucleotídica de secuencias portadoras de microsatélites con aquellas secuencias desprovistas de ellos, y reconstruir su historia evolutiva en una filogenia de la familia Poaceae.

### **Objetivos específicos:**

- 1) Hacer una búsqueda de los microsatélites presentes en los genomas mitocondriales de pastos disponibles.
- 2) Comparar la diversidad entre regiones que contengan microsatélites con respecto a secuencias desprovistas de éstos, utilizando diversos índices de diversidad nucleotídica.
- 3) Evaluar el modo en que han evolucionado el número de repeticiones de los microsatélites, a través de una reconstrucción de estados ancestrales en la historia evolutiva del grupo de los pastos.

## **Métodos.**

### **Búsqueda de los microsatélites.**

Se llevó a cabo una búsqueda exhaustiva de genomas mitocondriales completos disponibles para gramíneas en la base de datos del centro nacional de información de biotecnología (NCBI) hasta marzo de 2017 a través del buscador de genomas completos del NCBI. Se descargaron todas las secuencias disponibles y se sometieron al algoritmo Repeat Finder (Okonechnikov *et al.*, 2012) para identificar el número, posición y secuencia de todos los microsatélites presentes en estos genomas. Como criterio de discriminación, no se consideraron como microsatélites aquellas que tuvieran menos de cuatro unidades de repetición (para microsatélites mononucleótidos o dinucleótidos) o cuyas repeticiones tuvieran una longitud total menor a 12 pb (para trinucleótidos, tetranucleótidos, pentanucleótidos y hexanucleótidos). Este criterio es similar al utilizado en otros estudios (Subramanian *et al.* 2002), y busca evitar la elección de elementos formados al azar. Con los genomas encontrados (mostrados en la tabla M.1) se cubrieron las subfamilias Pooideae, Bambusoideae y Ehrhartoideae del clado BEP, y la subfamilia Panicoideae del clado PACMAD.

Se seleccionaron en total 20 regiones con microsatélites que cumplieron con el criterio de discriminación y que estuvieron conservadas entre las especies. Estas secuencias se sometieron al algoritmo BLAST (Altschul *et al.*, 1990) para encontrar similitudes con regiones homólogas en otras especies de plantas en la base de datos del NCBI. Se buscó que las secuencias presentaran una cobertura del 100%, con respecto a las especies de pastos depositadas, y que no mostraran coincidencias con regiones del cloroplasto o del núcleo que sugirieran algún tipo de transferencia horizontal (desde o hacia la mitocondria).

Finalmente se escogieron tres regiones candidatas (dos regiones intergénicas y un intrón), que fueron aquellas con el mayor número de coincidencias en las búsquedas realizadas.

Mitogenomas descargados de la página del centro nacional de información de biotecnología (NCBI).		
Genoma mitocondrial	Subfamilia	Código de acceso en GeneBank
<i>Aegilops speltoides</i>	Pooideae	AP013107.1
<i>Bambusa oldhamii</i>	Bambusoideae	EU365401.1
<i>Ferocalamus rimosivaginus</i>	Bambusoideae	JQ235167.1
<i>Hordeum vulgare</i>	Pooideae	AP017301.1
<i>Lolium perenne</i>	Pooideae	JX999996.1
<i>Oryza sativa (indica)</i>	Ehrhartoideae	JN861112.1
<i>Oryza sativa (japónica)</i>	Ehrhartoideae	BA000029.3
<i>Oryza rufipogon</i>	Ehrhartoideae	AP012528.1
<i>Sorghum bicolor</i>	Panicoideae	DQ984518.1
<i>Tripsacum dactyloides</i>	Panicoideae	DQ984517.1
<i>Triticum aestivum</i>	Pooideae	AP008982.1
<i>Triticum timopheevii</i>	Pooideae	AP013106.1
<i>Zea mays (CMS-C)</i>	Panicoideae	DQ645536.1
<i>Zea mays (CMS-S)</i>	Panicoideae	DQ490951.2
<i>Zea mays (CMS-T)</i>	Panicoideae	DQ490953.1
<i>Zea luxurians (Teocinte anual)</i>	Panicoideae	DQ645537.1
<i>Zea mays (NA)</i>	Panicoideae	DQ490952.1
<i>Zea mays (NB)</i>	Panicoideae	AY506529.1
<i>Zea mays (parviglumis)</i>	Panicoideae	DQ645539.1
<i>Zea perennis (Teocintle perenne)</i>	Panicoideae	DQ645538.1

Tabla M.1 genomas mitocondriales completos disponibles para gramíneas en la base de datos del centro nacional de información de biotecnología (NCBI) hasta marzo de 2017.

Estas fueron: parte del intrón 3 del gen *nadh2* y las regiones entre los genes *trnY* y *rps4*, y *nadh4* y *atp4*. Las dos primeras presentaron un microsatélite dinucleótido, compuesto por las repeticiones (AG)<sub>n</sub> y (CT)<sub>n</sub> respectivamente. La tercera mostró un microsatélite pentanucleótido, con unidades de repeticion (ATAGT)<sub>n</sub>. Estas regiones de interés fueron caracterizadas a partir de los genomas mitocondriales de *Tripsacum dactyloides*, *Lolium perenne* y *Zea luxurians* respectivamente. Para cumplir con los fines comparativos se escogieron otras 3 regiones altamente conservadas dentro del genoma mitocondrial de las gramíneas y que estuvieran desprovistas de microsatélites en todas las secuencias analizadas (dos regiones intergénicas entre los genes *trnN-Cp* - *trnD* y *trnS-a* - *L-b*, así como un segmento del intron 2 del gen *nadh1*) y en otras especies de plantas. Esto último se determinó nuevamente a través de un BLAST utilizando los criterios descritos anteriormente.

## Reconstrucción filogenética

Las relaciones filogenéticas fueron establecidas a partir de una matriz concatenada de tres marcadores cloroplásticos: los genes *rbcL* (región larga de la ribulosa 1,5-bifosfato carboxilasa/oxidasa), *nadhF* (NADH deshidrogenasa subunidad F) y la región comprendida entre los genes *matK* (maturasa-k) con la región intrónica del gen *trnK* (RNA de transferencia para lisina). Estos tres marcadores han sido utilizados ampliamente para realizar las reconstrucciones filogenéticas en plantas, incluyendo a los pastos (Hilu *et al.*, 1999; Guissani *et al.*, 2001; Edwards *et al.*, 2011). El alineamiento de las secuencias fue obtenido y modificado a partir de Edwards *et al.* (2011). La reconstrucción filogenética fue llevada a cabo a través de inferencia bayesiana en el programa MrBayes versión 3.1.2 (Huelsenbeck & Ronquist, 2001). Primero se utilizó el programa JModeltest (Darriba *et al.*, 2012) para encontrar el modelo de sustitución nucleotídica que mejor se ajustaba a los datos de cada gen. Este se escogió a través del criterio de información de Akaike (Posada & Buckley, 2004). Para esto, la matriz fue particionada en los tres genes, para los cuales se encontraron los siguientes modelos: TVM+G, para el gen *matK-trnK*, el modelo TPM3+G para el gen *rbcL* y el modelo TrN+G para el gen *nadhF*. Para tener el mayor grado de aproximación al modelo durante la reconstrucción filogenética en MrBayes, se escogió un modelo GTR+G, con los valores de los parámetros para cada modelo calculado con JModeltest. La reconstrucción filogenética fue llevada a cabo utilizando dos corridas independientes de 4 cadenas de Markov-Monte Carlo con acoplamiento de Metrópolis de 10 millones de generaciones cada una, muestreando cada 10,000 árboles, luego de descartar los primeros 10,000 como "burn in". Los árboles muestreados se utilizaron para construir la probabilidad posterior y determinar el soporte de cada rama.



### **Colecta, extracción y amplificación.**

Con el fin de evaluar la presencia de microsatélites y su impacto en regiones vecinas en un contexto filogenético más amplio, se colectaron 20 especies de pastos en la Ciudad de México (Tabla M.2) para la amplificación y el análisis de las regiones de interés.

Se llevó a cabo la extracción de DNA utilizando el protocolo de Vázquez-Lobo (1996), a partir de tejido foliar de los especímenes colectados. Para la amplificación de las regiones de interés se desarrollaron cebadores específicos utilizando la función Prime3 del programa Ugene (Okonechnikov *et al.*, 2012). Para las regiones con microsatélites, se especificó que los cebadores tuvieran una astringencia perfecta a 200 pares de bases río arriba y río abajo del arreglo de repeticiones. Para las regiones sin microsatélite se diseñaron los cebadores a los extremos de las regiones de interés. En ambos casos se buscó que los cebadores no formaran dímeros entre sí. Los primers y la temperatura de alineamiento se muestra en la tabla M.3. La reacción en cadena de la polimerasa se realizó en un termociclador ABI (Applied Biosystems) en un volumen final de 12  $\mu\text{L}$  que contenía 1.25  $\mu\text{L}$  de buffer de reacción 5Xgreen GoTaq a pH 8.5 (Promega corporation©), 0.375  $\mu\text{l}$  de  $\text{MgCl}_2$  con una concentración de 25mM, 0.5  $\mu\text{l}$  de DNA total de las muestras colectadas, con un rango de concentración de 0.5 a 20 ng/mL; 0.5  $\mu\text{L}$  de cada primer, a una concentración de 10pM y 0.25  $\mu\text{l}$  de desoxydinucleótidos a 10 mM cada uno. Las condiciones de la reacción se muestran en la Tabla M.4. Los productos de la amplificación fueron corroborados con una electroforesis en gel de agarosa con una concentración al 1.5 % g/mL. En ningún caso se observaron productos inespecíficos.

Listado de especies colectadas		
Especie	Subfamilia	Zona de muestreo
<i>Lolium perenne</i>	Pooideae	19° 19' 31''N 99°10'54'' W
<i>Bromus Carinatus</i>	Pooideae	19°19'32''N 99°10'53''W
<i>Pennisetum villosum</i>	Panicoideae	19°19'32''N 99°10'46''W
<i>Eragrostis mexicana</i>	Chloridoideae	19° 19' 30''N 99°11'16'' W
<i>Melinis repens</i>	Panicoideae	19° 19' 32''N 99°10'54'' W
<i>Eleusine indica</i>	Chloridoideae	19° 19' 26''N 99°10'53'' W
<i>Chloris virgata</i>	Chloridoideae	19° 19' 38''N 99°11'01'' W
<i>Sporobolus sp</i>	Chloridoideae	19° 19' 38''N 99°11'02'' W
<i>Paspalum pubiflorum</i>	Panicoideae	19° 19' 31''N 99°10'17'' W
<i>Setaria parviflora</i>	Panicoideae	19° 19' 25''N 99°11'31'' W
<i>Digitaria velutina</i>	Panicoideae	19° 28' 43''N 99°05'34'' W
<i>Poa annua</i>	Pooideae	19° 28' 43''N 99°05'34'' W
<i>Cynodon dactylon</i>	Chloridoideae	19° 28' 43''N 99°05'34'' W
<i>Eleusine multiflora</i>	Chloridoideae	19° 28' 43''N 99°05'34'' W
<i>Cortaderia selloana</i>	Danthonioideae	19° 28' 43''N 99°05'34'' W
<i>Bambusa sp.</i>	Bambusoideae	19°19'18''N 99°10'34''W
<i>Zea mays</i>	Panicoideae	19°19'32''N 99°10'58''W
<i>Arundo donax</i>	Arundinoideae	19°19'32''N 99°10'58''W
<i>Pennisetum clandestinum</i>	Panicoideae	19°19'32''N 99°10'58''W
<i>Digitaria setigera</i>	Panicoideae	19° 28' 41''N 99°05'36'' W

Tabla M.2 Especies colectadas de pastos en varios puntos de la Ciudad de México. Se trató de cubrir la presencia de otras familias que no estuvieran representadas en las bases de datos de mitogenomas completos.

Cebadores utilizados para amplificar las regiones de interés.				
Nombre del cebador	Cebador "Foreward"	Cebador "Reverse"	Temperatura de alineamiento (°C)	Región que amplifica.
TrnN-Cp - TrnD	F: 5'- TTCGAATTGAT AGTTATAGCTT TTCTG-3	R: 5'- GCACGATCTTC CTCTTCCTG-3'	60	región intergénica entre los genes TrnN-Cp y TrnD
intrón 2 del gen <i>nad1</i>	F 5'- GAACCACACG TGCAAGTTTCC -3	R: 5'- CGCCTTGTGAG CACGTTT-3'	60	Region del intrón 2 del gen <i>nad1</i>
<i>TrnS-a - TrnL-b.</i>	F: 5'- GGGCGAAATT ACGTGAGAG-3'	R: 5'- CACTTGCCTAA AGTCTCTTTTC TTC-3'	60	región intergénica entre los genes <i>TrnS-a</i> y <i>TrnL-b.</i>
Intrón 3 del gen <i>nad2</i>	F: 5'- GAGTGTTTCGTT ATAGGGGATA GGA-3'	R: 5'- TCGCACTGCTC TTTCTCTCA-3'	60	región intrónica número 3 del gen <i>nad2</i>
<i>nad4 - atp4</i>	F: 5'- ATGCTTCCTCA GTCAATCAGC A-3	R: 5'- CTTTGTTACCA GAGCTGGATC G-3'	60	Región intergénica entre los genes <i>nad4</i> y <i>atp4</i>

Tabla M.3 Cebadores utilizados para la amplificación de las regiones de interés.

<b>Condiciones de amplificación de las regiones de interés.</b>		
<b>Etapas del ciclo de PCR</b>	<b>Temperatura</b>	<b>Tiempo (min: seg)</b>
<b>Desnaturalización</b>	95°C	5:00
<b>Desnaturalización*</b>	95°C	0:45
<b>Alineamiento*</b>	60°C	0:45
<b>Extensión*</b>	72°C	1:00
<b>Extensión final</b>	72°C	10:00
<b>Conservación</b>	04°C	∞

Tabla M.4 condiciones de la reacción de la polimerasa en cadena para las regiones de interés.

Los productos de las reacciones de PCR fueron secuenciados con el método Sanger en el servicio de Macrogen, Corea (10F, 254 Beotkkot-ro Geumcheon-gu, Seoul, 08511, Rep. of Korea). Las secuencias fueron alineadas utilizando el programa Mesquite (Maddisson & Maddisson, 2011) utilizando el algoritmo MAFFT (Kato & Standley, 2013). Estas fueron posteriormente revisadas de manera visual y corregidas cuando era necesario. Con las secuencias alineadas se construyeron un total de 6 matrices de datos correspondientes a cada una a las regiones de interés. Las matrices fueron completadas a través de la adición de secuencias de pastos depositadas en las bases de datos públicas, para las regiones de interés, lo que nos permitió tener la mayor cobertura taxonómica posible.

### **Análisis de la diversidad nucleotídica.**

Para cada una de las regiones de interés se estimaron el número total de mutaciones, el número de haplotipos y el número promedio de diferencias nucleotídicas entre pares de secuencias ( $\pi$ ) (Nei, 1987). Con este último estimador se determinó la tasa de mutación a partir del número de sitios segregantes ( $\theta$ ) (Watterson *et al.*, 1974). Por otro lado, también se calculó el índice de divergencia entre las secuencias ( $D_{xy}$ ) (Nei, 1987), así como el índice de inserción y deleción ( $I$ ) utilizando como outgroup dos organismos de la familia Arecaeae (*Phoenix dactylifera* y *Cocus nucifera*), grupo hermano del clado de las poales. Todos los cálculos se llevaron a cabo usando el programa DnaSP V.5 (Librado & Rozas, 2009). Utilizando este mismo programa se realizaron simulaciones coalescentes, basadas en un modelo de sitios infinitos y un tamaño poblacional constante (Hudson, 1990), con el fin de evaluar su distribución estadística dentro de un límite de confianza del 95%. En el caso particular de las regiones con microsatélites, la estimación de los índices de diversidad se llevó a cabo de forma independiente en cada región flanqueante (río arriba y río abajo), de manera similar al de los microsatélites encontrados en el cuarto intrón del gen *nadh5* para diversas especies del género *Abies* (Jaramillo-Correa *et al.*, 2013).

### **Reconstrucción de caracteres ancestrales**

La reconstrucción de los estados ancestrales se llevó a cabo utilizando el programa Mesquite mediante el método de parsimonia, el cual determina los estados de carácter, minimizando el número de cambios entre un estado de carácter y otro, dada una topología (Maddison y Maddison, 2000). En este caso, los estados de carácter correspondieron a cada una de las unidades de repetición de los microsatélites. Los estados fueron mapeados sobre la filogenia construida previamente, utilizando el programa Mesquite (Maddison & Maddison, 2011), calculando los índices de retención y consistencia (RI y CI respectivamente), para conocer el grado de homoplasia entre los diferentes estados de cada carácter (Farris, 1989; Yang, 2006).

## Resultados.

### Búsqueda de microsatélites en los genomas mitocondriales de pastos depositados en el NCBI.

Se encontraron 20 mitogenomas para pastos (Tabla R.1), 19 de los cuales se encontraban completos y anotados. El genoma incompleto fue el de *Ferrocalamus rimosivaginus*, y este no se tomó en cuenta para calcular el valor promedio de número de microsatélites por genoma, el valor promedio de la longitud de los mitogenomas, ni la prueba de correlación entre la longitud del genoma y el número de microsatélites. Los genomas mitocondriales descargados tuvieron una longitud promedio de 557,908 pares de bases, lo cual es bastante en comparación al genoma mitocondrial humano (16,559 pares de bases, Yang *et al.*, 2009) El genoma de mayor tamaño fue el del citotipo CMS-C de *Zea mays*, con 739,719 pares de bases (pb), mientras que el más pequeño fue el de *Triticum timopheevii* con 443,419 pb (Tabla R.1).

Se encontraron un total de 3,056 arreglos de repetición en todos estos genomas, con un número promedio de microsatélites por genoma de 157, siendo 280 el mayor número de microsatélites encontrados (citotipo CMS-C de *Zea mays*) y 70 el menor número (*Sorghum bicolor*). En la figura Fig. R.1 se observa la frecuencia del tipo de arreglo de microsatélites en cada uno de los genomas analizados. En promedio, los microsatélites con unidades de repetición tetranucleótidas son los más abundantes, seguidos por los pentanucleótidos; mientras que los mononucleótidos son los menos abundantes en los genomas mitocondriales muestreados (Figura R.2). Se encontró una correlación positiva ( $r=0.762$ , T de Student  $P<0.001$ ) entre el número de microsatélites y el tamaño del genoma mitocondrial para las especies analizadas (figura R.3).



Abundancia de los microsatélites en los distintos genomas mitocondriales.									
Especie	Acceso	Tipo de unidad de repetición						Tamaño del genoma (pb)	Microsatélites totales.
		mononucleótido	dinucleótido	trinucleótido	tetranucleótido	pentanucleótido	hexanucleótido		
<i>Aegilops speltoides</i>	AP013107.1	3	15	15	40	9	6	476,091	88
<i>Bambusa oldhamii</i>	EU365401.1	9	9	17	45	3	1	509,941	84
<i>Ferocalamus rimosivaginus</i>	JQ235167.1	3	8	15	37	5	0	78,941	68
<i>Hordeum vulgare</i>	AP017301.1	5	12	22	50	10	4	525,599	103
<i>Lolium perenne</i>	JX999996.1	4	9	16	55	11	5	678,580	100
<i>Oryza sativa (indica)</i>	JN861112.1	6	11	26	53	14	4	454,894	114
<i>Oryza sativa (japónica)</i>	BA000029.3	3	9	20	42	10	1	490,520	85
<i>Oryza rufipogon</i>	BA000029.3	1	6	25	49	12	2	502,250	95
<i>Sorghum bicolor</i>	DQ984518.1	5	5	16	39	4	1	468,628	70
<i>Tripsacum dactyloides</i>	DQ984517.1	5	4	26	60	67	49	704,100	211
<i>Triticum aestivum</i>	AP008982.1	7	17	29	57	13	5	452,528	128

<i>Triticum timopheevii</i>	AP013106.1	5	9	20	39	8	4	443,419	85
<i>Zea mays</i> (CMS-C)	DQ645536.1	2	10	41	81	96	50	739,719	280
<i>Zea mays</i> (CMS-S)	DQ490951.2	3	6	26	58	74	34	557,162	201
<i>Zea mays</i> (CMS-T)	DQ490953.1	3	6	22	54	76	31	535,825	192
<i>Zea luxurians</i>	DQ645537.1	5	5	24	54	80	31	539,368	199
<i>Zea mays</i> (NA)	DQ490952.1	3	9	35	73	105	43	701,046	268
<i>Zea mays</i> (NB)	AY506529.1	2	6	27	62	81	32	569,630	210
<i>Zea mays parviglumis</i>	DQ645539.1	3	9	37	70	105	40	680,603	264
<i>Zea perennis</i>	DQ645538.1	4	5	29	66	78	29	570,354	211
<b>Promedio de microsátélites por unidad de repetición</b>		4.11	8.53	24.89	51.11	45.05	19.58	557,908	157.26

Tabla R.1 Descripción de los genomas mitocondriales disponibles para especies de la familia Poaceae en la base de datos de Genbank.

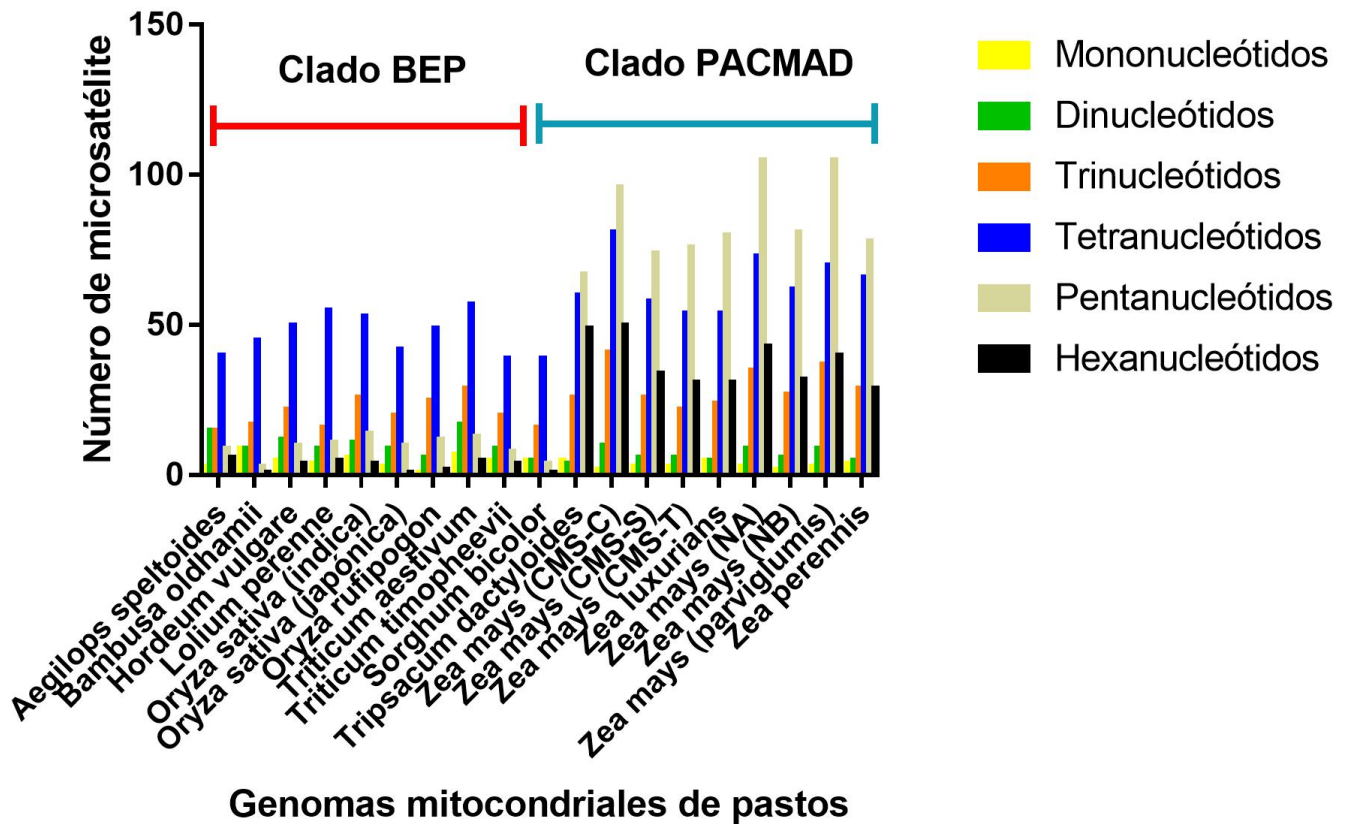


Figura R.1 Frecuencia del tipo de arreglo de microsatélite en los distintos genomas mitocondriales encontrados en Genbank. La frecuencia de los tipos de arreglo de los microsatélites se presenta en el gráfico entre los clados principales de la familia de las gramíneas: el clado BEP (representado por las subfamilias Pooideae, Bambusoideae y Ehrhartoideae), y el clado PACMAD (representado en el gráfico solo por organismos de la subfamilia Panicoideae)

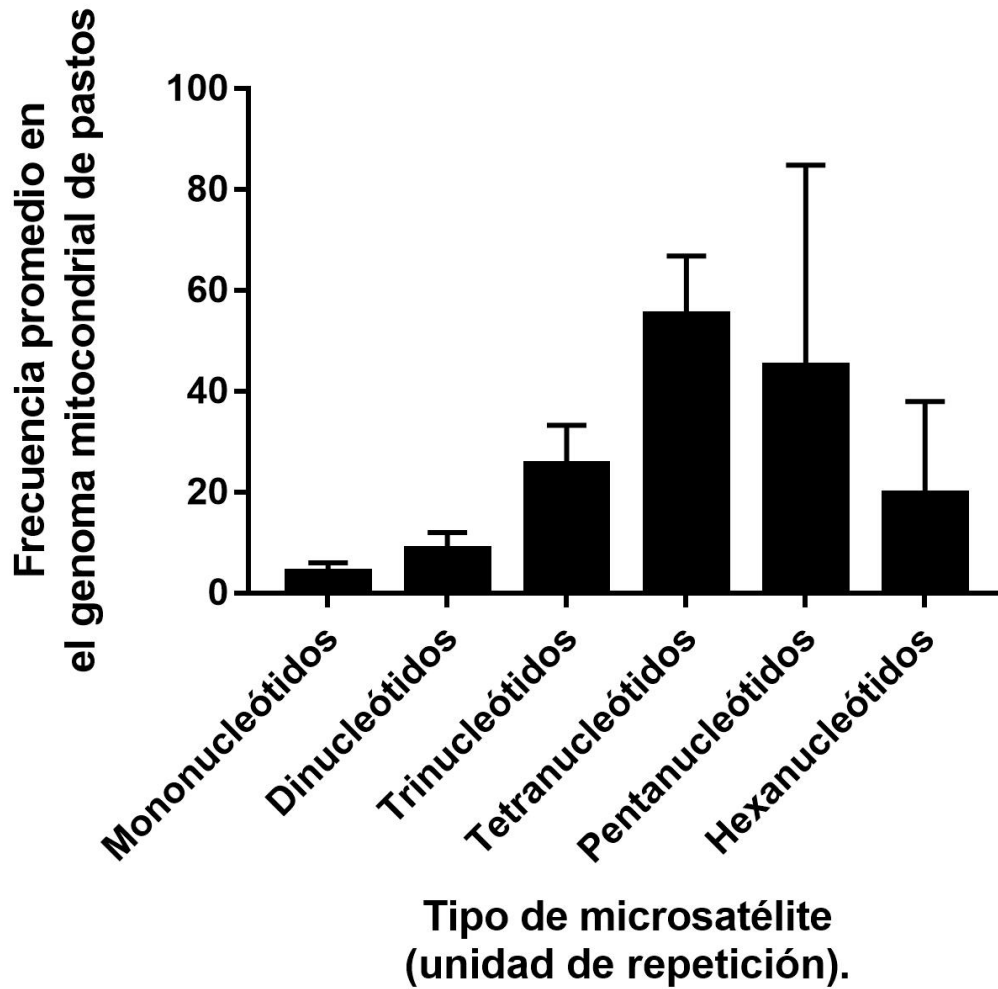


Figura R.2 Frecuencia del tipo de microsatélites en 19 genomas mitocondriales disponibles para la familia Poaceae en Genbank. Las barras en la gráfica representan la frecuencia promedio de cada tipo de microsatélite en el genoma mitocondrial vegetal.

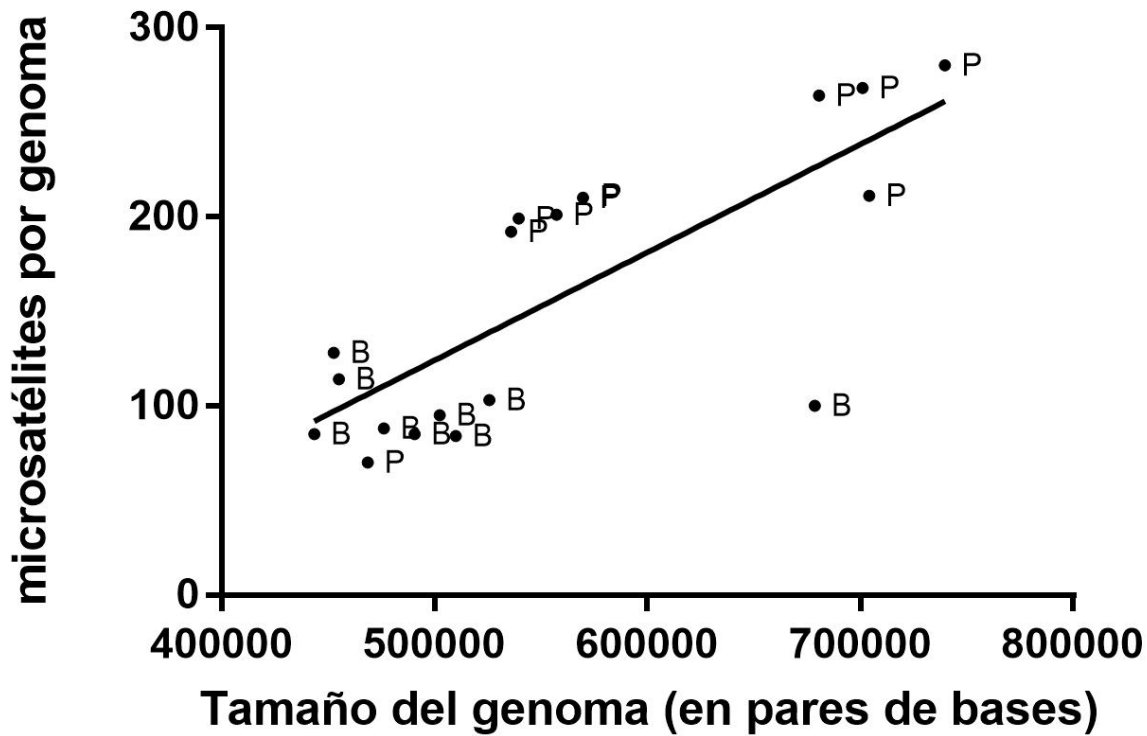


Figura R.3 Diagrama de dispersión entre el tamaño del genoma mitocondrial de los pastos, y el número total de microsatélites por genoma. Se realizó un análisis de correlación entre el tamaño del genoma mitocondrial de 19 especies de pasto (P = organismos del clado PACMAD; B = organismos del clado BEP) y el número total de microsatélites por genoma. La gráfica muestra una correlación positiva ( $r=0.762$ , t de Student  $P<0.001$ ) entre el tamaño del genoma mitocondrial y el número de microsatélites que presenta.

### **Descripción de las regiones de interés.**

En general, las regiones portadoras de microsatélites fueron más difíciles de amplificar que las regiones desprovistas de microsatélites, lo que sugiere que, entre otras cosas, están menos conservadas entre especies, por lo que las matrices correspondientes presentaron un número menor de secuencias (tabla R.3). Estas matrices se elaboraron no solo a través de la amplificación y secuenciación de los fragmentos de interés, sino de la descarga de secuencias depositadas en Genbank. En la tabla R.2 se muestran las especies utilizadas para la construcción de las diferentes matrices.

Dependiendo de la especie, el número de repeticiones por microsatélite varió en cada una de las regiones analizadas. En algunos casos, ciertas unidades de repetición presentaron algún tipo de interrupción por la inserción de un nucleótido. Además, salvo algunas excepciones, las regiones vecinas a los microsatélites generalmente estuvieron conservadas entre los taxa analizados.

El microsatélite encontrado en el intrón 3 del gen *nadh2* está compuesto por la unidad de repetición (AG)<sub>n</sub> y presentó un rango de 4 a 8 unidades de repetición, con *Aegilops speltoides*, *Bromus carinatus* y *Hordeum vulgare* mostrando el mayor número de unidades y *Rhynchelytrum repens* el menor. Las regiones vecinas presentaron pequeñas deleciones/inserciones de 5 a 7 pares de bases, compartidas entre los taxones en distintas regiones homólogas, tanto río arriba como río abajo del microsatélite (Figura R.4).

Especies utilizadas en las diferentes matrices construidas.						
Nombre de las especies	TrnN-Cp - TrnD	Intron 2 del gen nad1	TrnS-a - trnL-b	Intron 3 del gen nad2	nad4 - atp4	trnYrps4
<i>Aegilops speltoides</i>	1	1	1	1	1	1
<i>Arundo donax</i>	1	1	1	0	1	0
<i>Bambusa Oldhamii</i>	1	1	1	1	1	1
<i>Bambusa sp.</i>	1	1	1	0	1	0
<i>Bromus carinatus</i>	1	1	1	1	0	0
<i>Chloris virgata</i>	1	1	1	1	1	0
<i>Cortaderia sp</i>	1	1	1	0	1	0
<i>Cyodon dactylon</i>	1	1	1	1	1	0
<i>Digitaria setigera</i>	1	1	1	1	1	0
<i>Digitaria velutina</i>	1	1	1	1	0	0
<i>Eleusine indica</i>	1	0	1	0	1	0
<i>Eleusine multiflora</i>	0	1	1	1	1	0
<i>Eragrostis mexicana</i>	1	0	1	1	1	0
<i>Ferocalamus rimosivaginus</i>	1	1	1	1	1	1
<i>Hordeum vulgare</i>	1	1	1	1	1	1
<i>Lolium perenne</i>	1	1	1	1	1	1
<i>Oryza rufipogon</i>	1	1	1	1	0	1
<i>Oryza sativa indica</i>	1	1	1	1	0	1
<i>Oryza sativa japonica</i>	0	1	0	0	0	0
<i>Paspalum pubiflorum</i>	1	1	1	1	0	0
<i>Pennisetum clandestinum</i>	0	1	1	0	1	0
<i>Pennisetum villosum</i>	1	0	1	0	1	0
<i>Poa annua</i>	1	1	0	0	0	0
<i>Melinis repens</i>	1	0	1	1	1	0
<i>Sorghum bicolor</i>	1	1	1	1	1	1
<i>Sporobolus sp</i>	1	1	1	1	1	0
<i>Tripsacum dactyloides</i>	1	1	1	1	1	0
<i>Triticum aestivum</i>	1	1	1	0	1	1
<i>Triticum timopheevii</i>	1	1	1	0	1	1
<i>Zea diploperenis</i>	0	1	0	0	0	0
<i>Zea luxurians</i>	1	1	1	1	1	1
<i>Zea mays mays</i>	1	1	1	0	0	0
<i>Zea mays CMS C</i>	1	1	1	1	1	1
<i>Zea mays CMS S</i>	1	1	1	1	1	1
<i>Zea mays CMS T</i>	1	1	1	1	1	1
<i>Zea mays NA</i>	1	1	1	1	1	1
<i>Zea mays NB</i>	1	1	1	1	1	1
<i>Zea mays parviglumis</i>	1	1	1	1	1	1
<i>Zea perennis</i>	1	1	1	1	1	1
Total	35	34	35	26	29	17

Tabla R.2 Especies utilizadas para la construcción de las matrices de las regiones de interés. El número 1 indica la presencia de la secuencia en el alineamiento, mientras que el número 0 su ausencia. las columnas con los nombres: trnN-Cp - trnD, Intrón 2 del gen nad1 y rns-a - trnL-b corresponden a las regiones sin microsatélites. Las columnas con los nombres: Intron 3 del gen nad2, nad4 - atp4 y trnY-rps4 corresponden a las regiones con microsatélite, con las unidades de repetición (AG)<sub>n</sub>, (ATAGT)<sub>n</sub> y (CT)<sub>n</sub> respectivamente.

<b>Regiones mitocondriales</b>	<b>Tipo de región</b>	<b>Genoma utilizado para el diseño de los primers</b>	<b>Longitud del fragmento amplificado</b>	<b>Presencia/ Ausencia de microsatélites</b>	<b>No. de secuencias en el alineamiento.</b>
<i>TrnN-Cp-TrnD</i> 3156---4142 pb	<b>Región intergénica</b>	<i>Tripsacum dactyloides</i>	531 pb	Ausencia	35
Intron 2 del gen <i>nadh1</i> 70710..72110 pb	<b>Intrón</b>	<i>Tripsacum dactyloides</i>	978 pb	Ausencia	35
<i>TrnS-a - trnL-b</i> 650408 - 650652 pb	<b>Región intergénica</b>	<i>Tripsacum dactyloides</i>	178 pb	Ausencia	37
Intron 3 del gen <i>nadh2</i> (81300..83205 pb	<b>Intrón</b>	<i>Tripsacum dactyloides</i>	268pb	Presencia	27
<i>nadh4-atp4</i> 515435 – 515470 pb	<b>Region intergénica</b>	<i>Zea luxurians</i>	407 pb	Presencia	31
<i>trnY-rps4</i> 590057-590070 pb	<b>Región intergénica</b>	<i>Lolium perenne</i>	443 pb	Presencia	18

Tabla R.3 Descripción de las matrices construidas y origen de los primers para amplificar regiones con (intrón 3 del gen *nadh2*, *nadh4 - atp4* y *trnY-rps4*) y sin microsatélites (*trnN-Cp-trnD*, Intron 2 del gen *nadh1* y *trnS-a - trnL-b*) en 39 especies de Poaceae.





El microsatélite de la región intergénica *trnY - rps4* está compuesto por la unidad de repetición (CT)<sub>n</sub> y presentó un rango de 5 a 8 unidades de repetición. De manera similar al caso anterior, *Aegilops spetoides* y *Hordeum vulgare* mostraron el mayor número de repeticiones, mientras que las especies del género *Oryza* presentaron el menor. Por otro lado, a 96 pares de bases río arriba del microsatélite se observó una delección de 130 nucleótidos para las 8 especies del género *Zea* y *Sorghum bicolor*. El alineamiento presenta también varias delecciones cortas (de 5 a 8 pares de bases) en regiones homólogas compartidas entre este mismo grupo de especies río arriba y río abajo del microsatélite. El citotipo CMS-T de *Zea mays* presenta una inserción de 5 pares de bases (GATAG) río arriba del microsatélite, que no está presente en ningún otro taxón. Esta secuencia estuvo poco conservada dentro de la familia; su disponibilidad en las bases de datos fue limitada y no fue posible obtener amplificaciones de este segmento en muchas especies (por esta razón, no se reporta el diseño del cebador utilizado para la región) (Figura R.5).

El microsatélite presente en el espacio entre los genes *nadh4* y *atp4* está compuesto por la unidad de repetición (ATAGT)<sub>n</sub>. Este microsatélite fue caracterizado a partir del organismo *Zea luxurians* y mostró un rango de 3 a 12 unidades de repetición entre los organismos utilizados, teniendo un mayor número de repeticiones en las especies del género *Zea*. Éstas fueron las únicas que mostraron variación en cuanto al número de unidades de repetición, siendo *Zea perennis* la que tuvo el mayor número de las mismas. Las especies *Aegilops Speltoides*, *Bambusa oldhamii*, *Ferocalamus rimosivaginus*, *Lolium perenne*, *Sorghum bicolor*, *Tripsacum dactyloides*, *Triticum aestivum*, y *Triticum timopheevii* mostraron todas dos unidades de repetición interrumpidas por la inserción de una adenina (ATAGT)A(ATAGT). Por otro lado, se observaron delecciones cortas (de 2 a 6 pb) río arriba y río abajo del microsatélite compartidas entre varios taxones. *Zea luxurians*

presenta además una inserción de 5 pares de bases (TATGG) río abajo del microsatélite (Figura R.6).

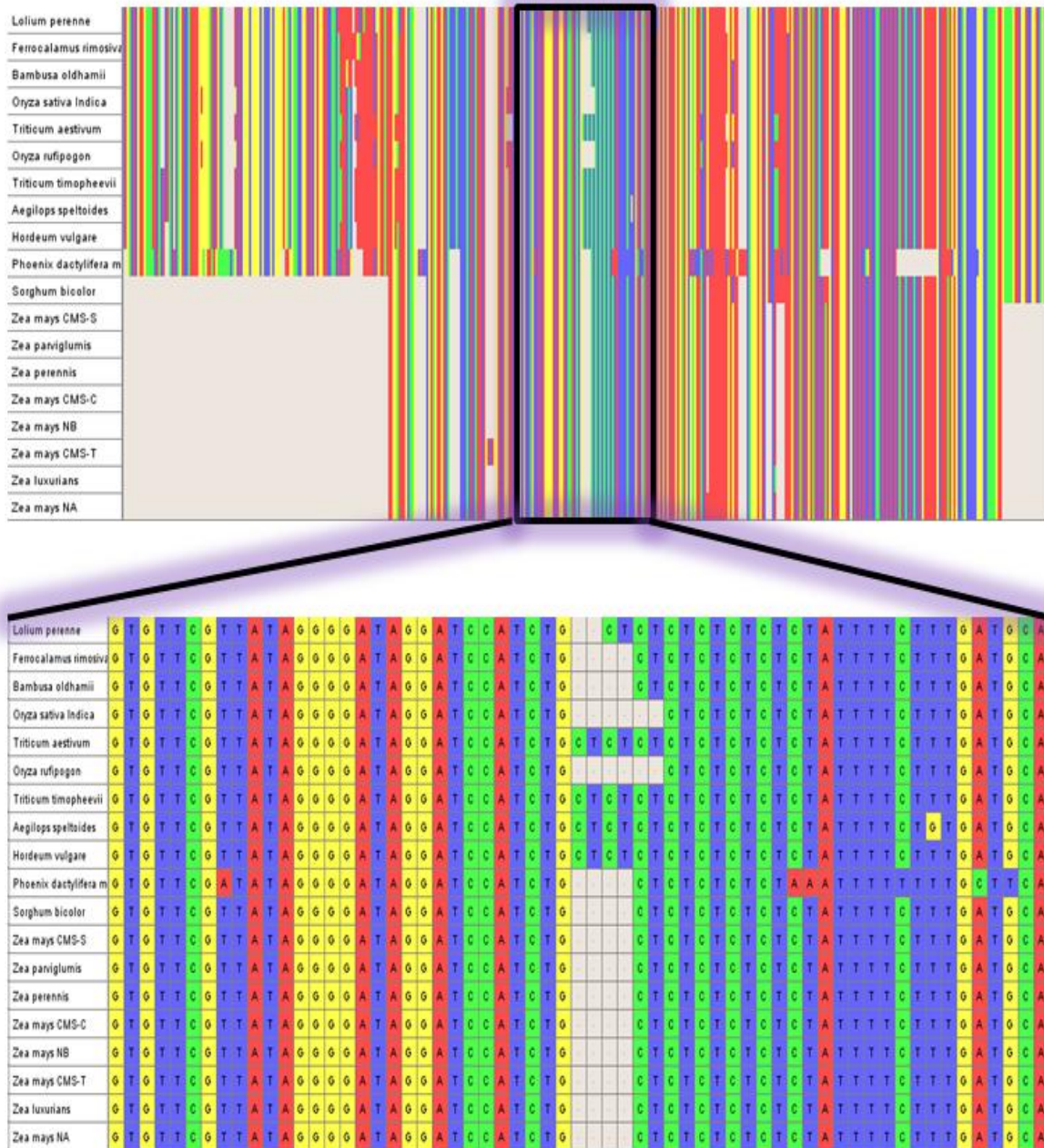


Figura R.5 Composición del microsatélite (CT)<sub>n</sub>, ubicado en la región entre los genes (trnY - rps4).

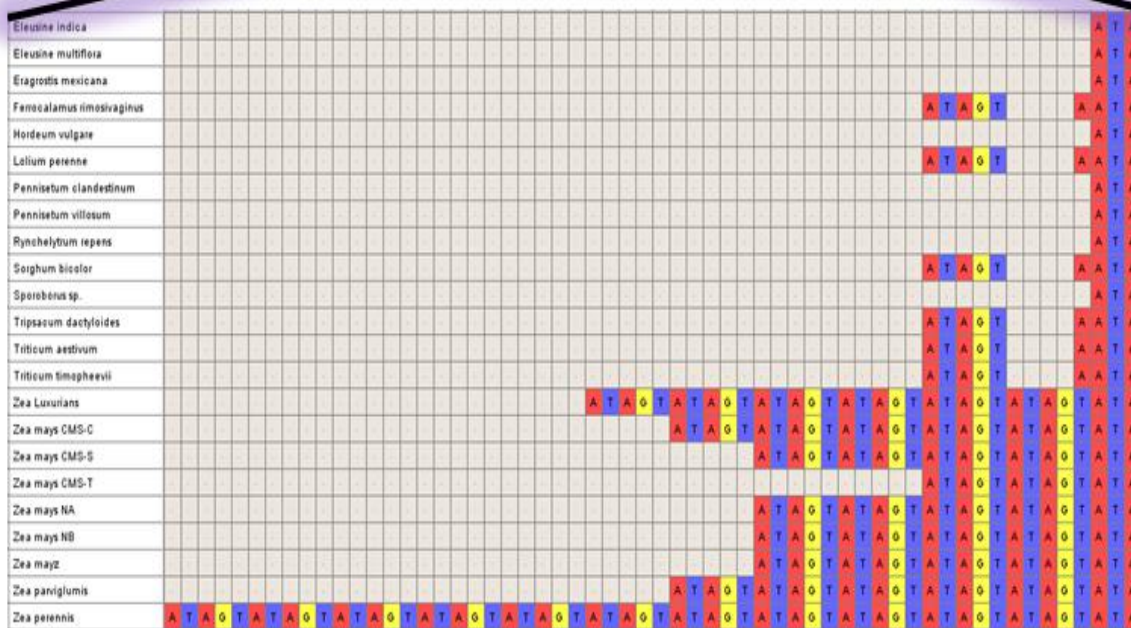
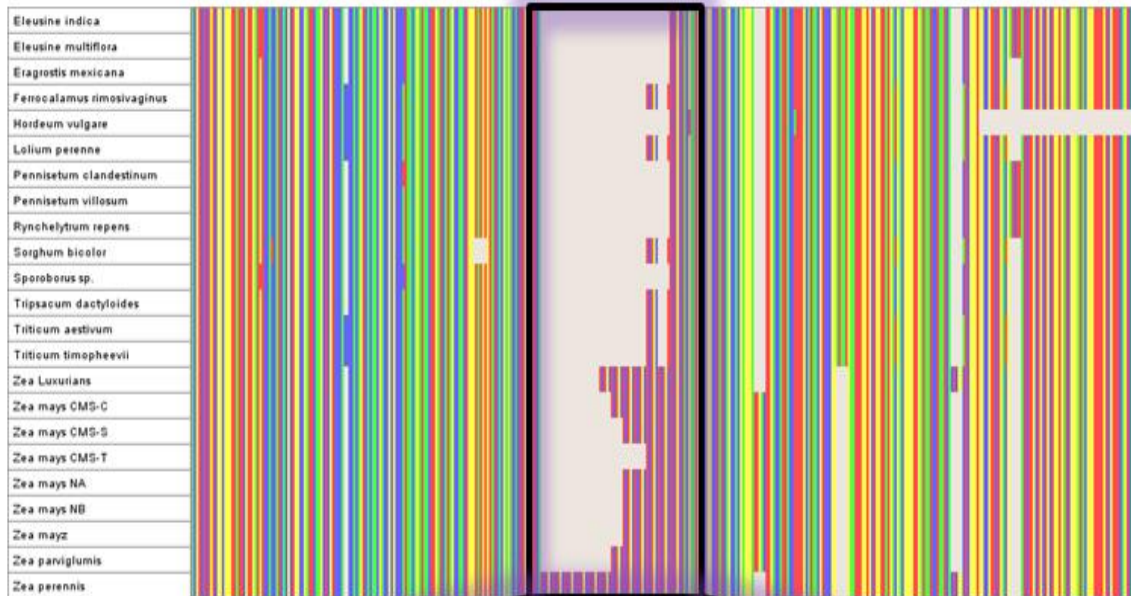


Figura R.6 Composición del microsatélite (ATAGT)<sub>n</sub>, ubicado en la región entre los genes nadh4 y atp4.

### **Análisis de la diversidad nucleotídica en las regiones de interés.**

Antes de hacer las comparaciones entre las secuencias portadoras y desprovistas de microsatélites, se les retiraron las unidades de repetición a los alineamientos de las primeras, y se analizaron las regiones vecinas por separado. Esto con el fin de no generar ningún tipo de sesgo al realizar el cálculo de los valores de la diversidad nucleotídica entre las secuencias (Tabla R.4).

Las regiones portadoras de microsatélites presentaron un número mayor de sitios polimórficos que las regiones desprovistas de ellos (fig R.7); sin embargo, esta diferencia no fue estadísticamente significativa ( $p = 0.601$ ,  $t$  de Welch = 0.6075, g.l. = 2.173.) Por otro lado, la comparación entre las regiones de interés también muestra un aumento en la tasa de inserción y delección por sitio en las regiones vecinas portadoras de microsatélites (fig. R.8). Sin embargo, este aumento tampoco es significativo estadísticamente ( $p=0.0633$ ,  $t$  de Welch = 2.929, g.l. = 2.912). De la misma manera, se observó un aumento en la divergencia de las secuencias al ser comparadas con los grupos externos (*Phoenix dactylifera* y *Cocus nucifera*) (fig. R.9), pero una vez más la prueba  $t$  de Welch no indicó un soporte estadístico suficiente ( $p=0.583$ ,  $t$  de Welch=0.7691 g.l. = 2.483).

Por otro lado y de manera general, los resultados muestran que las tasas de diversidad se elevan en aquellas regiones vecinas a un microsatélite (Tabla R.4). Un análisis de varianza (ANOVA) en donde se comparó el efecto de la tasa de mutación ( $\theta$ ) entre regiones, encontró diferencias significativas ( $F_{(5,177)}=32.39$ ,  $P<0.0001$ ) (fig R.10). Pruebas de Tukey *Ad Hoc* posteriores mostraron que había diferencias significativas al comparar las regiones *trnSa-trnLb* y el microsatélite (ATAGT)<sub>n</sub> ( $p=0.0182$ ), y las regiones *trnN-CP-trnD* y el microsatélite (CT)<sub>n</sub> ( $p=0.0432$ ). Análisis similares para el efecto en la diversidad

nucleotídica ( $\pi$ ) entre las regiones de interés indica que estos valores son estadísticamente significativos ( $F(5,177)=32.39$ ,  $p<0.0001$ ) (fig. R.11). Las pruebas *Ad Hoc* de Tukey mostraron diferencias significativas para todas las regiones comparadas, donde las diferencias más grandes se obtuvieron al comparar las regiones del intrón 2 del gen *nadh1* y el microsatélite (AG)<sub>n</sub>, (M = 0.01161, DE = 0.0003256,  $p<0.0001$ ), las regiones del intrón 2 del gen *nadh1* y el microsatélite (CT)<sub>n</sub> (M=0.01148, DE = 0.0003687,  $p<0.0001$ ) y la región intergénica *trnN-Cp-trnD* y el microsatélite (ATATG)<sub>n</sub> (M= 0.01148, DE=0.0003653,  $p<0.0001$ ).

Dado que los resultados anteriores apoyan la hipótesis sobre un aumento en la diversidad en las secuencias cercanas a microsatélites, se dividieron las secuencias portadoras en regiones río arriba y río abajo, se repitieron los análisis y se encontraron diferencias significativas entre regiones para  $\pi$  ( $p<0.0001$  en los tres casos) (fig. R.12) y para  $\theta$  (fig. R.13) en la región de microsatélite (AG)<sub>n</sub> ( $p=0.0268$ ). Este último resultado sugiere cambios en la tasa de mutación en ambos lados del microsatélite, lo que se vería reflejado en la acumulación diferencial de variación nucleotídica en las regiones vecinas a éstos.

Región de interés.		S	H	$\pi$	$\theta$	I	$D_{xy}$
Río arriba_(1- 164)	(AG)n	11	11	0.01918	0.0173	0.00294	0.13743
Río abajo (181-268)		7	5	0.0091	0.02522	0.01243	0.22685
Río arriba (1-149)	(ATAGT)n	8	7	0.00723	0.01298	0.00352	0.02886
Río abajo (210-407)		5	4	0.0194	0.01117	0.00718	NA*
Río arriba (1-222)	(CT)n	4	5	0.01307	0.01551	0.00864	0.02074
Río abajo_(239-443)		11	8	0.02194	0.017	0.00523	0.15326
SM	<i>TrnN-Cp-TrnD</i>	24	16	0.01001	0.01287	0.00396	0.07091
	<i>Intron 2 nadh1</i>	29	18	0.0043	0.00789	0.00448	0.04482
	<i>TrnSa-TrnLb</i>	6	8	0.00776	0.00876	0.00201	0.05035
CM	(AG)n	18	15	0.01591	0.01987	0.00635	0.0914
	(ATAGT)n	13	10	0.00892	0.01216	0.0052	0.02886
	(CT)n	15	10	0.0192	0.01655	0.00626	0.10122

Tabla R.4. Valores de diversidad nucleotídica obtenidos a partir de las regiones de interés. Los índices calculados fueron: **S** (Número de sitios segregantes), **H** (Número de haplotipos),  $\pi$  (diferencias nucleotídicas por sitio),  $\theta$  (estimador de la tasa de mutación escalado por el tamaño efectivo poblacional) y  $D_{xy}$ , (índice de divergencia). **Abreviaturas:** **CM** (Con microsatélite), **SM** (Sin microsatélite).

\* :El índice no pudo ser calculado debido a que no se encontró una región homóloga en los grupos externos.

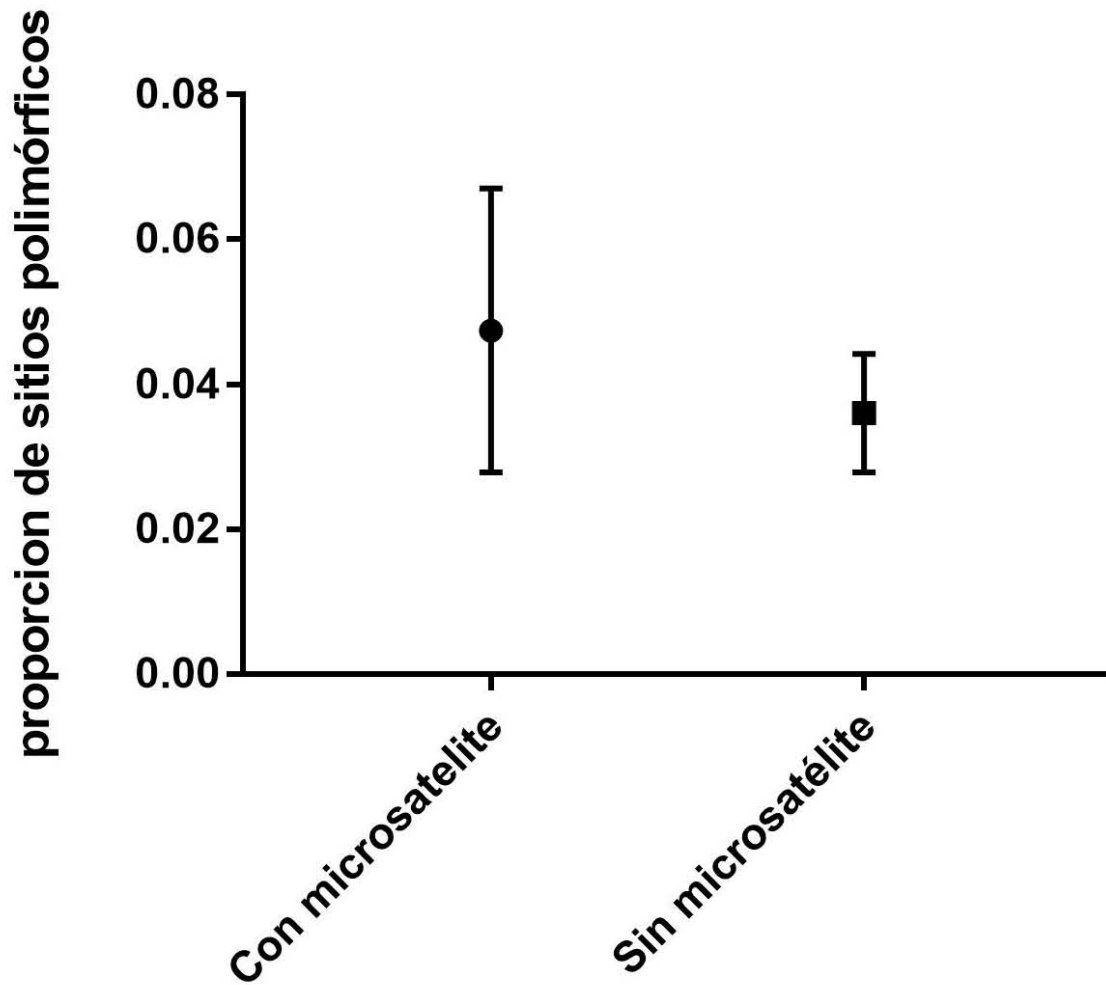


Figura R.7 Proporción de sitios polimórficos en regiones con y sin microsatélites calculadas para 39 especies de Poaceae. Las diferencias entre no son significativas ( $p = 0.601$ ,  $t$  de Welch = 0.6075, g.l. = 2.173).



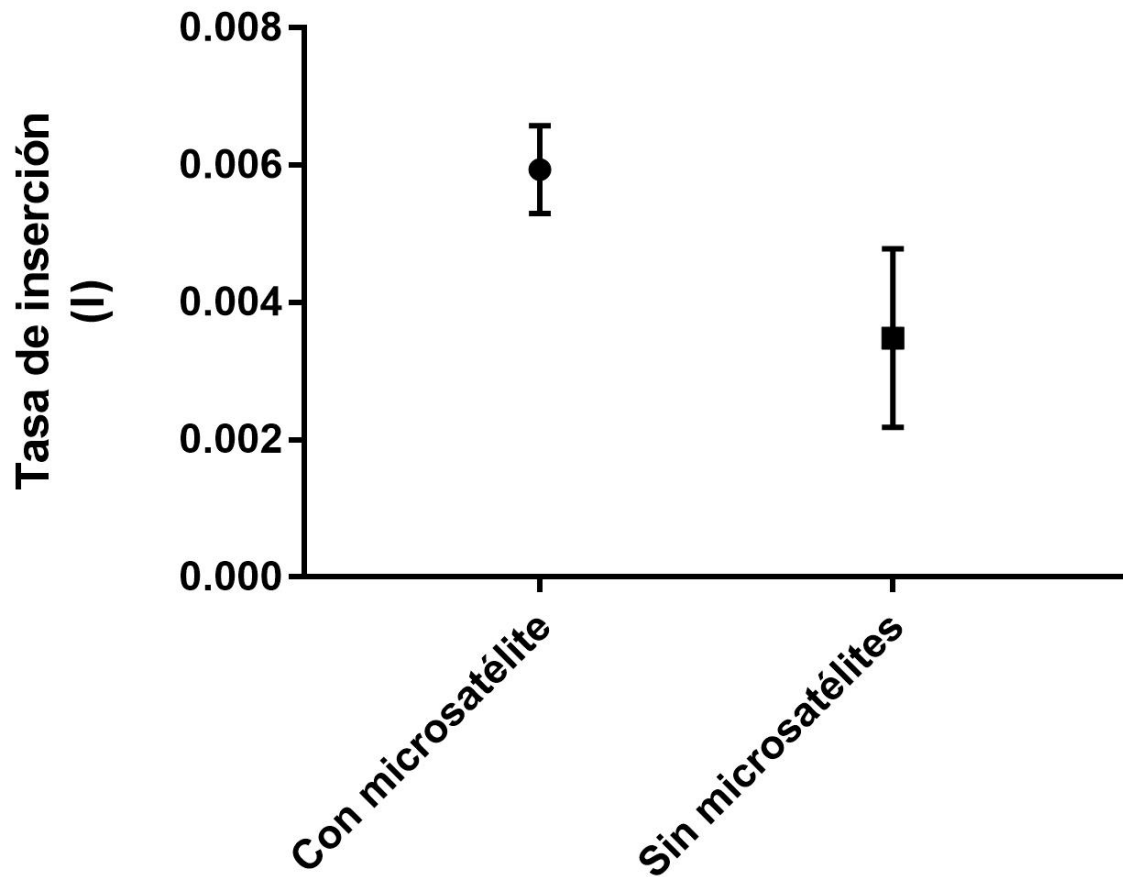


Figura R.8 Tasa de inserción y deleción (I) en regiones con y sin microsatélites calculadas para 39 especies de Poaceae. Las diferencias entre no son significativas ( $p=0.0633$ ,  $t$  de Welch = 2.929, g.l. = 2.912).

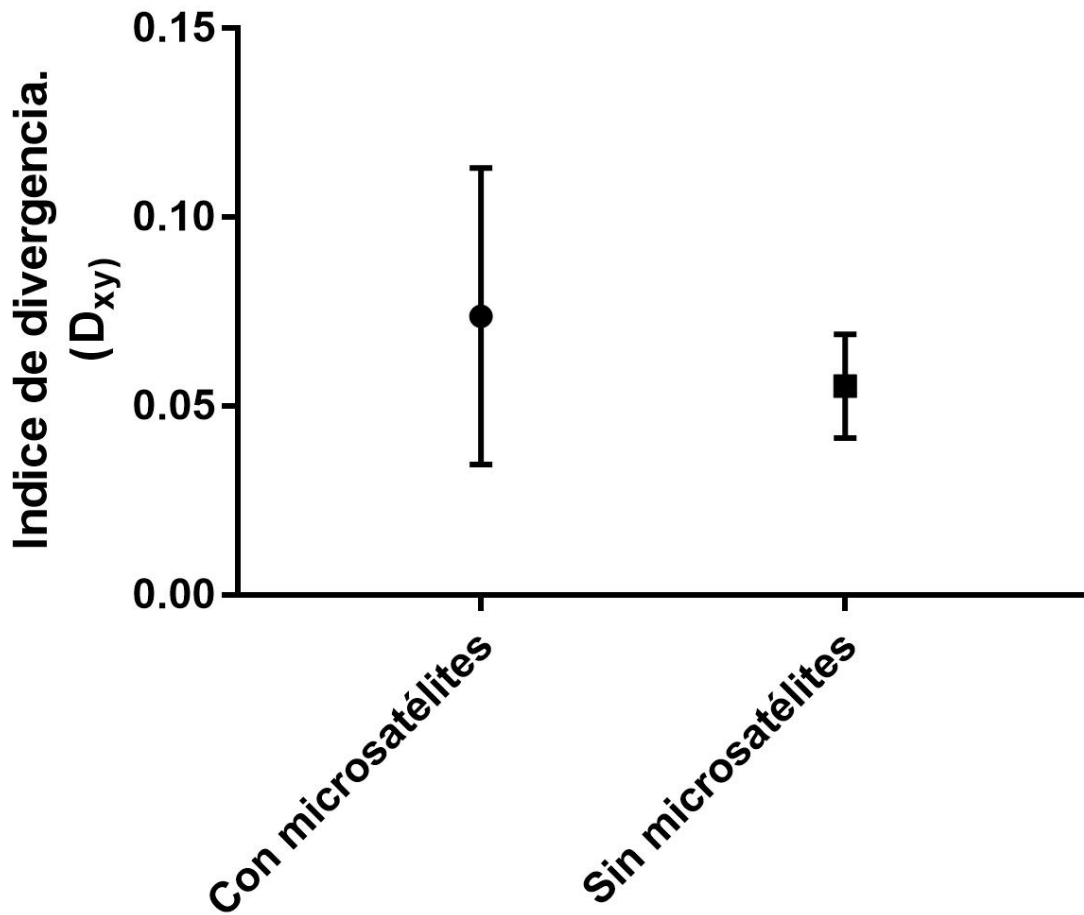


Figura R.9 Tasa de divergencia ( $D_{xy}$ ) entre regiones con y sin microsatélites calculadas para 39 especies de Poaceae. Las diferencias entre no son significativas ( $p=0.583$ , t de Welch = 0.7691 g.l. = 2.483). Se utilizó a *Phoenix dactylifera* y *Cocus nucifera* como grupos externos.

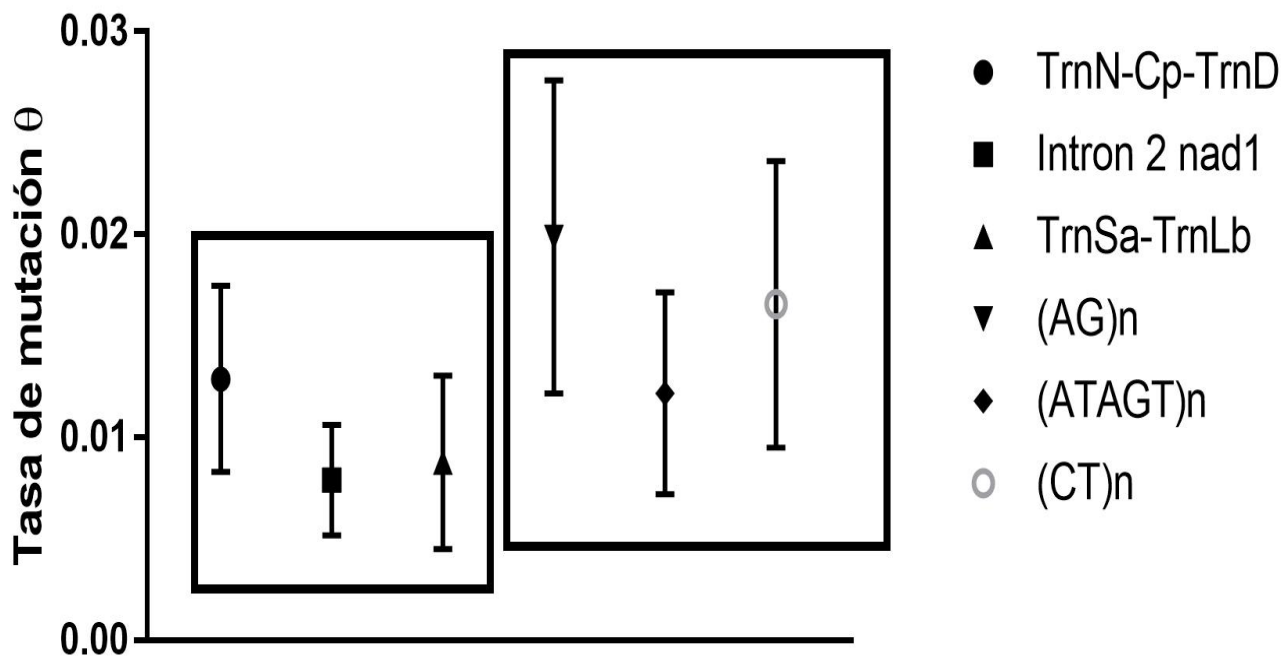


Figura R.10 Diferencias en la tasa de mutación ( $\theta$ ) entre regiones con y sin microsatélites calculadas para 39 especies de Poaceas. Se encontraron diferencias significativas al comparar las regiones *trnSa-trnLb* y el microsatélite (ATAGT)<sub>n</sub> ( $p=0.0182$ ), y las regiones *trnN-CP-trnD* y el microsatélite (CT)<sub>n</sub> ( $p=0.0432$ ).

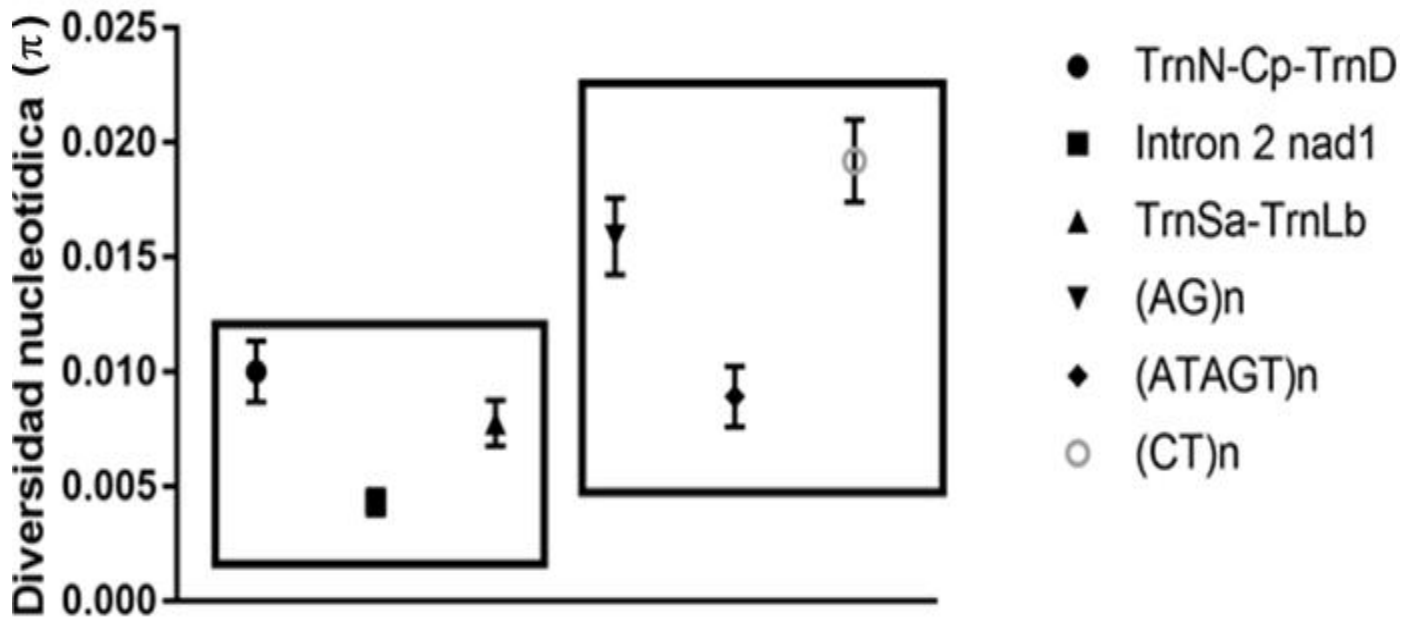


Figura R.11 Diferencias en la diversidad nucleotídica ( $\pi$ ) entre regiones con y sin microsatélites calculadas para 39 especies de Poaceae. donde las diferencias más grandes se obtuvieron al comparar las regiones del intrón 2 del gen *nadh1* y el microsatélite (AG)<sub>n</sub>, (M = 0.01161, DE = 0.0003256, p<0.0001), las regiones del intrón 2 del gen *nadh1* y el microsatélite (CT)<sub>n</sub> (M=0.01148, DE = 0.0003687, p<0.0001) y la región intergénica *trnN-Cp-trnD* y el microsatélite (ATAGT)<sub>n</sub> (M= 0.01148, DE=0.0003653, p<0.0001).

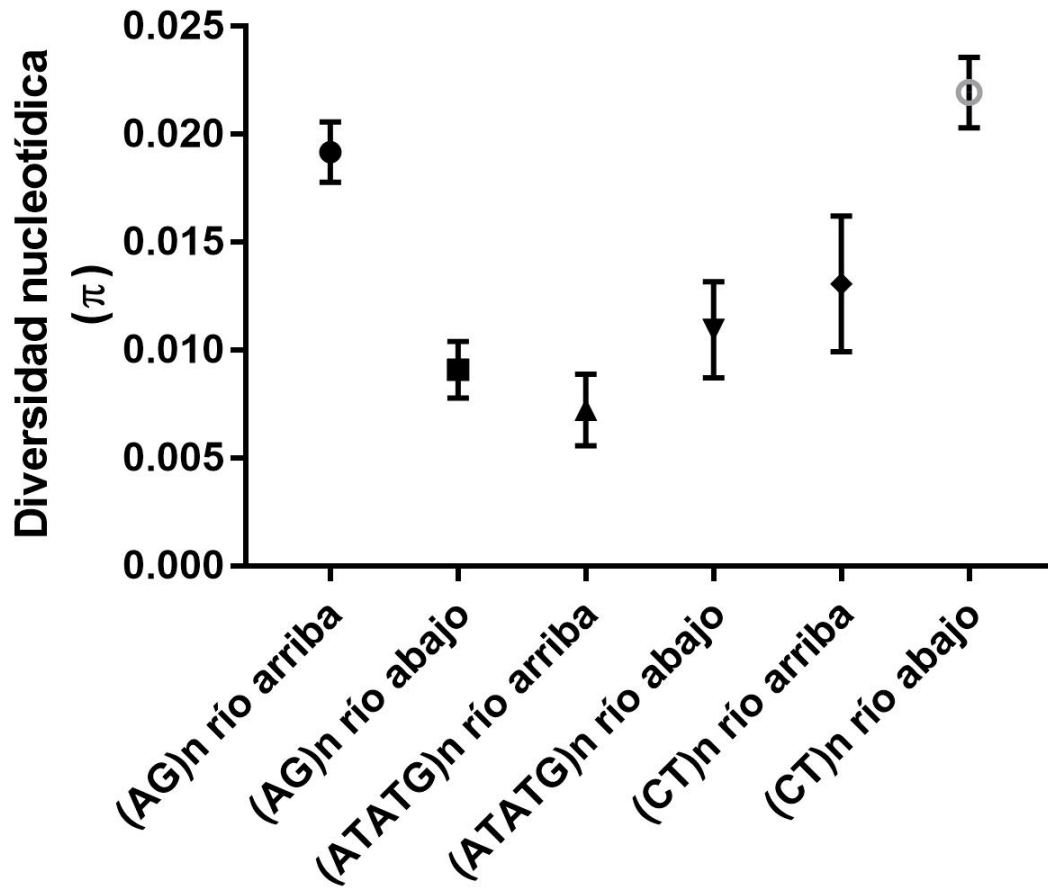


Figura R.12 Valores de diversidad  $\pi$  para segmentos río arriba y río abajo de las unidades de repetición en 39 especies de Poaceae. Existen diferencias significativas entre regiones ( $F(5,146) = 218.6, p < 0.0001$ ).

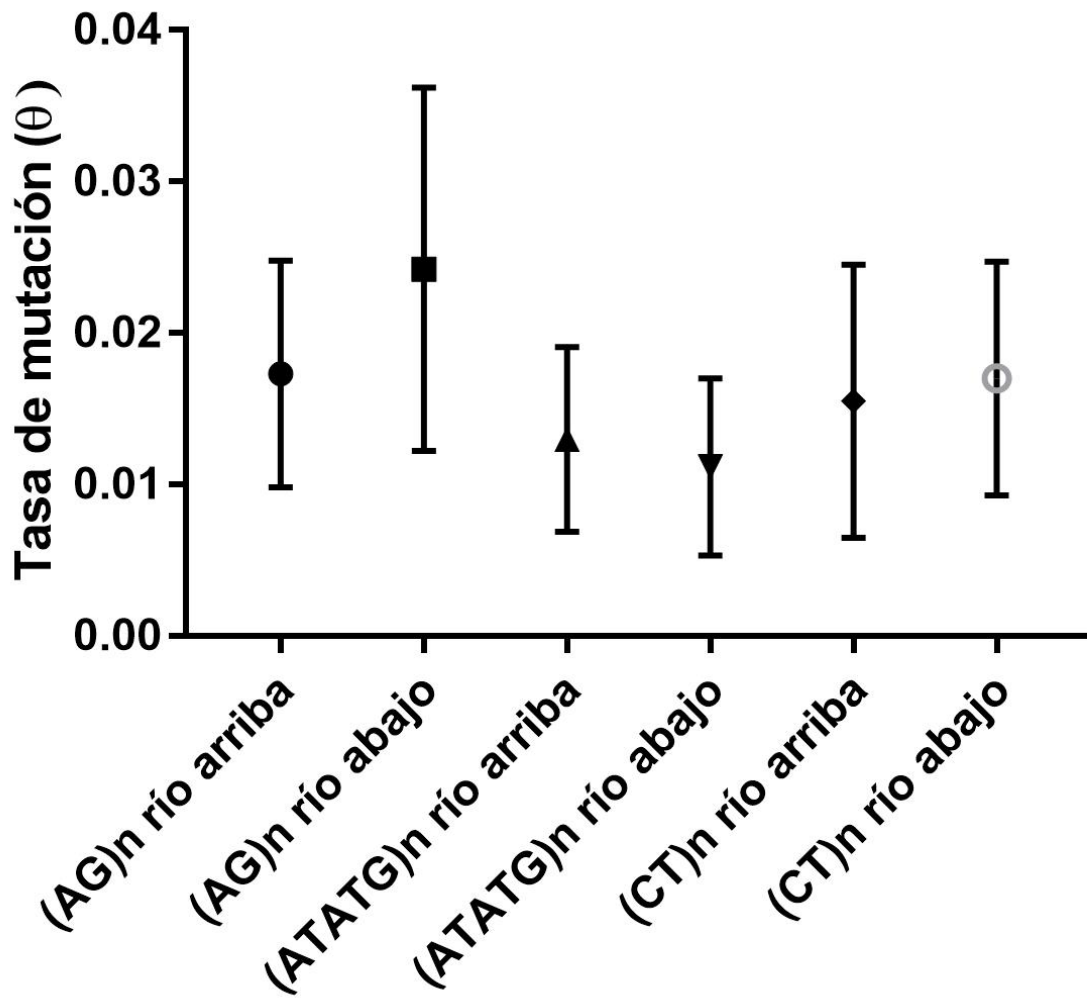


Figura R.13 Tasas de mutación para segmentos río arriba y río abajo de las unidades de repetición en 39 especies de Poaceae. Se encontró que existen diferencias significativas entre las regiones río abajo y río arriba ( $F(5,146) = 8.665, p < 0.0001$ ).

### **Reconstrucción filogenética.**

La figura R.14 representa la reconstrucción filogenética de las gramíneas. En esta se observan los dos clados principales descritos previamente para la familia Poaceae (Edwards *et al.*, 2011), BEP (subfamilias Bambusoideae, Ehrhartoideae y Pooideae) y PACMAD (subfamilias Panicoideae, Arundinoideae, Chloridoideae, Micraroideae, Aristidoideae y Danthonoideae). De manera general, el soporte fue elevado para la mayoría de los nodos, con valores de probabilidad posterior cercanos a 1, excepto para aquellos que contienen a las subfamilias Chloridoideae y Panicoideae.

Se realizó un análisis de la varianza de un factor utilizando los genomas mitocondriales completos de las subfamilias Pooideae, Ehrhartoideae y Panicoideae, para saber si existe una diferencia entre el tamaño del genoma de las diferentes subfamilias. No fue posible encontrar alguna relación entre el tamaño del genoma completo con algún grupo filogenético ( $P=0.0514$ ,  $F=3.761$ ). De hecho, se observó la existencia de un gran rango de tamaño de los mitogenomas. (figura R.15). Por ejemplo, los genomas disponibles para la subfamilia Pooideae (*Aegilops speltoides*, *Hordeum vulgare*, *Lolium perenne*, *Triticum aestivum* y *Triticum timopheevii*) mostraron un tamaño promedio de 515,243 pares de bases, donde la diferencia del tamaño del genoma más grande y del más pequeño fue de 235,161 pares de bases. Entre organismos del mismo género, los diferentes mitogenomas disponibles para los distintos citotipos de la especie *Zea mays*, así como las especies *Zea perennis*, *Zea luxurians* y *Zea parviglumis* mostraron un promedio de 611,713 pares de bases, con un rango de 203,894. Estos resultados señalan que el tamaño del genoma completo en las mitocondrias de los pastos es altamente variable, y que esta variación puede generarse en cortos periodos evolutivos.

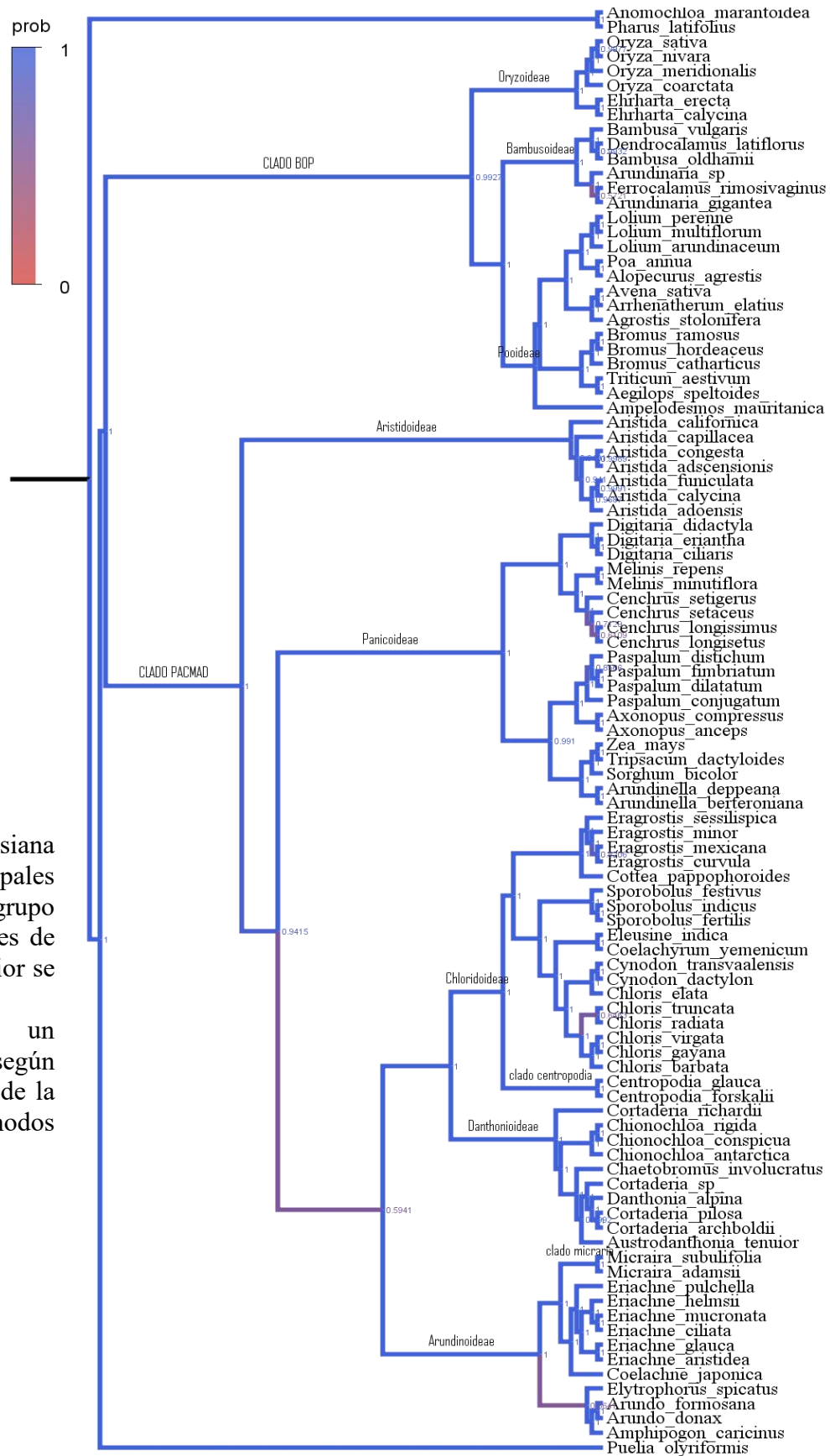


Figura R.14  
 Reconstrucción  
 filogenética bayesiana  
 de las principales  
 subfamilias del grupo  
 Poaceae. Los valores de  
 probabilidad posterior se  
 encuentran  
 representados en un  
 gradiente de color según  
 indica en la escala de la  
 izquierda y en los nodos  
 del árbol.



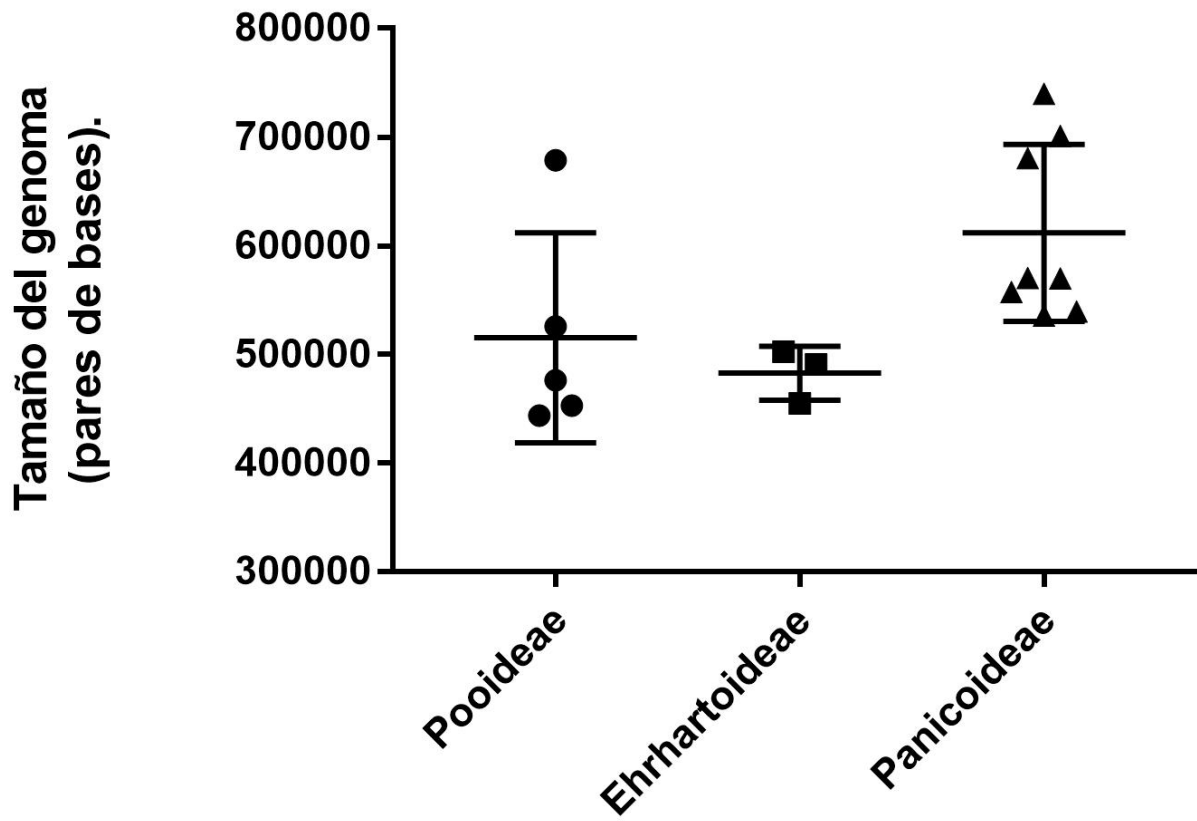


Figura R.15 Comparación del tamaño del genoma entre las subfamilias Pooideae, Ehrhartoideae y Panicoideae. No fue posible encontrar alguna relación entre el tamaño del genoma completo con algún grupo filogenético ( $P=0.0514$ ,  $F=3.761$ ).

### **Reconstrucción de caracteres ancestrales.**

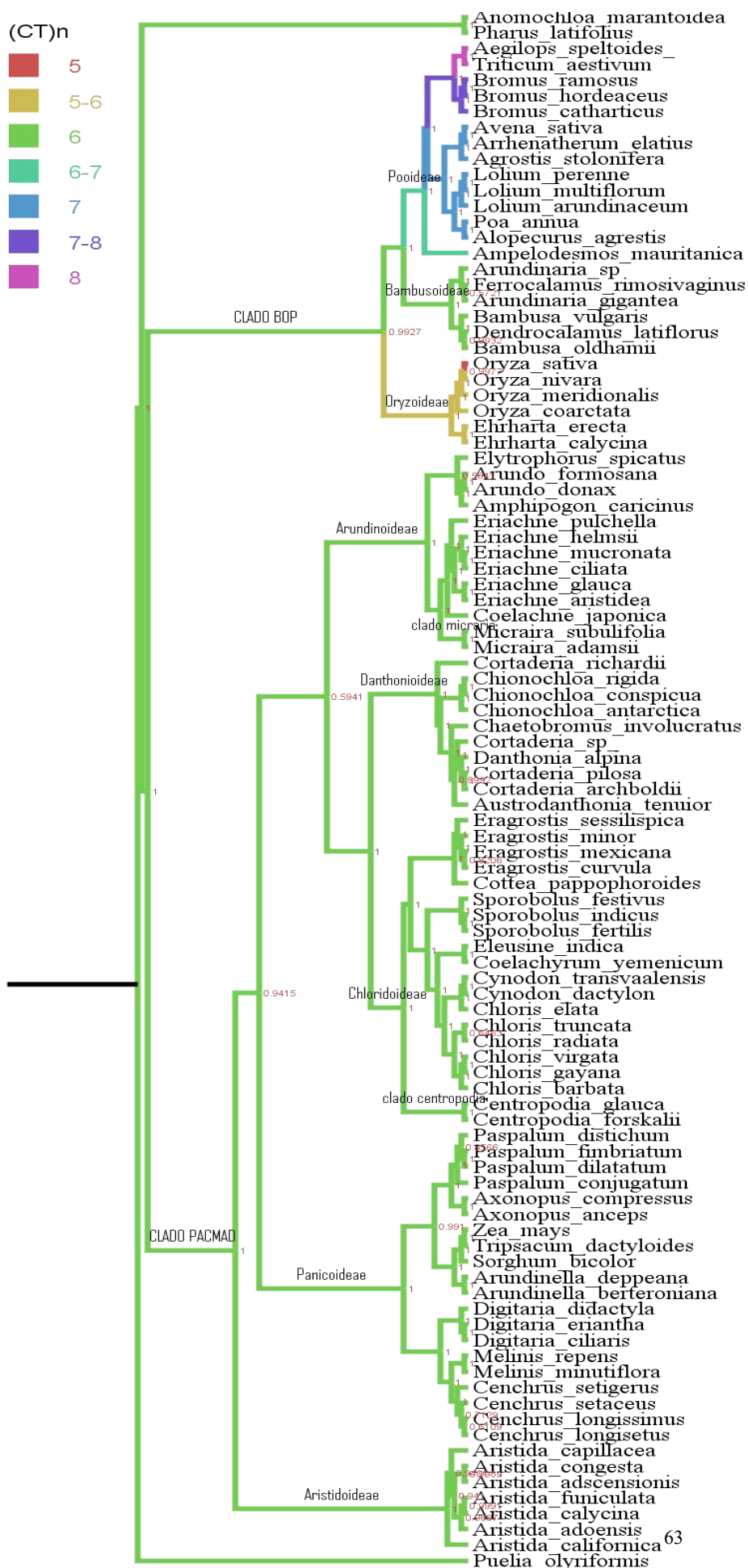
La reconstrucción de los caracteres ancestrales sugiere que los microsatélites  $(AG)_n$  y  $(CT)_n$  fueron adquiridos previamente a la diversificación del grupo de los pastos (figuras R.16 y R.17 respectivamente). Estos dos microsatélites poseían 6 unidades de repetición en el momento de la divergencia de la familia Poaceae; en ambos casos se encontró que el mayor número de repeticiones se encuentra dispuesta en el clado BEP (8 repeticiones). Sin embargo, en algunos clados es posible observar estados homoplásicos. Por ejemplo, para el microsatélite  $(AG)_n$  se observan 8 unidades de repetición que habrían aparecido de forma independiente en dos clados diferentes, Chloridoideae y BEP. Para tener una mejor idea del grado de homoplasia entre los caracteres reconstruidos se calculó los índices de consistencia y de retención. los valores calculados (Fig R.16) indican altos valores de homoplasia para la reconstrucción del microsatélite  $(AG)_n$ .

En el caso del microsatélite  $(ATATG)_n$ , la reconstrucción de caracteres ancestrales (fig. R.18) mostró que la inserción de esta región es reciente. Propiamente hablando, el microsatélite se encuentra presente sólo en los organismos del género *Zea*, presentando polimorfismos intraespecíficos, con un número de repeticiones que va de 3 a 12. Los valores de consistencia y retención para el microsatélite  $(ATATG)_n$  indican que la mayor cantidad de variación se encuentra presente en forma de autapomorfía (Fig. R.17 Los valores reportados para la repetición  $(CT)_n$  indican que no hay homoplasia para este carácter, pues sólo se presenta en los organismos del género *Zea* (Fig.R.18).



Figura R.17

Reconstrucción de estados ancestrales de la repetición (CT)<sub>n</sub> del espacio entre los genes *trnY* y *rps4*, sobre la filogenética de las principales subfamilias de Poaceae. Los distintos cambios de estado del carácter están representados con diversos colores según se indica en la escala del lado izquierdo. Los índices de retención y consistencia para este carácter son de 1, indicando ausencia de homoplasia.



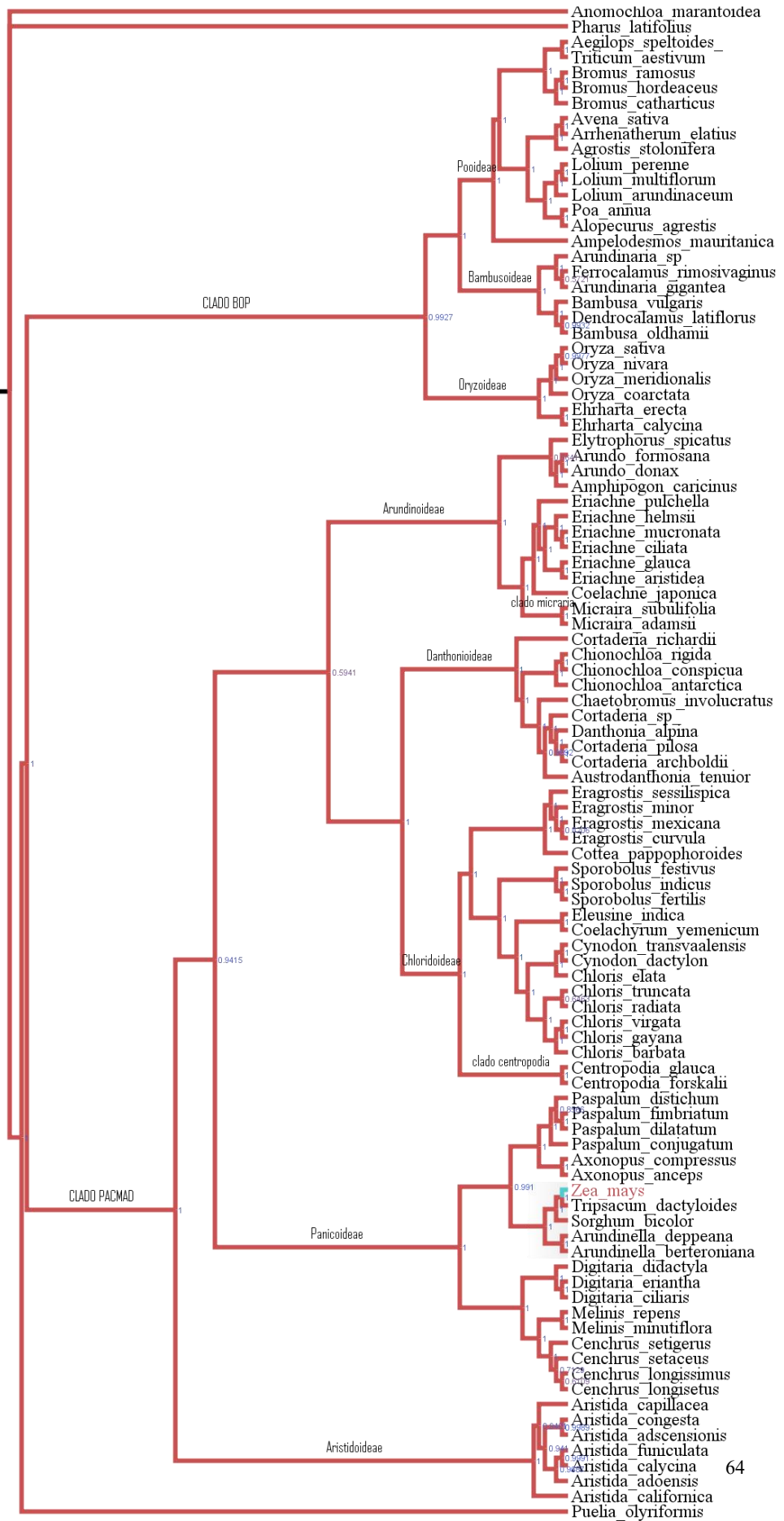


Figura R.18  
 Reconstrucción de estados ancestrales de la repetición (ATAGT)<sub>n</sub> presente en el espacio entre los genes *nadh4* y *atp4*, sobre la filogenética de las principales subfamilias del grupo Poaceae. Los distintos cambios de estado del carácter están representados con diversos colores según se indica en la escala del lado izquierdo. Se muestra una estasis general de la repetición, excepto para el género *Zea*, en donde existe un aumento significativo de repeticiones. El índice de consistencia es de 0, y el índice de consistencia es de 1 Indicando la presencia de autapomorfías, las cuales en este caso está presentes en el género *Zea* (el cual está señalado en el árbol con color gris).

## **Discusión.**

### **Distribución de elementos repetidos en las distintas regiones de los genomas mitocondriales.**

Las pruebas estadísticas realizadas para comparar la diversidad genética de las regiones vecinas a microsatélites con las de regiones sin microsatélites indican un aumento en la diversidad nucleotídica y las tasas de mutación en las primeras (Figuras R.8 y R.9, Tabla R.4). Si el aumento en la diversidad nucleotídica mantiene una relación con la inserción de microsatélites dentro del genoma mitocondrial vegetal, entonces el siguiente paso es verificar qué tan frecuente es la presencia de estos elementos en los genomas mitocondriales de otras especies de plantas, para conocer el efecto que en conjunto podrían tener dentro del genoma mitocondrial vegetal. Diversos estudios han encontrado la presencia de microsatélites en genomas mitocondriales de diferentes organismos vegetales (Jaramillo-Correa *et al.*, 2013; Guo *et al.*, 2017), con diferentes frecuencias con respecto a la unidad de repetición que presentan (Paterson *et al.*, 2016). En este estudio se encontraron un total de 3,056 arreglos en 19 especies de Poaceae. Esto indica que los microsatélites son elementos relativamente abundantes en los genomas mitocondriales de las plantas, y que su posible impacto a través del aumento de las tasas de evolución puede ser considerable a diferentes niveles taxonómicos.

Tanto en éste como en estudios precedentes se encontró que la mayoría de microsatélites se encuentra en regiones no codificantes (Nishikawa *et al.*, 2005; Alverson *et al.*, 2010; Gualberto *et al.*, 2014; Paterson *et al.*, 2016; Guo *et al.*, 2017), mientras que es poco frecuente encontrarlos en exones (Tóth *et al.*, 2000; Vieira *et al.*, 2016). Por ejemplo, se ha mostrado que en el genoma mitocondrial de *Capsicum annuum* (de la familia

Solanaceae), el 83% de los microsatélites se encuentran dentro de las regiones intergénicas (Cheng *et al.*, 2016), mientras que la gran mayoría de las 674 repeticiones encontradas para *Gossypium raimondii* (Malvaceae) también estuvieron mayormente concentradas en regiones no codificantes (Paterson *et al.*, 2016). El que haya una mayor frecuencia de microsatélites en regiones no codificantes puede deberse tanto a las presiones de selección a la que están sometidas las regiones codificantes, como a la proporción de las regiones no codificantes en el genoma mitocondrial vegetal. Al insertarse en exones, los microsatélites pueden generar alteraciones a través de cambios en el marco de lectura (Wang *et al.*, 2008), resultando en productos defectuosos y siendo eliminados por selección natural. Aquellos que aún se encuentran en exones suelen presentar unidades de repetición trinucleótidas o hexanucleótidas, que no afectan el marco de lectura de las proteínas (Morgante *et al.*, 2002; Vieira *et al.*, 2016).

En este estudio pudimos encontrar una correlación positiva entre el tamaño del genoma y el número de microsatélites presentes (tabla R.1, figura R.3). Esta correlación no sólo se ha reportado para genomas mitocondriales (Jaramillo-Correa *et al.*, 2013), sino también para genomas nucleares (Cheng *et al.*, 2016). Algunos procesos involucrados en la variación del tamaño del genoma mitocondrial, como los eventos de duplicación o la adquisición material genético exógeno (Palmer *et al.*, 2000; Clifton *et al.* 2004; Kitazaki & Kubo, 2010), pueden propiciar la proliferación de microsatélites. Por ejemplo, hay estudios que muestran que los intrones y los transposones pueden portar secuencias repetidas de diferentes longitudes (Scherer *et al.*, 2008; Cardoso *et al.*, 2017); en genes mitocondriales se ha reportado que algunos intrones de los genomas mitocondriales de los géneros *Psilotum* y *Ophioglossum* pueden llegar a tener tamaños mayores a 6 kb, lo que representa una extensión de un 60 a un 80% compuesta por elementos repetidos (Scherer,

2008; Guo *et al.*, 2017). La abundancia de intrones y elementos transponibles puede relacionarse directamente con el tamaño del genoma (Lynch & Conery *et al.*, 2003; Sona *et al.*, 2014), y si son portadores de microsatélites, estos últimos podrían proliferar dentro del genoma (Tomilin, 2008); haciendo posible que aquellos genomas con mayor número de elementos transponibles o intrones presentes, muestren una mayor variación en el tamaño del genoma, por la presencia de microsatélites.

Aun cuando en este estudio se demuestra la relación entre el número de elementos repetidos y el tamaño del genoma mitocondrial vegetal, es posible que no sea un fenómeno generalizable a otros genomas o a todos los grupos biológicos. Se ha encontrado que el número, distribución y tipo de microsatélites pueden presentar características taxa específicas (Tóth *et al.*, 2000), y en algunos casos los microsatélites pueden encontrarse en mayor abundancia en secuencias expresables (Morgante *et al.*, 2002). De ser cierta la posibilidad de que los resultados observados sean exclusivos para pastos, se hace necesaria la comparación con la presencia de microsatélites mitocondriales en otros organismos vegetales, animales y hongos. Una forma de abordar el tema sería la comparación de microsatélites entre genomas completos de organismos, con diferente grado de parentesco, para saber si la variación observada es exclusiva entre grupos, o resulta ser un fenómeno generalizado.



### **Comparación de la diversidad nucleotídica entre regiones con microsatélites y sin microsatélites.**

Los resultados obtenidos en este estudio sugieren un aumento en las tasas de diversidad nucleotídica en las regiones adyacentes a microsatélites. Las tasas de mutación usualmente son obtenidas a través del cálculo de la tasa de sustituciones sinónimas en genes del genoma mitocondrial vegetal, pues ofrecen una buena aproximación de las tasas de mutación (Palmer *et al.*, 2000; Sloan *et al.*, 2007; Sloan *et al.*, 2009). Utilizando esta aproximación se han reportado bajas tasas de sustitución en los genomas mitocondriales de angiospermas, siendo en algunos casos de 50 a 100 veces menor que en genomas de mitocondrias de vertebrados (Palmer, *et al.*, 2000). En este estudio, la evaluación de la diversidad a través de los estimados de  $\pi$  y  $\theta$  en regiones no codificantes encontró que la presencia de microsatélites en el genoma mitocondrial mantiene una relación con el aumento de la diversidad genética (tabla R.4, figuras R.8 y R.9). Esto último puede verse reflejado en los resultados obtenidos al comparar el índice de divergencia entre las distintas regiones del genoma mitocondrial escogidas en este estudio ( $D_{xy}$ , figura R.7), donde se aprecia un mayor grado de divergencia en regiones portadoras de microsatélites, al ser comparadas con el grupo externo. Si bien la comparación del estimado  $D_{xy}$  entre regiones no resultó estadísticamente significativo, si sugiere una relación entre la presencia de microsatélites y la acumulación de variación en distintas regiones del genoma. Cabe mencionar que es posible que la presencia misma de los microsatélites no sea la causa directa del aumento de la diversidad nucleotídica, sino el resultado de algún proceso o condición subyacente, permitiéndonos ver al microsatélite como un indicador de regiones de alta variabilidad nucleotídica.

Para explicar las bajas tasas de mutación en genomas mitocondriales vegetales, se ha utilizado la hipótesis de la carga mutacional (Mutational Burden Hypothesis o MBH por sus siglas en inglés), la cual señala que la probabilidad de que una mutación pueda interrumpir un elemento funcional dentro del genoma es directamente proporcional a la tasa de mutación, por lo que elevadas tasas de mutación pueden presentar un coste selectivo (Sloan et al., 2012). Aunque los genomas mitocondriales en plantas generalmente se caracterizan por presentar tamaños mayores y tasas de sustitución bajas cuando se les compara con hongos o animales, este y otros estudios han encontrado aumentos en las tasas de mutación, las cuales han estado asociadas con cambios dramáticos en el tamaño del genoma (Cho et al., 2004; Parkinson et al., 2005; Sloan et al., 2007; Sloan et al., 2012). La correlación entre el tamaño del genoma y el aumento en el número de microsatélites presentes, así como el aumento en la diversidad nucleotídica asociada a ellos, indica que pueden existir dentro del genoma elementos discretos relacionados con aumentos en las tasas de diversidad, abriendo el debate sobre el costo selectivo impuesto al aumento de las tasas de mutación asociado a genomas de gran tamaño, como propone la hipótesis de la carga mutacional.

A pesar de no ser estadísticamente significativos, los resultados también sugieren un aumento en las tasas de inserción y deleción en las regiones vecinas a los microsatélites (figura R.6). Lo anterior podría sugerir un posible papel en el agrandamiento y reducción del genoma mitocondrial. El aumento en las tasas de inserción y deleción puede atribuirse a la inestabilidad en las regiones repetidas, haciendo al genoma propenso a rearrreglos en su estructura, generando inserciones o delenciones en las zonas flanqueadas. Por ejemplo, a través del deslizamiento del DNA (por el cual los microsatélites van ganando o perdiendo unidades de repetición) (Bzymek & Lovett, 2001; Lovett, 2004). No obstante, las razones

del aumento o disminución de las tasas de inserción o deleción, pueden estar ligadas a diversos procesos como mutaciones en la polimerasa I, secuencias repetidas capaces de formar estructuras secundarias o por el desprendimiento de elementos transponibles (Bzymek & Lovett, 2001). De ser probado el aumento de las tasas de inserción y deleción, cabe preguntar si la longitud o el tipo del arreglo están correlacionados con la ganancia o pérdida de segmentos de DNA en las regiones colindantes. Se ha sugerido que la inestabilidad de los microsatélites podría estar afectada por su tamaño, composición o longitud. En particular, se ha encontrado en varios casos que la longitud y composición del microsatélite influyen en el proceso de deslizamiento de DNA (Lay & Sun, 2003), haciendo que secuencias repetidas con longitudes mayores sean más variables en el número de unidades de repetición (Wierdl *et al.*, 1997). Esto podría probarse a través de una comparación entre distintos tipos de microsatélites (di, tri, tetra, penta y hexanucleótidos) con distintas longitudes en el genoma mitocondrial vegetal, y su efecto en las tasas de inserción y deleción.

## **Evolución y conservación de los microsatélites seleccionados en el genoma mitocondrial de los pastos.**

La reconstrucción de estados ancestrales mostró una retención de los microsatélites (CT)<sub>n</sub> y (AG)<sub>n</sub> durante la historia evolutiva de las gramíneas, e incluso mucho antes de su divergencia (figuras R.13 y R.15). La conservación de microsatélites en los genomas mitocondriales de plantas se han encontrado entre grupos filogenéticamente cercanos (Nishikawa *et al.*, 2005; Ishii *et al.*, 2006; Jaramillo-Correa *et al.*, 2013). Sin embargo, al comparar las secuencias obtenidas de las especies de la familia de los pastos con las especies *Phoenix dactylifera* y *Cocos nucifera*, pertenecientes al grupo de las poales, se encontraron las regiones de los microsatélites (CT)<sub>n</sub> y (AG)<sub>n</sub>, con un número de unidades de repetición de 5 en ambos microsatélites. Tanto las gramíneas como las especies utilizadas como grupos externos pertenecen al clado de las Commelinidae, haciendo posible que la inserción de la región repetida en ambos microsatélites se haya dado hace más de 58.6 millones de años, antes de la divergencia de las gramíneas (Glémin & Bataillon, 2009, Magallón *et al.*, 2015). Por otra parte, la búsqueda de estas regiones en el genoma mitocondrial de la especie *Cycas taitungensis*, reveló que los microsatélites fueron insertados en etapas posteriores a la divergencia entre las gimnospermas y las angiospermas, hace aproximadamente 250 - 200 Ma (Condamine *et al.*, 2015), puesto que no se encontró la presencia de las regiones homólogas en su genoma mitocondrial. Una búsqueda final en la base de datos de NCBI mostró la presencia de ambas regiones en las especies *Beta vulgaris*, *Chunia bucklandioides*, *Helianthus annuus*, *Solanum lycopersicum*, *Nicotina silvestris* y *Capsicum annuum*. Esto sugiere que la inserción del microstélite se dio después de la separación de las angiospermas y gimnospermas y antes de la divergencia entre las monocotiledoneas y las dicotiledoneas (hace aprox. 135- 130 Ma) (Wolfe *et al.*, 1989,

Magallón et al, 2015), mostrando que algunos microsatélites pueden ser mantenidos durante largos periodos evolutivos dentro del genoma mitocondrial de las plantas.

También hay que tomar en cuenta que algunos organismos del mismo género pueden mostrar variaciones significativas en cuanto al número de repeticiones por microsatélite (Nishikawa *et al.*, 2005; Ishii et la, 2006), por lo que sería prudente hacer estudios entre organismos más emparentados, para poder cuantificar la variación de las unidades de repetición entre microsatélites. De existir polimorfismos intraespecíficos que no estén siendo detectados, éstos serían incapaces de ser recuperados durante la reconstrucción filogenética por el método de parsimonia (Joy *et al.*, 2016). Además, la ausencia de estados en las terminales del árbol para la reconstrucción de los microsatélites (CT)<sub>n</sub> (ATAGT)<sub>n</sub> y (AG)<sub>n</sub>, podría presentar un sesgo en el proceso de reconstrucción, puesto que el método de no podría encontrar la distribución de los estados ancestrales sobre la filogenia si no están siendo observados todos los caracteres en las puntas de árbol (Joy *et al.*, 2016).

La repetición (ATAGT)<sub>n</sub> muestra un ejemplo interesante sobre la rápida evolución de un microsatélite dentro de un clado específico. Las especies utilizadas en este estudio mostraron un motif con un par de unidades de repetición, interrumpidos por un nucleótido - lo cual no es un microsatélite, por definición. Sin embargo, en el género *Zea* alcanza hasta 10 unidades de repetición. Al encontrar un aumento en el número de repeticiones, se compararon las secuencias de los organismos del género *Zea* con el resto del alineamiento, para saber si había diferencias significativas en las tasas de mutación de las regiones adyacentes al microsatélite. Los resultados mostraron que no había diferencias entre las tasas de mutación de los organismos del género *Zea*, puesto que presentan un valor de diversidad  $\pi$  igual a 0 (tomando en cuenta sólo las regiones vecinas a los microsatélites,

todas las regiones pertenecientes a organismos del género *Zea* resultaron ser el mismo haplotipo). Esto quizá se deba a que el aumento en las tasas de mutación en las regiones colindantes al microsatélite sea un proceso subsecuente, pero no inmediato, a la formación de éste. También cabe señalar que es posible que tanto la presencia del microsatélite y el aumento en la diversidad de las regiones vecinas sean resultado de un proceso subyacente dentro del genoma mitocondrial vegetal. Al observar el rango del tiempo de divergencia de los microsatélites (CT)<sub>n</sub> y (AG)<sub>n</sub> - de 200 a 250 Ma - vemos que ha transcurrido una mayor cantidad de tiempo, en comparación con las edades estimadas para la divergencia del género *Zea* (100,000–300,000 años (Clark *et al.*, 2005; Ross-Ibarra *et al.*, 2009), por lo que es posible que el efecto de las mutaciones aún no se haya presentado y derive de un proceso relacionado a la inserción de los microsatélites.

### **Presencia de alelos nulos en el genoma mitocondrial vegetal.**

Durante la amplificación de las regiones portadoras de microsatélites, no fue posible amplificar un producto para algunos organismos. Existen varias razones por las que pudo haberse presentado este fenómeno:

#### **1) Eventos de recombinación.**

Los abundantes eventos de recombinación dentro del genoma mitocondrial (Fauron *et al.*, 1995, Kubo & Newton, 2008; Kitaazaki & Kubo, 2010; Sloan *et al.*, 2012; Gualberto, 2014; Guo *et al.*, 2016) podrían haber generado interrupciones en alguna de las zonas de ligamiento de los cebadores, perdiendo así la posibilidad de ser obtenidos a través de las reacciones de amplificación. La recombinación entre elementos repetidos cortos puede generar inversiones que alteran el orden de la secuencia; por ejemplo, en regiones génicas pueden fragmentar la región codificante casi por completo, incluso entre miembros del mismo género (Palmer *et al.*, 2000). De manera similar, un evento de recombinación de esta naturaleza en una región de reconocimiento del cebador no permitiría la amplificación y posterior secuenciación de la región de interés. Haciendo una analogía, la ausencia o presencia de los microsatélites por recombinación tendría un patrón de distribución entre grupos taxonómicos similar al de algunos genes mitocondriales que se han transportado al núcleo. Como por ejemplo el gen *rps7*, que está presente sólo en algunos genomas mitocondriales, incluso entre especies cercanamente emparentadas (Palmer, 1998; Palmer *et al.*, 2000). A través de un muestreo extenso de organismos cercanamente emparentados sería posible evaluar la presencia de un grupo de microsatélites, y así conocer el efecto en la recombinación en periodos evolutivos cortos.

## 2) **Altas tasas de evolución en las regiones vecinas.**

La alteración en las tasas de evolución en las regiones flanqueantes del microsatélite podrían disminuir la astringencia (especificidad) de los cebadores con el molde de DNA, bajando el rendimiento de amplificación o, en los casos más extremos, evitando la especificidad total de los cebadores (Chapuis & Estoup, 2006). A pesar de que las tasas de mutación pueden aumentar hasta tres veces en regiones vecinas a microsatélites (este estudio), los resultados obtenidos señalan que la presencia del microsatélite no es suficiente para aumentar las tasas de mutación, sino que tiene que haber un rango mínimo de tiempo en el que éstos hayan habitado el genoma mitocondrial vegetal. Dos de los microsatélites de este estudio tienen un intervalo de inserción de 25 a 58.6 Ma, permitiendo la proliferación de nuevas variantes a través del aumento en las tasas de mutación. Sin embargo esto no ocurre en microsatélites de origen reciente (como en el caso del microsatélite (ATAGT)<sub>n</sub>). La asociación entre la presencia de alelos nulos y regiones flanqueantes altamente variables ha sido demostrada en otros estudios de biología molecular, y se ha sugerido que éstas son menos estables que otras regiones genómicas (Chapuis & Estoup, 2006). De ser el caso, es posible que, en algún punto de la historia evolutiva del grupo, el microsatélite deje de ser detectable en linajes particulares.

## 3) **Eventos de transferencia horizontal.**

El genoma mitocondrial de las plantas puede experimentar altas tasas de transferencia horizontal, generalmente entre mitocondrias de organismos con relación parásito-hospedero (Archibald & Ruchards, 2010) y en algunos casos, con especies distantes filogenéticamente (Won y Renner, 2003; Berthorsson *et al.*,



2003; Richardson & Palmer, 2006). Esta transferencia de DNA permite que un organismo presente regiones de DNA que no comparte con algún otro con el que esté cercanamente emparentado. Es posible que esta transferencia de DNA ocurra a través de un sistema activo de captura generalista, permitiendo la incorporación tanto de DNA nativo, como de origen externo (Richardson & Palmer, 2006; Archibald & Richards, 2010). Dada la elevada tasa de transferencia horizontal, puede darse que ciertas regiones de DNA se hayan transmitido entre los clados durante la evolución de las plantas. En el caso de las regiones intrónicas, el rastreo de los intrones muestra tanto las ganancias como las pérdidas en un contexto filogenético, lo que indica que la variación de los intrones es un evento común durante la evolución de las plantas. Si una región transferida de manera horizontal transporta un microsatélite y éste comienza su ciclo de desarrollo, es de esperar que se presente homoplásico en algunos clados, o en casos extremos como una autapomorfía.

## **Perspectivas.**

En este estudio se encontró que la presencia de los microsátélites en diversas regiones del genoma mitocondrial vegetal puede aumentar la diversidad nucleotídica en secuencias vecinas a éstos. Se encontró también que estos elementos pueden generarse diferencialmente entre sus distintas regiones del genoma mitocondrial, mantener una correlación con el tamaño del genoma, y ser conservados durante largos periodos evolutivos.

A través de las técnicas de secuenciación de nueva generación sería posible tener un muestreo mayor de mitogenomas completos para los pastos y otras plantas, lo cual permitiría un análisis más detallado sobre la presencia de los microsátélites dentro y entre los genomas mitocondriales. Estos análisis permitirían saber el grado de polimorfismos intraespecíficos presentes para saber qué tan rápido se van acumulando o perdiendo unidades de repetición entre especies cercanamente emparentadas. De igual manera, estas técnicas podrían permitir rastrear en qué momento de la historia evolutiva un microsátélite se ha insertado en un clado a través de transferencia horizontal, o se ha perdido de entre los linajes. Un mayor número de microsátélites caracterizados permitiría saber si el aumento en las tasas de diversidad es un fenómeno generalizado, o está restringido a regiones específicas del genoma -como las regiones intergénicas o los intrones-, o a grupos taxonómicos específicos.

Con el propósito de saber si el efecto del aumento en la tasas de diversidad es causado por la sola presencia del microsátélite, o por el tipo de genoma que lo alberga, valdría la pena realizar estudios similares utilizando genomas de cloroplasto. Se han reportado altos niveles de variabilidad intraespecífica en microsátélites dentro del genoma del cloroplasto (Provan *et al.*, 2001), y podría resultar interesante saber si esta variabilidad

es significativamente distinta entre ambos genomas. Conocer estos niveles de variabilidad entre los genomas permitirá estudiar el efecto de su presencia en las secuencias vecinas de organismos cercanamente emparentados, así como evaluar el efecto en la diversidad al considerar el tipo de arreglo y la longitud del microsatélite entre los distintos tipos de genomas. Por otro lado, a través de la comparación entre las tasas de diversidad asociadas a microsatélites, entre los genomas mitocondrial y de cloroplasto, podría conocerse si el aumento es debido solamente a la presencia de éstos, o resulta una característica específica del genoma mitocondrial vegetal.

## Referencias.

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410.
- Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., & Palmer, J. D. (2010). Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution*, 27(6), 1436-1448.
- Amos, W., Flint, J., & Xu, X. (2008). Heterozygosity increases microsatellite mutation rate, linking it to demographic history. *BMC genetics*, 9(1), 72.
- Archibald, J. M., & Richards, T. A. (2010). Gene transfer: anything goes in plant mitochondria. *BMC biology*, 8(1), 147.
- Bergthorsson, U., Adams, K. L., Thomason, B., & Palmer, J. D. (2003). Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*, 424(6945), 197.
- Bi, C., Paterson, A. H., Wang, X., Xu, Y., Wu, D., Qu, Y., ... & Ye, N. (2016). Analysis of the Complete Mitochondrial Genome Sequence of the Diploid Cotton *Gossypium raimondii* by Comparative Genomics Approaches. *BioMed Research International*, 2016.
- Brohede, J., & Ellegren, H. (1999). Microsatellite evolution: polarity of substitutions within repeats and neutrality of flanking sequences. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1421), 825-833.
- Brown, G. G., des Francs-Small, C. C., & Ostersetzer-Biran, O. (2014). Group II intron splicing factors in plant mitochondria. *Frontiers in plant science*, 5.
- Buschiazzo, E., & Gemmell, N. J. (2006). The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*, 28(10), 1040-1050.

- Bzymek, M., & Lovett, S. T. (2001). Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences*, 98(15), 8319-8325.
- Cardoso, A., Pereira, R., Fonseca, M., & Leitão, J. (2017). A microsatellite sequence in the fifth intron provides a broad-spectrum SSR marker for multiple alleles of the *er1/PsMLO1* powdery mildew resistance gene in *Pisum sativum* L. *Molecular Breeding*, 37(7), 84.
- Chapuis, M. P., & Estoup, A. (2006). Microsatellite null alleles and estimation of population differentiation. *Molecular biology and evolution*, 24(3), 621-631.
- Charlesworth, B., & Barton, N. (2004). Genome size: does bigger mean worse?. *Current Biology*, 14(6), R233-R235.
- Cheng, J., Zhao, Z., Li, B., Qin, C., Wu, Z., Trejo-Saavedra, D. L., ... & Hu, K. (2016). A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. *Scientific reports*, 6, 18919.
- Cho, Y., Mower, J. P., Qiu, Y. L., & Palmer, J. D. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51), 17741-17746.
- Cho, Y., Mower, J. P., Qiu, Y. L., & Palmer, J. D. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51), 17741-17746.
- Parkinson, C. L., Mower, J. P., Qiu, Y. L., Shirk, A. J., Song, K., Young, N. D., & Palmer, J. D. (2005). Multiple major increases and decreases in

mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evolutionary Biology*, 5(1), 73.

· Clark, R. M., Tavaré, S., & Doebley, J. (2005). Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus. *Molecular biology and evolution*, 22(11), 2304-2312.

· Clifton, S. W., Minx, P., Fauron, C. M. R., Gibson, M., Allen, J. O., Sun, H., ... & Meyer, L. (2004). Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant physiology*, 136(3), 3486-3503.

· Condamine, F. L., Nagalingum, N. S., Marshall, C. R., & Morlon, H. (2015). Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC evolutionary biology*, 15(1), 65.

· Cupp, J. D., & Nielsen, B. L. (2014). Minireview: DNA replication in plant mitochondria. *Mitochondrion*, 19, 231-237.

· Darracq, A., Varré, J. S., Maréchal-Drouard, L., Courseaux, A., Castric, V., Saumitou-Laprade, P., ... & Touzet, P. (2011). Structural and content diversity of mitochondrial genome in beet: a comparative genomic analysis. *Genome biology and evolution*, 3, 723-736.

· Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8), 772.

· Davila, J. I., Arrieta-Montiel, M. P., Wamboldt, Y., Cao, J., Hagemann, J., Shedge, V., ... & Mackenzie, S. A. (2011). Double-strand break repair processes drive evolution of the mitochondrial genome in *Arabidopsis*. *BMC biology*, 9(1), 64

- Deng, W., Maust, B. S., Nickle, D. C., Learn, G. H., Liu, Y., Heath, L., ... & Mullins, J. I. (2010). DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques*, 48(5), 405.
- Eckardt, N.A. (2008) Grass Genome Evolution. *The plant cell*, 20(1), 3-4
- Edwards, E., Aliscioni, S., & Bell, H. (2011). New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist*, 193, 304-312.
- Edwards, E., Aliscioni, S., & Bell, H. (2011). New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist*, 193, 304-312.
- Farris, J. S. (1989). The retention index and the rescaled consistency index. *Cladistics*, 5(4), 417-419.
- Fauron, C. M. R., Moore, B., & Casper, M. (1995). Maize as a model of higher plant mitochondrial genome plasticity. *Plant Science*, 112(1), 11-32.
- Freeling, M. (2001) Grasses as a Single Genetic System. Reassessment 2001, *Plant Physiology*, 125(3), 1191-1197.
- Galtier, N. (2011). The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC biology*, 9(1), 61.
- Gaut B. S. (2002) Evolutionary dynamics of grass genomes. *New phytologist*, 154(1), 15-28
- Giussani, L. M., Cota-Sanchez, J. H., Zuloaga, F. O., & Kellogg, E. A. (2001). A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C4 photosynthesis. *American Journal of Botany*, 88(11), 1993-2012.
- Glémin, S., & Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. *New phytologist*, 183(2), 273-290.

- Godbout, J., JARAMILLO-CORREA, J. P., Beaulieu, J., & Bousquet, J. (2005). A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, 14(11), 3497-3512.
- Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., ... & Hadley, D. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*, 296(5565)
- Gray, M. W. (2012). Mitochondrial evolution. *Cold Spring Harbor perspectives in biology*, 4(9), a011403.
- Gualberto, J. M., Mileshina, D., Wallet, C., Niazi, A. K., Weber-Lotfi, F., & Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. *Biochimie*, 100, 107-120.
- Guo, W., Grewe, F., Fan, W., Young, G. J., Knoop, V., Palmer, J. D., & Mower, J. P. (2016). Ginkgo and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. *Molecular biology and evolution*, msw024.
- Guo, W., Zhu, A., Fan, W., & Mower, J. P. (2017). Complete mitochondrial genomes from the ferns *Ophioglossum californicum* and *Psilotum nudum* are highly repetitive with the largest organellar introns. *New Phytologist*, 213(1), 391-403.
- Hecht, J., Grewe, F., & Knoop, V. (2011). Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biology and Evolution*, 3, 344-358.
- Hecht, J., Grewe, F., & Knoop, V. (2011). Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria:



the root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biology and Evolution*, 3, 344-358.

· Hedrick, P. W. (2011). *Genetics of populations*. Jones & Bartlett Learning.

· Hilu, K. W., Alice, L. A., & Liang, H. (1999). Phylogeny of Poaceae inferred from matK sequences. *Annals of the Missouri Botanical Garden*, 835-851.

· Hisano, H., Tsujimura, M., Yoshida, H., Terachi, T., & Sato, K. (2016). Mitochondrial genome sequences from wild and cultivated barley (*Hordeum vulgare*). *BMC genomics*, 17(1), 824.

· Holá, D., Benešová, M., Fischer, L., Haisel, D., Hnilička, F., Hniličková, H., ... & Tůmová, L. (2017). The disadvantages of being a hybrid during drought: A combined analysis of plant morphology, physiology and leaf proteome in maize. *PloS one*, 12(4), e0176121.

· Ishii, T., Takahashi, C., Ikeda, N., Kamijima, O., & Mori, N. (2006). Mitochondrial microsatellite variability in common wheat and its ancestral species. *Genes & genetic systems*, 81(3), 211-214.

· Jaramillo-Correa, J. P., Aguirre-Planter, E., Eguiarte, L. E., Khasa, D. P., & Bousquet, J. (2013). Evolution of an ancient microsatellite hotspot in the conifer mitochondrial genome and comparison with other plants. *Journal of molecular evolution*, 76(3), 146-157.

· Joy, J. B., Liang, R. H., McCloskey, R. M., Nguyen, T., & Poon, A. F. (2016). Ancestral reconstruction. *PLoS Comput Biol*, 12(7), e1004763.

· Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.

- Kimura, M. (1983). The neutral theory of molecular evolution. Cambridge University Press.
- Kitazaki, K., & Kubo, T. (2010). Cost of having the largest mitochondrial genome: evolutionary mechanism of plant mitochondrial genome. *Journal of Botany*, 2010.
- Kitazaki, K., & Kubo, T. (2010). Cost of having the largest mitochondrial genome: evolutionary mechanism of plant mitochondrial genome. *Journal of Botany*, 2010.
- Kubo, T., & Newton, K. J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion*, 8(1), 5-14.
- Laroche, J., Li, P., Maggia, L., & Bousquet, J. (1997). Molecular evolution of angiosperm mitochondrial introns and exons. *Proceedings of the National Academy of Sciences*, 94(11), 5722-5727.
- Lai, Y., & Sun, F. (2003). The relationship between microsatellite slippage mutation rate and the number of repeat units. *Molecular biology and evolution*, 20(12), 2123-2131.
- Levings, C. S., & Brown, G. G. (1989). Molecular biology of plant mitochondria. *Cell*, 56(2), 171-179.
- Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11), 1451-1452.
- Lilly, J. W., & Havey, M. J. (2001). Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. *Genetics*, 159(1), 317-328.
- Lynch, M., & Conery, J. S. (2003). The origins of genome complexity. *science*, 302(5649), 1401-1404.
- Lynch, M., Koskella, B., & Schaack, S. (2006). Mutation pressure and the evolution of organelle genomic architecture. *Science*, 311(5768), 1727-1730.

- Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L., & Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist*, 207(2), 437-453.
- Maréchal, A., & Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytologist*, 186(2), 299-317.
- Masters, B. S., Johnson, L. S., Johnson, B. G., Brubaker, J. L., Sakaluk, S. K., & Thompson, C. F. (2010). Evidence for heterozygote instability in microsatellite loci in house wrens. *Biology letters*, rsbl20100643.
- Nei, M. (1987). *Molecular evolutionary genetics*.
- Nishikawa, T., Vaughan, D. A., & Kadowaki, K. I. (2005). Phylogenetic analysis of *Oryza* species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. *Theoretical and Applied Genetics*, 110(4), 696-705.
- Noyszewski, A. K., Ghavami, F., Alnemer, L. M., Soltani, A., Gu, Y. Q., Huo, N., ... & Kianian, S. F. (2014). Accelerated evolution of the mitochondrial genome in an alloplasmic line of durum wheat. *BMC genomics*, 15(1), 67.
- Okonechnikov K, Golosova O, Fursov M, the UGENE team. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 2012 28: 1166-1167. doi:10.1093/bioinformatics/bts091
- Oldenburg, D. J., Kumar, R. A., & Bendich, A. J. (2013). The amount and integrity of mtDNA in maize decline with development. *Planta*, 237(2), 603-617.
- Palmer, J. D., & Herbon, L. A. (1988). Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *Journal of Molecular evolution*, 28(1), 87-97.

- Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y. L., & Song, K. (2000). Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National Academy of Sciences*, 97(13), 6960-6966.
- Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y. L., & Song, K. (2000). Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National*
- Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y. L., & Song, K. (2000). Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National Academy of Sciences*, 97(13), 6960-6966.
- Parkinson, C. L., Mower, J. P., Qiu, Y. L., Shirk, A. J., Song, K., Young, N. D., & Palmer, J. D. (2005). Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evolutionary Biology*, 5(1), 73.
- Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology*, 53(5), 793-808
- Provan, J., Powell, W., & Hollingsworth, P. M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, 16(3), 142-147.
- Rajendrakumar, P., Biswal, A. K., Balachandran, S. M., Srinivasarao, K., & Sundaram, R. M. (2007). Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. *Bioinformatics*, 23(1), 1-4.

- Richardson, A. O., & Palmer, J. D. (2006). Horizontal gene transfer in plants. *Journal of experimental botany*, 58(1), 1-9.
- Richardson, A. O., Rice, D. W., Young, G. J., Alverson, A. J., & Palmer, J. D. (2013). The “fossilized” mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC biology*, 11(1), 29.
- Ross-Ibarra, J., Tenaillon, M., & Gaut, B. S. (2009). Historical divergence and gene flow in the genus *Zea*. *Genetics*, 181(4), 1399-1413.
- Salse, J., Bolot, S., Throude, M., Jouffe, V., Plegu, B., Quaraishi, U. M., ... & Feuillet, C. (2008) Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *The plant cell*, 20(1) 11-24
- Sang, T. (2009). Genes and mutations underlying domestication transitions in grasses. *Plant Physiology*, 149(1), 63-70.
- Scherer, S. (2008). short guide to the human genome. Cold Spring Harbor Laboratory Press.
- Schlötterer, C. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, 109(6), 365-371.
- Sena, J. S., Giguère, I., Boyle, B., Rigault, P., Birol, I., Zuccolo, A., ... & Bousquet, J. (2014). Evolution of gene structure in the conifer *Picea glauca*: a comparative analysis of the impact of intron size. *BMC plant biology*, 14(1), 95.
- Shankarappa, R. A. J., Margolick, J. B., Gange, S. J., Rodrigo, A. G., Upchurch, D., Farzadegan, H., ... & Huang, X. L. (1999). Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *Journal of virology*, 73(12), 10489-10502.

- Skippington, E., Barkman, T. J., Rice, D. W., & Palmer, J. D. (2015). Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all nad genes. *Proceedings of the National Academy of Sciences*, 112(27), E3515-E3524.
- Sloan, D. B. (2013). One ring to rule them all? Genome sequencing provides new insights into the ‘master circle’ model of plant mitochondrial DNA structure. *New Phytologist*, 200(4), 978-985.
- Sloan, D. B., Alverson, A. J., Chuckalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., & Taylor, D. R. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol*, 10(1), e1001241.
- Sloan, D. B., Alverson, A. J., Chuckalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., & Taylor, D. R. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS biology*, 10(1), e1001241.
- Sloan, D. B., Barr, C. M., Olson, M. S., Keller, S. R., & Taylor, D. R. (2007). Evolutionary rate variation at multiple levels of biological organization in plant mitochondrial DNA. *Molecular Biology and Evolution*, 25(2), 243-246.
- Sloan, D. B., Barr, C. M., Olson, M. S., Keller, S. R., & Taylor, D. R. (2007). Evolutionary rate variation at multiple levels of biological organization in plant mitochondrial DNA. *Molecular Biology and Evolution*, 25(2), 243-246.
- Sloan, D. B., Oxelman, B., Rautenberg, A., & Taylor, D. R. (2009). Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. *BMC Evolutionary Biology*, 9(1), 260.

- Subramanian, S., Mishra, R. K., & Singh, L. (2003). Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome biology*, 4(2), R13.
- Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, 30(14), 3059-3066.
- Tajima, F., Misawa, K., & Innan, H. (1998). The amount and pattern of DNA polymorphism under the neutral mutation hypothesis. In *Mutation and Evolution* (pp. 103-107). Springer Netherlands
- Tomilin, N. V. (2008). Regulation of mammalian gene expression by retroelements and non-coding tandem repeats. *Bioessays*, 30(4), 338-348.
- Tóth, G., Gáspári, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome research*, 10(7), 967-981.
- Tsunewaki, K. (2011). Interorganellar DNA transfer in wheat: dynamics and phylogenetic origin. *Proceedings of the Japan Academy, Series B*, 87(8), 529-549.
- Vázquez-Lobo, A. (1996). Evolución de hongos endófitos del género *Pinus* (Doctoral dissertation, Tesis de licenciatura, Facultad de Ciencias, Universidad Nacional Autónoma de México, México, DF)
- Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. D. F. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics and molecular biology*, (AHEAD), 0-0.
- Wang, M. L., Barkley, N. A., & Jenkins, T. M. (2009). Microsatellite markers in plants and insects. Part I: Applications of biotechnology. *Genes, Genomes and Genomics*, 3(1), 54-67.

- Wang, M. L., Barkley, N. A., & Jenkins, T. M. (2009). Microsatellite markers in plants and insects. Part I: Applications of biotechnology. *Genes, genomes and genomics*, 3(1), 54-67.
- Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical population biology*, 7(2), 256-276
- Wierdl, M., Dominska, M., & Petes, T. D. (1997). Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*, 146(3), 769-779
- Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M., & Li, W. H. (1989). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proceedings of the National Academy of Sciences*, 86(16), 6201-6205.
- Woloszynska, M. (2009). Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes—though this be madness, yet there's method in't. *Journal of experimental botany*, erp361.
- Won, H., & Renner, S. S. (2003). Horizontal gene transfer from flowering plants to *Gnetum*. *Proceedings of the National Academy of Sciences*, 100(19), 10824-10829.
- Yang, J., Zhu, Y., Tong, Y. I., Zhang, Z., Chen, L., Chen, S., ... & Ma, X. (2009). The novel G10680A mutation is associated with complete penetrance of the LHON/T14484C family. *Mitochondrion*, 9(4), 273-278.
- Yang, Z. (2006). *Computational molecular evolution*. Oxford University Press
- Yang, Z. (2006). *Computational molecular evolution*. Oxford University Press.
- Yang, Z., & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 13(5), 303-314.