



**UNIVERSIDAD DE BUENOS AIRES**  
Facultad de Ciencias Exactas y Naturales  
Departamento de Ecología, Genética y Evolución

**Estudios en expresión génica diferencial y genómica  
comparativa en yerba mate y yerba señorita**

Tesis presentada para optar al título de  
Doctora de la Universidad de Buenos Aires en el área Ciencias Biológicas

Lic. Jimena Cascales

**Directora de Tesis:** Dra. Alexandra M. Gottlieb

**Consejera de Estudios:** Dra. Viviana A. Confalonieri

Laboratorio de Citogenética y Evolución, IEGEBA (UBA-CONICET)

Ciudad Autónoma de Buenos Aires, 2018

Fecha de defensa: 28 de marzo de 2018

## RESUMEN

### Estudios en expresión génica diferencial y genómica comparativa en yerba mate y yerba señorita

*Ilex paraguariensis* (la yerba mate) e *I. dumosa* (la yerba señorita) son árboles dioicos nativos del sur de Sudamérica. La yerba mate posee enorme relevancia cultural y socio-económica para nuestro país. La yerba señorita exhibe características biológicas que la hacen atractiva para el mejoramiento de la yerba mate. En esta Tesis se propusieron dos objetivos principales. El primero involucra la secuenciación y caracterización completa de los plastomas. Esta información será de gran utilidad para aplicar tecnologías de mejoramiento modernas, y para proponer marcadores moleculares especie-específicos realmente variables. El segundo objetivo refiere al estudio comparativo de la epigenética y la expresión génica diferencial, a nivel inter- e intraespecífico, y entre distintos órganos y sexos. Además, se propone la identificación de genes o regiones con expresión o regulación diferencial.

Los plastomas obtenidos mostraron estructura conservada, 99% de identidad nucleotídica y sintenia. Se desarrollaron 30 marcadores microsatélite y seleccionaron 16 regiones de utilidad filogenética comprobada. Los perfiles epigenéticos de las muestras evidenciaron agrupamientos de acuerdo a la planta donante, basada en una correlación positiva significativa entre la variación genética y la epigenética. Los patrones de expresión génica diferencial tendieron a agrupar las muestras según el tipo de órgano considerado. No se encontraron evidencias de asociación con mecanismos de determinación sexual. Se identificaron 30 fragmentos diferenciales con funciones metabólicas, de biosíntesis de compuestos, regulación de la expresión, defensa y resistencia a estrés biótico y abiótico, entre otras.

**Palabras clave:** *Ilex*, yerba mate, yerba señorita, plastomas, filogenómica, epigenética, expresión génica diferencial.

## ABSTRACT

### **Studies in differential gene expression and comparative genomics in 'yerba mate' and 'yerba señorita'**

*Ilex paraguariensis* ('yerba mate') and *Ilex dumosa* ('yerba señorita') are dioecious trees native to southern South America. The Yerba mate has enormous cultural and socio-economic relevance for our country, whereas the Yerba señorita plants exhibit biological characteristics that make it attractive for the improvement of yerba mate. In this Thesis, two main objectives were proposed. The first involves the sequencing and the complete characterization of the species plastomes. This information will be useful to apply modern breeding technologies, and for the proposal of species-specific truly variable molecular markers. The second objective refers to the comparative study of epigenetics and differential gene expression at inter- and intraspecific levels, and between different organs and sexes. In addition, the identification of genes or regions with differential expression or regulation is proposed.

The species plastomes obtained herein showed conserved structure, 99% nucleotide identity and synteny. Thirty microsatellite markers were developed and 16 regions of proven phylogenetic utility were selected as phylogenetic markers. The epigenetic profiles of the samples evidenced groupings according to the donor plant, based on a significant positive correlation between the genetic variation and that from epigenetics. The patterns of differential gene expression tended to group the samples according to the type of organ considered. No evidence of association with mechanisms of sexual determination was found. Thirty differential fragments were identified showing various functions such as, metabolic, biosynthesis of compounds, regulation of gene expression, defense and resistance to biotic and abiotic stresses, among others.

**Key words:** *Ilex*, yerba mate, yerba señorita, plastomes, phylogenomics, epigenetics, gene expression.

## AGRADECIMIENTOS

Agradezco a toda la gente que me ayudó y acompañó durante todos estos años.

En primer lugar, a mi Directora, Alex, por su paciencia, por estar siempre presente, por sus consejos, y por haberme enseñado desde mis primeros pasos en esta profesión, siempre apuntando a mi futuro académico. Estoy feliz por poder seguir trabajando con vos, te admiro como directora y persona.

A la Dra. Lidia Poggio por darme un lugar en el laboratorio, y por confiar en mí para seguir mi vida profesional allí.

Al CONICET, a la Universidad de Buenos Aires y a la Agencia Nacional de Promoción Científica y Tecnológica por los subsidios otorgados a las Dras. AM Gottlieb y L Poggio, que permitieron el desarrollo de esta de Tesis, y al CONICET en particular por otorgarme las Becas de Doctorado y de Postdoctorado para seguir formándome en esta carrera tan linda.

A la Facultad de Ciencias Exactas y Naturales, por apostar por los científicos, y darnos lugar de trabajo. A la comunidad científica de Exactas, por su buena predisposición cuando pedí ayuda.

A toda la gente que me ayudó en el desarrollo de esta Tesis. Al Dr. PA Sansberro y al Dr. RA Scherer (Pindó SA), por enviarnos tan gentilmente muestras. Al personal del Jardín Botánico de Buenos Aires, en especial a los Ing. Agr. F Cano y G Benito, por recibirme siempre tan amablemente, e incluso compartir mates y charla. A la Dra. A Guberman y su grupo, por permitirme usar las instalaciones de su laboratorio. Al Dr. D de Pannis, por ayudarme a superar obstáculos en el análisis de datos. Al Dr. Ing. Agr. GC Giberti, que durante todos estos años contribuyó con nuestro trabajo en *Ilex*.

A mis compañeros, con los que compartí tantas tardes de mates, salidas, charlas y risas. Gracias por todos sus consejos. A Mari Bracco, que a pesar de estar lejos, me dejó muchas enseñanzas.

A mis amigos de la vida, por acompañarme y entenderme siempre, aconsejarme, y mantenerse firmes conmigo.

A Damí, por haber estado siempre conmigo y entenderme, en los buenos momentos y también en los difíciles. Gracias por haberme acompañado y apoyado desde el inicio de la carrera, y en especial por ser mi sostén estos últimos meses. Sin tu ayuda hubiera sido muy difícil llegar hasta acá, gracias por tanto amor y por hacerme feliz.

A toda mi familia, por preocuparse siempre por mí, e interesarse en mi carrera. Muy especialmente a mis papás y a mi hermano, que siempre fomentaron mi carrera y me ayudaron en todo lo que necesité. Gracias por estar presentes en todo momento, no tengo palabras para dimensionar todo lo que los quiero.

# ÍNDICE

<b>INTRODUCCIÓN.....</b>	<b>1</b>
<i>Características morfológicas .....</i>	1
<i>Distribución actual e histórica .....</i>	5
<i>Fitoquímica.....</i>	9
<i>Historia del consumo de la yerba mate: etnobotánica.....</i>	12
<i>Historia de la producción de yerba mate .....</i>	13
<i>Aspectos productivos .....</i>	14
<i>Antecedentes .....</i>	16
<b>OBJETIVOS GENERALES .....</b>	<b>21</b>
<b>CAPÍTULO 1.....</b>	<b>22</b>
Introducción.....	23
<i>Estructura de los plastomas .....</i>	23
<i>Contenido génico de los plastomas .....</i>	25
<i>Las regiones plastídicas como marcadores moleculares .....</i>	26
<i>Los plastomas como herramientas biotecnológicas .....</i>	28
Objetivos.....	29
Hipótesis .....	29
Predicciones.....	29
Materiales y Métodos.....	30
1. <i>Genotipado preliminar .....</i>	30
2. <i>Secuenciación de plastomas .....</i>	32
3. <i>Análisis de los plastomas .....</i>	35
Resultados.....	39
1. <i>Genotipado preliminar .....</i>	39
2. <i>Obtención de plastomas.....</i>	39
3. <i>Análisis de plastomas.....</i>	44
Discusión .....	57
<b>CAPÍTULO 2.....</b>	<b>62</b>
Introducción.....	63
<i>Orígenes del estudio de la epigenética.....</i>	63
<i>Relación entre epigenética, fenotipo y ambiente .....</i>	64
<i>Procesos epigenéticos .....</i>	66
<i>Herencia epigenética celular y transgeneracional.....</i>	67
<i>Metilación del ADN .....</i>	68
<i>Aproximación metodológica a la epigenética y la expresión génica diferencial, marco teórico.....</i>	72
Objetivos.....	76
Hipótesis .....	77

Predicciones.....	77
Materiales y Métodos.....	78
1. <i>Análisis genético</i> .....	79
2. <i>Análisis epigenético</i> .....	80
3. <i>Expresión génica diferencial</i> .....	88
Resultados.....	92
1. <i>Caracterización genética</i> .....	92
2. <i>Análisis de variación epigenética</i> .....	93
3. <i>Evaluación de la variación en la expresión génica diferencial</i> .....	116
Discusión.....	128
<b>CONCLUSIONES FINALES.....</b>	<b>141</b>
<b>PERSPECTIVAS .....</b>	<b>143</b>
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>144</b>
<b>APÉNDICES.....</b>	<b>157</b>
<i>Apéndice I – Tablas y figuras suplementarias</i> .....	157
<i>Apéndice II – Soluciones y protocolos</i> .....	180
<i>Soluciones utilizadas</i> .....	180
<i>Protocolos</i> .....	181

## INTRODUCCIÓN

El género *Ilex* (Aquifoliaceae) comprende aproximadamente 600 especies de arbustos o árboles, la mayoría de los cuales son perennes (Giberti 2001; Loizeau *et al.* 2005, 2016). Varían desde plantas rastreras hasta árboles de 15-30 metros de altura, habiendo también arbustos epífitos (Galle 1997). Muchas especies poseen valor ornamental y paisajístico por su morfología foliar, sus características de follaje y el color contrastante de sus frutos (por ej., *I. aquifolium* -conocida como acebo-, *I. cornuta* e *I. crenata*; **Figura 1**). Además, desde un punto de vista cultural, ciertas especies son asociadas con festividades como la Navidad o el Año Nuevo chino (por ej., *I. aquifolium*, *I. opaca*, *I. verticillata*, e *I. purpurea*). Es sabido que diversos grupos étnicos han utilizado las especies de *Ilex* en rituales o en medicina tradicional; hoy en día se conoce que el variado contenido de metabolitos secundarios les otorga propiedades medicinales y nutraceuticas (Keller 2008; Keller y Giberti 2011; Loizeau *et al.* 2016). En nuestra región se destaca por su enorme valor socio-cultural, económico y como alimento funcional la especie *I. paraguariensis*, más comúnmente conocida como el árbol de la yerba mate (Keller 2008; Filip 2011; Keller y Giberti 2011).

La familia Aquifoliaceae está bien situada dentro del orden Aquifoliales junto con Helwingiaceae y Phyllonomaceae. Las tres familias mencionadas son monogénicas, abarcando los géneros *Ilex*, *Helwingia* y *Phyllonoma*, respectivamente (Loizeau *et al.* 2016; Stevens 2017).

### Características morfológicas

La morfología foliar del género *Ilex* presenta gran variación interespecífica (**Figura 1**). El tamaño, la forma, y la textura de las hojas han sido características muy utilizadas en la clasificación tradicional de estas especies, aunque en algunos casos condujeron a confusiones taxonómicas (Galle 1997). El tamaño de las hojas puede ser de 5-15 mm hasta alcanzar más de 20 cm. Usualmente las hojas de las especies perennes son elípticas a oblongo-elípticas, mientras que las de las especies deciduas son ovadas u ovadas-elípticas. Aproximadamente el 60% de las especies tienen hojas con márgenes espinosos o aserrados, en grado variable (Galle 1997). Por ejemplo, *I. aquifolium* (**Figura 1A-B**) e *I. cornuta* (**Figura 1E**) presentan márgenes espinosos, y la última especie posee hojas sub-cuadrangulares a cuadrangular-oblongas. Es conocido que la edad y las condiciones ambientales pueden afectar el tamaño y la forma de las hojas. La

## Introducción

textura también presenta gran variación entre las especies: se reconocen ejemplos de hojas coriáceas (*I. dumosa*, *I. paraguariensis*), cartáceas o papiráceas (*I. hookeri*), glabras (*I. aquifolium*, *I. latifolia*), pubescentes (*I. brasiliensis*, *I. theezans*), rugosas (*I. intricata*) y lisas (*I. canariensis*), etc. (Galle 1997).



**Figura 1.** Variedad en la morfología foliar del género *Ilex* y ejemplos de especies ornamentales. **A.** *I. aquifolium*; **B.** *I. aquifolium* variegado (cultivado); **C.** *I. paraguariensis*; **D.** *I. latifolia*; **E.** *I. cornuta*; **F.** *I. crenata*; **G.** *I. dumosa*; **H.** *I. verticillata* (decidua).

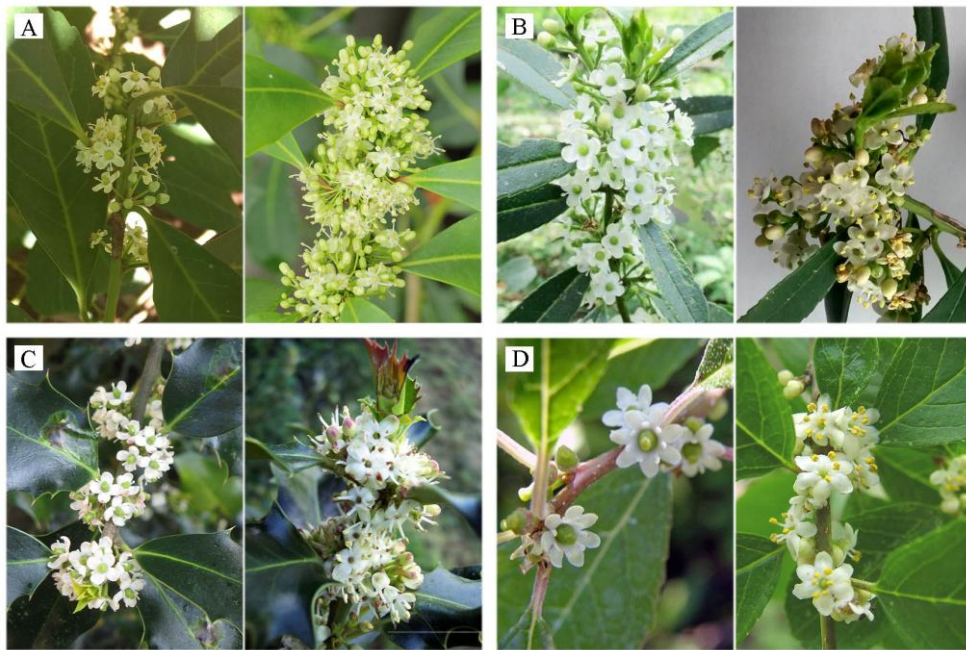
En contraposición, la morfología floral es bastante más conservada entre las especies (**Figura 2**). Por lo general, las flores son pequeñas (5-6 mm de diámetro), actinomorfas (con más de un plano de simetría), simpétalas (pétalos soldados), de 4-5 pétalos blancos (aunque existen de tonalidades rojizas a rosadas) de 3-10 mm, los estambres son libres y los carpelos están fusionados. Se encuentran agrupadas en inflorescencias cimosas (**Figura 3A**) en general del tipo dicasio (Giberti 2001; Giberti y Gurni 2008), aunque existen discrepancias entre los especialistas (Galle 1997; Loizeau *et al.* 2016).

Las especies de *Ilex* son funcionalmente dioicas, es decir que presentan flores unisexuales en individuos separados pero con rudimentos de órganos del sexo opuesto.

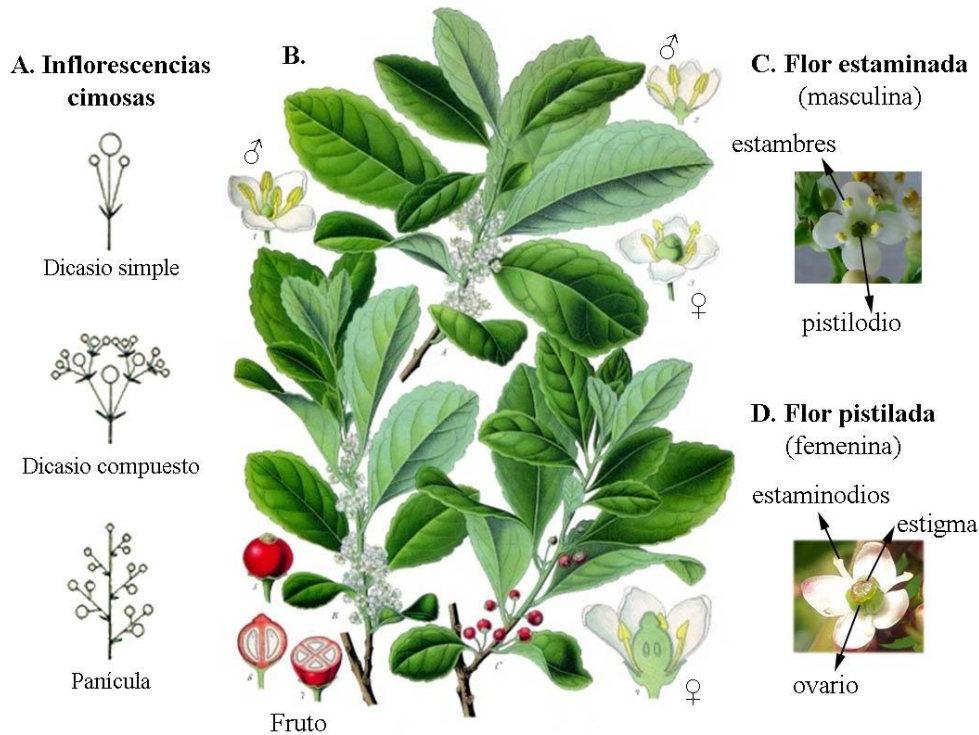


## Introducción

Así, las flores estaminadas (de ahora en más, para facilitar la lectura serán nombradas “masculinas”) poseen gineceos no funcionales llamados pistilodios, y las flores pistiladas (“femeninas”) poseen estambres estériles denominados estaminodios (**Figura 3B-D**; Giberti y Gurni 2008). En la mayoría de las especies las inflorescencias muestran dimorfismo sexual: las plantas productoras de flores masculinas poseen inflorescencias más prolíficas que las productoras de flores femeninas, debido a un mayor número de ramificaciones (Galle 1997).



**Figura 2.** Flores de especies del género *Ilex*. **A.** *I. paraguariensis*; **B.** *I. dumosa*; **C.** *I. aquifolium*; **D.** *I. verticillata*. En cada caso, la primera foto muestra flores pistiladas y la segunda flores estaminadas.



**Figura 3.** Tipos de inflorescencia y detalle de las estructuras funcionales y abortivas de las flores del género *Ilex*, y. **A.** Esquemas de tipos de inflorescencia cimosas encontradas en el género. **B.** Ilustración botánica de *I. paraguariensis* de F. E. Köhler (*Köhler's Medicinal-Pflanzen*, 1887), mostrando esquemas de las flores de ambos sexos y del pireno. **C.** Flor estaminada. **D.** Flor pistilada.

Las flores estaminadas poseen un androceo compuesto por estambres adheridos a la corola por las bases de los filamentos y alternados con los pétalos, presentando el mismo número de estambres y pétalos (**Figura 3B-C**). Los pistilodios son ovarios abortados casi siempre glabros, pequeños (Galle 1997; Giberti 2001). Las flores pistiladas tienen por lo general un gineceo pequeño de < 2 mm de altura, sésil, en el cual el ovario es usualmente ovoide, súpero, glabro, con un estigma discoide y lobulado o mamiforme (**Figura 3B, D**). Las anteras de los estaminodios no se desarrollan por completo y los sacos polínicos usualmente no están presentes o están vacíos (Galle 1997; Giberti 2001).

Una peculiaridad de estas flores es el desprendimiento de las corolas completas y el androceo adnato, después de la antesis. Esta característica colabora al hallazgo a campo de los árboles del género en época de floración, ya que las corolas desprendidas halladas en el suelo indican la presencia de individuos de *Ilex* en las cercanías (Giberti y Gurni 2008).

## *Introducción*

En estas plantas la polinización es entomófila, sin especificidad de polinizador, pudiendo ser mediada por hormigas, abejas, moscas, avispas, mariposas y escarabajos atraídos por el néctar producido en la base de los pétalos (Ferreira *et al.* 1983; Giberti 2001; Loizeau *et al.* 2016).

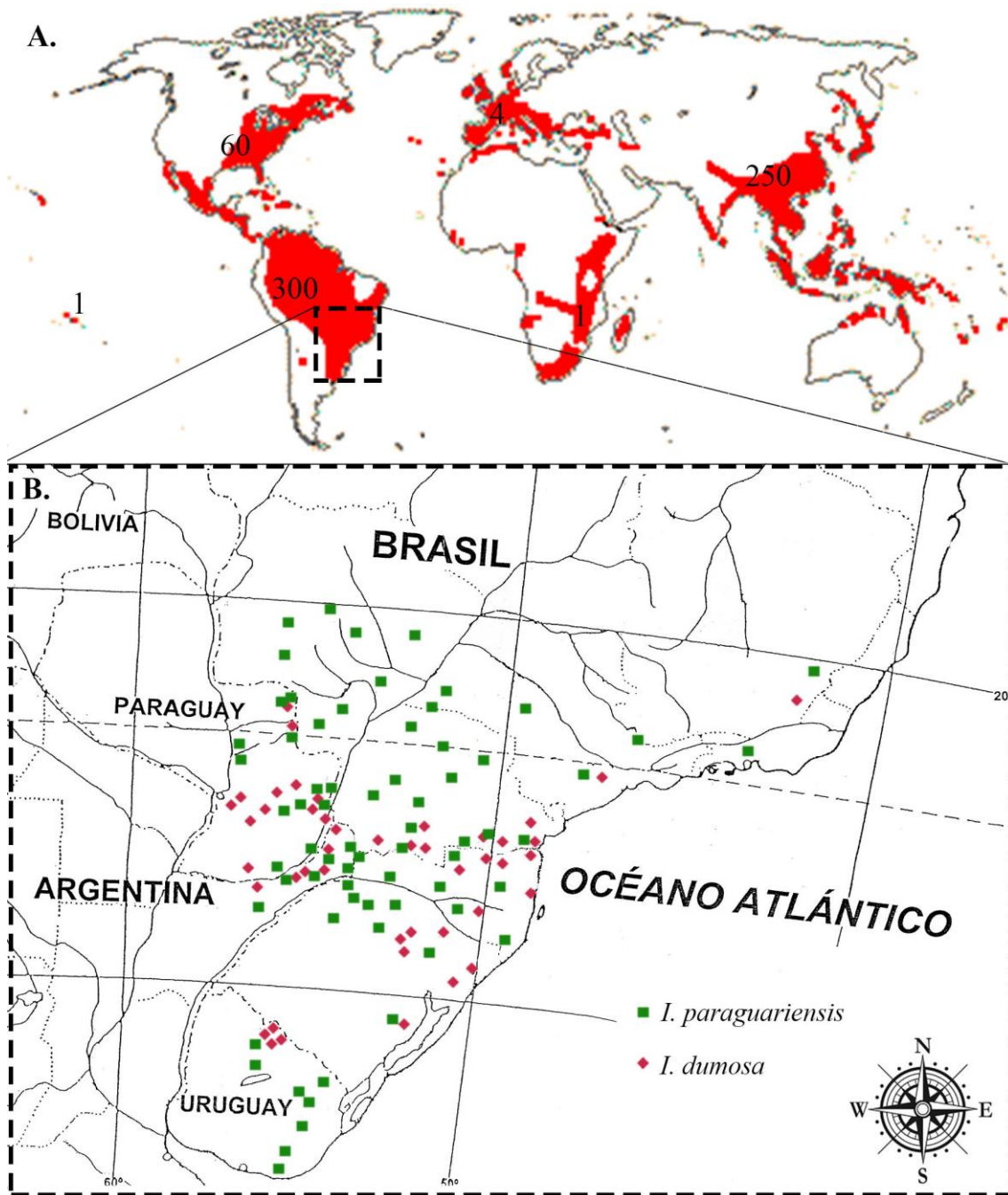
Los frutos, con cáliz y estigma persistentes, son del tipo drupa (**Figura 3B**). El epicarpo suele ser delgado y de color rojizo, resultando atractivo para las aves que actúan como agentes dispersivos. El mesocarpo es más desarrollado, blando y pulposo. El endocarpo esclerificado rodea completamente a la semilla en cada uno de los cuatro lóculos, constituyéndose en una unidad de propagación conocida botánicamente como pireno (**Figura 3B**; Giberti 2001).

### **Distribución actual e histórica**

El género *Ilex* es cosmopolita. Se encuentra en todas las áreas tropicales del mundo y se extiende en regiones templadas hasta los 63° latitud Norte (América, Eurasia) y los 35° latitud Sur (América, África), abarcando desde bosques de tierras bajas hasta zonas montañosas, a una altura máxima de 4000 m en los Andes (Loizeau *et al.* 2016). Se han documentado dos centros de diversificación: uno en las regiones tropicales de América (con aprox. 300 especies), y otro en el sudeste asiático (con aprox. 250 especies) (**Figura 4A**; Loizeau *et al.* 2005). Según un modelado, la distribución actual del género estaría asociada fundamentalmente a los niveles de precipitación, y en menor medida a la temperatura del mes más frío y al número de días de helada (Loizeau *et al.* 2005). El patrón de distribución mundial es disyunto (**Figura 4A**), probablemente como resultado de eventos de temperaturas frías durante el Mioceno y la glaciación del Cuaternario, que no habrían extinguido completamente a las especies sino que las habrían aislado en refugios. Esto pudo crear las condiciones para la especiación alopátrica, pero manteniendo el potencial de hibridación interespecífica (Loizeau *et al.* 2005). Se postula que al finalizar las condiciones climáticas adversas pudo haber intercambios florísticos entre el sudeste de Asia y Norteamérica a través del estrecho de Bering; se conoce que el género ya estaba presente en Sudamérica antes de la formación de Centroamérica (Loizeau *et al.* 2005). Los registros fósiles más antiguos son de granos de polen (muy característicos del género) de 90 millones de años hallados en Australia (Martin 1977); mientras que los encontrados en África, el oeste de América del Norte y América del Sur indican que el género ya era cosmopolita en el Cretácico tardío, hace 70-85 millones de años (Loizeau *et al.* 2005 y ref. allí citadas).

## Introducción

Particularmente en el Cono Sur de Sudamérica existen pocos representantes del género; entre 12 y 15 especies son nativas de la Argentina, sur de Brasil, Paraguay y Uruguay (Giberti 1979, 1994a,b, 2001, 2008, 2011a,b; Grela 2004; Giberti y Gurni 2008). Entre ellas se destacan *I. paraguariensis*, llamada “caá” (yerba mate, selva) o “caá-guazú” (yerba grande) por la cultura Guaraní, e *I. dumosa* y sus variedades botánicas, vulgarmente denominadas yerba señorita, “congonilla”, “erva piriquita”, “caá-miní” (yerba pequeña), “caá-berá” (yerba resplandeciente), etc. (Giberti 1979). Estas especies presentan una distribución natural simpátrica, abarcando el sur de Brasil (los Estados Mato Grosso do Sul, Paraná, Santa Catarina y Rio Grande do Sul), la región noreste de la Argentina (provincias de Corrientes y Misiones), el noreste del Paraguay (Departamentos del Alto Paraná, Amambay, Caaguazú, Canindeyú, Central, Guairá, Itapúa, Misiones y San Pedro) y la República Oriental del Uruguay (**Figura 4;** Giberti 1979, 1994a,b, 2001, 2008, 2011a,b; Grela 2004; Giberti y Gurni 2008). Este último país representa la zona más austral del rango de distribución de ambas especies; *I. dumosa* fue documentado en el Departamento de Rivera, al norte del Uruguay (Grela 2004), mientras que se han encontrado parches de yerbales silvestres de *I. paraguariensis* en los Departamentos de Cerro Largo, Lavalleja, Maldonado, Treinta y Tres y Tacuarembó (Giberti 1994a,b, 2011a,b; Grela 2004). La yerba mate es la especie más austral del género (Giberti 2001).



**Figura 4.** Distribución genérica y específica. **A.** Distribución del género a nivel mundial, tomado de *Angiosperm Phylogeny Website* versión 13. Los números indican la cantidad de especies presentes en cada región, según Cuénoud *et al.* (2000). **B.** Distribución de la yerba mate (cuadrados verdes) y la yerba señorita (rombos rojos), según Giberti (2001) y Grela (2004). Adaptado de Giberti (2001).

La yerba mate crece naturalmente en bosques subtropicales mixtos junto al pino Paraná (*Araucaria angustifolia*), conocidos como “mato preto” o selva oscura, debido a la coloración verde oscura de sus hojas. Estos bosques son mucho más extensos en el

## Introducción

sur brasilero que en nuestro país o en Paraguay (Giberti 2001). Si bien la yerba mate y el pino Paraná muestran una asociación estrecha en su distribución, el pino Paraná no llega a Uruguay (Giberti 2011a). También se han documentado individuos de *I. paraguariensis* creciendo aislados en el monte nativo, que Giberti (2001) atribuyó a la diseminación endozoica mediada por aves frugívoras. El avance de la frontera agrícola y la sobreexplotación de los yerbales nativos llevada a cabo antiguamente, pusieron en riesgo la persistencia de este germoplasma. No obstante, aún pueden encontrarse poblaciones aisladas en ciertas regiones de Brasil, Paraguay y Uruguay (Dr. G.C. Giberti y Dr. P. Speranza, com. pers.).

Las plantas de *I. dumosa* son arbustos o árboles pequeños, de 1 a 8 metros de altura (Coulleri 2010). En contraste, *I. paraguariensis* puede alcanzar los 18 metros de altura, pero es mantenido como arbusto cuando se lo cultiva con fines comerciales (Buduba 2011) (**Figura 5**). Ambas especies son perennes, con hojas coriáceas de margen aserrado, las de yerba mate con mayor área foliar (Giberti 1979); florecen entre octubre y noviembre y fructifican entre marzo y junio. Requieren precipitaciones anuales uniformes y superiores a los 1200 mm (Giberti 2011a) y suelos tipo areno-arcillosos y arcillo-arenosos, ricos en ácido fosfórico, potasio y hierro, profundos y permeables (<http://www.agropecuario.org/agricultura/yerba-mate>).





**Figura 5. A, B:** Árboles silvestres de *I. paraguariensis* (Lavalleja, Uruguay; fotos gentilmente cedidas por el Ing. Agr. S. Prat Kricun). **C, D:** Arbustos de *I. dumosa*, de uso comercial (Establecimiento Las Marías S.A., Corrientes, fotos amablemente cedidas por el Dr. PA Sansberro).

### Fitoquímica

Muchas especies de *Ilex* son productoras de metabolitos secundarios, entre los cuales pueden encontrarse polifenoles, metilxantinas, flavonoides, saponinas, etc. (Nahar *et al.* 2005; Filip 2009, 2011; Bracesco *et al.* 2011; Hao *et al.* 2013; L Li *et al.* 2013; Chen *et al.* 2016; Thea 2016). En 1987, Alikaridis estudió los constituyentes naturales de aproximadamente 50 especies, y detectó cantidades muy elevadas de alcaloides purínicos (cafeína, teobromina, teofilina o adenina), vitaminas (del tipo C, B1, B2), ácido nicotínico y carotenos.

Las hojas de más de 60 especies de *Ilex* son usadas para preparar bebidas. *Ilex guayusa*, que crece en la región Amazónica de Ecuador, Colombia y Perú, tiene la

## Introducción

concentración más alta de cafeína y metilxantina conocida (Lewis *et al.* 1991; Radice y Vidari 2007). La especie *I. aquifolium* (nativa del sur, oeste y centro de Europa y el Mediterráneo) contiene esteroides, terpenoides, glucósidos cianogénicos, antocianinas y flavonoides (Nahar *et al.* 2005). Las asiáticas *I. cornuta*, *I. latifolia*, *I. kaushue* e *I. pentagona* contienen triterpenoides, ácidos fenólicos, flavonoides y aceites esenciales que les otorgan propiedades antioxidantes, antiinflamatorias, antidiabéticas, hepatoprotectoras y antitumorales, entre otras, a la vez que promueven el metabolismo de los lípidos (Hao *et al.* 2012, 2013; L Li *et al.* 2013; Chen *et al.* 2016). Con *I. cornuta* se prepara un té denominado Gougucha en el sudoeste de China, mientras que *I. latifolia* e *I. kaushue*, conocidas como las especies “Kudingcha” de hoja grande, son utilizadas para preparar tés tradicionales en el sur de China, cuyas propiedades medicinales son aprovechadas desde hace más de 2000 años (Hao *et al.* 2013; L Li *et al.* 2013). Entre las especies sudamericanas, la yerba mate es la más apreciada y relevante a nivel productivo, y por ende ha sido la más estudiada a nivel fitoquímico (para una revisión, ver Heck y de Mejía 2007; Filip 2009; Bracesco *et al.* 2011; Thea 2016), aunque recientemente *I. dumosa* ha ganado cierta relevancia (<http://intainforma.inta.gov.ar/?p=40204>).

En la yerba mate la concentración de los metabolitos secundarios varía según el tipo de material estudiado (hojas frescas, secas, o industrializadas), y el tipo de infusión acuosa preparada (mate o mate cocido). Los extractos acuosos de yerba mate poseen alta concentración de metilxantinas (cafeína, teobromina y trazas de teofilinas), polifenoles (cafeoil-derivados y flavonoides), una gran cantidad de saponinas triterpénicas, y minerales (Heck y de Mejía 2007; Filip 2009, 2011; Buduba 2011; Rafiee Seyed 2012; Cardozo Junior y Morand 2016; entre muchos otros). Por su parte, la yerba señorita ha sido menos estudiada que la yerba mate. Se sabe que tiene cantidades significativamente menores de metabolitos secundarios (Filip *et al.* 1998, 2001; Maiocchi *et al.* 2016), y se caracteriza por contener saponinas del tipo oleanano y altos niveles de fenilpropanoides (Choi *et al.* 2005; Kim *et al.* 2010), lo que incrementa su importancia como alimento funcional. Tanto *I. dumosa* como *I. paraguariensis*, a diferencia de otras especies cogenéricas simpátricas, carecen de arbutina, compuesto con propiedades antioxidantes, diuréticas y antisépticas urinarias (Rafiee Seyed 2012).

Entre los metabolitos secundarios hallados en la yerba mate, las metilxantinas son una clase de alcaloides purínicos que también se encuentran en el té, café y el chocolate. La cafeína –la sustancia psicoactiva más popular y más tolerada a nivel mundial– es la



## Introducción

que se encuentra en mayor concentración en la yerba mate, otorgándole propiedades estimulantes del sistema nervioso central (Heck y de Mejía 2007; Anesini *et al.* 2012). El mayor contenido de cafeína y teobromina se ha detectado en las hojas jóvenes (Soares Lopes *et al.* 2007; Filip 2009). En contraposición, yerba señorita posee cantidades no cuantificables de cafeína (Filip *et al.* 1998), por lo que se la postula como una alternativa descafeinada (<http://intainforma.inta.gov.ar/?p=40204>).

La alta concentración de polifenoles en las hojas de yerba mate muestra una gran correlación con las propiedades antioxidantes, dadas primariamente por los cafeoil-derivados (fenilpropanoides) como el ácido cafeico y el ácido clorogénico (Filip *et al.* 2000; Heck y de Mejía 2007; Anesini *et al.* 2012). La yerba mate y la yerba señorita se caracterizan por poseer altos niveles de fenilpropanoides, aunque debidos a distintos compuestos (Kim *et al.* 2010). Los flavonoides rutina, quercetina y canferol, presentes en altas cantidades en la yerba mate, también poseen actividad antioxidante y antitumoral mediante la inhibición de enzimas involucradas en la formación de radicales libres y en la proliferación y diferenciación celular (Filip 2011). En la yerba señorita se detectaron niveles similares de canferol, pero menores de rutina y quercetina (Filip *et al.* 2001).

Se sabe que las saponinas son las responsables del sabor amargo característico de la yerba, y además brindan propiedades antiinflamatorias e hipocolesterolémicas (Filip 2011; Cardozo Junior y Morand 2016). Por otra parte, muchas de las saponinas encontradas en la yerba mate mostraron propiedades antiparasitarias; también se confirmó que los triterpenoides son antitripanosómicos (Heck y de Mejía 2007; Buduba 2011).

Tanto la yerba mate como la yerba señorita contienen altas concentraciones de compuestos inorgánicos, como aluminio, cromo, cobre, hierro, manganeso, níquel, potasio y zinc. Estos minerales son importantes en el metabolismo y desarrollo humano (Buduba 2011; Maiocchi *et al.* 2016).

Por lo expuesto, es lógico que las hojas de la yerba mate se emplearan en medicina popular para combatir los más variados síntomas (por ejemplo, la artritis, la constipación, el reumatismo, la obesidad, la fatiga, etc.). Esta composición fitoquímica diferencia a la yerba mate de otras especies co-genéricas simpátricas (Filip *et al.* 1998, 2000, 2001), y le otorga un valor agregado como alimento funcional (Filip 2011).

**Historia del consumo de la yerba mate: etnobotánica**

El consumo de mate o tereré es un ritual social, principalmente en el sur de Brasil, Paraguay, Argentina y Uruguay (Rau 2009). En nuestro país la yerba mate está presente en casi la totalidad de los hogares, denotando el fuerte símbolo cultural que representa su consumo. De hecho, en 2010, año del Bicentenario de la República Argentina, el mate fue declarado Infusión Nacional (ley 26871, 2013). Este rito costumbrista tan arraigado tiene siglos de antigüedad, siendo un legado de los Guaraníes, quienes cotidianamente utilizan la yerba mate como infusión digestiva y estimulante, para el tratamiento de diferentes patologías y sintomatologías, así como en diversos rituales (Keller 2013a).

La cultura Guaraní cuenta con distintas versiones sobre el origen de la yerba mate, unas involucrando la transformación de una mujer en la planta (Keller 2007, 2013a), y otras presentándola como el regalo de los dioses a los humanos en agradecimiento por su hospitalidad (<https://rescatando-mi-cultura-guaran.blogspot.com.ar/2013/07/la-caa-yari-diosa-de-la-yerba-mate.html>; <http://www.fansdelmate.com/historias-y-anecdotas/que-leyenda-hay-sobre-el-mate.html>; <https://www.rosamonte.com.ar/la-leyenda-del-mate-y-la-luna/>). La mitología Guaraní habla también de un colosal árbol de yerba mate, denominado *ka'a guachu* o *caa guazú*, traducido como “yerba mate gigante” o “selva inmensa”, y que refiere a la idea de “la selva contenida en un solo árbol”: un inmenso árbol de yerba mate en cuyas ramas crecen todos los árboles de la selva, el viento transporta sus semillas, y de ese modo germinan en el resto de la selva. Esta creencia asegura que, si se hiciera una tala completa de la selva, ésta se alzaría nuevamente, ya que el viento dispersaría las semillas desde el *caa guazú*, garantizando así la regeneración y reproducción de la selva (Keller 2013a,b).

Actualmente, los sacerdotes guaraníes continúan utilizando la yerba mate para adquirir inspiración religiosa, aminorar las tormentas y tornar inofensivos los “dardos” de brujería. También para aliviar cefaleas, tratar verrugas, otitis, afecciones cardíacas y renales, fiebre, hemorragias y epidemias severas como el sarampión (Keller 2008). Además, utilizan al mate como una infusión curativa, al agregarle una variedad de productos vegetales (Keller 2013a). Los niños recién nacidos son consagrados al “creador” en la ceremonia del bautismo que se lleva a cabo en primavera, utilizando ramas de yerba mate como uno de los elementos vinculantes entre las deidades y los niños (Keller 2013b).

**Historia de la producción de yerba mate**

Según expone Burtnik (2006), en la época precolombina la yerba mate era consumida por los Guaraníes como una infusión de hojas secas y trituradas, utilizando una bombilla rudimentaria confeccionada de tacuara y fibras vegetales. Los españoles adoptaron rápidamente la costumbre, y crearon un sistema de “encomiendas” para buscar la materia prima. Este sistema era inhumano, por lo que en el siglo XVII se trató de luchar contra la costumbre de beber mate, considerándola un vicio. Sin embargo, los Jesuitas de la región percibieron que los indígenas que tomaban mate se mantenían más tiempo despiertos y eran más trabajadores, por lo que en el siglo XVIII comenzaron la domesticación de esta planta. Contaron con sus propios huertos yerbateros, aunque se continuó con la explotación de los yerbales naturales, especialmente en el límite entre Paraguay y Brasil (Burtnik 2006). Así, en las reducciones jesuíticas descubrieron cómo germinar las semillas y sentaron las bases tecnológicas de la producción bajo cultivo, convirtiéndolo en el principal bien mercantil con demanda en el Virreinato del Río de la Plata y el Alto Perú, producido con mano de obra aborigen (Rau 2009). Luego de la expulsión de los Jesuitas la tradición del cultivo se fue perdiendo (Burtnik 2006).

Durante el siglo XIX la Argentina consumió yerba mate importada de Brasil y especialmente de Paraguay. A comienzos del siglo XX se lograron las primeras plantaciones en San Ignacio (provincia de Misiones) y recién en las décadas de 1920 y 1930 se estimuló el cultivo yerbatero (Burtnik 2006). En el año 1935 se produjo la primera gran crisis de sobreproducción yerbatera que trajo como consecuencia la creación de la Comisión Reguladora de la Yerba Mate, ente que controló la implantación y cosecha con el propósito de ajustar la producción a la demanda, lo que casi nunca se logró (Burtnik 2006; Rau 2009). Entre 1935 y fines de los '80, se implantaron 100.000 hectáreas en Misiones. Incentivados por el alto precio, a finales de esa década los productores se interesaron en mejorar el manejo del cultivo y del suelo, en el uso racional de agroquímicos y en el manejo integrado de plagas y enfermedades. En 1991 se desreguló totalmente el cultivo aumentando el número de hectáreas de yerbales, y la sobreoferta del producto trajo como consecuencia una gran baja del precio de la materia prima (hoja verde). A principios del siglo XXI el principal obstáculo a la tecnificación de la producción era la escasa retribución que recibían los productores primarios (tareferos) por su labor (Burtnik 2006). En la actualidad, la problemática sigue vigente ya que, ante una nueva sobreoferta, los yerbateros protagonizaron protestas que derivaron en un acuerdo con las autoridades del Instituto Nacional de la

## Introducción

Yerba Mate (INYM; <https://www.cronista.com/negocios/Yerbateros-llegaron-a-un-acuerdo-y-levantaron-la-protesta-en-Misiones-20170323-0028.html>).

A pesar de las fluctuaciones mercantiles, Argentina se mantiene como el primer productor mundial de yerba mate (64%), seguido por Brasil (34%) y Paraguay (4%). El 87% de la producción nacional se destina al mercado interno y el resto a países vecinos, aunque en los últimos años se logró cierta penetración comercial en EE.UU., Europa y Medio Oriente, a partir del consumo generado por las colonias de residentes sudamericanos. Nuestro país es además el mayor consumidor, alcanzando las 250 mil toneladas producidas por año (INYM 2015, 2017), ofreciendo una gran cantidad de marcas y variadas presentaciones. La incorporación al mercado de *I. dumosa* es relativamente reciente, saliendo a la venta en el año 2013 luego de 10 años de investigaciones e intentos frustrados (<http://www.fansdelmate.com/estudios-e-investigacion>) ya que solía ser considerada como un adulterante de la yerba mate (Giberti 1989). Particularmente puede encontrarse a *I. dumosa* en una mezcla con hojas de *I. paraguariensis* que es publicitada como una versión con menor contenido de cafeína (**Figura 6**).

La yerba mate se comercializa también en forma de extractos que se emplean en formulaciones cosméticas, fitoterapéuticas y otros alimentos funcionales.



**Figura 6.** Variedad de yerba mate para el consumo nacional; se enfatiza en la presentación que incorpora *I. dumosa*. Imágenes tomadas de Google imágenes.

## Aspectos productivos

El cultivo de la yerba mate presenta algunas dificultades de naturaleza biológica que en gran medida han sido resueltas, permitiendo los altos niveles productivos actuales. Aún en su ambiente natural las semillas poseen muy bajas tasas de germinación (Prat

## Introducción

Kricun 1995) y su reproducción vía esquejes es problemática (Giberti 2011b). Por este motivo, en nuestro país se han desarrollado trabajos pioneros sobre metodologías de micropropagación *in vitro* (Bernasconi *et al.* 1996; Sansberro *et al.* 1998, 1999, 2000; Mroginski *et al.* 1999; Rey *et al.* 2002; Tarragó *et al.* 2005; Luna *et al.* 2013, 2017), y posteriormente se extendieron al Uruguay (Ross *et al.* 2017). También se han realizado estudios en la yerba señorita (Sansberro *et al.* 2001a; Sansberro *et al.* 2001b; Luna *et al.* 2013, 2017).

En la Estación Experimental Agropecuaria del INTA Cerro Azul (EEA-INTA CA), Misiones, se desarrolla desde hace más de cuatro décadas un plan de mejoramiento genético del cultivo de la yerba mate. Los viajes de recolección de semillas de especies de *Ilex* de la Argentina, Brasil, Paraguay y Uruguay, permitieron crear en la EEA-INTA CA un Banco de Germoplasma de yerba mate y de especies afines, único en Latinoamérica. Allí se han realizado estudios de germinación, generación de progenies controladas, multiplicación y regeneración vegetativa, producción de plantines en condiciones semi-controladas y de mejoramiento genético; también se tratan aspectos de índole netamente productivo como mejoras en el sistema de poda, de la cosecha mecanizada, y de buenas prácticas de manufactura, entre otros muchos aportes (Mayol 1997; Prat Kricun 2001; Burtnik 2003; Prat Kricun y Belingheri 2003; Mayol y Molina 2011).

Los yerbales implantados, principal fuente de materia prima en la Argentina, se pueden comenzar a cosechar recién hacia el cuarto año, entrando en plena producción a los diez años; a partir de entonces se considera que pueden mantenerse en buenas condiciones productivas durante tres décadas, aunque se conocen plantas que alcanzan los 80 años con excelente producción (Burtnik 2006). En nuestro país hay experiencias de selección masal y mejoramiento genético tradicional para la obtención de nuevas variedades en base al contenido fitoquímico o al rendimiento de hoja verde por hectárea (Scherer *et al.* 2002, 2006; Marx *et al.* 2003; Prat Kricun 2011). Las características biológicas de la especie ralentizan los tiempos de generación de nuevas variedades, y acentúan la necesidad de utilizar otras herramientas que aceleren dicho proceso.

La elaboración de la yerba mate comprende tres etapas fundamentales: el sapecado, el secado y el canchado (**Figura 7**; Burtnik 2006). Dentro de las 24 horas de la cosecha, y para evitar la fermentación y degradación de los tejidos, se deshidrata casi totalmente el material cosechado. El sapecado se realiza en un horno rotativo cilíndrico y consiste en la exposición de las hojas a la acción directa de fuego vivo (400-460°C), durante 20-

## *Introducción*

30 segundos, perdiendo entre 55 y 75% de agua. Luego, el material es sometido a un proceso de secado hasta reducir su contenido de humedad al 5-6%. Posteriormente, con el fin de facilitar el embolsado y transporte, se somete la yerba a un proceso de trituración grosera, denominado canchado. La yerba canchada se estaciona para su maduración, ya sea suelta o envasada en bolsas de arpillera, de nueve meses a un año; durante este período desarrolla su aroma y sabor. También se puede hacer un estacionamiento acelerado en cámaras de ambiente controlado, que reduce el período a 30 días, pero cambiando notablemente la calidad del producto (color, sabor y aroma) (Burtnik 2006).



**Figura 7.** Etapas de la elaboración de la yerba mate. Imágenes tomadas de <http://www.rutadelayerbamate.org.ar/yerba-mate/proceso-de-elaboracion>.

## *Antecedentes*

La presente Tesis se centra en las dos especies de *Ilex* con mayor relevancia para nuestro país. Como se mencionara, la yerba mate es la especie más emblemática, con mayor historia e importancia socio-cultural y económica, y más rica en términos nutricionales. La yerba señorita presenta mayor tolerancia a condiciones de sequía y a

## Introducción

las bajas temperaturas, y resistencia al psílido *Gyropsylla spegazziana* (Hemiptera, Psyllidae), considerado la principal plaga del cultivo (Prat Kricun 2009). Estas características agronómicas junto con las metabólicas posicionan a la yerba señorita como una fuente potencial de genes y/o de regiones reguladoras de la expresión de enorme interés para futuros planes de mejoramiento de la yerba mate.

Respecto a otros cultivos industriales, los programas de mejoramiento tradicionales de estas especies están atrasados debido a la lenta de germinación de las semillas y los extensos períodos hasta la floración (7-10 años desde la germinación). Bajo estas condiciones, los retrocruzamientos necesarios para seleccionar rasgos deseables, como la resistencia a las enfermedades o la calidad orientada al gusto del consumidor, requieren un tiempo y esfuerzos considerables. Para superar estos inconvenientes, la tecnología de transformación de cloroplastos se ofrece como una metodología alternativa de mejoramiento que, comparada con los procedimientos convencionales, se caracteriza por una mayor contención de los transgenes, por eliminar el efecto de posición a través de la integración sitio-específica, y también por eliminar efectos pleiotrópicos ya que los productos transgénicos quedan contenidos mediante compartimentación subcelular (Koop *et al.* 2007; Verma y Daniell 2007). La transformación plastídica ha demostrado ser una vía de recombinación homóloga eficiente, y se ha logrado en diferentes especies cultivadas (Verma y Daniell 2007). El conocimiento del plastoma completo, su estructura, variación de secuencia y evolución es un requisito previo para el desarrollo de dicha tecnología (Maier y Schmitz-Linneweber 2004; Daniell *et al.* 2005). Adicionalmente, las regiones plastómicas no codificantes proporcionan fuentes de datos muy valiosas para estudios filogenéticos (Borsch y Quandt 2009) y materia prima para desarrollar marcadores genéticos polimórficos útiles en los análisis de evaluación genética y de variabilidad poblacional.

Por otra parte, para *I. paraguariensis* e *I. dumosa*, y para el género *Ilex* en general, aún se desconocen los mecanismos que, en los primordios florales, desencadenan el desarrollo de un conjunto de órganos sexuales y el arresto del otro, así como el sistema operativo de determinación del sexo. Las plantas dioicas (ca. 6% de las angiospermas) ofrecen un sistema particularmente interesante para estudiar la genética y evolución de los sistemas de determinación sexual, ya que las distintas especies exhiben diversos sistemas genéticos de determinación sexual. Los mismos varían desde sistemas con un *locus* simple hasta múltiples *loci*, e involucran desde un grupo de *loci* ligados a un cromosoma autosómico, hasta cromosomas sexuales homomórficos y heteromórficos

## Introducción

(Grant 1999). Estos mecanismos de diferenciación sexual involucran no sólo la presencia y expresión de genes codificados en el genoma, sino también señales ambientales y hormonales, y no se descarta la posible participación de mecanismos de control epigenético (Vyskot y Hobza 2004, 2015; Kejnovsky y Vyskot 2010). En particular para *I. paraguariensis* hasta la fecha no se ha probado la existencia de cromosomas sexuales, aun cuando se ha detectado un bivalente fuera de placa que migra anticipadamente (con frecuencia del 5% según Niklas 1987, y del *ca.* 20% según Realini *et al.* 2014) y que fue interpretado como un cromosoma sexual heteromórfico (Barral *et al.* 1995). Posteriormente, otros estudios citogenéticos clásicos en yerba mate informaron meiosis masculinas regulares sin bivalentes heteromórficos (Greizerstein *et al.* 2004, Realini *et al.* 2014). Para esta especie se planteó la posibilidad de que la expresión sexual podría ser modificada o regulada por factores hormonales y/o ambientales (Winge 1997), a la fecha esto no ha sido probado empíricamente. Tomando como cierta la afirmación de Winge (1997), Giberti y Gurni (2008) reportaron que las condiciones térmicas pueden llegar a modificar la expresión del sexo, al constatar la inducción de trioecia (ocurrencia de flores femeninas, masculinas y hermafroditas en distintas plantas) por las altas temperaturas. Cabe destacar que esta última afirmación carece de documentación y experimentación que así lo acrediten. El estudio realizado para indagar acerca de la base genética de la determinación sexual y la falta de evidencias concretas (Gottlieb y Poggio 2010) pone de manifiesto la importancia de estudiar mecanismos de determinación sexual mediante otras aproximaciones.

Además de ser interesante desde un punto de vista biológico, el conocimiento de los mecanismos implicados en la determinación sexual de estas especies podría servir para desarrollar sistemas de sexado basados en marcadores moleculares. Estos podrían ser aplicados en plantines (provenientes de semilla), sin necesidad de esperar a la primera floración, ya que el sexado a campo sólo puede hacerse observando la flor. El sexado tendría aplicación práctica al acelerar los tiempos de mejoramiento y permitir discriminar los plantines, ya que las plantas femeninas invierten energía en la floración y fructificación (Rakocevic *et al.* 2009) y la cosecha se detiene durante estos meses. Asimismo, se ha registrado una mayor concentración de cafeína y teobromina, y mayor producción de ramas y brotes en los individuos masculinos (Scherer *et al.* 2002, Rakocevic *et al.* 2012).

Dadas las características de estas plantas es comprensible que gran parte de los trabajos científicos publicados sobre la yerba mate y la yerba señorita se hayan centrado



## Introducción

en la caracterización y cuantificación de los compuestos fitoquímicos (por ejemplo, Filip *et al.* 1998, 2000, 2001; Anesini *et al.* 2012; Cardozo Junior y Morand 2016; Maiocchi *et al.* 2016; entre muchos otros).

En cambio, recién en las últimas dos décadas se fue acumulando información sobre aspectos de la genética básica de estas dos especies, en especial de la yerba mate. A nivel poblacional Gauer y Cavalli-Molina (2000) y Neumann Wendt (2005) detectaron una base genética restringida en cuatro y seis poblaciones naturales de *I. paraguariensis* de Brasil usando marcadores RAPDs e isoenzimas, respectivamente. En contraposición, Pereira *et al.* (2013) informaron altos valores de variabilidad microsatélite en dos poblaciones silvestres brasileras. Contrariamente, las poblaciones naturales de yerba mate que aún persisten en el Uruguay mostraron bajos niveles variabilidad microsatélite (Cascales *et al.* 2014), lo cual no resulta sorprendente al considerar que se trata del límite más austral de la distribución de la especie. Este último estudio sugirió además cierta adaptación al ambiente. Para *I. dumosa* no se conocen estudios genético-moleculares a nivel poblacional; el único trabajo centrado en esta especie estudió la dinámica de seis poblaciones naturales de Argentina y Brasil y las respuestas a los cambios ambientales, empleando información morfológica (Coulleri 2010). Gottlieb *et al.* (2005, 2011) compararon diversas accesiones de yerba mate, yerba señorita y otras ocho especies de *Ilex* del Banco de Germoplasma de la EEA-INTA CA, mediante la genotipificación por AFLP. En dichos trabajos, hallaron en *I. dumosa* un nivel de variación molecular dos veces mayor respecto de *I. paraguariensis*, aunque este resultado podría estar sesgado por el origen diverso de las muestras de yerba señorita. En un estudio que apuntaba a detectar diferencias entre plantas de yerba mate de ambos sexos, se obtuvo secundariamente el primer aislamiento y caracterización de elementos móviles en la especie (Gottlieb y Poggio 2010). Un estudio posterior y comparativo puso de manifiesto la mayor abundancia de ADN repetitivo en el genoma de yerba mate respecto de la yerba señorita, y permitió conocer los tamaños del genoma de cada especie con citometría de flujo (yerba mate ~1670 Mbp y yerba señorita ~1848 Mbp; Gottlieb y Poggio 2014). Como se mencionó anteriormente, trabajos relativamente recientes intentaron dilucidar si la diferenciación sexual en yerba mate tiene una base genética, mediante metodologías de hibridación sustractiva empleando ADN genómico como punto de partida (Gottlieb y Poggio 2010). Algunos fragmentos aislados resultaron altamente ligados a *loci* relacionados con el sexo o a regiones localizadas en cromosomas sexuales de especies distantemente relacionadas con las Aquifoliaceae; a

## Introducción

pesar de estos resultados, no se halló asociación efectiva con regiones relacionadas con el sexo.

Recientemente, investigadores argentinos lograron hallar transcritos relacionados con las respuestas fisiológicas frente al estrés hídrico, expresados diferencialmente entre grupos de plantas sometidas a distintos niveles de estrés (Acevedo *et al.* 2013, 2016); también se obtuvo el primer borrador del transcriptoma de la yerba mate (Debat *et al.* 2014). Además, en 2015-2016 se llevó a cabo el proyecto genoma de la yerba mate, que se encuentra en etapa de anotación.

El estudio de los genomas nucleares, plastomas, epigenomas y transcriptomas, permite una aproximación a la comprensión de procesos biológicos, como la variación fenotípica, la microevolución, el potencial de respuesta a distintos factores ambientales, etc. En el caso de las especies cultivadas, estos conocimientos también son importantes para generar nuevas variedades que sumen resistencias a patógenos y/o que contribuyan mejorar a la calidad nutracéutica o la palatabilidad. En el caso de las especies estudiadas en esta Tesis, la yerba mate es un cultivo industrial de innegable interés nacional y regional, mientras que la yerba señorita se encuentra en sus primeros años de desarrollo comercial. Aún si la yerba señorita no alcanzara la relevancia socio-económica de la yerba mate, sus características bio-agronómicas la posicionan como un reservorio génico de potencial interés para el mejoramiento biotecnológico de la yerba mate. En este contexto, aquí se analizan estas dos especies a través de un enfoque genético-molecular que explora diferentes aspectos. Por un lado, la obtención de la secuencia completa de los genomas cloroplastídicos (plastomas) y su caracterización exhaustiva, como plataforma para el desarrollo de herramientas moleculares destinadas al mejoramiento genético de la yerba mate. Por el otro, la caracterización de las marcas epigenéticas del genoma y de la expresión génica diferencial que contribuyen a distinguir la identidad de los distintos órganos y de los sexos, como herramientas para la detección de genes de interés para dicho mejoramiento.

## **OBJETIVOS GENERALES**

- 1) Caracterizar los genomas plastídicos (plastomas) de la yerba mate y la yerba señorita (**Capítulo 1**).
- 2) Detectar diferencias genéticas entre distintos órganos, sexos y especies (**Capítulo 2**) a través del estudio de patrones de metilación diferencial del ADN (epigenética) y de la expresión génica diferencial.

Los objetivos específicos se detallan en los capítulos correspondientes.

# CAPÍTULO 1

## Caracterización de los plastomas de la yerba mate y la yerba señorita



## Introducción

El genoma plastídico, también llamado plastoma, contiene la información genética de los cloroplastos y existe en miles de copias dentro de cada célula fotosintética. Debido a que los ancestros de los cloroplastos fueron bacterias fotosintéticas probablemente relacionadas con las cianobacterias actuales, el plastoma ha retenido numerosas características procariontas (Bock 2007). Se estima que la endosimbiosis ocurrió hace 600-1500 millones de años, según reportan McFadden y Van Dooren (2004). El plastoma es un genoma circular de tipo bacteriano, con empaquetamiento en nucleoides (grandes cuerpos de nucleoproteínas), posee organización de genes en operones, y una maquinaria de expresión génica procarionta (Bock 2007). Se conoce que hubo transferencia de genes desde el genoma plastídico al núcleo, y que esto redujo la capacidad de codificación del plastoma a menos del 5 % de la capacidad ancestral (Ruhlman y Jansen 2014). Entre los genes plastídicos que son co-transcritos desde operones, encontramos aquellos relacionados con las subunidades de la ARN polimerasa (*rpoB* - *rpoC1* - *rpoC2*) (Jansen y Ruhlman 2012). Algunos ARN policistrónicos plastídicos son multifuncionales, tales como el operón ribosómico que contiene secuencias de ARN ribosomal y de transferencia (ARNr y ARNt). Otros, como el *cluster* del gen *rpl23*, codifican polipéptidos pertenecientes a diferentes complejos funcionales (Ruhlman y Jansen 2014).

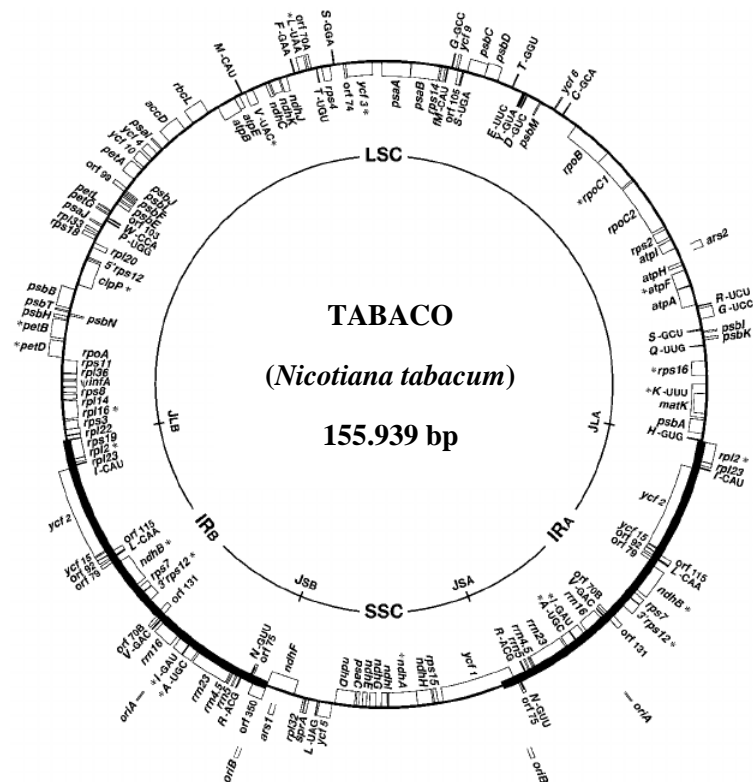
Al ser funcionalmente haploides y tener herencia uniparental (maternal en angiospermas), los plastomas poseen la mitad del tamaño efectivo del genoma nuclear, y se excluyen de la recombinación sexual (Bock 2007). Estas características los convierten en una fuente valiosa de datos para estudios filogenéticos, poblacionales y biogeográficos (Provan *et al.* 2001, 2004; Borsch y Quandt 2009). Los mecanismos de replicación y segregación en los plástidos previenen la persistencia de dos genomas plastídicos diferentes en las células (heteroplasmia). La heteroplasmia solo se puede mantener durante largos períodos de tiempo mediante la selección de ambos genomas de plástidos (Drescher *et al.* 2000; Day y Madesis 2007).

### Estructura de los plastomas

Los plastomas de las plantas terrestres tienen alta densidad génica en relación con los genomas nucleares o mitocondriales. Poseen de 100 a 130 genes, y en general su tamaño varía entre 120-170 kilopares de bases (Kpb) (Bock 2007; Ruhlman y Jansen 2014). El plastoma más pequeño registrado comprende 21 Kpb (*Sciaphila densiflora*;

Lam *et al.* 2015) y el más grande 217 Kpb (en el híbrido *Pelargonium* × *hortorum*; Chumley *et al.* 2006).

La mayoría de los plastomas de angiospermas exhiben una región repetida e invertida grande (~25 Kpb; llamada IR, del inglés *Inverted Repeat*), presente en dos copias idénticas en secuencia, designadas IR<sub>A</sub> e IR<sub>B</sub>. Por lo tanto, los genes contenidos en los IR están duplicados. Por lo general, los genes del operón ribosomal (*rrn16*, *rrn23*, *rrn4.5* y *rrn5*) y siete genes ARNt se encuentran duplicados en los IR. Además de los genes de las proteínas ribosómicas (dos mayores *-rpl-* y una menor *-rps-*), los IR también contienen los genes *ndhB* e *ycf2* (ver más adelante) y el segundo y tercer exón del gen *rps12* que sufre *trans-splicing*. Las dos copias IR están separadas por regiones de copia única y de distinto tamaño, la región mayor (LSC, *Large Single-Copy region*, ~80 Kpb) y la menor (SSC, *Small Single-Copy region*, ~20 Kpb). Esta disposición cuadripartita se observó en el genoma plástido de *Nicotiana tabacum*, primero en ser completamente secuenciado (Shinozaki *et al.* 1986), y se considera que es la forma ancestral y típica de las angiospermas (**Figura 8**) (Raubeson *et al.* 2007). La presencia de los IR no sólo en plantas terrestres, sino también en linajes de clorofitas, sugiere que es una característica antigua (Ruhlman y Jansen 2014).



**Figura 8.** Plastoma típico del tabaco mostrando la estructura cuadripartita. Tomado de Wakasugi *et al.* (1998).

Si bien los plastomas no sufren recombinación sexual, la recombinación intramolecular entre los IR podría ser un mecanismo para evitar la divergencia entre las copias, teniendo el potencial de invertir la polaridad del segmento entre los IR. Además, la presencia de los IR puede jugar un papel en la estabilidad estructural del plastoma, aunque no la garantiza, ya que experimenta expansiones y contracciones (Ruhlman y Jansen 2014).

Los primeros estudios relacionados con las características estructurales de los plastomas utilizaron la digestión con enzimas de restricción de moléculas de ADN aisladas de cloroplastos purificados. Con el desarrollo de tecnologías que facilitan la secuenciación directa de plastomas completos, muchas de las predicciones acerca de su estructura se confirmaron, proporcionando una imagen razonablemente clara del plastoma típico de las angiospermas. Los mapas circulares presentados en la literatura representan un solo monómero, aunque una serie de estudios han identificado formas alternativas incluyendo círculos multiméricos así como moléculas lineales y ramificadas (Ruhlman y Jansen 2014).

### **Contenido génico de los plastomas**

El complemento genético de la gran mayoría de plastomas de angiospermas está altamente conservado y la disposición de las secuencias es co-lineal (Bock 2007; Ruhlman y Jansen 2014; **Tabla AI-1 del Apéndice I**). Los genes se agrupan en tres clases generales según la función de sus productos: (a) componentes fotosintéticos, (b) componentes del sistema genético, y (c) otras funciones. En esta última categoría se agrupan los genes que no encajan bien en las otras dos categorías funcionales, como por ejemplo, *clpP*, *accD*, *ccsA* y *cemA* (Ruhlman y Jansen 2014).

Los marcos de lectura abiertos (ORFs, del inglés *Open Reading Frames*) son designados como *ycf* (*hypothetical chloroplast reading frames*, es decir, marcos de lectura hipotéticos de cloroplastos) hasta tanto se pueda atribuir alguna función a su producto. Como excepción, se sabe que los productos de los genes *ycf3* e *ycf4* forman parte del fotosistema I (FSI), pero aún no se han renombrado (Ruhlman y Jansen 2014). Si bien todavía no se ha atribuido ninguna función definitiva a los dos ORFs más grandes, *ycf1* e *ycf2*, han sido designados como genes esenciales debido a que su delección resulta en fenotipos mutantes inestables (Drescher *et al.* 2000).

Más de un tercio del plastoma típico comprende genes codificantes de proteínas implicadas en la fotosíntesis y en reacciones relacionadas (**Tabla AI-1**),

mayoritariamente contenidos dentro de las regiones de copia única. Con excepción de la subunidad grande de la enzima RuBisCo (*rbcL*), todos los genes de esta clase codifican subunidades de complejos de membrana que se ensamblan junto con proteínas codificadas en el núcleo. El gen de la subunidad pequeña de la RuBisCo (*rbcS*) se encuentra también en el genoma nuclear, pero la holoenzima se reúne en el estroma plastidial y no se asocia con las membranas (Ruhlman y Jansen 2014).

El sistema genético incluye los ARNt, ARNr, proteínas asociadas con las subunidades ribosomales, los subtipos de ARN polimerasa (PEP, *plastid-encoded RNA polymerase*), la maturasa K (*matK*) y el factor de iniciación de la traducción 1 (*infA*) (**Tabla AI-1**; Ruhlman y Jansen 2014).

Las regiones codificantes representan el ~50 % del total de los plastomas; el resto se divide entre intrones, regiones reguladoras y espaciadores intergénicos (IGS, del inglés *InterGenic Spacers*). El contenido general de GC (guanina-citosina) es típicamente del 30-40%, y en algunas regiones no codificantes el contenido AT (adenina-timina) excede el 80% (Ruhlman y Jansen 2014).

### **Las regiones plastídicas como marcadores moleculares**

A pesar de que la organización estructural y las secuencias nucleotídicas están altamente conservadas en la mayoría de los plastomas de las plantas con flores, una fracción destacada muestra una marcada variación en sus tasas de evolución, tanto de secuencia como en su arquitectura genómica (Ruhlman y Jansen 2014). Sin embargo se han documentado varios casos de expansiones y contracciones de los IR, inversiones, e inserciones y deleciones (*indels*) solapadas, que dan como resultado una reconfiguración del orden génico. Las inversiones simples pueden ser útiles para la resolución de relaciones filogenéticas, especialmente en los nodos profundos (Ruhlman y Jansen 2014). En general, las regiones no codificantes como los intrones e IGS ofrecen mayor información a la hora de realizar estudios filogenéticos, ya que las regiones codificantes no poseen suficiente variación nucleotídica (Borsch y Quandt 2009).

Como ya fue expuesto, las características del genoma plastídico (haploide, no recombinante, transmisión materna en angiospermas), sumadas al menor flujo génico en comparación con los genes nucleares (de herencia biparental), hacen de los marcadores plastídicos buenos indicadores de mecanismos evolutivos como la deriva génica y de procesos demográficos como los cuellos de botella y los efectos fundadores; por lo



tanto, constituyen herramientas valiosas aplicables a estudios poblacionales, filo- y biogeográficos (Provan *et al.* 2001, 2004; Rendell y Ennos 2003; Borsch y Quandt 2009).

A diferencia de los marcadores nucleares, los datos de ADN plastídico pueden resultar útiles para identificar eventos de hibridación pasados (Péros *et al.* 2011). Por lo tanto, estos marcadores son preferidos para análisis filogenéticos especialmente ante evidencia de eventos de hibridación interespecífica, como en el caso del género *Ilex* (Galle 1997; Cuénoud *et al.* 2000; Setoguchi y Watanabe 2000; Manen *et al.* 2002, 2010; Manen 2004; Lee *et al.* 2006; Shi *et al.* 2016).

En la literatura abundan los estudios que se centran en detectar marcadores plastídicos variables para estudiar las relaciones filogenéticas y la filo/biogeografía de varios géneros de angiospermas (Dong *et al.* 2012; Jansen *et al.* 2007; Kim y Lee 2004; Prince 2015; Shaw *et al.* 2005, 2014, entre otros). Sin embargo, las Aquifoliaceae han sido escasamente representadas (Bremer *et al.* 2002; Moore *et al.* 2010). Recientemente, Yao *et al.* (2016b) secuenciaron y caracterizaron los plastomas de seis especies de *Ilex* asiáticas (*I. delavayi*, *I. latifolia*, *I. polyneura*, *I. pubescens*, *I. wilsonii* e *I. szechwanensis*) y de una *Ilex* sp., y propusieron una serie de marcadores, aunque sin probar su utilidad filogenética. La falta de marcadores desarrollados específicamente para las especies de *Ilex* se hizo evidente ya desde el trabajo de Rendell y Ennos (2003) quienes utilizaron marcadores microsatélite plastídicos heterólogos, originalmente desarrollados para *N. tabacum*, para analizar poblaciones de *I. aquifolium*.

Los estudios filogenéticos en *Ilex* se han basado en unos pocos marcadores nucleares universales (como ITS, *nepGS*) y/o plastídicos (*matK*, *trnL-trnF*, *psbA-trnH*, *atpB-rbcL* y *rbcL*), estos últimos con un poder resolutivo dudoso al ser regiones muy conservadas y poco informativas (Cuénoud *et al.* 2000; Setoguchi y Watanabe 2000; Manen *et al.* 2002, 2010; Manen 2004; Gottlieb *et al.* 2005; Selbach-Schnadelbach *et al.* 2009; Shi *et al.* 2016). En consecuencia, la mayoría de las filogenias moleculares publicadas muestran inestabilidad en la posición de los taxones, pobre resolución de los nodos internos e incongruencias, estableciendo relaciones inciertas (Setoguchi y Watanabe 2000; Manen *et al.* 2002, 2010; Gottlieb *et al.* 2005; Selbach-Schnadelbach *et al.* 2009). A excepción de la filogenia plastídica de Setoguchi y Watanabe (2000), las filogenias moleculares publicadas muestran discrepancia con el sistema de clasificación infragenérico de Loesener (1901, 1908) que reconoce tres subgéneros, con secciones y series (Cuénoud *et al.* 2000; Manen *et al.* 2002, 2010; Gottlieb *et al.* 2005; Selbach-

Schnadelbach *et al.* 2009). Por otro lado, las filogenias discordantes entre datos nucleares y plastídicos (Setoguchi y Watanabe 2000; Manen *et al.* 2002, 2010; Manen 2004) apuntan a la ocurrencia de eventos frecuentes de introgresión interespecífica e hibridación (Setoguchi y Watanabe 2000; Manen *et al.* 2002, 2010; Manen 2004; Shi *et al.* 2016). Por lo expuesto, los marcadores plastídicos serían las mejores herramientas para estudiar las relaciones entre las especies de este género. De acuerdo con Cuénod *et al.* (2000) y Manen *et al.* (2002), las filogenias plastídicas correlacionan parcialmente con la distribución geográfica de las especies actuales, a diferencia de las nucleares.

Gottlieb *et al.* (2005) usando datos nucleares reportaron que *I. paraguariensis* e *I. dumosa* estarían distantemente relacionadas con las demás especies sudamericanas, las cuales forman un grupo monofilético; *I. dumosa* estaría relacionada con las asiáticas *I. crenata* e *I. mutchagara*; mientras que las afinidades de *I. paraguariensis* permanecieron inciertas, pudiendo estar relacionada con otras especies asiáticas. Por su parte, Selbach-Schnadelbach *et al.* (2009) y Manen *et al.* (2010) reportaron que, en base a datos plastídicos, *I. dumosa* tendría afinidad genética con especies sudamericanas no simpátricas, y que *I. paraguariensis* estaría relacionada con especies norteamericanas o centroamericanas. Por tanto, los datos sugieren que *I. dumosa* e *I. paraguariensis* no serían taxones hermanos, y sus relaciones con otras especies co-genéricas del sur de Sudamérica son dudosas (Manen *et al.* 2002, 2010; Gottlieb *et al.* 2005; Selbach-Schnadelbach *et al.* 2009). Más recientemente, y en concordancia con lo anterior, Cascales *et al.* (2017), utilizando numerosos marcadores plastídicos y nucleares obtuvieron topologías en donde *I. paraguariensis*, *I. dumosa* e *I. argentina* ocuparon posiciones intermedias entre las demás especies del sur de Sudamérica y las especies euroasiáticas.

### **Los plastomas como herramientas biotecnológicas**

Si bien no es el eje primordial de esta Tesis, la caracterización de los plastomas sienta las bases para el desarrollo de novedosas metodologías de mejoramiento. Se han obtenido resultados exitosos al aplicar la tecnología de transformación de cloroplastos en diversas especies de interés económico, como por ejemplo en coliflor, repollo, lechuga, papa, tabaco, tomate, zanahoria, algodón, arroz y soja (Verma y Daniell 2007; Daniell *et al.* 2016). Considerando que los vectores plastídicos especie-específicos aumentan la eficiencia de transformación (Verma y Daniell 2007; Daniell *et al.* 2016), el conocimiento del plastoma de la especie a mejorar –en este caso, la yerba mate– es un

prerrequisito que cobra mayor relevancia. Las características beneficiosas de esta tecnología mencionadas en la Introducción general la configuran como una herramienta atractiva para su implementación en yerba mate.

## Objetivos

El objetivo general de este capítulo es caracterizar los genomas plastídicos de la yerba mate y la yerba señorita.

Los objetivos específicos planteados fueron los siguientes:

- ✓ Secuenciar y anotar los plastomas completos de ambas especies.
- ✓ Analizar la conservación de las secuencias nucleotídicas y de la organización genómica.
- ✓ Evaluar la utilidad de las distintas regiones como marcadores filogenéticos.
- ✓ Desarrollar marcadores microsatélite plastídicos (cpSSR, *chloroplast simple sequence repeats*).

## Hipótesis

- ✓ Los plastomas de ambas especies poseen diferencias sólo a nivel de secuencia nucleotídica, teniendo orden y contenido génico conservado (sintenia).

## Predicciones

- ✓ Los plastomas de la yerba mate y la yerba señorita tendrán alta identidad de secuencia.
- ✓ La mayor variación de secuencia se encontrará en las regiones no codificantes, y tendrán mayor utilidad como marcadores moleculares.

## Materiales y Métodos

### 1. Genotipado preliminar

Se realizó un genotipado preliminar de 16 plantas de *I. dumosa* y 50 de *I. paraguariensis* para seleccionar los plasmomas a secuenciar. Estas plantas provienen del Banco de Germoplasma de la EEA-INTA CA, del Jardín Botánico de Buenos Aires, del Establecimiento Las Marías S.A. (donación mediada por el Dr. PA Sansberro; IBONE-CONICET), y de PINDÓ S.A. (donado por el Dr. RA Scherer) (**Tabla 1**).

Dada la falta de marcadores cpSSR específicos, se utilizaron 6 marcadores heterólogos (**Tabla AI-2**) pertenecientes a la serie *ccmp* generada para tabaco (Weising y Gardner 1999). Los *loci* se encuentran ubicados en la región LSC, con excepción de *ccmp10* que se extiende sobre un IR. Las amplificaciones por PCR se realizaron utilizando aproximadamente 30 ng de ADN genómico, 60 ng de cada cebador, *buffer* de PCR 1X, MgCl<sub>2</sub> 2 mM, 0,2 mM de cada dNTP y 0,5 unidades de Taq DNA polimerasa (*Thermo Fisher Scientific*), en un volumen final de 25 µl. El programa de termociclado consistió en una desnaturalización inicial a 94°C por 4 min, seguida por 30 ciclos de 1 min a 94°C, 1 min a 50°C, 1 min a 72°C, y una extensión final a 72°C por 7 min, en un aparato GenePro TC-E (Bioer Technology Ltd.). Los productos de PCR fueron chequeados mediante electroforesis en gel de agarosa al 1,5 % (p/v), usando *buffer* TAE 1X (ver **Apéndice II**). Los geles fueron teñidos con bromuro de etidio y las bandas fueron visualizadas y fotografiadas sobre un transiluminador UV. Posteriormente, las muestras fueron mezcladas con 10 µl de *buffer* formamida 98 % (v/v) (**Apéndice II**), desnaturalizadas por 8 min a 95°C y sometidas a electroforesis en geles de poliacrilamida (PAGE, *Polyacrylamide Gel Electrophoresis*) al 6 % (p/v) desnaturalizante (urea 8 M). La electroforesis se realizó utilizando *buffer* TBE 1X (**Apéndice II**) a 60 W por aprox. 2 hs en un aparato Model S2 (Gibco BRL Sequencing System, Life Technologies). Las bandas se visualizaron por tinción con nitrato de plata utilizando el *Silver Sequence kit* (Promega). Los tamaños alélicos se estimaron a partir de imágenes digitalizadas de los geles, por comparación con el peso molecular estándar 30–330 AFLP DNA Ladder (*Thermo Fisher Scientific*) y usando marcadores internos en cada corrida.

**Tabla 1.** Material vegetal analizado con marcadores cpSSR heterólogos.

Origen geográfico	Accesiones <sup>1</sup>
<b><i>I. dumosa</i></b>	
<b>EEA-INTA CA</b>	
Argentina, Misiones, Iguazú	<b><u>222</u></b> ♂
Argentina, Misiones, Campo Viera	<b>49</b> ♂, 49♀,
Argentina, Misiones, Oberá	71♀, 7♂, 8♀
Brasil, Paraná	<b>55</b> ♂
Paraguay, Alto Paraná	<b><u>227</u></b> ♀, 235♂
Argentina, Misiones, Oberá, Campo Viera <sup>2</sup>	CV0♂, CV4♀, CV5♀
Argentina, Misiones, Oberá, Campo Ramón <sup>2</sup>	CR2♂
<b>Establecimiento Las Marías</b>	
Argentina, Corrientes, Gobernador Virasoro <sup>2</sup>	YS s/n, <b><u>CRV7♂/4</u></b> , <b>CRV48/10</b>
<b><i>I. paraguariensis</i></b>	
<b>EEA-INTA CA</b>	
Argentina, Misiones	<b>CA1-74</b> ♀, CA 8-74 ♂, CA 11-75 ♀, CA 51-75♀
Argentina, Misiones, Candelaria	195
Argentina, Misiones, General Belgrano	47
Argentina, Misiones, Iguazú	<b>220</b>
Argentina, Misiones, Oberá	51, 51-8♀
Argentina, Misiones, San Ignacio	50♂
Argentina, Misiones, San Pedro	39
Brasil, Paraná	<b><u>217</u></b> ♂, 92, 100
Brasil, Santa Catarina	<b>30</b> ♀, 28♂,
Paraguay, Alto Paraná	<b><u>224</u></b> ♂
Paraguay, Caaguazú	254
Paraguay	684♀
Uruguay, Lavalleja	261
Uruguay, Maldonado	257
Uruguay, Tacuarembó	258
Uruguay, Treinta y Tres	255, 256, 260
<b>Establecimiento Las Marías</b>	
Argentina, Corrientes, Gobernador Virasoro	SI-19, SI-49, SI-67, YM s/n, <b><u>JB 1837</u></b> ♂ <sup>3</sup>
<b>Pindó S.A.</b>	
Material comercial	A205-2, A205-5, A205-11, A205-14, A205-20, A205-21, A205-31, A205-44, A205-71, A205-119, A205-136, A205-139, A217-44, A217-45, A217-140, L86-27, M159-39, M190-37, S39-60, U137-105

<sup>1</sup> para los casos disponibles, se incluye la determinación a campo del sexo de las plantas. Los individuos que fueron efectivamente secuenciados se indican subrayados y en negrita.

<sup>2</sup> material donado por el Establecimiento Las Marías S.A.

<sup>3</sup> muestra tomada del Jardín Botánico de Buenos Aires, cedida a esa entidad por el Establecimiento Las Marías.

EEA-INTA CA: Estación Experimental Agropecuaria Cerro Azul (INTA, Instituto Nacional de Tecnología Agropecuaria).

## 2. Secuenciación de plastomas

A partir de los resultados obtenidos en el genotipado preliminar, y teniendo en cuenta los distintos orígenes geográficos para abarcar con mayor probabilidad la variación plástica, se seleccionaron seis individuos de cada especie (**Tabla 1**, en negrita) para la extracción de ADN plástico y posterior secuenciación.

En agosto de 2013 se coleccionó material fresco de dichos individuos (**Tabla 1**). Se tomaron muestras de ramas con hojas en crecimiento activo y visualmente saludables, las cuales se mantuvieron en oscuridad por 24 hs. para disminuir el contenido de almidones (Shi *et al.* 2012) y luego fueron conservadas a -70°C hasta el momento de la extracción.

### 2.1. Extracción de cloroplastos y de ADN plástico

Aproximadamente 30 g de hojas por individuo fueron procesadas para aislar cloroplastos enteros. A fin de minimizar la contaminación con el ADN nuclear, se utilizó el *Chloroplast Isolation Kit* (Sigma-Aldrich) siguiendo el protocolo propuesto para tabaco por el fabricante. Para la extracción del ADN plástico se adaptaron los protocolos de Diekmann *et al.* (2008) y Shi *et al.* (2012). El procedimiento completo se detalla en el **Apéndice II**. Las 12 muestras procesadas se cuantificaron en un espectrofotómetro Nanodrop, y fueron enviados a la Plataforma de Genómica y Bioinformática del INDEAR (Instituto de Agrobiotecnología Rosario, Santa Fe, Argentina), para ser secuenciados mediante *Whole Genome Sequencing* (WGS) *single-end*.

### 2.2. Secuenciación WGS

De las muestras enviadas al INDEAR sólo 8 se ajustaron a los estándares requeridos y por tanto fueron secuenciadas (4 de cada especie; **Tabla 1**, subrayadas). Brevemente, en INDEAR prepararon bibliotecas RL FLX+ (Roche), que fueron unidas en un *pool* equimolar. Se realizó una PCR en emulsión del *pool* de bibliotecas, y se secuenció en un cuarto de placa de titanio en un aparato 454 GS FLX+ (Roche). Además, INDEAR llevó a cabo un primer análisis bioinformático: se ensamblaron contigs *de novo* y usando plastomas como referencia (indicados en la **Tabla AI-3**) con los ensambladores Celera Assembler 6.1 (Myers *et al.* 2000) o Newbler 2.6 y 2.8 (Roche), y se realizó una separación preliminar de las secuencias pertenecientes al cloroplasto de otras contaminantes (nucleares y/o mitocondriales).

### 2.3. Obtención de genomas consenso

Como parte del procesamiento de los datos enviados por el servicio, se procedió a verificar que todos los *contigs* ensamblados en INDEAR estuvieran formados exclusivamente por secuencias plásticas mediante herramientas BLAST (*Basic Local Alignment Search Tool*). Luego de la eliminación de las secuencias mitocondriales y nucleares y de baja calidad, se alinearon los *contigs* y las secuencias plásticas de yerba mate y yerba señorita disponibles en GenBank (yerba mate: accesiones AF471684, EF590404, EF590456, EF590537, EF590624, EF590705, EF590879, EF590943, FJ394634, FJ394776, FR849983, GQ247982, GQ248141, GQ248322, GQ248477, GQ248625, GQ248796, GQ248958; yerba señorita: AF471630, AJ492571, EU359316, X98725), utilizando el programa Sequencher v4.1.4 (GeneCodes Corp.) con una superposición mínima de 100 pb y un porcentaje mínimo de identidad del 90 %. Los alineamientos fueron inspeccionados para corregir posibles errores. Luego se generaron los plastomas consenso de cada especie, ensamblando los *contigs* alineados. La secuencia lineal final se obtuvo reordenando según el sentido de las referencias (por ejemplo, *Camellia sinensis*).

En el caso de *I. dumosa*, cuyo genoma consenso quedó parcialmente incompleto (8,7 % de datos faltantes), el orden y sentido de los 21 *contigs* no solapantes fue identificado en Sequencher por comparación con el plastoma consenso completo de *I. paraguariensis*. Las regiones faltantes en *I. dumosa* (cuyos tamaños oscilaron entre los 150 y 1.785 pb), fueron extraídas del plastoma de yerba mate para buscar lecturas “perdidas” en la base de datos de lecturas crudas de *I. dumosa*, mediante BlastN. Para las regiones que no pudieron ser completadas (*gaps*), se generaron pares de cebadores de las regiones flanqueantes con los programas Primer3Plus (Untergasser *et al.* 2012) y Primer-BLAST (Ye *et al.* 2012) (**Tabla 2**). Las reacciones de PCR se llevaron a cabo como fue explicado anteriormente, con ADN de los mismos individuos secuenciados. Los amplicones se purificaron con el AccuPrep® *Gel purification kit* (BIONEER) y fueron secuenciados en un instrumento ABI 3730xl (Applied Biosystems) en la Unidad de Genómica del Instituto de Biotecnología del INTA Castelar. Las secuencias así generadas fueron ensambladas a los *contigs* usando Sequencher. A continuación, se verificó la existencia de las dos copias de IR, y que el ensamblado hubiera sido realizado correctamente. Para esto, se amplificaron por PCR las uniones entre las distintas regiones (es decir, IRa-LSC, LSC-IRb, IRb-SSC, SSC-IRa), utilizando los pares de cebadores que se indican en la **Tabla 2**. Las reacciones de PCR y la

secuenciación Sanger se llevaron a cabo como se detalló anteriormente. Finalmente, las secuencias fueron alineadas contra los plastomas consenso.

**Tabla 2.** Cebadores utilizados para cerrar *gaps* y verificar las uniones entre regiones del genoma plastídico.

Región amplificada	Secuencias		Tamaño esperado (pb)	Tm (°C)
	Directo (5'-3')	Reverso (5'-3')		
<b>Cerrado de <i>gaps</i><sup>a</sup></b>				
Intrón <i>rps16</i>	ATCCACAACGGATCGTGTCC	ACAGTGACAATTGTTTCAGCCTA	717	55
IGS <i>rpoB-trnC</i>	CGTATGAACGGAGGAATAAAGA	CCCCAGAACAGAAATTGTTACAC	656	55
IGS <i>petN-psbM</i>	GCAACGCGTTTTGAACTTTT	GTGACTCGCCCAAGTCAAGA	578	55
IGS <i>psbM-trnD</i> - IGS <i>trnD-trnY</i>	GGGATTCCTGAACCAATCAA	TTCGCCAATCTACCATGAGA	631	55
<i>psaA</i> - IGS <i>psaA-ycf3</i>	GTAGCCCCGCTAAATGGTGA	ATGGTTGCTCGCGTGATTTG	986	55
<b>Verificación de uniones</b>				
<i>rpl2</i> - <i>psbA</i> <sup>b</sup> (IRa-LSC)	GCAACAGTCGGACAAGTGGGGA	TGGTGTATTCGGCGGCTCCCT	1210	58
<i>rpl22</i> - <i>rpl2</i> <sup>b</sup> (LSC-IRb)	ACGGTAGTACCCTCGTTGACTTCG	GCAACAGTCGGACAAGTGGGGA	936	58
<i>ycf1</i> - <i>ndhF</i> <sup>a</sup> (IRb-SSC)	TCAATGCGTAATCTCAGCGT	CAAATCTGTAGGGGGCGGG	809	58
<i>ycf1</i> - <i>ycf1</i> <sup>a</sup> (SSC-IRa)	GATTGTTCACGATATGGCCC	TCAATGCGTAATCTCAGCGT	1182	58

<sup>a</sup> cebadores diseñados en este trabajo; <sup>b</sup> cebadores obtenidos de Yang *et al.* (2013). IGS, espaciador intergénico; LSC, región mayor de copia simple; SSC, región menor de copia simple; IR, repetición invertida.

#### 2.4. Ensamblado del genoma borrador de *I. cornuta*

Aprovechando la información disponible públicamente en GenBank se ensambló un plastoma borrador de la especie asiática *I. cornuta*, la cual se empleó para realizar comparaciones con las especies aquí estudiadas (al momento de realizarse, no se contaba con datos de secuencia de ninguna otra especie de *Ilex*). Para esto, se descargaron del GenBank 126 secuencias anotadas (AF471624-25, FJ394600-01, FJ394742-43, FR849938, FR849944, FR849946, FR849950, FR849953, FR849958, FR849962, FR849965, FR849969, FR849971, FR849976, FR849988, FR849989, FR849992, GQ434981-82, GQ436364-65, GQ997298-380, HQ664579, JF941993-99, JF954087-95, KF255704, KF255721, totalizando 105 Kpb). El ensamblado se realizó con el programa Geneious 7.1.7 (Biomatters, <http://www.geneious.com>) usando *I. paraguariensis* como referencia. Al no disponer de material vivo, las regiones faltantes del genoma de *I. cornuta* fueron rellenadas con bases ambiguas (N).



### 2.5. Anotación de los genomas

Los genomas consenso de la yerba mate y la yerba señorita fueron anotados utilizando CpGAVAS (Liu *et al.* 2012) con parámetros *default*. Los inicios y finales de todos los genes, así como los marcos de lectura abiertos, fueron verificados manualmente y corregidos cuando fue necesario, por comparación con las secuencias de referencia y las disponibles de *I. cornuta*. Adicionalmente, y con el mismo fin, se realizó un alineamiento de las regiones codificantes a nivel de aminoácidos, en MSWAT (Cai 2010). Los genes ARNt fueron identificados usando el programa tRNAscan-SE 1.21 (Schattner *et al.* 2005), y se corroboraron por comparación vía alineamiento con los genomas de referencia y búsquedas BLAST. Las anotaciones finales fueron ajustadas en Geneious.

Los plastomas completos y anotados fueron depositados en GenBank (KP016927-KP016928). Los mapas circulares se generaron con OGDRAW (OrganellarGenomeDRAW; Lohse *et al.* 2013) usando parámetros *default*.

## 3. Análisis de los plastomas

### 3.1. Conservación nucleotídica intraespecífica

Se realizaron alineamientos entre los plastomas individuales (4 de *I. dumosa* y 4 de *I. paraguariensis*) usando el programa MAFFT *multiple aligner* 1.3.3 (Kato y Standley 2013) implementado en Geneious, y se calcularon los porcentajes pareados de identidad.

### 3.2. Repeticiones

Se estudió el número y localización de pequeños rearrreglos estructurales como las repeticiones directas, reversas y palindrómicas usando el programa REPuter (Kurtz 2001), una distancia de Hamming = 3 (identidad  $\geq 90$  %) y un tamaño mínimo de repetición de 30 pb, siguiendo a Yang *et al.* (2010).

Para verificar que los cebadores cpSSR heterólogos utilizados en el análisis preliminar amplificaran *loci* microsatélite válidos en estas especies, se realizaron PCR *in silico* sobre los plastomas individuales en Geneious.

Además, se buscaron potenciales repeticiones microsatélite de  $\geq 10$  pb y se diseñaron cebadores usando el programa WebSat (Martins *et al.* 2009). Los resultados fueron filtrados manualmente para evitar redundancia. El funcionamiento de 14 de los 30

marcadores generados fue verificado en 8 individuos de *I. dumosa* (CR2; CV0; CV5; 7♀, 55, 222, 227, 235) y 8 de *I. paraguariensis* (30, 50, 195, 217, 228, 252, 258, 261). Las reacciones de PCR se llevaron a cabo como anteriormente, con una temperatura de *annealing* de 60°C; los productos de amplificación se separaron por medio de electroforesis en PAGE 6 % (p/v) desnaturalizante (urea 8 M), como fue descrito previamente. De los geles digitalizados, se contabilizó la cantidad de alelos por especie y por *locus* cpSSR.

### 3.3. Conservación nucleotídica interespecífica

Se estudió el grado de conservación nucleotídica a lo largo de los tres plastomas generados usando el programa mVISTA (Frazer *et al.* 2004) en modo LAGAN (Brudno *et al.* 2003) y el plastoma de *Panax ginseng* (AY582139) como referencia. Además se realizó un alineamiento global entre las tres especies ensambladas en este trabajo y se calcularon los porcentajes de identidad pareados.

### 3.4. Utilidad filogenética de secuencias plastídicas

En primer lugar, para evaluar la capacidad como marcadores filogenéticos de las distintas particiones de datos plastídicos –genes codificantes de proteínas (CDS), ARNr, ARNt, intrones e IGS–, se realizaron análisis usando cada partición separadamente, e incluyendo taxones representantes de grupos monofiléticos bien establecidos, a saber: euastéridos I (Solanales: *Solanum tuberosum* NC\_008096 y *Nicotiana tabacum* NC\_001879.2; Gentianales: *Coffea arabica* NC\_008535) y euastéridos II (Aquifoliales: *Ilex paraguariensis*, *I. dumosa*, *I. cornuta*; Apiales: *P. ginseng* NC\_006290 y *Daucus carota* NC\_008325; Asterales: *Lactuca sativa* NC\_007578 y *Helianthus annuus* NC\_007977); además se incluyeron taxones de Ericales (*Camellia sinensis*) y Malvales (*Theobroma cacao*) como *outgroups*. Las regiones exploradas fueron extraídas de los plastomas completos de los taxones mencionados.

La partición de datos de CDS comprendió las secuencias de los 79 genes. El gen *ycf15* fue excluido del análisis debido a que ya no se considera un gen funcional (Ruhlman y Jansen 2014) y a la gran divergencia detectada entre los taxones muestreados. Los genes fueron concatenados y alineados en Geneious usando el subprograma MUSCLE (Edgar 2004) con los parámetros *default*. El mismo procedimiento se aplicó a los datos de ARNr (genes *rrn23*, *rrn16*, *rrn5* y *rrn4.5*) y de ARNt (30 genes). El conjunto intrónico abarcó todos los intrones (20) presentes en 18

genes (listados en **Tabla AI-4**); los dos intrones de los genes *clpP* e *ycf3* fueron concatenados y considerados como un intrón por gen. Luego de una inspección de los alineamientos, los intrones de los genes *rps12* y *trnA* fueron excluidos por no poseer sitios variables. El conjunto final de intrones consistió en 16 regiones. La partición de datos de IGS consistió en 35 regiones (*accD-psaI*, *atpB-rbcL*, *atpH-atpI*, *infA-rps8*, *ndhB-rps7*, *ndhC-trnV*, *ndhF-rpl32*, *petA-psbJ*, *psbE-petL*, *rpl2-rpl23*, *rpl23-trnI*, *rpl32-trnL*, *rpl36-infA*, *rps7-rps12*, *rps8-rpl14*, *rps12-trnV*, *rrn4.5-rrn5*, *rrn5-trnR*, *rrn16-trnI*, *rrn23-rrn4.5*, *trnA-rrn23*, *trnF-ndhJ*, *trnH-psbA*, *trnI-trnA*, *trnI-ycf2*, *trnK-rps16*, *trnL-ndhB*, *trnL-trnF*, *trnN-ycf1*, *trnR-trnN*, *trnS-trnG*, *trnT-psbD*, *trnV-rrn16*, *ycf2-ycf15* e *ycf15-trnL*), elegidas en base a la disponibilidad de secuencias de *I. cornuta* (es decir, con información para 3 especies de *Ilex*).

Los análisis filogenéticos se realizaron utilizando algoritmos probabilísticos bajo el criterio de Máxima Verosimilitud en MEGA 6.0 (Tamura *et al.* 2013), eliminando las posiciones con menos del 50 % de cobertura por sitio. El modelo de sustitución más ajustado para cada partición (matriz multilocus alineada) fue determinado usando los criterios de información Bayesiano, la corrección de Akaike, y el valor de Máxima Verosimilitud (lnL) implementado en MEGA. Los filogramas, donde los largos de rama son proporcionales al número de sustituciones por sitio, se visualizaron con el mismo programa. Los valores de soporte de *bootstrap* se estimaron con 500 pseudoréplicas en MEGA; en las figuras correspondientes se indican los valores > 50 %.

En base a los resultados obtenidos (ver más adelante) se continuó trabajando con las secuencias no codificantes para seleccionar aquellas regiones particulares con mayor utilidad en análisis filogenéticos del género. Para esto, se realizaron alineamientos independientes para cada intrón e IGS, incorporando las secuencias de especies adicionales de *Ilex*, según la disponibilidad en GenBank. Las especies adicionales incluidas (sin plastoma completo) y los números de entrada o accesoión están indicados en las **Tablas AI-4 y AI-5**.

Cada intrón e IGS se alineó independientemente en MEGA usando MUSCLE, registrándose la cantidad de sitios variables y conservados. Con estos valores se calculó una tasa, normalizada por el número de especies, que describe el nivel de variación de cada región (número de sitios variables / número de sitios constantes / número de especies consideradas por región). Además, se obtuvieron valores de distancias pareadas (distancia *p* sin corregir = número de sitios variables / número total de sitios). Los promedios de las tasas normalizadas y de las distancias se utilizaron como valores

umbral de selección. Así, aquellos alineamientos que presentaron valores mayores a los umbrales establecidos fueron considerados como marcadores promisorios.

## Resultados

### 1. Genotipado preliminar

Se contabilizaron 1-3 alelos por *locus* cpSSR heterólogo (**Tabla 3**), tanto para cada especie, como considerando ambas.

Se hallaron 3 *loci* bialélicos y uno trialélico en *I. dumosa*, que formaron 4 haplotipos distintos. Las muestras de *I. paraguariensis* solo mostraron variación para un *locus*, con 2 alelos formando 2 haplotipos (**Tabla 3**). Las muestras provenientes de PINDÓ S.A. no mostraron variación ni representaron un haplotipo distinto, por lo que no se las siguió analizando.

Los individuos seleccionados (**Tabla 1** en negrita) de aquí en adelante serán denominados: Id55, Id222, Id227, IdCRV7, Ip30, Ip217, Ip224 e IpJB (Id: *I. dumosa*; Ip: *I. paraguariensis*).

**Tabla 3.** Número de individuos genotipados por especie, y número de alelos observados por especie y en total.

	<i>ccmp2</i>	<i>ccmp3</i>	<i>ccmp4</i>	<i>ccmp6</i>	<i>ccmp7</i>	<i>ccmp10</i>
<b><i>I. dumosa</i></b>						
N° observado	16	14	15	16	16	14
Tamaños alélicos	186, 194	108, 110, 112	126, 128	100	133	100, 107
N° alelos	2	3	2	1	1	2
<b><i>I. paraguariensis</i></b>						
N° observado	26	26	27	25	27	28
Tamaños alélicos	194	108	126, 128	100	121	103
N° alelos	1	1	2	1	1	1
<b><i>I. paraguariensis</i> Pindó S.A.</b>						
N° observado	19	20	20	18	18	19
Tamaños alélicos	194	108	126	100	121	103
N° alelos	1	1	1	1	1	1
<b>N° total de alelos</b>	<b>2</b>	<b>3</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>3</b>

### 2. Obtención de plastomas

#### 2.1. Secuenciación WGS y ensamblado

A modo de resumen, en el INDEAR se produjeron en total 209.464 lecturas de tamaño promedio 617,85 pb (desviación estándar aproximada 230 pb), totalizando 129,4 Mb secuenciadas. El valor de modal de la corrida fue de 709 pb.

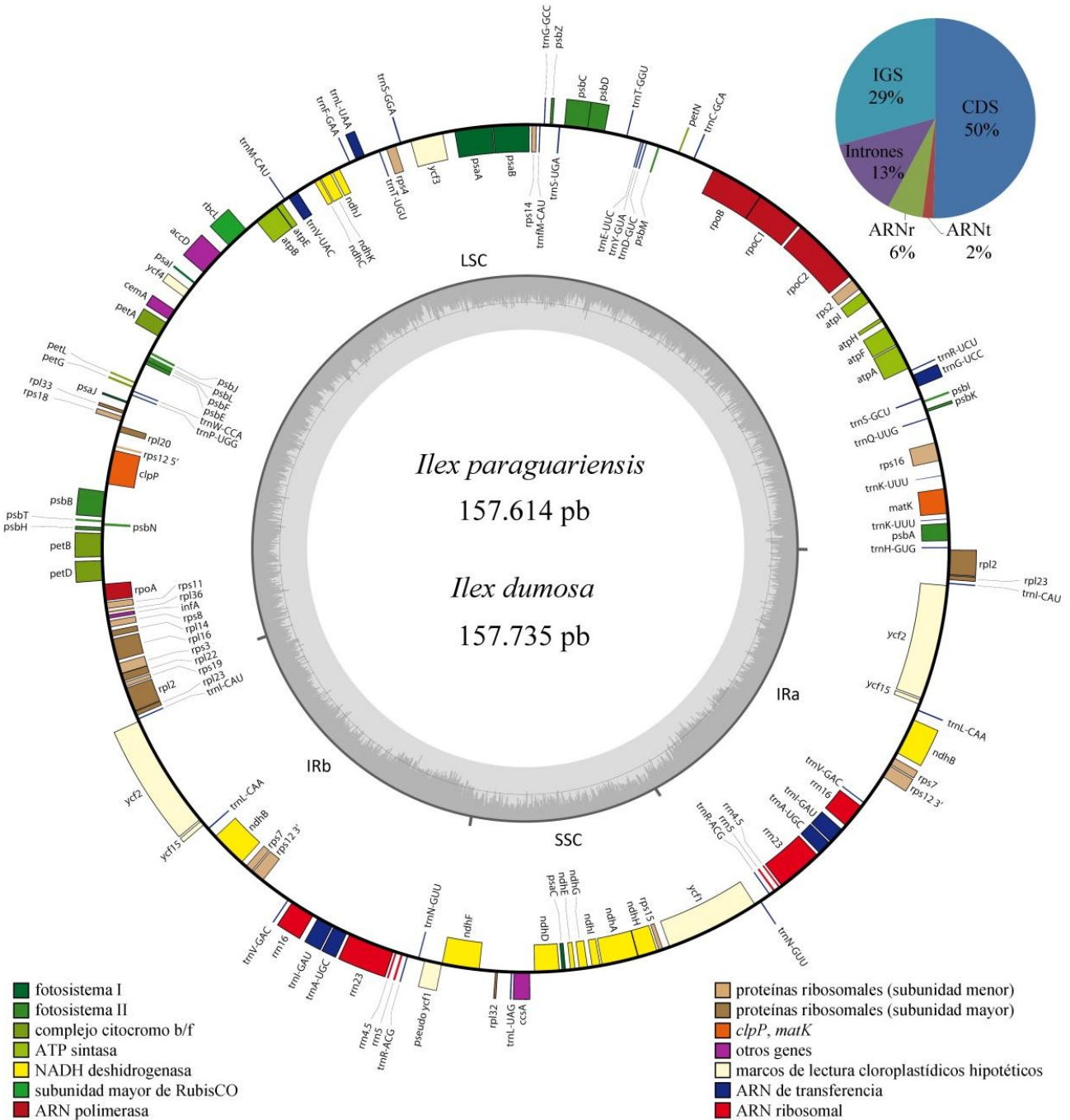
El servicio de bioinformática del INDEAR identificó 8.496 lecturas pertenecientes a cloroplastos (7.042 de *I. paraguariensis* y 1.454 de *I. dumosa*), representando un 4,056 % del total de lecturas iniciales, y sumando 5,4 Mb. De éstos, 724.533 pb pudieron ser ensamblados en 157 contigs, 101 corresponden a *I. dumosa* (232.018 pb) y 56 a *I. paraguariensis* (492.515 pb) (**Tabla AI-6**).

Los *contigs* ensamblados por el servicio fueron re-evaluados en nuestro laboratorio en búsqueda de secuencias contaminantes que fueron eliminadas. Con los datos así procesados se generaron los consensos, como se indicó en Materiales y Métodos.

## 2.2. Caracterización de los plastomas consenso

El plastoma consenso completo de *I. paraguariensis* consta de 157.614 pb, con una profundidad de secuenciación de 38X; mientras que el de *I. dumosa* consiste en 157.735 pb, con una profundidad de 8X.

Las regiones plastómicas mostraron tamaños similares: la región LSC consistió en 87.144 y 87.134 pb, la SSC en 18.307 y 18.415 pb, y los IR en 26.077 y 26.087 pb, para *I. paraguariensis* e *I. dumosa*, respectivamente (**Figura 9**). Los plastomas generados muestran sintenia (igual orden de los genes) e igual número de CDS, ARNt, ARNr, intrones e IGSs (**Tabla 4, Figura 9**); el 58 % del plastoma está representado por genes, mientras que el 42 % son secuencias no codificantes. De las 132 secuencias codificantes, 18 están duplicadas por estar ubicadas en los IRs. Se detectaron dos pseudogenes, el  $\psi$  *ycf1* (1.055 pb) localizado en la unión IRb-SSC, y el  $\psi$  *rps19* (5 pb) en la unión IRa-LSC. Los 20 intrones identificados están distribuidos en 10 CDS y 6 ARNt, cada uno con un intrón, y los genes *clpP* e *ycf3* con dos intrones cada uno.



**Figura 9.** Mapa de los plastomas de *I. paraguariensis* e *I. dumosa*. Se muestran las repeticiones invertidas (IRa e IRb) y las regiones de copia simple (LSC y SSC). Los genes dibujados en la parte interna son transcritos en sentido horario, y los dibujados por fuera del círculo se transcriben en sentido opuesto. El círculo interno representa el contenido GC (en gris oscuro). Las categorías funcionales de los genes se indican con colores. En la esquina superior derecha se muestra la proporción de CDS, ARNt, ARNr, intrones e IGS presente en ambas especies.

**Tabla 4.** Genes detectados en los genomas plastídicos de *I. dumosa*, *I. paraguariensis* e *I. cornuta*

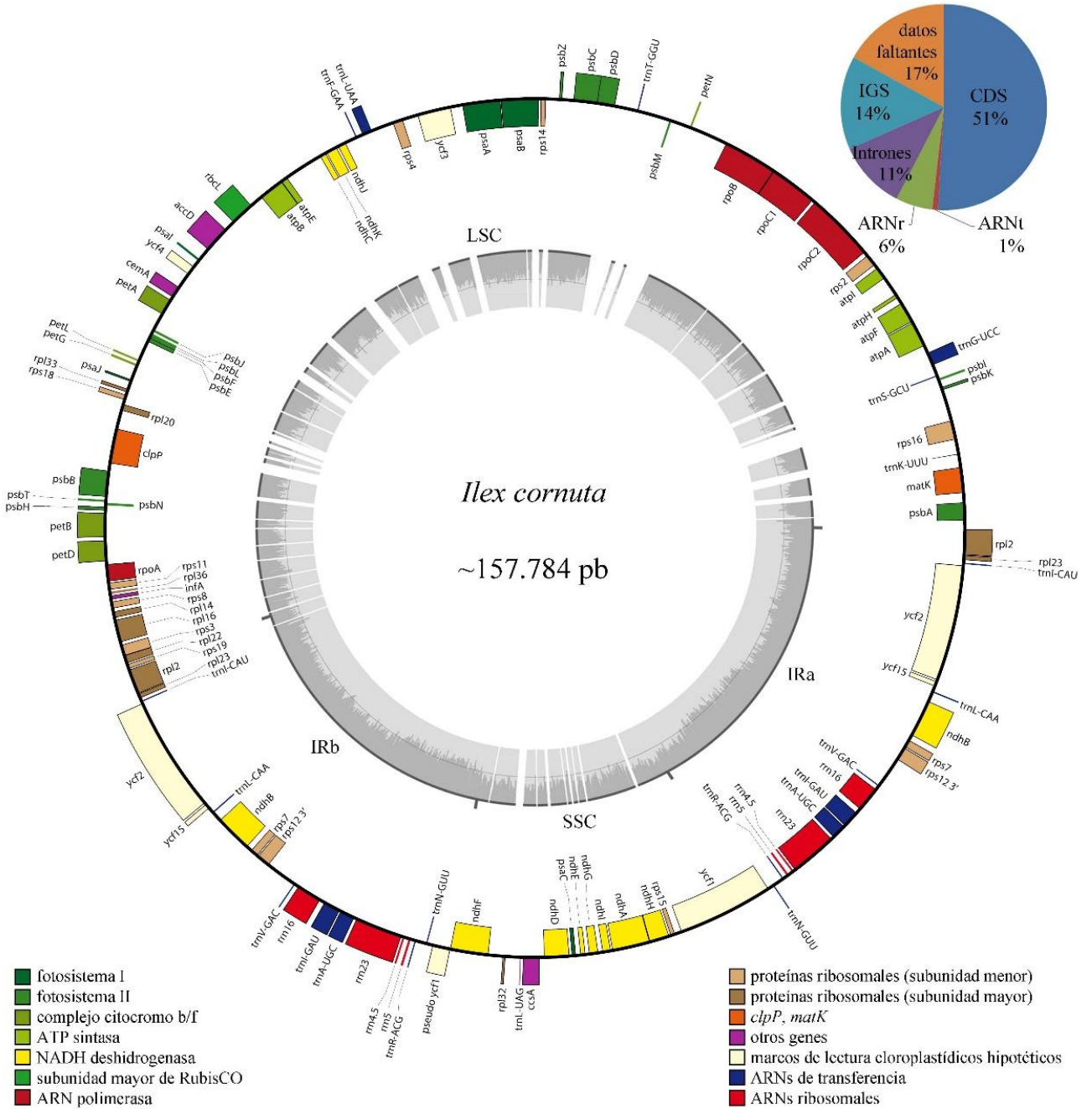
Grupo de genes	Identificación
Subunidad ribosomal menor	<i>rps2</i> , 3, 4, 7, 8, 11, <b>12<sup>1</sup></b> , 14, 15, 16 <sup>1</sup> , 18, 19
Subunidad ribosomal mayor	<b>rpl2<sup>1</sup></b> , 14, 16 <sup>1</sup> , 20, 22, <b>23</b> , 32, 33, 36
Subunidades de la ARN polimerasa	<i>rpoA</i> , B, C1 <sup>1</sup> , C2
NADH deshidrogenasa	<i>ndhA</i> <sup>1</sup> , <b>B<sup>1</sup></b> , C, D, E, F, G, H, I, J, K
Fotosistema I	<i>psaA</i> , B, C, I, J, <i>ycf3</i> <sup>2</sup> , <i>ycf4</i>
Fotosistema II	<i>psbA</i> , B, C, D, E, F, H, I, J, K, L, M, N, T, Z
Complejo citocromo b/f	<i>petA</i> , B <sup>1</sup> , D <sup>1</sup> , G, L, N
ATP sintasa	<i>atpA</i> , B, E, F <sup>1</sup> , H, I
Subunidad grande de RuBisCo	<i>rbcL</i>
Factor de iniciación de la traducción	<i>infA</i>
Maturasa	<i>matK</i>
Proteasa	<i>clpP</i> <sup>2</sup>
Proteína de la membrana interna	<i>cemA</i>
Subunidad acetil-CoA carboxilasa	<i>accD</i>
Gen de síntesis de citocromo de tipo C	<i>ccsA</i>
Marco de lectura abierto conservado (ORF) de función desconocida	<b><i>ycf1</i>, 2, 15</b>
ARN de transferencia (ARNt)	<b><i>trnA</i>-UGC<sup>1</sup></b> , C-GCA, D-GUC, E-UUC, F-GAA, G-GCC, G-UCC <sup>1</sup> , H-GUG, I-CAU, I-GAU <sup>1</sup> , K-UUU <sup>1</sup> , L-CAA, L-UAA <sup>1</sup> , L-UAG, M-CAU, fM-CAU, N-GUU, P-UGG, Q-UUG, R-ACG, R-UCU, S-GCU, S-GGA, S-UGA, T-GGU, T-UGU, V-GAC, V-UAC <sup>1</sup> , W-CCA, Y-GUA
ARN ribosomales (ARNr)	<b><i>rrn4.5</i>, 5, 16, 23</b>

Los genes localizados en las regions IR se indican en negrita. <sup>1</sup> gen con un intrón; <sup>2</sup> gen con dos intrones. Los ARNt con subrayado simple no están disponibles para *I. cornuta*, y aquellos con secuencias parciales se muestran con subrayado doble.

### 2.3. *Plastoma borrador de I. cornuta*

El plastoma esquemático construido para *I. cornuta* posee el 100 % de los CDS y ARNr encontrados en *I. dumosa* e *I. paraguariensis*, el 95 % de los intrones, el 46,7 % de los ARNt, y el 30 % de las regiones IGS. Este plastoma tiene 16,8 % de datos faltantes pero composición relativa similar a la yerba mate y yerba señorita (**Tabla 4, Figura 10**).





**Figura 10.** Mapa del plastoma borrador de *I. cornuta*. Los genes dibujados en la parte interna del mapa son transcritos en sentido horario, y los de afuera se transcriben en sentido contrario. El círculo interno representa el contenido GC (en gris oscuro), las regiones en blanco indican las zonas faltantes. Las categorías funcionales de los genes son indicadas con colores. En la esquina superior derecha se muestra un gráfico de torta con la proporción de CDS, ARNt, ARNr, intrones, IGS y datos faltantes con respecto a *I. dumosa* e *I. paraguariensis*.

### 3. Análisis de plastomas

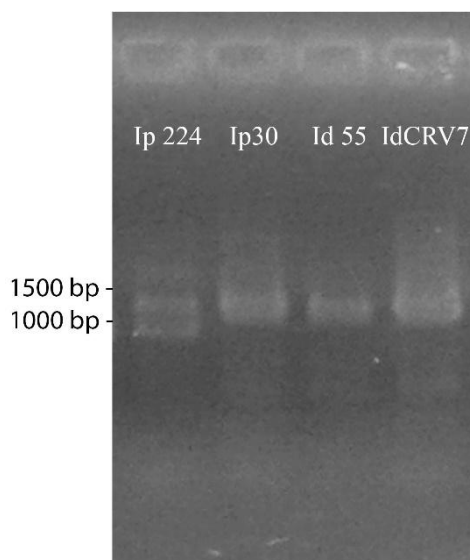
#### 3.1. Conservación nucleotídica intraespecífica

La distancia genética promedio calculada entre individuos de *I. paraguariensis* (distancia  $p$  media=  $1,61 \times 10^{-3}$ ) resultó ser aproximadamente la mitad de la media calculada para los individuos de *I. dumosa* (distancia  $p$  media=  $3,11 \times 10^{-3}$ ) (**Tabla 5**).

**Tabla 5.** Distancias genéticas intraespecíficas pareadas (distancias  $p$ ) calculadas entre plastomas individuales, incluyendo una sola región repetida invertida. Por encima de la diagonal y en itálicas, se indican los valores obtenidos entre individuos de *I. dumosa*. Debajo de la diagonal, se muestran los valores entre individuos de *I. paraguariensis*.

<i>I. dumosa</i>					
	<i>IdCRV7</i>	<i>Id55</i>	<i>Id222</i>	<i>Id227</i>	
<b>Ip30</b>	---	<i>1,01 x 10<sup>-3</sup></i>	<i>2,91 x 10<sup>-3</sup></i>	<i>4,3 x 10<sup>-3</sup></i>	<i>IdCRV7</i>
<b>Ip217</b>	<i>1,56 x 10<sup>-3</sup></i>	---	<i>2,86 x 10<sup>-3</sup></i>	<i>4,61 x 10<sup>-3</sup></i>	<i>Id55</i>
<b>Ip224</b>	<i>1,99 x 10<sup>-3</sup></i>	<i>9,4 x 10<sup>-4</sup></i>	---	<i>2,94 x 10<sup>-3</sup></i>	<i>Id222</i>
<b>IpJB</b>	<i>2,34 x 10<sup>-3</sup></i>	<i>1,41 x 10<sup>-3</sup></i>	<i>1,42 x 10<sup>-3</sup></i>	---	<i>Id227</i>
	<b>Ip30</b>	<b>Ip217</b>	<b>Ip224</b>	<b>IpJB</b>	
<i>I. paraguariensis</i>					

Al verificar la unión entre las regiones IRa y LSC, de manera azarosa se encontró variación en el tamaño del fragmento amplificado en el individuo Ip224, que mostró 2 amplicones: uno esperado de *ca.* 1200 pb y otro más pequeño de *ca.* 1000 pb (**Figura 11**). El resto de los individuos mostraron la banda de 1200 pb. La existencia en Ip224 del amplicón más corto fue corroborada mediante aislamiento del fragmento y secuenciación Sanger tradicional, y también por el hallazgo en los datos crudos de secuenciación WGS de la secuencia de ~1200 pb y de la más corta. La diferencia entre los dos fragmentos involucra un *indel* de 247 pb. Al investigar su ocurrencia en otras especies mediante la búsqueda BLAST en bases de datos públicas, no se hallaron resultados (todos los aciertos poseían los 247 pb).



**Figura 11.** Gel de agarosa donde se observan los productos de amplificación de la región que une IRa-LSC.

### 3.2. Repeticiones

Los plastomas de *I. dumosa* e *I. paraguariensis* poseen nueve repeticiones directas de entre 30 y 60 pb. *I. dumosa* posee además dos repeticiones palindrómicas y tres repeticiones invertidas; mientras que *I. paraguariensis* tiene tres palindrómicas y dos invertidas (33-48 pb). En *I. cornuta* solamente se detectaron nueve repeticiones directas (30-60 pb).

Al inspeccionar los *loci* heterólogos de cpSSR usados inicialmente, se comprobó en la existencia de homopolímeros  $T_n$  en *ccmp2*, *ccmp3*, *ccmp4* y *ccmp10* (este último no posee secuencia disponible en *I. cornuta*), y de un homopolímero  $A_n$  en el producto de *ccmp7*. Para las tres especies se detectó que el *locus ccmp6* posee la secuencia “TTTCTTTTCTTTTTCTTTCT”, no ajustándose a un microsatélite.

Aplicando herramientas bioinformáticas, se pudieron detectar un total de 30 potenciales marcadores cpSSR en las regiones no génicas de *I. dumosa* e *I. paraguariensis* (81 % se localizan en regiones IGS y 19 % en intrones). El 91 % de los cpSSR detectados tienen un motivo de repetición mononucleótido, y el resto tienen motivo dinucleótido. La validación de 14 de dichos *loci* permitió contabilizar entre 1 y 4 alelos por *locus*, promediando 2,5 alelos por marcador para *I. dumosa* y 1,86 para *I. paraguariensis* (**Tabla 6**).

**Tabla 6.** Marcadores microsatélite cloroplastídicos (cpSSR) diseñados para *I. paraguariensis* e *I. dumosa*. Los cpSSR validados se presentan en la parte superior de la tabla.

Localización del microsatélite <sup>a</sup>	Cebador directo <sup>c</sup>	Cebador reverso <sup>c</sup>	Motivo de repetición <sup>d</sup>	Tamaño (pb) <sup>e</sup>	Ao Id <sup>f</sup>	Ao Ip <sup>f</sup>
<i>rps16</i>	ATAGACGGCTCATTGGGATAGA	GAAAGGATTAAGAAGCACCGAA	(A)n	467	4	2
<i>trnS-psbZ</i>	AGGTGTGAATATACTGGGCGAT	AATCCTCCAAGCTACCAACAAA	(A)n	263	2	2
<i>trnV(UAC)</i>	CGATGAACCAACCAGTCAGTTA	CTAGTCCAGAGGCAGACCTACG	(T)n	390	1	2
<i>ycf4-cemA</i>	TCTGTCTTGTTTTGCCACTCAT	TGATTTGGTGGTTAGTCCTTGA	(A)n	352	3	1
<i>rps18-rpl20</i>	AAGAATCGGGGAAGAAGAAATC	TGTGGATCACTCGAATAAATGC	(T)n	351	4	3
<i>trnR-atpA</i>	CGACTATAACCCCTAGCCTTCC	AATCAAGCGTTTCCGATTCA	(AT)n	462	4	1
<i>ndhH-rps15</i>	<b>TTCTTG TAGCTGGTGCAGTCAT</b>	<b>TGGTCAGTTGGATATTCGAGA</b>	<b>(A)n</b>	<b>153</b>	4	2
<i>atpH-atpI</i>	<b>CGTTCTTCCATTTCTTTGTTCC</b>	<b>CTCAACTCTTGTGGATGTTTCG</b>	<b>(TA)n</b>	<b>304</b>	1	1
<i>psbM-trnD</i>	CGCTCATTTCAATTAAGACGC	CCCATTTTGGATTCCGGTATG	(T)n	420	1	1
<i>psbI-trnS</i>	GCACTTCATTTCTTGGTGTCAA	AATCCCTCTCTTCCGTTTCC	(T)n	351	3	3
<i>trnP-psaJ</i>	CCACATCGTTATTTCTCCATT	CCTGCTAAAGACCCAAACCATA	(T)n	428	2	1
<i>accD-psaI</i>	<b>ATTATTT CAGCTCCACGCTTTC</b>	<b>CTAGTCAAAATGGGATTCTCGG</b>	<b>(T)n</b>	<b>381</b>	2	2
<i>rpl14-rpl16</i>	<b>CATACCGTTGTCACGTTTGAGT</b>	<b>TGAAATGGGTGGAGTAACAGAA</b>	<b>(T)n</b>	<b>447</b>	3	4
<i>infA-rps8</i>	<b>AACTACGTCGGATCTTTCCTGA</b>	<b>CGGGATAGGGATTGTAATTCTG</b>	<b>(T)n</b>	<b>351</b>	1*	1*
<i>trnH-psbA</i>	CCTTGATCCACTTGGCTACATC	CGTGCTAACCTTGGTATGGAAG	(T)n	486	nt	nt
<i>rps16-trnQ</i>	TGTCGAGTCAAGAGCACCTTTA	GGGAGCTGGGTAAATCTATCCT	(A)n	412	nt	nt
<i>trnG(UCC)</i>	CAATGACAGATTCGAGGAACAA	GTTTCCACCGAGCTAAAACAAA	(A)n	305	nt	nt
<i>psbC-trnS</i>	CCTGTTCTTTCCATGACTCCTC	ATCGCCCAGTATATTCACACCT	(A)n	447	nt	nt
<i>trnG-trnfM</i>	GTTGGCCGTGATAGTGTAGTGA	TACTTCGTGAAATGGTTCATGC	(A)n	413	nt	nt

Localización del microsatélite <sup>a</sup>	Cebador directo <sup>c</sup>	Cebador reverso <sup>c</sup>	Motivo de repetición <sup>d</sup>	Tamaño (pb) <sup>e</sup>	Ao Id <sup>f</sup>	Ao Ip <sup>f</sup>
<i>ycf3-trnS</i>	TTTCTCCTGAAGTTGTCTCGGAAT	CCTCCTATGCACACAAACACAA	(A)n	496	nt	nt
<i>trnM-atpE</i>	GGATKAGTTGCGTTTGGTTTTTC	AGTAAATGATGCGGAGAAGAGC	(T)n	497	nt	nt
<i>psaI-ycf4</i>	TTTAGTTTCGACATTCCTTGTGC	TTTTATGTTCGGTGGAAAGCC	(T)n	487	nt	nt
<i>rps12-trnV</i>	AGGATTGTACCGTGAGAGAAGC	TTCGATACGGGTAGGAGAAGAA	(T)n	496	nt	nt
<i>rrn5-trnR</i>	CTCTACTGCGGTGACGATACTG	CCAAGTTCCATAGCTTTTGGTC	(A)n	475	nt	nt
<i>trnK-rps16</i>	<b>GACTTGTGTTGGATTGGCACTA</b>	<b>GGAACTCTTAACAACCGAGGAA</b>	<b>(A)n</b>	<b>158</b>	nt	nt
<i>trnG(UCC)</i>	<b>TAGCTCGGTGAAAACAAAATTC</b>	<b>GGAAGGCTAGGGGTTATAGTCG</b>	<b>(T)n</b>	<b>320</b>	nt	nt
<i>atpA-atpF</i>	<b>ATCAAGACCATGAATACGAGCA</b>	<b>GAAAGTGGAAATGGAAGCAGAG</b>	<b>(A)n</b>	<b>497</b>	nt	nt
<i>trnM-atpE</i>	<b>GGATKAGTTGCGTTTGGTTTTTC</b>	<b>CTAGGACACGAGTAGAGGCCAT</b>	<b>(T)n</b>	<b>362</b>	nt	nt
<i>clpP</i> <sup>b</sup>	<b>TCAAATTACCACCCTTCCTTTC</b>	<b>TACCAAACGTCTAGCATTCCT</b>	<b>(T)n</b>	<b>467</b>	nt	nt
<i>trnR-atpA</i>	<b>GAATGAAAAGCGTCCATTGTCT</b>	<b>CAATCCCAATCAGCTCCTCTTA</b>	<b>(TA)n</b>	<b>477</b>	nt	nt

<sup>a</sup> las regiones intrónicas se muestran subrayadas. <sup>b</sup> el locus cpSSR se encuentra en el segundo intrón de *clpP*. <sup>c</sup> todos los pares de cebadores tienen una Tm de 60°C; K= G/T. Los cebadores diseñados para *I. dumosa* se marcan en negrita, y aquellos para *I. paraguariensis* están en formato plano. <sup>d</sup> el número de repeticiones es n >10 pb. <sup>e</sup> los tamaños fueron estimados a través de PCR *in silico*. <sup>f</sup> alelos observados en *I. paraguariensis* (Ao Ip) y en *Ilex dumosa* (Ao Id). Ocho individuos de cada especie fueron usados para la validación (ver Materiales y Métodos). \* Ambas especies muestran el mismo alelo, al menos en tamaño. nt, no testeado.

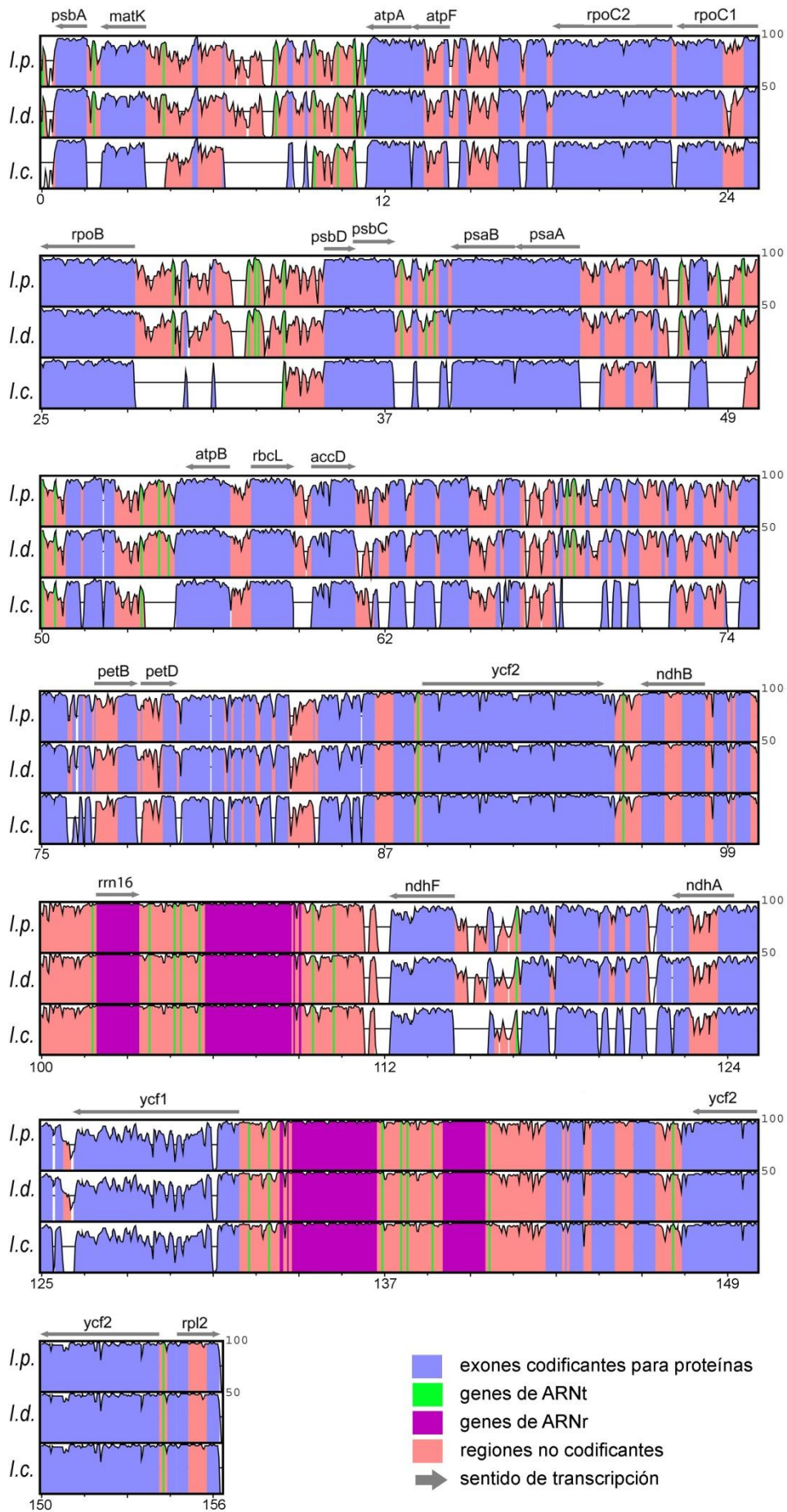
### 3.3. Conservación nucleotídica interespecífica

El alineamiento global entre los tres plastomas consistió en 158,3 Kpb, con 0,4 % de *gaps*. Como se esperaba, las regiones más conservadas fueron, en orden decreciente, los genes ribosomales, los ARNt, los CDS, los intrones y por último las regiones IGS (**Figura 12**).

La identidad estimada entre los plastomas de *I. dumosa* e *I. paraguariensis* es del 98,93 %. La similitud global promedio entre las tres especies es del 90,78 % (**Tabla 7**).

**Tabla 7.** Porcentaje de identidad pareada entre los plastomas de *Ilex* disponibles.

	<i>I. paraguariensis</i>	<i>I. dumosa</i>	<i>I. cornuta</i>
<i>I. paraguariensis</i>	-		
<i>I. dumosa</i>	98,93	-	
<i>I. cornuta</i>	86,66	86,76	-



**Figura 12.** (página anterior) Identidad entre los plastomas de *I. dumosa*, *I. paraguariensis* e *I. cornuta*, usando *P. ginseng* como referencia. Los picos y valles indican la identidad de secuencia entre el 50 % y el 100 % (línea inferior y superior, respectivamente) estimada respecto a la referencia. El eje x muestra la posición lineal en Kpb. En cada bloque se muestran algunos genes como guía, las flechas denotan el sentido de transcripción. Las regiones con datos faltantes respecto de la referencia se indican con espacios en blanco. Las regiones están coloreadas según función.

### 3.5. Utilidad filogenética de secuencias plastídicas

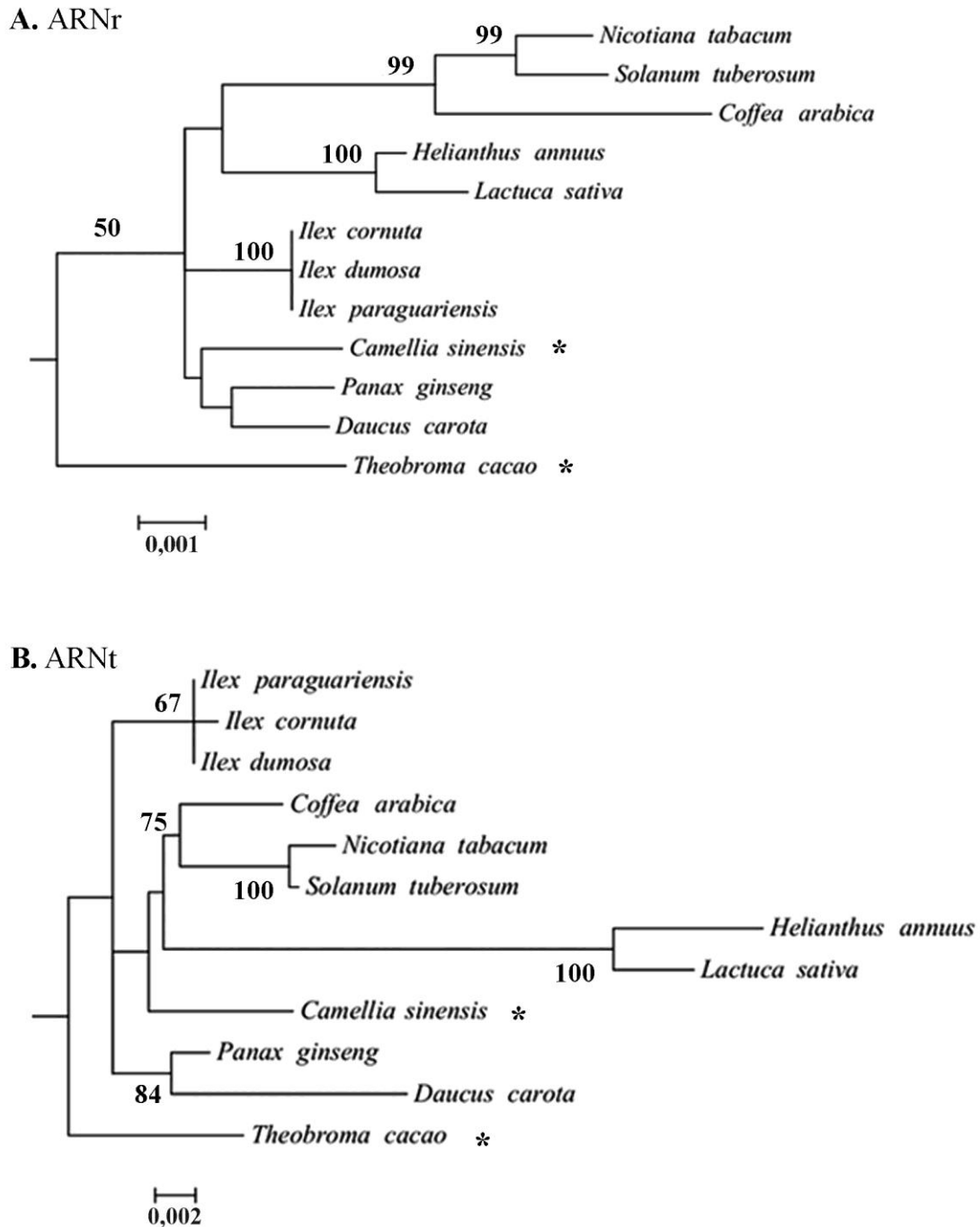
Los niveles de conservación nucleotídica de cada partición (**Tabla 8**) concuerdan con el orden cualitativo establecido a partir de la **Figura 12**.

**Tabla 8.** Características de los alineamientos y modelos de sustitución seleccionados para cada conjunto de datos.

Partición <sup>a</sup>	Nº reg. <sup>b</sup>	Alineamiento			Modelo de sust. nucleotídica		
		largo (pb)	cons. <sup>c</sup>	var. <sup>c</sup>	nombre <sup>d</sup>	G <sup>e</sup>	I <sup>f</sup>
ARNr	4	4536	97,55	2,23	HKY	0,1	--
ARNt <sup>g</sup>	30	2306	93,15	6,62	K2P	0,05	--
CDS	79	70687	76,40	22,08	GTR	0,67	35,84
Intrones <sup>h</sup>	16	16309	61,82	29,74	GTR	0,49	--
IGS <sup>h</sup>	35	27757	46,28	35,50	GTR	0,74	--

<sup>a</sup> cada partición consistió en secuencias de 3 especies de euastéridos I, 8 de euastéridos II, y dos *outgroups*. <sup>b</sup> número de regiones consideradas. <sup>c</sup> porcentaje sobre el largo total; cons., sitios conservados; var., sitios variables. <sup>d</sup> HKY, Hasegawa, Kishino y Yano (Hasegawa *et al.* 1985); K2P, Kimura 2-parámetros (Kimura 1980); GTR, *General Time Reversible* (Tavaré 1986). <sup>e</sup> G, parámetro Gamma = tasa de diferencias entre sitios. <sup>f</sup> I, proporción de sitios no variables. <sup>g</sup> para esta partición, *I. cornuta* tiene 53,3 % de datos faltantes. <sup>h</sup> las regiones se detallan en Materiales y Métodos.

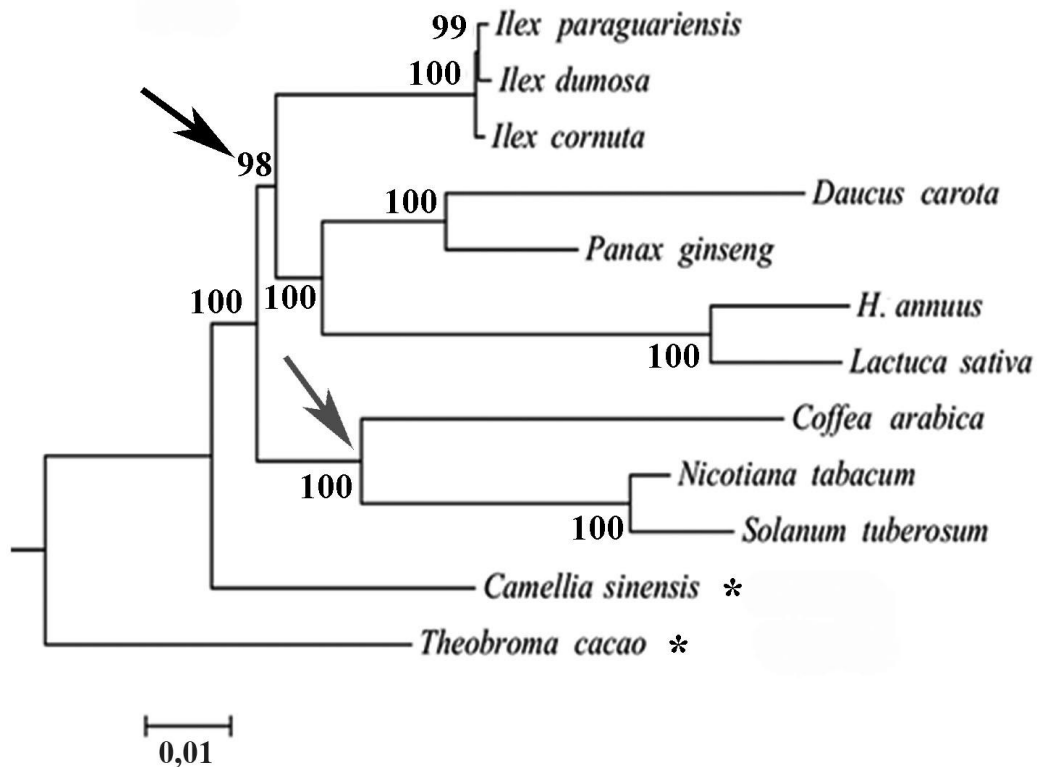




**Figura 13.** Análisis filogenómicos de Máxima Verosimilitud de 12 especies de angiospermas. **A)** Filograma derivado de datos de ARNr concatenados (lnL = -7157,76). **B)** Filograma obtenido usando ARNt concatenados (lnL = -4354,91). En ambos casos, todas las posiciones fueron consideradas. Los asteriscos indican los *outgroups*.

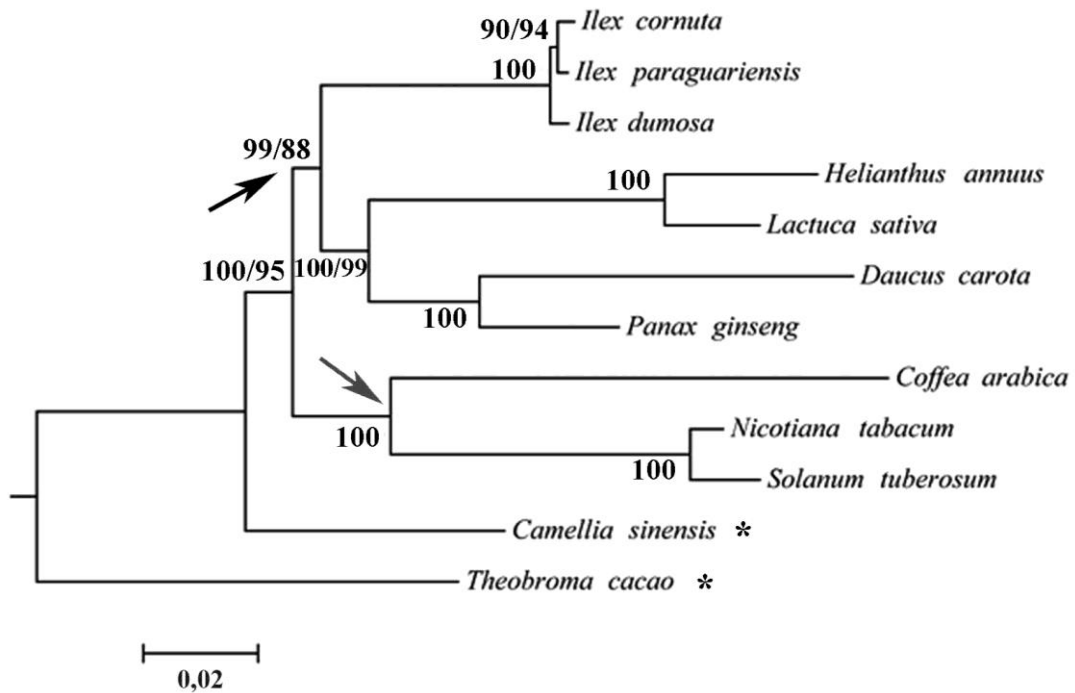
Los filogramas derivados de los datos de 4 ARNr y 30 ARNt muestran un clado *Ilex* (*bootstrap*= 100 % y 67 %, respectivamente) sin resolución y que los clados euastéridos I y II no se recuperan (**Figura 13**).

El filograma obtenido a partir de los 79 CDS muestra que el clado (*I. paraguariensis* + *I. dumosa*) (*bootstrap* = 99 %) sería el grupo hermano de *I. cornuta* (*bootstrap* = 100 %). Los clados euastéridos I y II están claramente definidos (BSV  $\geq$  98 %; **Figura 14**).

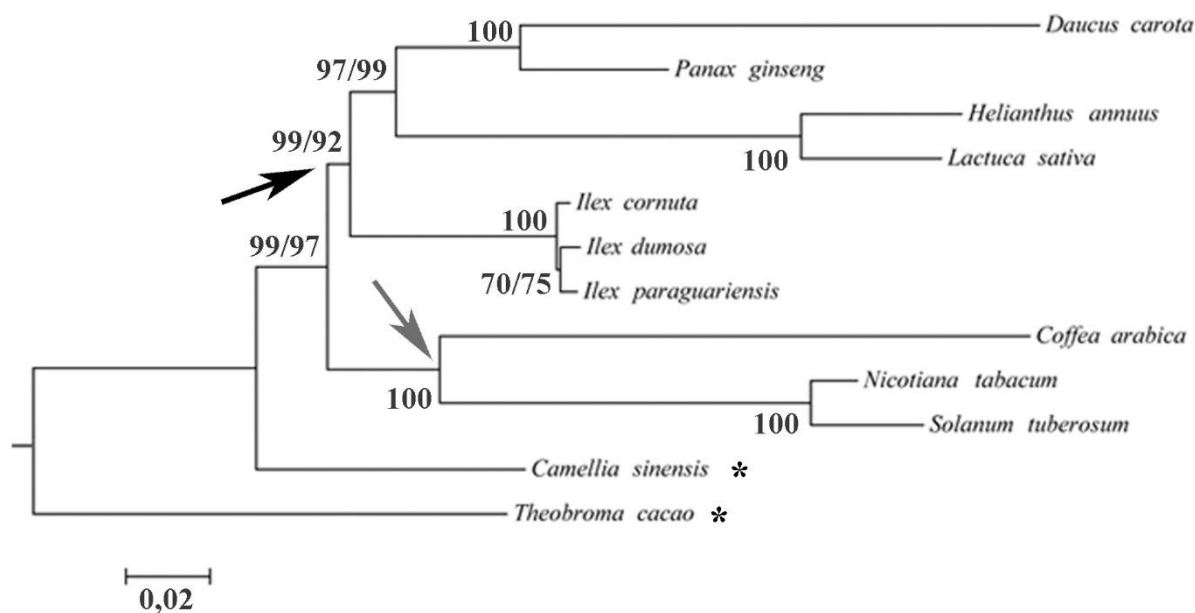


**Figura 14.** Filograma de Máxima Verosimilitud derivado de CDS concatenados (lnL = -221060,94); se consideraron todas las posiciones. La flecha gris indica el clado euastéridos I y la negra los euastéridos II. Los asteriscos indican los *outgroups*.

Los filogramas inferidos para los 16 intrones (**Figura 15**) y los 35 IGS (**Figura 16**) muestran un patrón de ramificación similar al derivado de los CDS. El clado de *Ilex* exhibe altos valores de *bootstrap*; sin embargo, el filograma derivado de intrones presenta una relación de grupos hermanos entre *I. paraguariensis* e *I. cornuta*, no obtenida con los CDS ni los IGS.



**Figura 15.** Filograma de Máxima Verosimilitud obtenido del análisis de intrones concatenados. En cada nodo, el primer valor corresponde al soporte *bootstrap* de la topología obtenida para los 16 intrones ( $\ln L = -53386,52$ ; 13.966 posiciones fueron consideradas por presentar  $\geq 50\%$  de cobertura), el segundo valor al obtenido para la topología de 7 intrones ( $\ln L = -25718,91$ ; todas las posiciones fueron consideradas), los valores únicos son valores concordantes. La flecha gris indica el clado euastéridos I y la negra los euastéridos II. Los asteriscos indican los *outgroups*.



**Figura 16.** Filograma de Máxima Verosimilitud obtenido del análisis de IGS concatenados. En cada nodo, el primer valor corresponde al soporte *bootstrap* de la topología obtenida para los 35 IGS ( $\ln L = -83248,29$ ; se consideraron 18.239 posiciones con  $\geq 50\%$  de cobertura), el segundo valor corresponde al obtenido para la topología de 9 IGS ( $\ln L = -43146,55$ ; 8.070 posiciones consideradas), los valores únicos son valores concordantes. La flecha gris indica el clado euastéridos I y la negra los euastéridos II. Los asteriscos indican los *outgroups*.

Al analizar los 16 intrones por separado, se calcularon tasas de variación nucleotídica normalizada de entre 0,0003 y 0,0058, con un promedio de 0,0020; el rango de distancia promedio fue 0,001 – 0,012, con una media de 0,0045. Para los intrones de los genes *atpF*, *clpP*, *ndhA*, *petD*, *rpl16*, *rps16* y *trnG* se calcularon niveles de variación por encima de los umbrales establecidos, por lo que se consideraron marcadores promisorios (**Figura 17A**). La topología del filograma derivado de los siete intrones selectos y concatenados concuerda con la de los 14 intrones, con soportes internos ligeramente menores (*bootstrap*  $\geq 88\%$ ; **Figura 15**), a excepción de la relación *I. paraguariensis*-*I. cornuta* cuyo valor aumentó levemente.

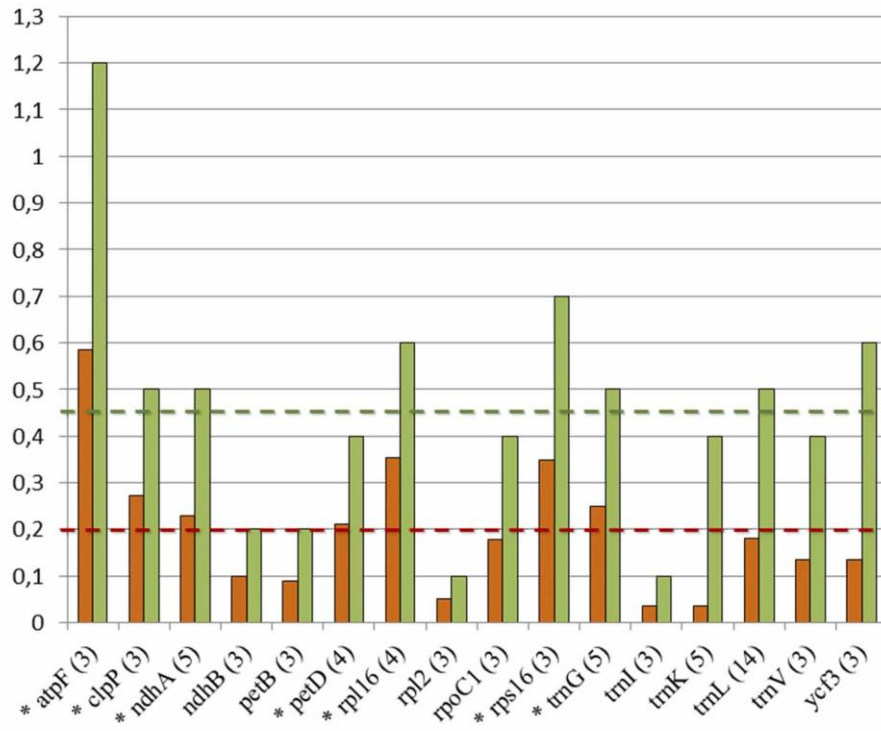
Las tasas de variación normalizada de los IGS oscilaron entre 0,0010 y 0,0105, promediando 0,0040. El rango de distancias *p* promedio fue 0,0013–0,0228 registrándose un promedio de 0,0078 (**Figura 17B**). Nueve regiones (*atpB-rbcL*, *atpH-atpI*, *ndhC-trnV*, *ndhF-rpl32*, *rpl32-trnL*, *trnH-psbA*, *trnK-rps16*, *trnS-trnG*, y *trnT-psbD*) mostraron valores por encima de al menos un umbral. El IGS *trnH-psbA* resultó

la región más variable (**Figura 17B**). La topología del filograma derivado de los nueve IGS selectos coincide con la de los 35 IGS originales (*bootstrap*  $\geq 75$  %; **Figura 16**).

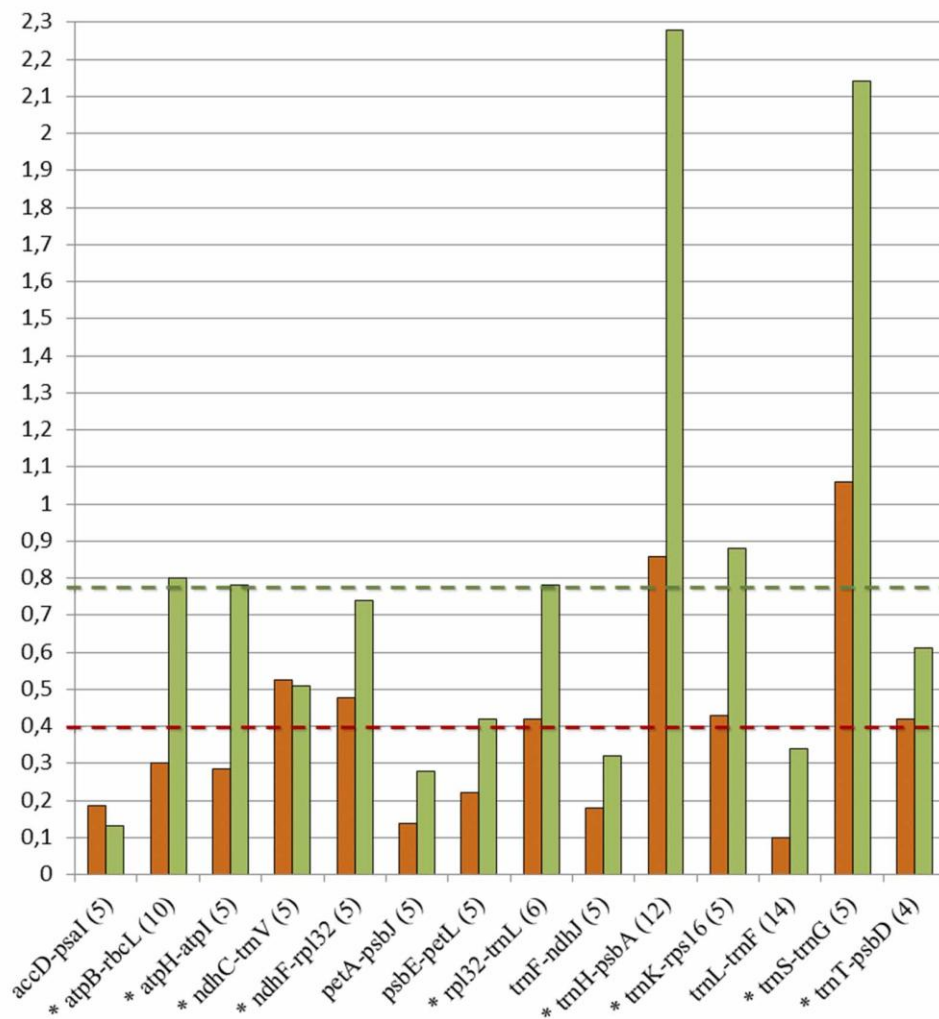
Además, al inspeccionar los alineamientos, se detectaron rearrreglos estructurales pequeños (>200 bp). Por ejemplo, en el IGS *ycf15-trnL* se encontró una región de 4 pb flanqueada por dos regiones complementarias de 21 pb, con potencialidad de formar una estructura de tipo “tallo-bucle” (“*stem-loop*” en inglés), y que posee la misma orientación en *I. dumosa* e *I. paraguariensis*. Otra estructura del mismo tipo está presente en el IGS *trnS-trnG*, en donde una región de 19 pb está flanqueada por dos regiones complementarias de 10 pb; en este caso, *I. paraguariensis* presenta la orientación contraria con respecto a *I. dumosa* e *I. cornuta*. También se hallaron *indels*, por ejemplo, exclusivamente en *I. paraguariensis* se encontró en el IGS *rpl32-trnL* una delección de 101 pb de largo; el BlastN de la región deleccionada sólo obtuvo resultados con *Ilex*.

**Figura 17.** (*Página siguiente*) Estudio de la variación nucleotídica de las regiones intrónicas e IGS. Los histogramas muestran en color naranja la tasa entre sitios variables y constantes, normalizada por el número de especies y multiplicada por 100; y en verde la distancia *p* promedio multiplicada por 100. El número de especies de *Ilex* usadas en cada alineamiento está indicado entre paréntesis. Las líneas punteadas marcan los valores umbrales. **A)** Valores estimados a partir de alineamientos de intrones individuales; **B)** Valores estimados a partir de alineamientos de IGS individuales. Los asteriscos indican las regiones seleccionadas.

**A.**



**B.**



## Discusión

En este capítulo se presentan los primeros plastomas completos de especies sudamericanas de *Ilex*, junto con un plastoma borrador de la especie ornamental asiática *I. cornuta*.

La profundidad de secuenciación obtenida fue menor a la esperada para tecnologías NGS, y en comparación con estudios que utilizaron tecnologías similares (Tangphatsornruang *et al.* 2010; Middleton *et al.* 2014) o con Yao *et al.* (2016b). Por tanto, es particularmente prudente considerar el plastoma de *I. dumosa* como un borrador (*draft*). Cabe destacar que se siguieron los mismos protocolos experimentales y bioinformáticos para ambas especies, y que el número de lecturas crudas fue similar (28.441 para *I. dumosa* y 27.958 para *I. paraguariensis*). Sin embargo, se obtuvieron pocas lecturas plastídicas (< 7 % en promedio; 0,79-2,19 % para los individuos de *I. dumosa* y 3,21-12,15 % para los de *I. paraguariensis*). Middleton *et al.* (2014) que secuenciaron mediante Roche-454 ADN genómico total y extrajeron las lecturas pertenecientes a cloroplastos, obtuvieron entre 1,96 % y 4,27 % de lecturas plastídicas, porcentajes similares a los obtenidos en este trabajo, pero con valores de cobertura más altos (22 - 92X, promedio 45,75X). Aquí, la baja cobertura podría deberse a artefactos experimentales combinados con la abundante presencia de metabolitos citosólicos, que pudieron interferir con el intento de aislamiento de los cloroplastos intactos y/o con los procedimientos posteriores. La presencia de metabolitos citosólicos interfirió en las mediciones de contenido de ADN por citometría de flujo en estas especies (Gottlieb y Poggio 2014), lo cual también fue reportado para *Coffea* (Noirot *et al.* 2003). En *Quercus spinosa* también se obtuvo contaminación con ADN nuclear y mitocondrial al aislar cloroplastos enteros, debido a la unión de los ácidos nucleicos de otros compartimentos a la membrana plastídica (Du *et al.* 2015). A pesar de estos aspectos negativos, los datos obtenidos y curados permitieron generar mapas plastómicos coherentes, realizar análisis filogenómicos, y contribuir con información molecular valiosa para el género *Ilex*.

En concordancia con la naturaleza conservada de los cloroplastos de las angiospermas (Raubeson y Jansen 2005; Jansen *et al.* 2007; Moore *et al.* 2010), los plastomas analizados muestran sintenia incluso con *P. ginseng* (Araliaceae), con la que poseen relaciones filogenéticas distantes. Además, muestran una alta conservación nucleotídica global, alta identidad de secuencia interespecífica (promedio = 90,78 %), y

largos de rama relativamente cortos en los análisis filogenómicos. Incluso esto se comprobó posteriormente al compararlos con los 6 plastomas de *Ilex* generados recientemente por Yao *et al.* (2016b): los plastomas de *I. dumosa*, *I. paraguariensis* y las especies asiáticas presentan un 96,2 % de identidad incluyendo el borrador de *I. cornuta*, y 99 % excluyéndolo (Cascales *et al.* 2017).

Las especies de *Ilex* estudiadas poseen el orden de genes ancestral, observado en plastomas de angiospermas que no sufrieron grandes reestructuraciones, como los de *Nicotiana* (Solanaceae) (Kim y Lee 2004; R Li *et al.* 2013; Ruhlman y Jansen 2014). En contraposición, para *Lactuca* y *Helianthus* (Asteraceae) Timme *et al.* (2007) mapearon sobre la región LSC dos inversiones sinapomórficas conspicuas (comprendiendo una región de 22,8 Kpb y una inversión anidada de 3,3 Kpb). El plastoma borrador de *I. paraguariensis* realizado por Debat *et al.* (2014), usando datos transcriptómicos y con 41,5 % de datos faltantes, fue ensamblado usando *Lactuca* como referencia; en consecuencia, su región LSC muestra erróneamente las dos grandes inversiones superpuestas. Por lo tanto, se debe tener cuidado de no confundir estos rearrreglos como características convergentes entre *I. paraguariensis* y las Asteraceae.

Se sabe que hubo eventos de reestructuración a distintas escalas (duplicaciones, inversiones, *indels*) y pseudogenizaciones que diversificaron los plastomas de las angiospermas (Raubeson y Jansen 2005; Ruhlman y Jansen 2014; Shaw *et al.* 2014). En este trabajo se detectaron únicamente pequeños rearrreglos estructurales. Particularmente la delección en el IGS *rpl32-trnL* podría ser considerada provisionalmente una autapomorfía de *I. paraguariensis*.

El individuo de *I. paraguariensis* Ip224 (único procedente de Paraguay) presenta variación de 247 pb en el tamaño de la región IRa-LSC, particularmente en el IGS *trnH-psbA*. En esta región se reportaron variaciones interespecíficas en *Ilex* (Selbach-Schnadelbach *et al.* 2009) debidas a *indels* de menor tamaño (31-74 pb), sin mencionar la variación intraespecífica ni intraindividual. Como Ip224 posee ambas secuencias (con y sin el *indel* de 247 pb), en este trabajo se interpretó como un caso de heteroplasmia plastídica. Se conocen casos de este fenómeno en un gran número de taxones, como *Gossypium*, *Medicago*, *Oenothera*, *Oryza*, *Pelargonium*, *Actinidia*, *Senecio* y *Phoenix* (Sabir *et al.* 2014); siendo más frecuente en las secuencias no codificantes y por variación en el largo de *indels* (Frey *et al.* 2005), como en el presente caso de Ip224. En *Senecio vulgaris* se halló una gran variación en el grado de heteroplasmia entre cloroplastos de la misma hoja, de distintas hojas de la misma planta y entre plantas,



sugiriendo que es un fenómeno común en esa especie (Frey *et al.* 2005). Aquí se encontró un caso en un individuo de los cuatro analizados, por lo que la heteroplasmia podría ser común en yerba mate, o estar confinada a una región geográfica. Desde un punto de vista filogeográfico, resultaría interesante investigar si es un caso aislado o si sucede en más individuos de *I. paraguariensis*, analizando muestras con distintos orígenes geográficos.

El promedio de las distancias pareadas intraespecíficas para los plastomas de *I. dumosa* resultó el doble del calculado para los plastomas de los individuos de *I. paraguariensis*. En concordancia, el estudio preliminar realizado aquí sobre la variación cpSSR intraespecífica aún con cebadores heterólogos permitió la detección del doble de haplotipos en *I. dumosa* respecto de *I. paraguariensis*. Rendell y Ennos (2003), en el único estudio preexistente que evalúa la variabilidad cpSSR en *Ilex*, también hallaron un único alelo para *ccmp6* en *I. aquifolium*, reforzando la propuesta hecha aquí de que no se trata de un microsatélite en *Ilex*. De hecho, recientemente hemos podido constatar que este *locus* posee secuencia idéntica en *I. cornuta*, *I. wilsonii*, *I. delavayi*, *I. polyneura*, *I. pubescens*, *I. latifolia* e *I. aquifolium*; mientras que *I. szechwanensis* muestra una única sustitución. Además, Rendell y Ennos (2003) encontraron baja variabilidad plastídica: solo 8 haplotipos en 16 poblaciones distribuidas por Europa. También usando los marcadores *ccmp*, Mohanty *et al.* (2002) analizaron poblaciones europeas de *Prunus spinosa* hallando solo tres haplotipos; mientras que la genotipificación de las mismas poblaciones con otros marcadores plastídicos permitió detectar 24 haplotipos. Estos ejemplos resaltan la necesidad de desarrollar marcadores moleculares específicos con el nivel de variación acorde al rango taxonómico de interés. En este sentido, se espera que los marcadores microsatélite generados en este estudio contribuyan a la realización de análisis poblacionales incluso en otras especies de *Ilex*.

Como se mencionó anteriormente, los estudios filogenéticos llevados a cabo en el género usando marcadores universales y con baja variación nucleotídica derivaron en relaciones filogenéticas inestables y topologías pobremente resueltas (Cuénoud *et al.* 2000; Setoguchi y Watanabe 2000; Manen 2004; Gottlieb *et al.* 2005; Loizeau *et al.* 2005; Selbach-Schnadelbach *et al.* 2009; Manen *et al.* 2010; Shi *et al.* 2016). Aquí se examinaron cuidadosamente grandes cantidades de datos plastídicos para identificar marcadores realmente informativos en *Ilex*. Debido a su alto nivel de conservación, las particiones de datos de ARNr y ARNt resultaron ineficientes para recuperar los linajes euastéridos I y II. De las particiones de datos de CDS, intrones e IGS se recuperaron

hipótesis filogenómicas concordantes en cuanto a que las Aquifoliales aparecieron como grupo hermano de los (Apiales + Asterales). Un resultado similar fue obtenido por Yao *et al.* (2016a), aunque interpretando erróneamente como “basal” la ubicación de las Aquifoliales.

En este trabajo se encontró concordancia topológica entre los datos de 16 intrones y el subconjunto de 7, implicando que futuros estudios podrían optar por usar este subconjunto. Sorprendentemente, el intrón *trnL*, que es usado rutinariamente en una variedad de especies de plantas (ej., Dangi *et al.* 2016), no agrega sitios segregantes incluso cuando el número de especies de *Ilex* aumentó (Cascales *et al.* 2017). En cambio, los intrones de *petD*, *rpl16*, *rps16* y *trnK*, frecuentemente utilizados en otras angiospermas (Borsch y Quandt 2009), también fueron seleccionados en este trabajo, siendo que la región intrónica *rpl16* es considerada una de las más variables entre las angiospermas (Shaw *et al.* 2014).

Asimismo, se halló concordancia topológica entre los análisis derivados de los 35 IGS y los 9 IGS seleccionados aquí. Un análisis paralelo a este trabajo de Tesis, centrado en estos últimos marcadores, probó la aptitud de éstos para realizar análisis filogenéticos más profundos en *Ilex*, con la inclusión de varias especies sudamericanas y euroasiáticas (Cascales *et al.* 2017). En concordancia con los resultados de otros autores, los IGS *ndhC-trnV*, *ndhF-rpl32*, *rpl32-trnL* y *trnT-psbD* se ubican entre las regiones más variables (Shaw *et al.* 2007, 2014), precedidos por *trnH-psbA*. Este último IGS también resultó el más variable para Shaw *et al.* (2007) y Timme *et al.* (2007). En cambio, Selbach-Schnadelbach *et al.* (2009), analizando 87 especies de *Ilex* concluyeron que el *trnH-psbA* muestra una baja tasa evolutiva. Por su parte, Shi *et al.* (2016) investigando la naturaleza híbrida de *Ilex dabiieshanensis* y usando este marcador, encontraron solo dos sustituciones entre los parentales propuestos (*I. cornuta* e *I. latifolia*). En esta Tesis, el análisis independiente de *trnH-psbA* produjo topologías no resueltas (no mostrado), por lo que, al menos en *Ilex*, es necesario analizar varias regiones en simultáneo para ganar resolución, como lo realizado en Cascales *et al.* (2017). Una mención especial requiere el IGS *trnL-trnF*, que fue usado por Setoguchi y Watanabe (2000) para estudiar 20 especies de *Ilex* asiáticas, por Manen (2004) para analizar 2 especies de las Islas Canarias y de África, y por Manen *et al.* (2010) para analizar filogenéticamente 108 especies de distintos orígenes geográficos. Sin embargo, este IGS, al igual que el intrón *trnL*, no reúne la variación necesaria, ni siquiera con un muestreo más amplio (Cascales *et al.* 2017). Otro IGS muy empleado en *Ilex* es *atpB-*

*rbcL* (Cuénoud *et al.* 2000; Manen *et al.* 2002, 2010; Manen 2004), que, si bien fue seleccionado en este trabajo de Tesis, sus valores de variación fueron los más bajos entre las regiones seleccionadas.

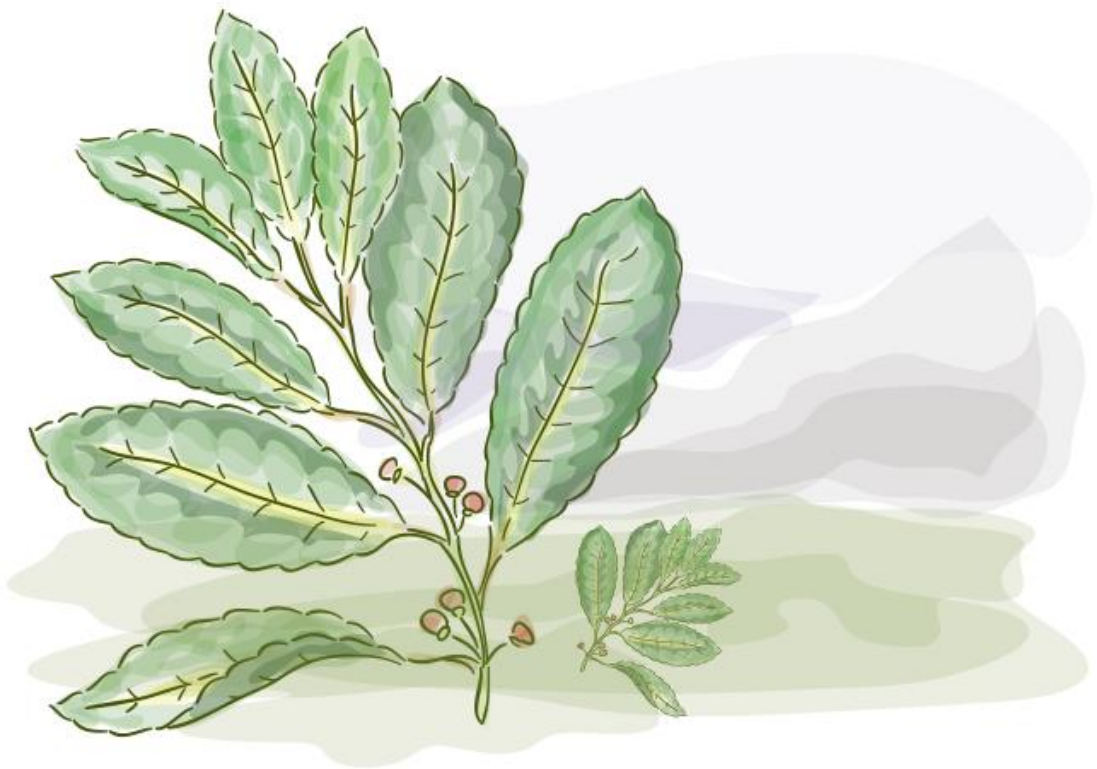
Yao *et al.* (2016b) recientemente propusieron 11 regiones plásticas distintas a las seleccionadas aquí como marcadores filogenéticos potenciales. Dichas regiones son compuestas, es decir contienen secuencias codificantes y no codificantes. Un análisis de esas regiones realizado paralelamente detectó que, aunque la exclusión de los genes codificantes aumentaba la variación de cada marcador, las tasas de variación normalizadas de las regiones fueron menores a los valores mínimos obtenidos para los intrones y los IGS seleccionados aquí (Cascales *et al.* 2017). En conclusión, la secuenciación de las regiones no codificantes como las seleccionadas en este trabajo sería preferible (más informativa y económica), al menos para filogenias a nivel de especie.

Dos de los análisis filogenéticos aquí realizados (con datos CDS e IGS) indicaron que los linajes maternos de *I. dumosa* e *I. paraguariensis* están más relacionadas entre sí que con el de *I. cornuta*. En contraposición, con secuencias intrónicas se obtuvo que los plastomas de *I. cornuta* e *I. paraguariensis* tendrían un ancestro común más reciente. Estas discrepancias pueden deberse al restringido muestreo de este trabajo y a que es sabido que las inferencias filogenéticas son condicionales al muestreo taxonómico. Un reciente estudio empleando un mayor número de especies del sur de Sudamérica y euroasiáticas evidenció al menos tres linajes maternos, uno sudamericano y dos euroasiáticos bien definidos, y que *I. dumosa*, *I. paraguariensis* se ubicaron en posiciones intermedias junto a *I. argentina* (Cascales *et al.* 2017).

Como ya fue expuesto en la Introducción de esta Tesis, se espera que el conocimiento generado facilite la implementación de la tecnología de transformación de cloroplastos (Koop *et al.* 2007; Verma y Daniell 2007; Ruhlman y Jansen 2014), particularmente en aquellas con atributos nutraceuticos y/u ornamentales. Esto, a su vez, aceleraría el desarrollo de variedades modernas que puedan satisfacer a los productores y a las exigencias del mercado.

## CAPÍTULO 2

### Estudio de patrones epigenéticos y expresión génica diferencial en la yerba mate y la yerba señorita



## Introducción

Las hojas de la yerba mate y de la yerba señorita son la principal materia prima que se emplea en la industria yerbatera. Estos órganos, cuya función primordial es fotosintética, son también centros de síntesis de los metabolitos secundarios tan apreciados por los consumidores del mate.

En el presente capítulo se aborda la detección de genes con algún tipo de participación en rutas metabólicas de potencial interés comercial, indagando la expresión génica en distintos estadios del desarrollo foliar. Tomando como premisa que en *Ilex* aún se desconocen los mecanismos de determinación sexual y aquellos que desencadenan el desarrollo o arresto de uno u otro conjunto de órganos en los primordios florales (Gottlieb y Poggio 2010), también se encara la detección de genes con algún tipo de participación en dichos procesos. Estos interrogantes son explorados metodológicamente a nivel epigenético y a nivel transcripcional.

### Orígenes del estudio de la epigenética

El material genético no contiene por sí solo toda la información necesaria para determinar qué genes se expresarán en un momento dado, sino que existen capas de información adicionales, englobadas bajo el término “epigenética”, que regulan la expresión como respuesta a distintas señales ambientales (Henderson y Jacobsen 2007; Jablonka y Raz 2009; Felsenfeld 2014). Es en este marco que en primer lugar nos proponemos explorar si existen mecanismos epigenéticos involucrados en el desarrollo de los órganos y en la diferenciación sexual de las especies en estudio.

La historia de la epigenética está vinculada con el estudio de la evolución y el desarrollo (Felsenfeld 2014). El sentido del término “epigenética” ha cambiado con el tiempo, y resulta interesante considerar cómo sus orígenes históricos se relacionan con el uso actual (Richards 2006). En el siglo XIX, la herencia y el desarrollo se estudiaban juntos; pero en el siglo XX, con el redescubrimiento de los estudios de Gregor Mendel (1822-1884) la genética comenzó a ser estudiada como una disciplina aparte. Sin embargo, hacia mediados del mismo siglo se volvió a considerar que ambas disciplinas estaban relacionadas, y es en este contexto que Conrad Waddington (1905-1975) utilizó por primera vez la palabra “epigenética” como un derivado de la palabra “epigénesis”, teoría del desarrollo que proponía que el embrión temprano no estaba diferenciado (Holliday 2006). Según exponen Goldberg *et al.* (2007), Waddington la definió en 1942

como “la rama de la biología que estudia las interacciones causales entre los genes y sus productos, que hacen que el fenotipo se convierta”.

Hasta la década de los cincuenta, la palabra epigenética se usó ampliamente para categorizar todos los eventos de desarrollo que llevaban desde el cigoto fecundado al organismo maduro (Waddington 1953). Inherente al uso original fue la visión de que los mecanismos epigenéticos son borrados y restablecidos en un punto en el ciclo de vida de un organismo (Richards 2006). Durante los últimos 50 años el significado del término ha sufrido una evolución paralela al conocimiento de los mecanismos moleculares subyacentes a la regulación de la expresión génica en eucariotas. Esto, sumado a la comprensión de que todas las células de un organismo llevan el mismo ADN, llevó a un cambio en la definición para centrarse en las formas en que los rasgos hereditarios pueden no estar asociados con cambios en la secuencia de nucleótidos, sino con modificaciones químicas del ADN o de las proteínas estructurales y reguladoras unidas a él (Felsenfeld 2014). Así se llegó a una definición operativa de la epigenética como “el estudio de los cambios en la función génica mitótica y/o meióticamente hereditarios que no pueden ser explicados por cambios en la secuencia del ADN” (Riggs y Porter 1996; Felsenfeld 2014). De manera más gráfica, Goldberg *et al.* (2007) definió que “la epigenética, en un sentido amplio, es un puente entre el genotipo y el fenotipo, un fenómeno que cambia el resultado final de un *locus* o cromosoma sin cambiar la secuencia de ADN subyacente”. La definición de Waddington en su sentido básico, sigue siendo pertinente al concepto moderno de la epigenética (McKeown y Spillane 2014).

En sentido amplio, la epigenética es una rama de la ciencia en la que se estudian los procesos moleculares que regulan la actividad genómica, independientemente de la secuencia del ADN, que son mitóticamente estables, reversibles, potencialmente influenciados por el medio ambiente (Angers *et al.* 2010; Skinner *et al.* 2010), y que estarían implicados en la determinación y diferenciación celular. En niveles más altos de organización biológica, los mecanismos epigenéticos generan las interacciones intercelulares que conducen a la homeostasis fisiológica y morfológica (Jablonka y Raz 2009).

### **Relación entre epigenética, fenotipo y ambiente**

Durante muchos años se consideró que la variación genética (es decir, la que surge de las diferencias alélicas) era la única fuente de variación fenotípica hereditaria. El

conocimiento de la epigenética ha aumentado dramáticamente en las últimas décadas; la evidencia acumulada indica que la variación epigenética también contribuye a explicar las diferencias fenotípicas entre individuos de la misma especie, y puede ser transmitida verticalmente de forma estable (Jablonka y Raz 2009; Medrano *et al.* 2014). Esta subdisciplina ha ganado atención para el estudio de variados procesos biológicos (Burggren y Crews 2014; Burggren 2016), especialmente como una conexión entre el medio ambiente y el fenotipo.

Cuando es posible, los individuos escapan de condiciones ambientales desfavorables. Alternativamente, el mantenimiento de la homeostasis fisiológica y/o la variabilidad genética individual pueden amortiguar los cambios ambientales. Se ha documentado que la regulación de la expresión de un gen específico o la activación de una vía génica alternativa a través de cambios epigenéticos pueden tener profundas repercusiones sobre el comportamiento, la fisiología, la morfología y el desarrollo (Angers *et al.* 2010). Los alelos epigenéticos (epialelos) son estados alternativos de metilación de un mismo gen, es decir, alelos que muestran diferencias hereditarias en la expresión como consecuencia de modificaciones epigenéticas y no de cambios en la secuencia del ADN (Henderson y Jacobsen 2007; Angers *et al.* 2010; Masuelli y Marfil 2011).

Las marcas epigenéticas pueden ser influenciadas por el medio ambiente (dieta, condiciones estresantes, etc.), ser independientes del fondo genético, y mostrar diferentes niveles de herencia mitótica y/o meiótica (Henderson y Jacobsen 2007; Angers *et al.* 2010). Al poder ser afectadas por el ambiente, por el genotipo y por su interacción, pueden considerarse rasgos fenotípicos (Richards 2006; Angers *et al.* 2010) que complejizan la partición de los componentes de la varianza fenotípica (varianza genética, varianza ambiental e interacción genotipo  $\times$  ambiente). En plantas, la evidencia disponible sugiere que una fracción importante de la variabilidad fenotípica de las poblaciones tiene origen epigenético (Varotto y Forestan 2010), sin embargo la contribución exacta de la variación epigenética a la varianza fenotípica sigue siendo difícil de determinar (Avramidou *et al.* 2015a).

Los epialelos, ya sean naturales o inducidos, pueden generar diversidad fenotípica heredable que ayude a la adaptación y sea objeto de la evolución (Masuelli y Marfil 2010; Fulneček y Kovařík 2014), confiriendo plasticidad al genoma, mejorando la amplitud ecológica y el rango de distribución de las especies (Sides *et al.* 2014). Por tanto, deberían ser considerados en los programas de mejoramiento genético de cultivos (Varotto y Forestan 2010). A su vez, contribuyen a aumentar la diversidad funcional de

las comunidades vegetales, un componente clave de la biodiversidad con importantes implicaciones para la coexistencia de especies y el funcionamiento de los ecosistemas (Medrano *et al.* 2014 y ref. allí citadas).

### **Procesos epigenéticos**

Actualmente se reconocen cuatro tipos de procesos moleculares epigenéticos que incluyen: (i) marcas en la cromatina (por ej., metilación de ADN, modificaciones post-traduccionales de histonas); (ii) moléculas de ARN no codificantes (ARNnc); (iii) estructuras particulares de proteínas, como los priones; y (iv) bucles (*loops*) metabólicos autosuficientes (es decir, circuitos metabólicos que por retroalimentación positiva pueden mantener diferentes patrones de actividad, y la perpetuación de estados celulares alternativos). Estos procesos, que pueden ser heredados a la siguiente generación celular, a menudo se interrelacionan mecánica y funcionalmente en una red epigenética interactiva compleja que, sin cambiar el código genético subyacente, modula y controla la expresión génica (Richards 2006; Jablonka y Raz 2009; Masuelli y Marfil 2011; Schulz *et al.* 2013; Skinner 2014). Existe evidencia de que las marcas en la cromatina y los ARNnc también pueden transmitirse a través de las gametas (Jablonka y Raz 2009).

El término cromatina incluye el ADN genómico y todos sus factores asociados físicamente: proteínas histónicas y no histónicas, moléculas de ARN, y pequeños grupos químicos unidos covalentemente al ADN (grupos metilo) y a histonas (por ej., grupos metilo y acetilo, grupos fosfato, ubiquitina). La actividad génica implica factores reguladores que reconocen y se unen a secuencias particulares en *enhancers*, promotores y otros sitios de control. Estos factores de transcripción son proteínas con dominios de unión a secuencias específicas del ADN acopladas a dominios que reclutan cofactores, que afectan directa o indirectamente a la expresión génica, incluyendo en muchos casos complejos de modificación de histonas o remodelación de nucleosomas. Una vez que esos estados iniciales de regulación se establecen, existen mecanismos epigenéticos que los mantienen (Felsenfeld 2014). A una escala mayor, la organización de la cromatina y los cromosomas, su localización en el núcleo y las interacciones dinámicas entre los diversos componentes determinan cuándo, dónde y en qué medida se transcriben los genes, entre otras funciones. Al proporcionar acceso diferencial a la información genética subyacente, regulan la expresión génica de manera pre-transcripcional (Richards 2006; Jablonka y Raz 2009; Masuelli y Marfil 2011; Felsenfeld 2014).



Las características de la heterocromatina (cromatina histoquímicamente densa y transcripcionalmente inactiva; Grewal y Jia 2007) dependen de las modificaciones post-traduccionales de las histonas y la metilación de bases de citosina. Las modificaciones de las colas de las histonas controlan el empaquetado del ADN, regulando por lo tanto el acceso a los genes para la transcripción (Angers *et al.* 2010), como se mencionara anteriormente. El silenciamiento de las secuencias de elementos transponibles (TEs, por *transposable elements*) dentro de la heterocromatina es probablemente una estrategia de defensa del genoma (Henderson y Jacobsen 2007; McKeown y Spillane 2014; Ikeda y Nishimura 2015).

Los procesos epigenéticos tienen, además, la capacidad de promover mutaciones: la mutación puntual más frecuente es una conversión de C a T facilitada por la metilación; las variaciones en número de copias de los TEs son controladas por hipermetilación; los eventos de translocación y las inversiones están influenciados por modificaciones de las histonas, metilación del ADN y ARNnc. Los índices de mutación, transposición y recombinación son más bajos en la heterocromatina que en la eucromatina. Además, las secuencias repetidas consideradas “ADN basura” podrían ser portadoras de marcas epigenéticas, que pueden comunicarse a otras regiones del genoma. Por lo tanto, una marca alterada podría dar lugar a cambios hereditarios coordinados en la expresión de varios genes de manera simultánea, acelerando así la evolución adaptativa (Jablonka y Raz 2009). La epigenética se postula como un medio por el que el ambiente puede influir, en última instancia, sobre el origen de la variación genotípica (Jablonka y Raz 2009; Skinner 2015, y ref. allí citadas).

### **Herencia epigenética celular y transgeneracional**

Una propiedad importante de los fenotipos inducidos ambientalmente es que las variaciones asociadas a la regulación de genes no son necesariamente heredables: el gen siempre se transmite pero no necesariamente su estado de expresión (Angers *et al.* 2010). Sin embargo, cuando las variantes epigenéticas se heredan, pueden hacerlo por “herencia epigenética celular”, en referencia a la transmisión mitótica; o por “herencia epigenética transgeneracional”, mediante transmisión meiótica en el caso de organismos con reproducción sexual (Jablonka y Raz 2009; Skinner *et al.* 2010).

Cuando se hereda de manera transgeneracional, la variación en la metilación del ADN puede ser blanco de la selección natural, la deriva y la migración, incluso en ausencia de variación genética (Jablonka y Raz 2009; Angers *et al.* 2010). Esto

representa un desafío a la Síntesis Neo-Darwiniana, vigente desde la década de 1930 (Richards 2006). Casi todos los modelos de evolución actuales consideran que el ambiente es importante para el proceso de selección natural, pero que no altera los procesos moleculares de evolución. Sin embargo, la inclusión de la epigenética puede ayudar a explicar la discordancia entre la velocidad evolutiva y la frecuencia extremadamente baja de mutación (Skinner 2014, 2015): las variantes epigenéticas son usualmente inducidas cuando cambian las condiciones ambientales, por lo que muchos individuos en la población podrían adquirir modificaciones similares al mismo tiempo (Jablonka y Raz 2009).

Si el ambiente físico o comportamental de la célula u organismo influye en la formación de epialelos, existiría una fundamentación para la herencia “suave o blanda” de Mayr (1980). En plantas, los linajes celulares reproductivos (o germinales) derivan de linajes vegetativos (o somáticos) y el “borrado” epialélico es menos extenso, por lo que pueden darse casos de herencia suave, con epialelos transmitidos por mitosis o meiosis a la siguiente generación como “caracteres adquiridos” (Richards 2006; Angers *et al.* 2010; Skinner 2014, 2015). Según Skinner (2014) la transmisión transgeneracional proporciona “un concepto neo-lamarckiano que facilita la evolución darwiniana”. Por lo tanto, los cambios epigenéticos tendrían un papel importante en la microevolución (Day y Bonduriansky 2011) y contribuirían a los procesos macroevolutivos (Rebollo *et al.* 2010; Flatscher *et al.* 2012). Sobre esta base se ha propuesto una teoría extendida, que contempla estudios de desarrollo y herencia epigenética, y que incorpora marcos darwinistas, lamarckianos y saltacionales, pretendiendo reemplazar a la teoría sintética de la evolución (Jablonka y Raz 2009; Skinner 2015).

### **Metilación del ADN**

En el presente capítulo se estudian los patrones de metilación del ADN, por lo que el marco teórico se centrará en este tipo de marca epigenética, que se da en los tres dominios de la vida (bacterias, arqueas y eucariotas) (Jablonka y Raz 2009). Sus defectos son letales en la etapa embrionaria de mamíferos, y en las plantas pueden conducir a defectos morfológicos pleiotrópicos (Law y Jacobsen 2010).

En organismos superiores está involucrada en diferentes actividades celulares, como la regulación de la expresión génica durante el desarrollo y la diferenciación celular, la impronta genómica (*imprinting*), el mantenimiento de la diferenciación entre tejidos, la

formación de heterocromatina, el tiempo de replicación, la inactivación del cromosoma X (lionización) y la estabilización del genoma por represión de TEs (Reyna-Lopez *et al.* 1997; Bird 2002; Jablonka y Raz 2009; Fulneček y Kovařík 2014). La impronta genómica en mamíferos y en ciertas plantas se establece durante la gametogénesis, donde la metilación específica de un sexo altera la expresión de cientos de genes. En la cigota, los genes que poseen esta marca se expresan únicamente a partir del alelo heredado de la madre o del padre, según corresponda (Angers *et al.* 2010).

La 5-metilcitosina (5mC) es la principal y más estudiada base metilada en eucariotas multicelulares (Angers *et al.* 2010; Schulz *et al.* 2013; Luo y He 2017). La evidencia también propone a la 6-metiladenina (6mA) como marca epigenética en eucariotas, aunque aún se encuentra en evaluación (Luo y He 2017). En plantas, la distribución de las 5mC no es aleatoria, y los organismos difieren en el tipo de secuencias que están metiladas. A diferencia de los vertebrados, para los que las metilaciones ocurren en contextos CG (mCG, donde la **m** indica que la citosina de la derecha está metilada), en las plantas pueden ocurrir en cualquier contexto de secuencia, pero con un orden de frecuencia descendiente para los motivos mCG, mCHG y mCHH (donde H = A, T o C) (Henderson y Jacobsen 2007; Jablonka y Raz 2009; Angers *et al.* 2010; Schulz *et al.* 2013; Fulneček y Kovařík 2014; Niederhuth *et al.* 2016).

Las plantas y los animales poseen rutas moleculares similares para controlar y mantener la metilación (Law y Jacobsen 2010). La metilación de citosinas está mediada por citosina-metiltransferasas. En las plantas, la metilación *de novo* es catalizada por *Domains Rearranged Methyltransferase 2* (DRM2), y es mantenida por tres vías diferentes: la metilación mCG por la *DNA Methyltransferase 1* (MET1 o DMT1); la mCHG por la *Chromomethylase 3* (CMT3, específica de plantas); y la mCHH por la DRM2 (Law y Jacobsen 2010; Masuelli y Marfil 2010; Varotto y Forestan 2010, y ref.). Las metiltransferasas de mantenimiento operan por la propagación clonal y semi-conservativa de los patrones 5mC después de la replicación, restaurándose la metilación completa con una fidelidad del 95-99 % (Richards 2006; Jablonka y Raz 2009). Así, estas marcas son heredables, mientras que la desmetilación puede ser activamente catalizada por las enzimas ADN glicosilasas *Repressor Of Silencing 1* (ROS1) / *Demeter* (DME), probablemente en combinación con la vía de reparación por escisión de bases (*Base Excision Repair*, BER), u ocurrir pasivamente durante la replicación (Law y Jacobsen 2010; Kumar *et al.* 2016). La desmetilación ocurre siempre antes del inicio de cualquier proceso de diferenciación tisular como la floración, la embriogénesis

o el crecimiento activo. Por el contrario, en árboles se reportó un aumento gradual en la metilación con la edad (Baránek *et al.* 2010). La metilación *de novo* es fundamentalmente inducida por señales intrínsecas del desarrollo, aunque también puede aparecer por azar o ser inducida por las señales ambientales (Angers *et al.* 2010).

En vegetales, el contenido global de 5mC varía ampliamente entre especies, órganos, estados de desarrollo y tipos de estrés ambiental (Henderson y Jacobsen 2007; Baránek *et al.* 2010; Alonso *et al.* 2016). Las diferencias interespecíficas se correlacionan con el tamaño del genoma y muestran una fuerte señal filogenética (Alonso *et al.* 2016; Niederhuth *et al.* 2016); mientras que la metilación global puede variar entre individuos y poblaciones (Alonso *et al.* 2016).

Se ha documentado que el nivel de metilación en los promotores está negativamente correlacionado con los niveles de expresión, ya que interfiere estéricamente con la unión de la maquinaria transcripcional (Angers *et al.* 2010; Stark 2010; Varotto y Forestan 2010). La metilación en los tres contextos (mCG, mCHG y mCHH) dentro de las regiones codificantes también está asociada a supresión de la expresión génica (Law y Jacobsen 2010; Niederhuth *et al.* 2016); así, por ejemplo, las secuencias repetidas y los transposones típicamente están metilados (Alonso *et al.* 2016). En las angiospermas, el enriquecimiento de metilación mCG dentro de regiones génicas transcritas y el empobrecimiento de metilación en los sitios de inicio y terminación de la transcripción, estaría asociada con la expresión constitutiva (Bewick *et al.* 2016; Bewick y Schmitz 2017).

### Regulación epigenética de la determinación sexual

En plantas dioicas, se denomina “regulación epigenética de la determinación del sexo” a los procesos que conducen a diferencias entre los fenotipos sexuales de las flores y que no implican diferencias genéticas, sino en la regulación epigenética a nivel transcripcional o post-transcripcional. Factores ambientales como la temperatura, el fotoperíodo, la nutrición y la disponibilidad de agua, así como las hormonas, afectarían la expresión del sexo de las flores, incluso otorgando plasticidad a una amplia gama de grupos filogenéticos (Adam *et al.* 2011). Por ejemplo, en poblaciones androdioicas de ortiga mansa, *Mercurialis annua* (Euphorbiaceae), se encontró que la proporción de sexos de la progenie depende de la densidad de las plantas parentales (Dorken y Pannell 2008; Pannell *et al.* 2014). La orgaza europea, *Atriplex halimus* (Chenopodiaceae), es una especie monoica que muestra un aumento de flores femeninas en períodos de días

cortos y con poca intensidad de luz (Talamali *et al.* 2001, 2003). La palmera aceitera africana, *Elaeis guineensis* (Arecaceae), es considerada “dioica facultativa” ya que alterna ciclos de inflorescencias femeninas y masculinas funcionalmente unisexuales (Adam *et al.* 2011). Recientemente se han registrado patrones de metilación sexo-específicos que han sido implicados en la determinación del sexo de dos especies de álamo (*Populus*, Salicaceae) (Song *et al.* 2012a; Bräutigam *et al.* 2017).

Se ha propuesto que la epigenética, y particularmente la metilación, ha tenido un rol importante en los orígenes de la diferenciación sexual y en la formación de cromosomas heteromórficos. Gorelick (2003) modificó el modelo tradicionalmente aceptado sobre la evolución de los cromosomas sexuales (“el trinquete de Muller”), agregando los efectos de la metilación. Así, la dioecia y los cromosomas sexuales se habrían originado en hermafroditas ancestrales diploides como un par de autosomas en los que uno tenía más grupos metilo unidos a una región controladora del sexo que su homólogo. La posterior acumulación de mutaciones en el homólogo metilado (incipiente cromosoma Y o W), que podría estar relacionada con la mayor tasa de mutación de los nucleótidos metilados, habría proporcionado una fuerte presión selectiva para suprimir dichos *loci*, reduciendo su tamaño con respecto al incipiente cromosoma X o Z. Además, la metilación habría tenido un papel importante en la formación diferencial de la heterocromatina y en la supresión de la recombinación, esta última por bloqueo de los sitios de unión de las enzimas que la median. Las últimas etapas de la evolución del par sexual podrían haber ocurrido, por ejemplo, a través de reordenamientos estructurales cromosómicos. Según este modelo, la impronta genómica sería la causa más que el efecto de tener dos sexos. Además, el autor propuso que todos los linajes dioicos tendrían cromosomas sexuales, que podrían diferenciarse en los patrones de metilación (Gorelick 2003).

En el ciclo de vida de las angiospermas los productos de la meiosis experimentan proliferación mitótica para formar gametofitos multicelulares, es decir, el saco embrionario y el grano de polen (Henderson y Jacobsen 2007). Durante la gametogénesis masculina, se producen granos bicelulares de polen que contienen un núcleo vegetativo y uno espermático; mientras que en la gametogénesis femenina se producen tres células antípodas, dos células sinérgidas, una ovocélula, y dos núcleos centrales (Law y Jacobsen 2010). En la planta modelo *Arabidopsis thaliana* se ha visto que durante la gametogénesis masculina hay desmetilación activa del ADN en el núcleo vegetativo del polen, limitada a solo unas pocas secuencias de TEs; a la vez, se generan

ARNnc que podrían viajar a la célula espermática y reforzar el silenciamiento de sus TEs. En *Arabidopsis*, y en otros modelos como maíz y arroz, se ha demostrado que el genoma materno está globalmente hipometilado en el endosperma triploide, probablemente como consecuencia de la desmetilación activa en las células núcleo central (Law y Jacobsen 2010; Bouyer *et al.* 2017). Similarmente, los ARNnc potencialmente generados en las células núcleo central podrían reforzar el silenciamiento de TEs en la ovocélula y posiblemente también en el embrión en desarrollo (Law y Jacobsen 2010; Bouyer *et al.* 2017). La impronta genómica estaría restringida al endosperma, y sólo el alelo materno se expresaría, teniendo implicancias para el desarrollo apropiado del embrión (Law y Jacobsen 2010).














### Aproximación metodológica a la epigenética y la expresión génica diferencial, marco teórico

#### Epigenética

El MSAP (*Methylation Sensitive Amplified Polymorphism*; Reyna-Lopez *et al.* 1997) es una modificación del método AFLP (*Amplified Fragment Length Polymorphism*; Vos *et al.* 1995) en la que se reemplaza una enzima de restricción por dos o más isoesquizómeros que reconocen el mismo sitio de corte pero con sensibilidad diferencial a la metilación. Su empleo en combinación con una enzima “indiferente” a la metilación, y la comparación de los perfiles de fragmentos resultantes permite la discriminación de estados de metilación particulares en los sitios de restricción (Schulz *et al.* 2013; Fulneček y Kovařík 2014). Este método fue desarrollado para el estudio en hongos, y luego adaptado para su uso en plantas por Xiong *et al.* (1999).

Las enzimas sensibles a la metilación *HpaII* y *MspI* permiten la comparación de grandes cantidades de regiones 3'-CCGG-5' anónimas abarcando la metilación en los contextos más frecuentemente metilados, CG y CHG (**Figura 18**). Ambas enzimas cortan cuando el sitio no está metilado (condición I); si solo la citosina interna está hemi- o completamente metilada (en una o en las dos hebras, respectivamente) solamente *MspI* corta (condición II); *HpaII* corta si solo la citosina externa está hemimetilada (condición III); la actividad de ambas se bloquea cuando las dos citosinas están metiladas “hipermetilación” (condición IV). La falta de corte también puede ser consecuencia de una alteración del sitio de restricción (Schulz *et al.* 2013); esta posibilidad, sumada al enmascaramiento de diferentes tipos de metilación que no

pueden ser identificados, hacen que la condición IV sea conflictiva (**Figura 18**). En la literatura es frecuente encontrar interpretaciones contradictorias sobre la calificación y el significado de las condiciones I-IV (Schulz *et al.* 2013; Fulneček y Kovařík 2014; Alonso *et al.* 2016).

Estado de metilación	<i>HpaII</i>	<i>MspI</i>	Tipo de información	Bandas MSAP	
				<i>E+H</i>	<i>E+M</i>
 No metilado	+	+	Condición I (1/1)		
 Metilación completa de la citosina interna	-	+	Condición II (0/1)		
 Hemimetilación de la citosina interna	-	+	Condición II (0/1)		
 Hemimetilación de la citosina externa	+	(-) 10%	Condición III (1/0)		
 Metilación completa de la citosina externa	(-) 5%	-	Condición IV (0/0)		
 Metilación completa de ambas citosinas	-	-	Condición IV (0/0)		
 Hemimetilación de ambas citosinas	-	-	Condición IV (0/0)		
 Desconocido	-	-	Condición IV (0/0)		

**Figura 18.** Sensibilidad de los isoesquizómeros *HpaII* y *MspI*. Las citosinas metiladas en el sitio de restricción están coloreadas en gris. El símbolo “+” indica corte; el “-” no corte; “(-)” indica reportes de corte en la frecuencia expresada en porcentaje. El tipo de información “(1/1), 0/1, 1/0, (0/0)” indica la presencia/ausencia de bandas en código binario (ver Materiales y Métodos). Adaptado de Schulz *et al.* (2013).

Esta metodología de epigenotipificación permite identificar variaciones en la metilación de muchos *loci* epigenéticos simultáneamente y abarcando todo el genoma, de manera económica y sin requerir conocimiento previo del genoma, posibilitando la investigación de organismos no modelo y las comparaciones a nivel poblacional, individual, o incluso entre distintos órganos de un mismo individuo (Schrey *et al.* 2013).

Como muchas metodologías, el MSAP tiene ciertos aspectos restrictivos. Por caso, detecta *loci* anónimos de localización desconocida; resulta difícil establecer una conexión obvia entre la metilación del ADN y la expresión génica dado que se carece de modelos predictivos de las consecuencias de los cambios en la metilación. Por tratarse

de marcadores dominantes no distingue epigenotipos heterocigotas; luego cabe recordar la interpretación compleja de la condición IV (Schrey *et al.* 2013).

Con todo, en los últimos años el MSAP fue utilizado para responder una variedad de preguntas biológicas en distintas especies vegetales. Por ejemplo, en poblaciones naturales de cerezo silvestre (*Prunus avium*) se analizó la existencia de estructura epigenética (Avramidou *et al.* 2015b). En papa silvestre (*Solanum ruiz-lealii*) se estudió si ciertas anomalías morfológicas podían ser explicadas por diferencias epigenéticas (Marfil *et al.* 2009; Lira-Medeiros *et al.* 2010). En manzano (*Malus x domestica*) se analizó la existencia de una conexión entre procesos fisiológicos y cambios en los estados de metilación (Kumar *et al.* 2016). En el banano malayo (*Musa acuminata*) se estudió la respuesta epigenética a interacciones con patógenos fúngicos (Luo *et al.* 2016). En varias especies cultivadas como la uva (*Vitis vinifera*; Baránek *et al.* 2010), el cacao (*Theobroma cacao*; Rodríguez López *et al.* 2010) y el café (*Coffea arabica*; Bobadilla Landey *et al.* 2013) se analizaron los cambios epigenéticos luego de la aplicación de distintos tratamientos. Así, Baránek *et al.* (2010) estudiaron la metilación después del cultivo *in vitro* y termoterapia, y concluyeron que el origen de las plantas individuales, y no el tipo de estrés aplicado, era el factor más influyente. Rodríguez López *et al.* (2010) sugirieron que la edad de los callos usados para regenerar las plantas influye en la variación somaclonal (genética y epigenética). Bobadilla Landey *et al.* (2013) detectaron una desmetilación general como consecuencia de la embriogénesis somática *in vitro*. Kumar *et al.* (2016) evidenciaron cambios de metilación durante la latencia invernal de las yemas. Lira-Medeiros *et al.* (2010) estudiaron comparativamente dos poblaciones de mangle blanco (*Laguncularia racemosa*) localizadas en ambientes marcadamente contrastantes, que no mostraban diferenciación genética significativa pero sí diferencias morfológicas notables. Los autores hallaron que la diversidad epigenética a nivel intra- e interpoblacional fue mayor que la genética, sugiriendo que los cambios en la metilación podrían estar asociados con la heterogeneidad ambiental. Este, y otros estudios, han puesto de manifiesto que las poblaciones silvestres a menudo albergan mayor diversidad epigenética que genética (Li *et al.* 2008; Lira-Medeiros *et al.* 2010), y que la variación intraespecífica en la metilación puede estar asociada con diferencias en rasgos morfológicos y/o funcionales (Marfil *et al.* 2009; Alonso *et al.* 2014; Kumar *et al.* 2016). Otra serie de estudios concluyeron que la diversidad epigenética puede mejorar la capacidad colonizadora, la



productividad y la estabilidad de las poblaciones vegetales (Richards *et al.* 2012; Latzel *et al.* 2013; Herrera *et al.* 2014; Medrano *et al.* 2014).

### Expresión génica diferencial

La metodología cDNA-AFLP fue descrita por Bachem *et al.* (1996), quienes la emplearon por primera vez para comparar los perfiles de expresión de tubérculos de papa en diferente grado de desarrollo, e identificar fragmentos derivados de transcritos diferencialmente expresados. Esta metodología permite detectar y aislar transcritos específicos de tejidos, órganos o procesos particulares en una amplia gama de sistemas biológicos, de manera robusta y reproducible (Bachem *et al.* 1998). Además, supera limitaciones de otros métodos para la amplificación selectiva de ADNc basados en PCR, que presentan problemas de reproducibilidad, dificultad para representar transcritos de baja frecuencia, y generan falsos positivos debido al uso de cebadores aleatorios y a las bajas temperaturas de *annealing* (Bachem *et al.* 1998). El uso de condiciones más astringentes le otorgan al cDNA-AFLP mayor fidelidad de amplificación (Bachem *et al.* 1996), no estando exenta de fallas. El método consta de cuatro etapas: (1) síntesis de ADNc doble cadena, usando un oligonucleótido poli-dT para la retrotranscripción de los ARNm; (2) digestión con dos enzimas de restricción y ligado de adaptadores; (3) preamplificación con cebadores complementarios a los adaptadores, con una base selectiva; y (4) amplificación selectiva con cebadores que tengan dos o más bases selectivas (Bachem *et al.* 1998).

El cDNA-AFLP se usó en *Cannabis sativa* (especie dioica con cromosomas heteromórficos) para estudiar la diferenciación sexual; los ARNm hallados sugirieron la activación de procesos metabólicos en las plantas femeninas, o su represión en las masculinas; una secuencia relacionada con la transducción de señales podría actuar en respuesta a auxinas (Moliterni *et al.* 2004). En la planta medicinal *Angelica sinensis* se analizaron las diferencias entre los botones florales y el meristema apical, para comprender mejor el mecanismo de floración, obteniendo secuencias expresadas específicamente en los botones, entre las que se cuentan TEs, factores de transcripción, y enzimas metabólicas (Yu *et al.* 2011). Otros estudios se focalizaron en las respuestas a factores ambientales. Por ejemplo, en *Populus hopeiensis* se identificaron transcritos relacionados con la tolerancia a estrés hídrico (Song *et al.* 2012b); y en hojas de naranjo (*Citrus sinensis*) se detectó un menor número de cambios en la expresión génica como

respuesta a la toxicidad del manganeso, respecto al pomelo *C. grandis* (Zhou *et al.* 2013).

El único estudio comparativo de expresión génica desarrollado en *I. paraguariensis* empleó la técnica *Differential Display* (Liang y Pardee 1992) para explorar la respuesta de las hojas al estrés hídrico, identificando proteínas relacionadas con la tolerancia al estrés y homeostasis de metales, la fotosíntesis, el metabolismo secundario, la biosíntesis de lignina y pared celular, y la estabilización del ARN o la actividad chaperona (Acevedo *et al.* 2016). Particularmente una subunidad del complejo proteico mitocondrial succinato deshidrogenasa resultó de gran interés, por estar probablemente asociada al mecanismo de respuesta a deshidratación (Acevedo *et al.* 2013, 2016).

## Objetivos

El objetivo general de este capítulo es indagar las diferencias entre órganos, sexos y especies a nivel epigenético y de expresión génica. Las preguntas que guiaron nuestro trabajo son: ¿cuál es el grado de diferenciación epigenética entre las especies?, ¿qué genes son los responsables de la alta diferenciación metabólica de las hojas de estas especies, y cómo interviene la epigenética en la regulación de su expresión?, ¿existen patrones de metilación asociados con las características fisiológicas particulares de los sexos y los distintos órganos, que los definan como tales?, ¿ocurren cambios en los patrones de metilación o de expresión génica (a nivel de ARNm circulante) durante el desarrollo floral y foliar?, ¿qué genes están diferencialmente metilados y/o expresados entre sexos, órganos y especies?

En primer lugar y con el propósito de detectar clones entre las plantas de yerba mate del Jardín Botánico de Buenos Aires, se propuso evaluar la variación genética de dichas plantas.

En segundo lugar, se propusieron los siguientes objetivos específicos:

A nivel interespecífico:

- ✓ Estudiar diferencias en los perfiles de metilación del ADN y de expresión génica entre la yerba mate y la yerba señorita.

A nivel intraespecífico:

- ✓ Estudiar diferencias en los patrones de metilación del ADN y de expresión génica entre órganos y sexos para cada especie.
- ✓ Detectar:

- cambios en los estados de metilación en distintos estadios de desarrollo de los órganos reproductivos (botones florales - flores) y vegetativos (hojas).
- *loci* particulares diferencialmente metilados o expresados entre sexos, órganos y especies, e identificar los genes asociados.

### **Hipótesis**

- ✓ Los sitios metilados y la expresión génica difieren entre especies y entre órganos de cada especie.
- ✓ Los mecanismos de diferenciación sexual en estas especies involucran metilaciones en el ADN, en contextos mCG y mCCG, y expresión génica diferencial.

### **Predicciones**

- ✓ La yerba mate y la yerba señorita formarán grupos diferenciados en todas las comparaciones.
- ✓ Los órganos mostrarán patrones diferenciales de metilación y de expresión génica.
- ✓ Los botones florales se diferenciarán según el género (sexo) de las plantas dadoras.
- ✓ Se encontrarán genes o regiones potencialmente implicados en la diferenciación sexual entre los fragmentos diferenciales aislados de botones florales.
- ✓ Se detectarán genes o regiones con funciones en biosíntesis de metabolitos secundarios entre los fragmentos aislados de hoja.

## Materiales y Métodos

Cabe destacar que, lamentablemente y por motivos ajenos a nuestra voluntad, no se pudo tener acceso a las plantas del Banco de Germoplasma de la EEA-INTA CA como se había planteado originalmente, por tanto, se obtuvieron muestras de las plantas existentes en el Jardín Botánico “Carlos Thays” de Buenos Aires y de ciertas plantas del Establecimiento Las Marías, gentilmente cedidas por el Dr. PA Sansberro.

Los materiales utilizados en este capítulo se detallan en la **Tabla 9**. Los individuos serán denominados empleando un código interno para facilitar la lectura.

**Tabla 9.** Material empleado en el presente capítulo. Se indican el código identificador formal, el sexo y las metodologías de tipificación aplicadas a cada planta.

Especie	Identificador formal	Código interno	Sexo	Metodologías		
				ISSR	MSAP	cDNA-AFLP
<i>I. dumosa</i> (YS)	JB1091	D91	Femenino	-	+	-
	CRV7/4*	D4	Femenino	-	+	+
	CRV48/10*	D13	Masculino	-	+	+
	CRV7/12*	D7	Femenino	-	-	+
	42/92/19*	D21	Masculino	-	-	+
<i>I. paraguariensis</i> (YM)	JB1856	A1	Masculino	+	-	-
	JB1853	A2	Masculino	+	-	-
	JB1854	A3	Masculino	+	-	-
	JB1840	B1	Masculino	+	-	-
	JB1851	B2	Femenino	+	+	+
	JB1852	B3	Femenino	+	-	+
	JB1855	B4	Masculino	+	-	-
	JB1835	C1	Masculino	+	-	+
	JB1849	C2	Femenino	+	+	+
	JB1850	C3	Masculino	+	-	-
	JB1847	C4	Femenino	+	-	-
	JB1848	D1	Femenino	+	-	+
	JB1837	D2	Masculino	+	-	+
	JB1836	D3	Masculino	+	-	-
	JB1839	E1	Femenino	+	-	+
	JB1841	E2	Masculino	+	+	+
	JB1842	E3	Masculino	+	+	+
	JB1845	F1	Masculino	+	-	-
	JB1844	F2	Masculino	+	-	-
JB1843	F3	Femenino	+	-	-	

JB, Jardín Botánico de Buenos Aires; \* Establecimiento Las Marías.

### 1. Análisis genético

Exclusivamente para las plantas de yerba mate del Jardín Botánico se realizó una genotipificación con marcadores ISSR (Zietkiewicz *et al.* 1994) con el objetivo de determinar el grado de variación y las relaciones genéticas entre ellas. Los individuos de yerba señorita fueron excluidos de la genotipificación ya que al momento del análisis se disponía de una sola muestra.

Aproximadamente 100 mg de hoja preservada en sílica-gel fueron procesadas utilizando el *DNeasy Plant Quick Extraction kit* (Qiagen Inc.), según indicaciones del fabricante. La calidad del ADN fue verificada siguiendo el procedimiento descrito en el Capítulo 1. Las reacciones de PCR se realizaron según Realini *et al.* (2015), empleando los cebadores (AC)<sub>8</sub>G, (AG)<sub>8</sub>YC y (AG)<sub>8</sub>YT. Los productos ISSR fueron separados por electroforesis en PAGE al 5 % (p/v) y en condiciones desnaturalizantes (8M de urea), como en Cascales *et al.* (2014). Los patrones de bandas se visualizaron por tinción con nitrato de plata usando el *kit* comercial *Silver Sequence Staining Reagents* (Promega) según especificaciones del fabricante. Los geles fueron digitalizados y, por tratarse de marcadores dominantes, los patrones de bandas registrados como datos de presencia (1), ausencia (0) o como dato faltante (cuando la interpretación fue ambigua), generando una matriz binaria.

El porcentaje de *loci* polimórficos y de datos faltantes fueron computados con el programa FAMD 1.31 (Schlüter y Harris 2006); el índice de diversidad promedio de Shannon (I; Shannon y Weaver 1949) según la fórmula  $S = -\sum P_i \log_2 (P_i)$  (donde  $P_i$  es la frecuencia relativa de la presencia o ausencia de la banda;  $i = 0, 1$ ) se calculó manualmente en Excel. Además, en FAMD se calculó una matriz de distancias genéticas pareadas con el complemento del coeficiente de similitud de Dice (Sneath y Sokal 1973)  $1-S_{ij}$ , según:

$$S_{ij} = \frac{2a}{2a+b+c}$$

donde  $a$  = número de bandas presentes simultáneamente en los individuos  $i$  y  $j$ ;  $b$  = número de bandas presentes en  $i$  y ausentes en  $j$ ;  $c$  = número de bandas presentes en  $j$  y ausentes en  $i$ . Este coeficiente omite la consideración de pares negativos (0-0), ya que la ausencia compartida de una banda no es, necesariamente, una indicación de similitud entre dos individuos, y da doble peso a las presencias compartidas (Sneath y Sokal 1973).

A partir de la matriz de distancias se realizó un análisis de agrupamiento mediante el algoritmo de *Neighbour-Joining* (NJ); los valores de soporte de las ramas se obtuvieron mediante remuestreo por *bootstrapping* con 1.000 pseudoréplicas, usando el mismo programa. El retículo de NJ se visualizó con el programa FigTree v1.4.3 (Rambaut 2016). Además, se realizó un análisis discriminante de componentes principales (DAPC, *Discriminant Analysis of Principal Components*; Jombart *et al.* 2010) usando el paquete *adegenet* (Jombart 2008) del entorno de R 3.3.1 (R Core Team 2016), teniendo en cuenta los grupos definidos con la función *find.clusters* (es decir, sin brindar información previa del origen o tipo de muestra). El número óptimo de componentes principales se determinó con la función *optim.a.score*. La probabilidad de pertenencia (Pbp) de cada muestra al grupo asignado automáticamente se graficó usando la función *compoplot*; de ahora en más, este gráfico se llamará “compoplot”. Adicionalmente, considerando los agrupamientos identificados por DAPC, se realizó un análisis molecular de la varianza (AMOVA, *Analysis of Molecular Variance*; Excoffier *et al.*, 1992) con 999 permutaciones, en GenAlex 6.501 (Peakall y Smouse 2012).

## 2. Análisis epigenético

### 2.1. Materiales

Las plantas utilizadas fueron seleccionadas según la disponibilidad de los distintos órganos al momento de las colecciones. Las muestras (en total 34 UTOs -unidades taxonómicas operativas) fueron colocadas en nitrógeno líquido para evitar cambios en el estado de metilación, y luego conservadas a -80°C hasta su uso. Se tomaron muestras de botones florales a intervalos de 15-30 días, representando cuatro estadios de desarrollo denominados temprano (bI), medio (bII), maduro (bIII) y pimpollo (bIV), y muestras de flores abiertas, hojas jóvenes y hojas adultas, de individuos femeninos y masculinos de ambas especies (**Tabla 10**). Se consideró que todos los estadios de los botones florales representan órganos reproductivos inmaduros por estar envueltos por los pétalos, en comparación con la flor completamente abierta. Los estadios fueron clasificados según su apariencia externa, suponiéndolos ontogenéticamente equivalentes. Sin embargo, cabe notar que estas especies presentan ontogenia asincrónica, por lo que el grado de madurez relativa de los botones sería desparejo incluso entre aquellos de una misma inflorescencia (Giberti 2001, Giberti y Gurni 2008), haciendo que la clasificación propuesta aquí pueda resultar artificial. En otro orden, dado que durante el desarrollo de

este trabajo las plantas femeninas disponibles de *I. dumosa* florecieron escasamente, el número de muestras de botones florales femeninos fue bajo. A fin de salvar esta situación y obtener la cantidad de ADN necesaria para realizar los análisis posteriores, para la yerba señorita no se hizo la distinción entre estadios.

Por otro lado, se consideraron como “hojas jóvenes” a los brotes foliares nuevos de la temporada, de coloración rojiza en el caso de la yerba mate, y verde muy claro en el caso de la yerba señorita, mientras que se consideraron “hojas adultas” aquellas emergidas de brotes de años anteriores, completamente expandidas, y de coloración oscura. Ambos estadios foliares fueron coleccionados el mismo día que los botones florales bIII, y de las mismas plantas.

**Tabla 10.** Individuos muestreados por órgano y por sexo para cada especie.

	Botones florales				Flores	Hoja joven	Hoja adulta	N°
	Temprano (I)	Medio (II)	Maduro (III)	Pimpollo (IV)				
<i>I. dumosa</i> ♀	D4, D91				D91	D4, D91	D4, D91	7
<i>I. dumosa</i> ♂	D13	D13	D13	D13	D13	D13	D13	7
<i>I. paraguariensis</i> ♀	C2	C2	C2, B2	-	C2, B2	C2, B2	C2, B2	10
<i>I. paraguariensis</i> ♂	E2	E2	E2, E3	-	E2, E3	E2, E3	E2, E3	10
<b>N° muestras</b>	<b>14</b>				<b>6</b>	<b>7</b>	<b>7</b>	<b>34</b>

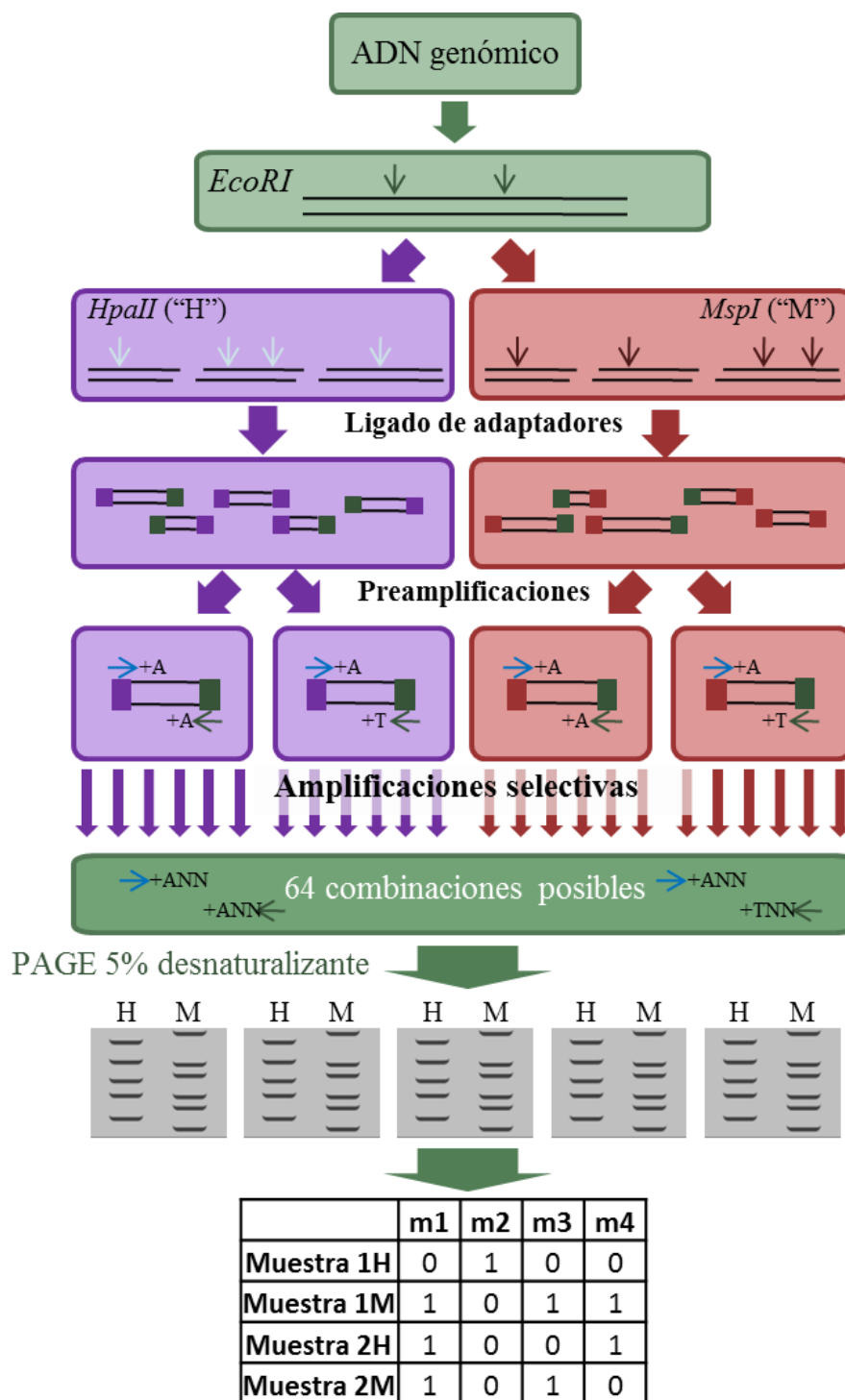
## 2.2. Extracción del ADN genómico

Durante la manipulación de las muestras, se tuvo la precaución de evitar que se descongelaran. La extracción de ADN genómico y la verificación de su integridad mediante electroforesis en gel de agarosa al 1% (p/v) se realizaron como fue descripto previamente (Capítulo 1). Se cuantificó el ADN mediante un espectrofotómetro Nanodrop.

## 2.3. Método MSAP

El protocolo se detalla en el Apéndice II; en la **Figura 19** se muestran esquemáticamente los pasos seguidos. En síntesis, 500 ng de ADN de cada muestra se digirieron con la enzima *EcoRI*, verificando la digestión completa por electroforesis en gel de agarosa al 1,5% (p/v) como se explicó anteriormente. Luego, cada muestra fue dividida en dos tubos, uno se restringió con *HpaII* (de ahora en más, denominado “H”), y el otro con *MspI* (llamado “M”). Seguidamente se ligaron oligonucleótidos

adaptadores doble cadena específicos para cada enzima, y se realizaron preamplificaciones con 2 combinaciones de cebadores preselectivos (+A y +T) (Tabla AI-7). Estos productos fueron verificados mediante electroforesis en gel de agarosa al 1,5% (p/v) y fueron utilizados como molde en las amplificaciones selectivas subsiguientes.



**Figura 19.** Esquema del protocolo MSAP aplicado. Los detalles pueden consultarse en el Apéndice II.



Con el fin de seleccionar las mejores combinaciones de cebadores selectivos, se realizó un ensayo preliminar sobre un subgrupo de 6 muestras (3/especie). Las bases selectivas de los cebadores *EcoRI*+3 y *HpaII*/*MspI*+3 se detallan en la **Tabla 11**. Los productos de amplificación selectiva fueron sometidos a PAGE desnaturizante al 5% (p/v), digitalizados y procesados como fue descrito más arriba. A partir de los patrones de bandas, se contabilizó el número total de bandas y se registró la ocurrencia de variación entre los patrones. De esta manera, se seleccionaron las 13 combinaciones que produjeran mayor cantidad de bandas y con mayor variación cualitativa entre muestras, y se procedió a amplificar la totalidad de las muestras (**Tabla 11**).

**Tabla 11.** Combinaciones de cebadores selectivos ensayadas. En verde, combinaciones probadas pero no seleccionadas. En celeste y cruzadas, combinaciones probadas, seleccionadas y efectivamente analizadas. En blanco, combinaciones no ensayadas.

		<i>HpaII/MspI</i> +3							
		TTG	TTC	TAG	TAC	ATG	ATT	AAC	AAG
<i>EcoRI</i> +3	AAC				X	X			
	AAG		X				X	X	
	ACA					X			
	ACC								
	ACG							X	
	ACT					X			X
	AGC				X		X		
	AGG			X					X

Los geles fueron digitalizados, y los patrones de bandas fueron registrados y volcados en matrices binarias de presencia/ausencia en las que cada muestra individual (UTO, unidad taxonómica operativa) está representada por dos filas, una para cada enzima (H y M, en ese orden). De esta manera, una UTO que para determinado *locus* es 1/1 se consideró no metilada, una UTO registrada con 1/0 se consideró hemimetilada en la citosina externa, una 0/1 se consideró con la citosina interna hemi- o completamente metilada, y la ausencia de banda para ambas enzimas, se interpretó como 0/0, reuniendo las condiciones de hipermetilación (ambas citosinas metiladas en una o ambas hebras) y/o de mutaciones en el sitio de restricción (falta de secuencia diana).

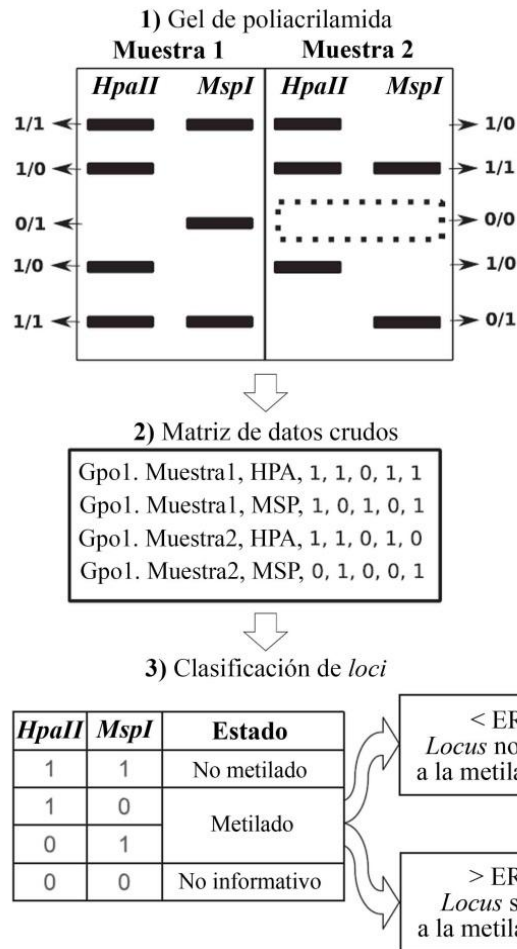
#### 2.4. Análisis de datos

Las matrices derivadas de las correspondientes combinaciones de cebadores selectivos fueron concatenadas en una única matriz (denominada YM+YS), que se utilizó para los análisis interespecíficos.

Para realizar los análisis intraespecíficos, se generó una matriz para cada especie (denominadas YM y YS). Las tres matrices fueron analizadas usando el paquete *msap* 1.1.8 (Pérez-Figueroa 2013) del entorno de R, que interpreta las bandas H/M por individuo y por *locus*, y clasifica cada *locus* según su comportamiento ante la metilación: susceptibles a la metilación (MSL) o no susceptibles (NML) (**Figura 20**). Si la proporción observada de estados metilados en todas las muestras superan el valor umbral de la tasa de error (por defecto = 5%), el programa considera un sitio susceptible a la metilación. De lo contrario, se registra como sitio no susceptible. Luego, a partir de cada matriz original (YM+YS, YM y YS) se generan dos matrices binarias reducidas (denominadas MSL y NML), que se analizan independientemente. Aunque no se trate de un estudio poblacional, aquí se emplearán los términos polimorfismo o variación para referirse a los *loci* que poseen al menos dos veces cada estado (polimórficos o variables). Únicamente estos *loci* son considerados en los análisis subsiguientes.

Schulz *et al.* (2013) describieron ocho formas de análisis de datos MSL, y concluyeron que no existe un método correcto. En esta Tesis los datos se analizaron según dos métodos, el denominado “*Methylation Scoring*” (MS), que agrupa las condiciones II y III (estados metilados) y las compara contra la condición I (no metilado); y el “*Mixed Scoring 2*” (MS2), que distingue entre las condiciones I, II y III, otorgando la posibilidad de analizar la metilación en contextos CG (condición II) y CHG (condición III) por separado, y dando una imagen más completa. Por defecto, el paquete *msap* analiza cada matriz MSL según la conversión de datos *Methylation Scoring* (MSL-MS, que se describe más abajo); adicionalmente se usó el *script* MSAP\_calc.r, para la conversión de los datos tipo *Mixed Scoring 2* (MSL-MS2, ver más adelante) propuesta por (Schulz *et al.* 2013).

La matriz NML también es transformada por *msap* presentando la información de presencia (1/1) como 2 y de ausencia (0/0) como 1. En este caso, los registros discordantes entre H y M (1/0 y 0/1) que no superan el umbral definido *a priori* también se registran como 2, es decir, presencia (**Figura 20**) (Pérez-Figueroa 2013).



4) Transformación de datos

A. MSL

a) Matriz binaria

<i>HpaII</i>	<i>MspI</i>	Estado	MSL-MS	MSL-MS2
1	1	No metilado	1	1 0 0
1	0	Metilado	2	0 1 0
0	1			0 0 1
0	0	No informativo	NA	0 0 0

b) Análisis de tipos de metilación

<i>HpaII</i>	<i>MspI</i>	Estado	Transformación
1	1	No metilado	<b>u</b> (Cond. I)
1	0	Cit ext hemimetilada	<b>h</b> (Cond. II)
0	1	Cit interna metilada	<b>i</b> (Cond. III)
0	0	Otras metilaciones mutación	<b>f</b> (Cond. IV)

B. NML

<i>HpaII</i>	<i>MspI</i>	Estado	Transformación
1	1	Presencia de fragmento	2
1	0		
0	1	Ausencia de fragmento	1
0	0		

**Figura 20.** Diagrama de flujo del análisis de datos MSAP. Transformación de las matrices susceptibles (MSL) o no susceptibles (NML) a la metilación, las primeras según el criterio “Methylation Scoring” (MSL-MS) o “Mixed Scoring 2” (MSL-MS2). Adaptado de (Pérez-Figueroa 2013). Cit, citosina. Cond., condición. ERT, *error rate-based threshold*, valor umbral de la tasa de error utilizado para clasificar loci.

La variación genética contenida en la matriz NML, y la variación epigenética de las matrices MSL-MS y MSL-MS2, se estimaron utilizando el índice de diversidad de Shannon, calculado como se indicó anteriormente. Además, para indagar si la variación epigenética puede verse influenciada por el fondo genético, se evaluó la existencia de correlación entre las matrices MSL-MS, MSL-MS2 y NML. Para esto, se calcularon las matrices de distancia Dice en el programa FAMD, y luego se realizaron *tests* de Mantel pareados (999 permutaciones) en GenAlex.

### 2.5. Análisis por Methylation Scoring, MSL-MS

Para la matriz MSL, este análisis considera los estados metilados (1/0 y 0/1) como un marcador dominante codificado como 2, el estado no metilado (1/1) se codifica como 1, y falta de dato para un estado no informativo (0/0) se codifica como NA (no aplicable) (**Figura 20**).

Adicionalmente, para la matriz MSL no transformada, el programa *msap* crea una nueva matriz (“MSL multiestado”) en la que para cada *locus* se codifica el tipo de metilación con cuatro estados posibles (**Figura 20**) de la siguiente manera:

u = 1/1, no metilado

h = 1/0, hemimetilado

i = 0/1, citosina interna metilada

f = 0/0, hipermetilación o ausencia de secuencia blanco por mutación.

Con la matriz MSL multiestado se calcularon las frecuencias relativas de cada tipo de metilación dentro de los grupos de muestras (predefinidos *a priori*), que fueron visualizadas con histogramas realizados en Excel.

Los análisis de coordenadas principales (PCoA, *Principal Coordinate Analysis*; Gower 1966) se realizaron con la configuración por defecto implementada en el paquete *msap*. Con el mismo software, se realizaron AMOVA significando estadísticamente cada componente de la varianza en base a 999 permutaciones. Los gráficos PCoA se obtuvieron en Excel, importando la matriz de coordenadas principales. Además, se realizaron DAPC, del mismo modo que se indicó anteriormente; los agrupamientos detectados fueron evaluados mediante AMOVA.

### 2.6. Análisis por Mixed Scoring 2, MS2

La matriz MSL fue transformada según el método *Mixed Scoring 2* (MS2) usando el *script MSAP\_calc.r* (Schulz *et al.* 2013). La transformación MS2 se realiza triplicando

la matriz MSL en tándem: en la primera copia se le da valor “1” a la condición I (es decir, a los patrones 1/1), en la segunda copia se anota como “1” solo a la condición II (patrones 0/1), y en la tercera copia se anota como “1” únicamente a la condición III (patrones 1/0). En cada caso, los restantes patrones los anota como “0”. Por tanto, esta transformación crea de una a tres condiciones por *locus* (denominadas *subepiloci*).

El *script* calcula, entre otros, los siguientes parámetros: número de marcadores (*subepiloci*), porcentaje de marcadores polimórficos, y el índice de diversidad promedio de Shannon. También calcula para cada estado de metilación (condiciones I-III) el número de marcadores y el índice de diversidad.

Además, se realizó un PCoA usando el paquete *labdsv* 1.8-0 (Roberts 2016), según los comandos especificados en el manual de *MSAP\_calc.r* (Schulz *et al.* 2013), y basándose en una matriz de distancias de Dice derivada de la matriz MSL-MS2. Como se explicó, este coeficiente excluye el estado no informativo de los fragmentos de MSAP doblemente ausentes “0/0”, que podrían indicar otros tipos de metilación o falta de secuencia diana de corte. Finalmente, se realizó un DAPC en *adegenet*, como fue indicado anteriormente.

### 2.7. Selección e identificación de fragmentos

Se escrutó la matriz MSL multiestado de YM+YS para identificar entre los grupos de muestras fragmentos diferencialmente metilados, esto es, aquellos que presentaran un mismo estado en la totalidad o en un porcentaje de las muestras. Las bandas (fragmentos) seleccionadas fueron aisladas de los geles PAGE según protocolo de Stumm *et al.* (1997), y los fragmentos fueron reamplificados por PCR usando los cebadores originales, y temperatura de *annealing* de 60°C. Los productos fueron verificados en agarosa al 1,5% (p/v); aquellos que mostraron una única banda fueron enviados al servicio de MacroGen INC, donde fueron purificados y secuenciados. La identidad de las secuencias se indagó comparando con bases de datos públicas utilizando los algoritmos BLASTn, BLASTx y tBLASTx. Además, se contrastó con las secuencias depositadas en la base de datos *Gypsy DataBase 2.0* (GyDB; Llorens *et al.* 2011) de retrotransposones tipo LTR (*Retroelements with long Terminal Repeats*). En todas las búsquedas se consideró como umbral de significación las comparaciones con valor  $E \leq 10^{-5}$ .

### 3. Expresión génica diferencial

#### 3.1. Materiales

Para esta parte del estudio se emplearon 29 muestras en total (**Tabla 12**), las cuales fueron originalmente sumergidas en nitrógeno líquido o en la solución *RNAlater* (Thermo Fisher Scientific), y luego conservadas a -80°C.

**Tabla 12.** Individuos muestreados por órgano y por sexo de cada especie.

	Órgano reproductivo	Órgano vegetativo		Nº
	Botones florales	Hoja joven	Hoja adulta	
<i>I. dumosa</i> ♀	D4, D7	D4, D7	D4, D7	6
<i>I. dumosa</i> ♂	D13, D21	D13, D21	D13, D21	6
<i>I. paraguariensis</i> ♀	B2, B3, C2, D1, E1	D1, C2	D1, C2	9
<i>I. paraguariensis</i> ♂	C1, D2, E2, E3	D2, E3	D2, E3	8
Nº	13	8	8	29

Nº, número de muestras analizadas por grupo.

#### 3.2. Extracción de ARN

Aproximadamente 100 mg de muestra fueron molidos en mortero con nitrógeno líquido, y se extrajo el ARN total según el protocolo del *Trizol Reagent* (Thermo Fisher Scientific) sugerido por el fabricante (Apéndice II). Los ARN resuspendidos fueron re-purificados utilizando el *RNeasy Plant Mini Kit* (Qiagen) o el *EasyPure® Plant RNA Kit* (Transgen) y luego fueron cuantificados en un Nanodrop. La integridad fue verificada mediante electroforesis en gel de agarosa al 1% (p/v), como anteriormente.

Para eliminar una posible contaminación con ADN, se realizaron tratamientos con *DNasa* (Promega) siguiendo el protocolo del fabricante (ver Apéndice II). Para la obtención del ADN copia (ADNc) doble cadena, se utilizó el *kit* comercial *SuperScript III First-Strand Synthesis System* (Invitrogen) siguiendo el protocolo del fabricante, y usando como cebador un oligo dT<sub>20</sub>. Luego, se sintetizó la segunda cadena del ADNc usando las enzimas ADN ligasa, ADN polimerasa I, y RNasa H de *E. coli* (New England Biolabs) (Apéndice II). El ADNc doble cadena se cuantificó en un *Nanodrop*.

#### 3.3. AFLP sobre ADNc doble cadena, cDNA-AFLP

En una primera instancia, se intentó poner a punto la técnica *Differential Display* para analizar 12 muestras de ADNc de distintos órganos de yerba mate y yerba señorita.

Se probaron 48 combinaciones de cebadores. Si bien se ensayaron distintas condiciones de PCR, no se lograron obtener resultados confiables, debido a fallas en la amplificación, falta de repetibilidad, y/o a un bajo número de bandas observadas, por lo que se optó por cambiar a la metodología cDNA-AFLP.

Se inició el protocolo AFLP (Apéndice II), partiendo de 300 ng de ADNc; para las dobles digestiones se utilizaron las enzimas de restricción *EcoRI* y *MseI* (Figura 21). Los adaptadores y cebadores utilizados se listan en la **Tabla AI-8**.

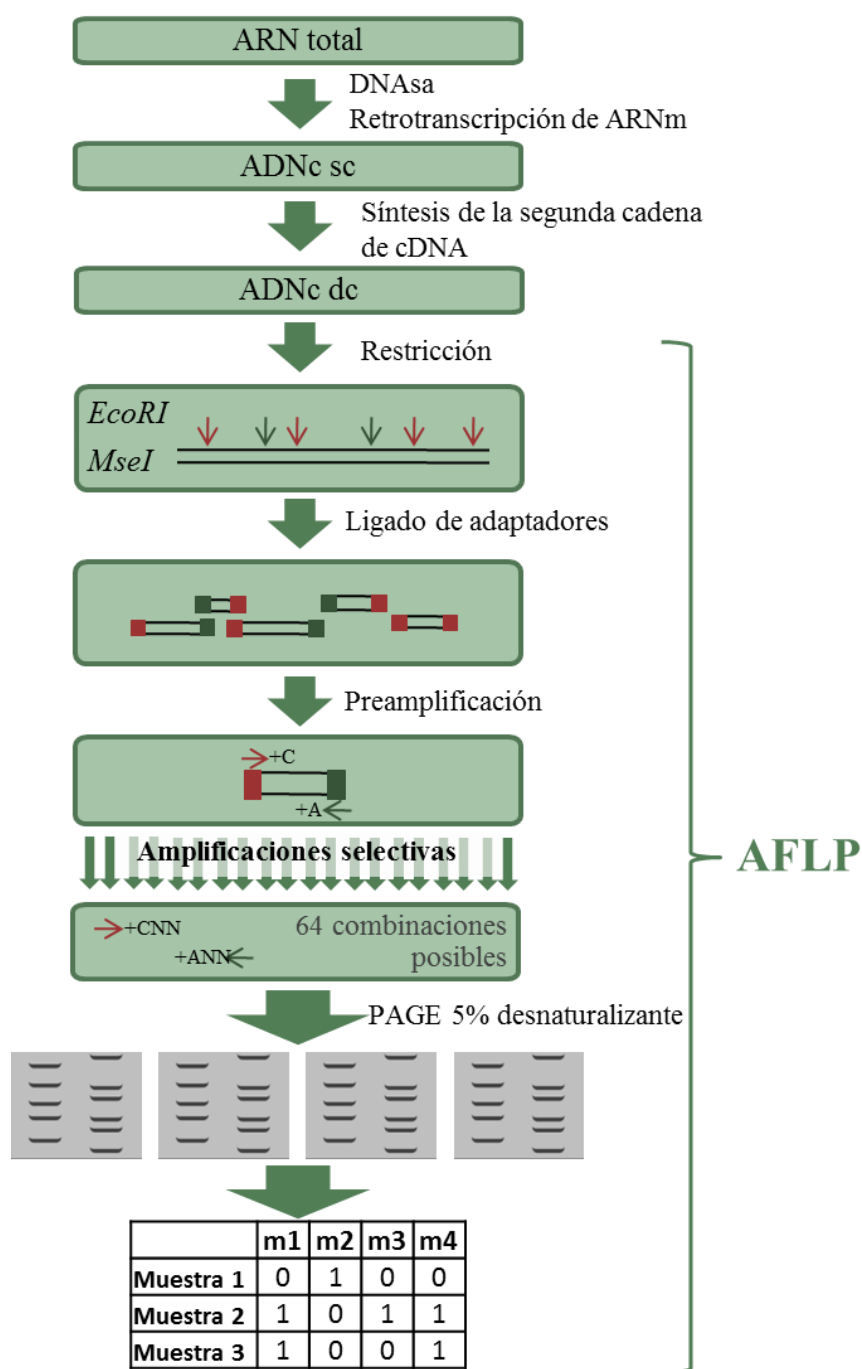


Figura 21. Esquema del protocolo cDNA-AFLP aplicado.

Se ensayaron 31 combinaciones de cebadores selectivos de las 64 posibles (**Tabla 13**), usando un subgrupo de 6 muestras. Se seleccionaron las 15 combinaciones que generaron los mayores números de bandas, y fueron utilizadas para amplificar todas las muestras. Los patrones de bandas visualizados en los PAGE fueron registrados y codificados para construir matrices binarias, como se detalló anteriormente.

**Tabla 13.** Combinaciones de cebadores ensayadas en cDNA-AFLP. En verde, combinaciones probadas pero no seleccionadas. En celeste y cruzadas, combinaciones probadas, seleccionadas y efectivamente analizadas. En blanco, combinaciones no ensayadas.

		<i>MseI</i> + 3							
		CAA	CAC	CAG	CAT	CTA	CTC	CTG	CTT
<i>EcoRI</i> + 3	AAC		X						
	AAG	X		X		X		X	
	ACA				X				
	ACC						X		
	ACG	X							X
	ACT		X		X				
	AGC			X		X			
	AGG	X							X

### 3.4. Análisis de patrones de bandas cDNA-AFLP

Las matrices obtenidas para cada combinación de cebadores selectivos fueron concatenadas. Se generaron tres matrices multicebador, YM+YS, YM y YS, a partir de las cuales se calculó número de *loci* y porcentaje de polimorfismo en FAMD, y el índice de diversidad de Shannon. También se realizaron análisis PCoA y AMOVA (con 1000 permutaciones) usando el paquete *msap*, adaptando el *script* para indicar que se trata de datos genéticos (*meth=F*) no obtenidos por la técnica MSAP. Las coordenadas de los PCoA fueron importadas en GenAlex para obtener los gráficos respectivos. También se realizaron análisis DAPC como ya fue explicado. Para la yerba mate se generó además la matriz “YM-bot”, incluyendo únicamente las muestras de botones, para indagar si existe diferenciación entre sexos.

### 3.5. Selección e identificación de fragmentos diferencialmente expresados

Escrutando las 15 matrices binarias, se seleccionaron 148 fragmentos derivados de transcritos, potencialmente expresados de manera diferencial entre los grupos de muestras. Las bandas correspondientes a dichos *loci* se aislaron de los geles PAGE, y se



reamplificaron por PCR usando la combinación de cebadores original. Las reamplificaciones exitosas fueron enviadas para su purificación y secuenciación, y los fragmentos fueron identificados por comparación con bases de datos públicas, procediendo como se explicó en la sección 2.7.

## Resultados

### 1. Caracterización genética

El individuo D1♀ fue eliminado de este análisis ya que no se logró amplificación efectiva. Para las 19 plantas restantes de yerba mate se registraron un total de 267 *loci* ISSR, 90,6% mostrando variación. La matriz construida presentó 5% de datos faltantes. Se observó que, para estas muestras, los tres cebadores se comportaron de manera similar en cuanto a la variación genética evidenciada (**Tabla 14**). El índice de Shannon promedio fue de  $0,726 \pm 0,245$ .

**Tabla 14.** Número de *loci*, porcentaje *loci* con variación e índice de diversidad de Shannon (I) calculados por cebador y para la matriz total de datos ISSR de yerba mate.

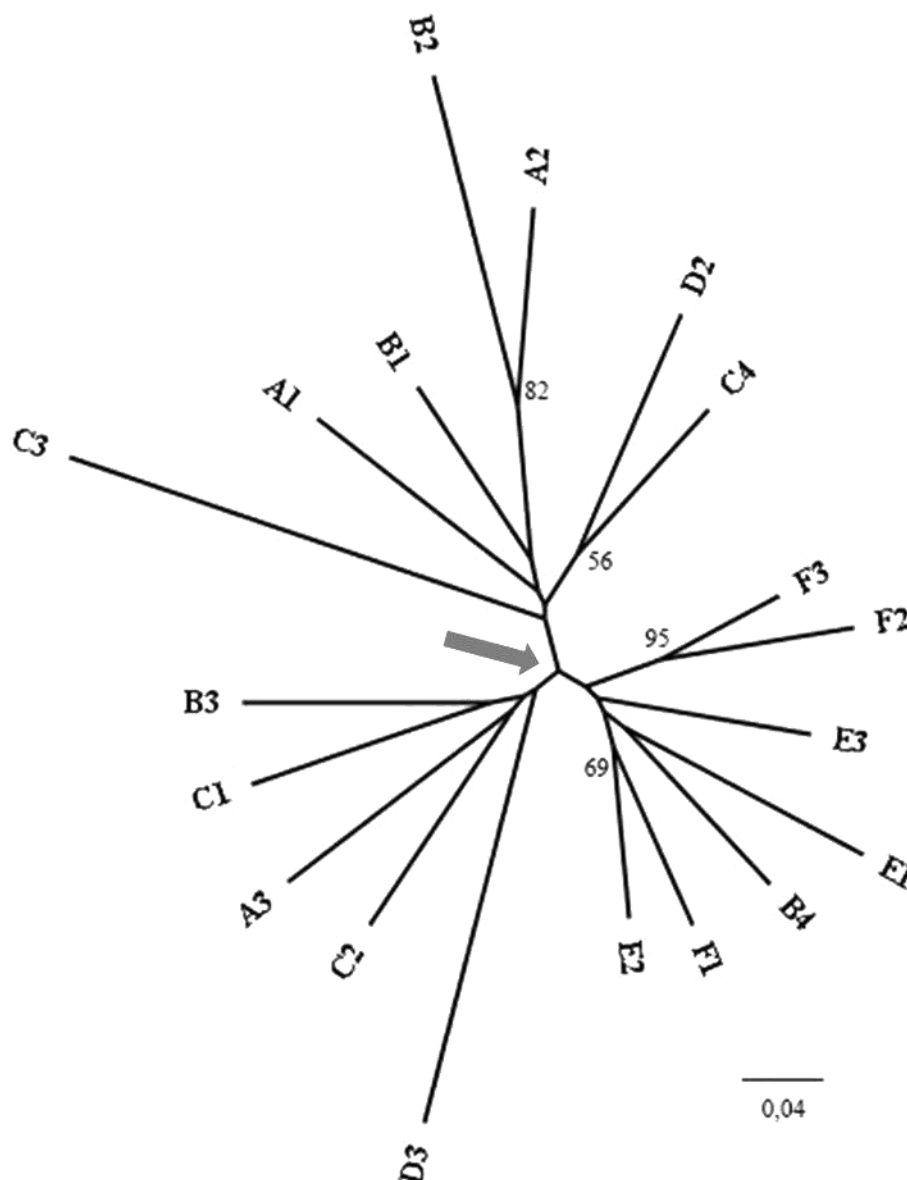
Cebador	Nº <i>loci</i>	% <i>loci</i> variables	I ± DE
(AC) <sub>8</sub> G	76	93,42	$0,738 \pm 0,234$
(AG)8YC	102	88,24	$0,683 \pm 0,253$
(AG)8YT	89	91,01	$0,764 \pm 0,240$
<b>promedio</b>	89	90,89	-
<b>Matriz total</b>	267	90,6	$0,726 \pm 0,245$

DE: desvío estándar

El retículo NJ mostró solo cuatro grupos con apoyo > 50% (los pares de plantas F2-F3, A2-B2, E2-F1 y C4-D2), que no habría plantas idénticas en su fondo genético, y que A2-B2, C4-D2, C3, A1, B1 se distinguen de las restantes, aunque para esa rama no se obtuvo apoyo de *bootstrap*. Además, los individuos D3 y C3 y el par A2-B2 son los que presentan los mayores largos de rama, evidenciando la mayor acumulación de cambios (**Figura 22**).

La distancia genética promedio es de 0,320; los individuos E3 y F3 muestran la menor distancia (0,183), mientras que A3 y B2 son los más distantes (0,496).

El DAPC muestra dos grupos levemente solapados, definidos por una única variable que retiene el 24,7% de la varianza (probabilidad de pertenencia al grupo asignado  $\geq 72\%$ ). Las muestras A1, A2, B1, B2, C3, C4 y D2 con más afinidad genética, forman un grupo, y las demás conforman el otro (**Figura AI-1**). El AMOVA indica que esta diferenciación es significativa ( $\phi_{ST} = 0,154$ ;  $p = 0,001$ ), y que el 15% de la varianza se encuentra entre estos grupos.



**Figura 22.** Retículo de *Neighbour-Joining* de individuos de yerba mate, analizados con marcadores ISSR. La flecha indica el nodo que define los grupos apoyados por otros métodos.

## 2. Análisis de variación epigenética

Se analizaron 13 combinaciones de cebadores selectivos MSAP. La muestra correspondiente al estadio I de botón floral del individuo C2♀ (C2♀bI) fue excluida de las matrices YM+YS y YM ya que en ensayos preliminares se observó un comportamiento anómalo (*outlier*). La matriz YM+YS (1014 *loci* y 33 UTOs, aquí muestras de órganos provenientes de siete plantas; 1,6 % dato faltante) presentó mayor porcentaje de *loci* no metilados (NML, 55,6%) que de metilados (MSL, 44,4%). En promedio se obtuvieron 78 *loci* por combinación de cebadores (**Tabla 15**).

La matriz YM consistió en 19 UTOs de cuatro individuos, para los que se contabilizaron 677 *loci* (1,1 % dato faltante); mientras que la matriz YS incluyó 14 UTOs de tres individuos, que produjeron 547 *loci* (2,5 % dato faltante). La matriz YM promedió 52 *loci* por combinación de cebadores, mientras que la matriz YS registró un promedio de 42 *loci* por combinación; en ambos casos la mayor proporción consistió en *loci* MSL (**Tabla 15**). En todos los casos, los NML mostraron mayor polimorfismo (rango: 94-98%) que los MSL (rango: 37-45%).

**Tabla 15.** Número de *loci* analizados para cada combinación de cebadores y por matriz. Se indican los porcentajes de sitios susceptibles a la metilación (MSL) y no susceptibles (NML).

Combinación de cebadores	YM + YS (33 UTOs)			YM (19 UTOs)			YS (14 UTOs)		
	<i>loci</i>	MSL (%)	NML (%)	<i>loci</i>	MSL (%)	NML (%)	<i>loci</i>	MSL (%)	NML (%)
E-AAC + H/M-TAC	73	45,8	54,2	43	58,1	41,9	45	60,0	40,0
E-AAC + H/M-ATG	55	60,0	40,0	36	80,6	19,4	33	63,6	36,4
E-AAG + H/M-TTC	74	58,1	41,9	45	77,8	22,2	50	58,0	42,0
E-AAG + H/M-ATT	60	54,2	45,8	36	77,8	22,2	35	82,9	17,1
E-AAG + H/M-AAC	74	51,4	48,6	48	62,5	37,5	41	58,5	41,5
E-ACA + H/M-ATG	70	50,7	49,3	39	69,2	30,8	42	59,5	40,5
E-ACG + H/M-AAC	88	11,4	88,6	57	33,3	66,7	27	40,7	59,3
E-ACT + H/M-ATG	95	46,1	53,9	69	75,4	24,6	29	75,9	24,1
E-ACT + H/M-AAG	79	46,8	53,2	35	48,6	51,4	56	64,3	35,7
E-AGC + H/M-TAC	120	48,7	51,3	79	69,6	30,4	55	65,5	34,5
E-AGC + H/M-ATT	71	43,7	56,3	61	68,9	31,1	32	50,0	50,0
E-AGG + H/M-TAG	125	34,5	65,5	84	53,6	46,4	66	53,0	47,0
E-AGG + H/M-AAG	64	36,5	63,5	45	55,6	44,4	36	55,6	44,4
<b>Total absoluto</b>	<b>1014</b>	<b>450</b>	<b>564</b>	<b>677</b>	<b>429</b>	<b>248</b>	<b>547</b>	<b>331</b>	<b>216</b>
<b>(% loci totales)</b>		<b>(44)</b>	<b>(56)</b>		<b>(63)</b>	<b>(37)</b>		<b>(60,5)</b>	<b>(39,5)</b>
<b>Promedio (absoluto)</b>	<b>78</b>	<b>34,6</b>	<b>43,4</b>	<b>52,1</b>	<b>33</b>	<b>19,1</b>	<b>42,1</b>	<b>25,5</b>	<b>16,6</b>

YM+YS: matriz interespecífica. YM: matriz de yerba mate; YS: matriz de yerba señorita. UTOs: unidades taxonómicas operativas. E= *EcoRI*; H= *HpaII*; M= *MspI*. Se indican con color las combinaciones para las que se hallaron marcadas diferencias entre los porcentajes de *loci* MSL y NML. En gris claro, relación aproximada 2:1; en gris oscuro, relación aproximada 3:1 o más.

Con la información obtenida se construyeron 12 matrices transformadas (**Tabla 16**). La matriz YM+YS-MSL-MS2 (interespecífica y derivada de los sitios MSL polimórficos) consistió en 538 *subepiloci*, la mayor parte de yerba mate; mientras que

las matrices YM y YS consistieron en 407 y 335 *subepiloci*, respectivamente. En los tres conjuntos de datos analizados (YM+YS, YM y YS), los *tests* de Mantel pareados indican correlación positiva significativa entre las matrices MSL-MS, MSL-MS2 y NML (**Tabla AI-9**).

**Tabla 16.** Dimensiones (UTOs y *loci* polimórficos) de las matrices analizadas.

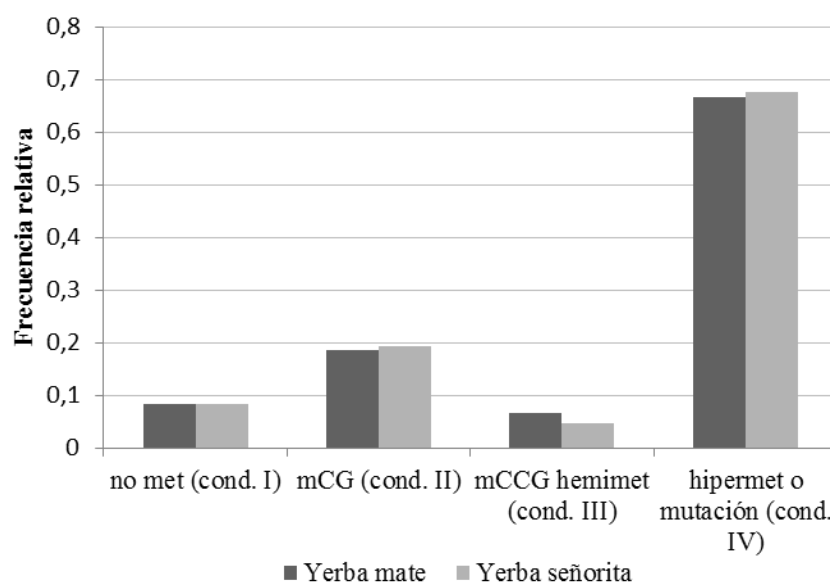
	UTOs	Datos crudos	Matrices			
			NML	MSL-MS	MSL-MS2	MSL multiestado
<b>YM+YS</b>	33	1014	558	203	538	203
<b>YM</b>	19	677	242	162	407	162
<b>YS</b>	14	547	203	121	335	121

UTOs: unidades taxonómicas operativas.

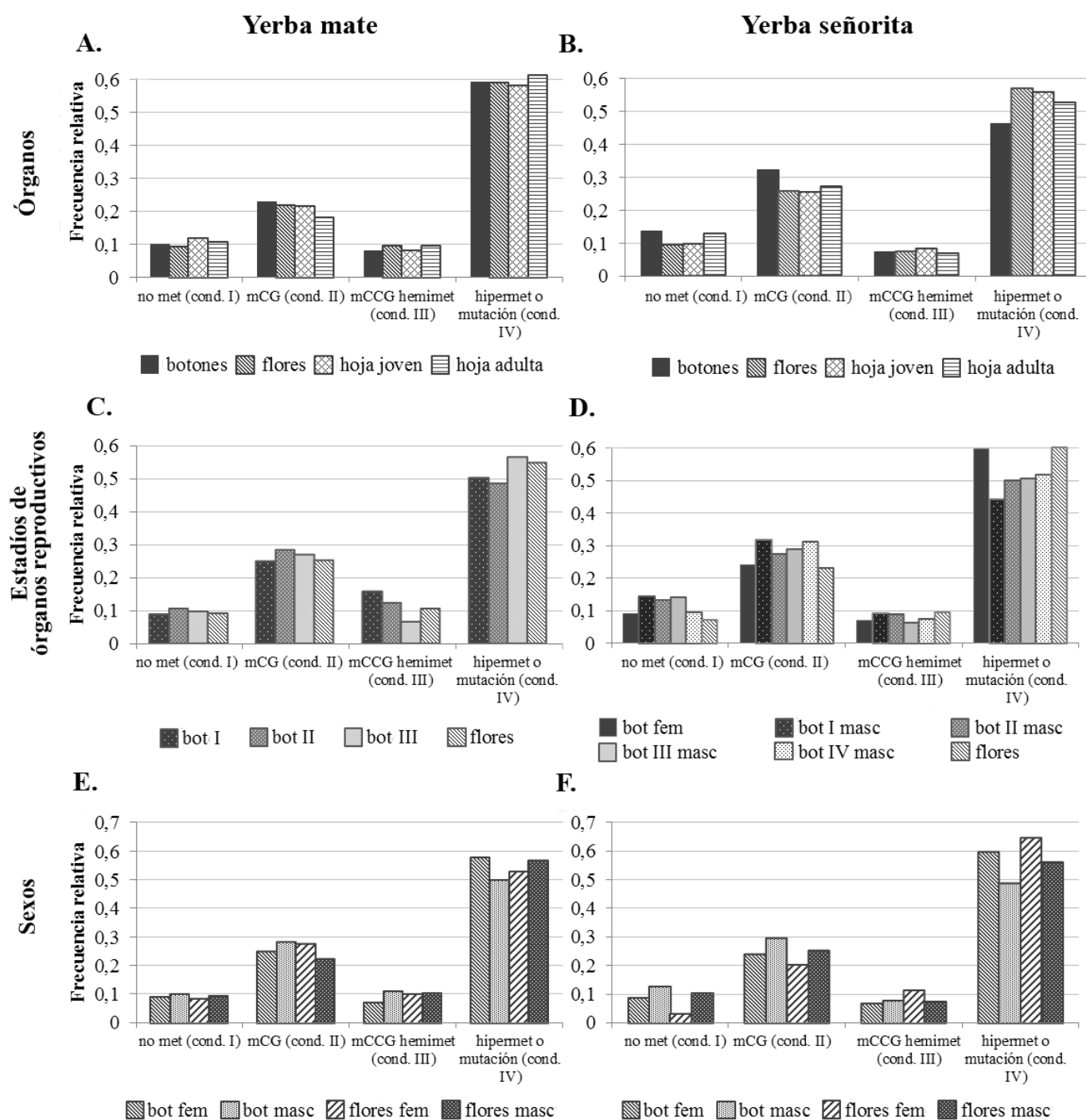
### 2.1. *Tipos de metilación: frecuencias*

En los análisis inter- e intraespecíficos, las frecuencias relativas de las condiciones de metilación (I-IV) mostraron valores y patrones similares. A nivel interespecífico, se observó una predominancia de la condición IV, seguida de la II, I y III. Si bien las diferencias entre las especies son leves, la frecuencia de sitios mCCG hemimetilados (condición III) en la yerba mate es ligeramente mayor, mientras que en yerba señorita se detectó una proporción levemente mayor de sitios de condición II y IV (**Figura 23**). Al analizar cada grupo de muestras, en términos generales, se mantiene la predominancia de la condición IV, y la consecución de las condiciones II, I y III. Para la yerba mate, la frecuencia de los tipos de metilación en las hojas jóvenes respecto de las adultas, y en los botones respecto a las flores, fue mayor para las condiciones I y II, y menor para las condiciones III y IV, aunque entre botones y flores la condición IV fue idéntica (**Figura 24A**). Se observaron cambios relativos similares en las frecuencias de las condiciones I y II a medida que avanza el desarrollo del botón floral, y se detectó una tendencia a la disminución en las frecuencias de metilación mCCG (III) durante la maduración del pimpollo, y que ésta aumenta en las flores abiertas (**Figura 24C**). Los botones femeninos tuvieron menor frecuencia de condiciones I, II y III que los botones masculinos, pero mayor de hipermetilación (IV). Las flores femeninas solo se diferenciaron de las masculinas en las frecuencias de las condiciones II y IV (**Figura 24E**).

En cuanto a la yerba señorita, las hojas adultas respecto a las jóvenes y los botones respecto a las flores, exhibieron mayor frecuencia de sitios de condición I y II, mientras que se invierte para las condiciones III y IV, más notoriamente (**Figura 24B**). También en esta especie se detectó una tendencia a la disminución de la frecuencia de sitios mCCG (III) conforme avanza el desarrollo de los botones (estadios bot I-III), seguida por un aumento en etapas posteriores (bot IV y flor abierta) (**Figura 24D**). Los botones femeninos y masculinos exhibieron la misma relación entre tipos de metilación que los de la yerba mate, mientras que las flores mostraron una relación contraria, además, se observó en las flores femeninas la menor frecuencia de sitios no metilados (I) (**Figura 24F**).



**Figura 23.** Histograma de las frecuencias relativas de cada tipo de metilación. En gris oscuro, yerba mate; en gris claro, yerba señorita. no met: no metilado; mCG: citosina interna metilada; mCCG hemimet: citosina externa hemimetilada; hipermet o mutación: metilaciones no reconocidas por *HpaII* y *MspI* o falta de corte enzimático por diferencias genéticas; cond.: condición.



**Figura 24.** Histogramas de frecuencias relativas de los tipos de metilación en cada grupo de muestras analizado. **A, B:** distinción entre tipo de órgano; **C, D:** entre estadios de órganos reproductivos; **E, F:** entre sexos. Las abreviaturas son las mismas que en la Figura 23.

## 2.2. *Diversidad y variación epigenética*

Los índices de Shannon de las matrices transformadas indican que los conjuntos de datos YM+YS, YM y YS alojan diversidades “fenotípicas” similares. En todos los casos, el índice fue mayor para las matrices MSL-MS respecto a las MSL-MS2 y NML (**Tabla 17**). Este resultado podría reflejar artefactos metodológicos producto del exceso de hipermetilación y/o diferencias genéticas, registrados como dato faltante en las matrices MSL-MS. Por esto, y dado que la transformación MS2 discrimina mejor los grupos (ver más adelante), es que se le prestará mayor atención al análisis de la matriz MSL-MS2.

**Tabla 17.** Valores del índice de diversidad ( $I \pm DE$ ) de cada matriz transformada.

<b>Matriz</b>	<b>MSL-MS</b>	<b>MSL-MS2</b>	<b>NML</b>
<b>YM+YS</b>	0,829 $\pm$ 0,180	0,438 $\pm$ 0,256	0,360 $\pm$ 0,208
<b>YM</b>	0,825 $\pm$ 0,178	0,616 $\pm$ 0,243	0,362 $\pm$ 0,159
<b>YS</b>	0,848 $\pm$ 0,142	0,400 $\pm$ 0,103	0,469 $\pm$ 0,184

YM+YS: matriz interespecífica. YM: matriz de yerba mate; YS: matriz de yerba señorita. DE: desvío estándar.

En general, se observó que las plantas de yerba mate albergan un mayor porcentaje de *loci* variables que las de yerba señorita (**Tabla 18**). En ambas especies, los botones presentaron los mayores números de *subepiloci* y porcentajes de polimorfismo, mientras que las flores presentaron los menores valores, y los órganos vegetativos tuvieron valores intermedios. En los grupos de muestras vegetativas y reproductivas se detectó una mayor abundancia de marcadores no metilados (condición I), seguidos en menor medida por los metilados de condición II y III (**Tabla 18**).

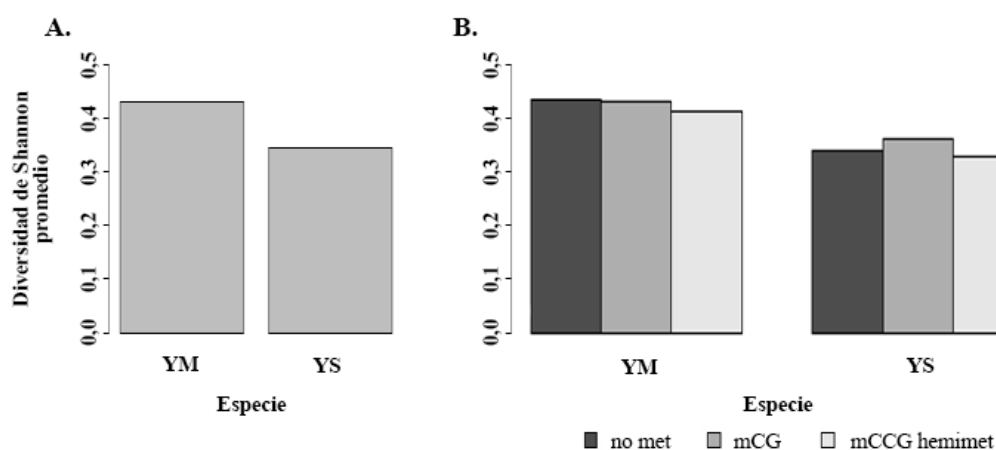


**Tabla 18.** Estimaciones estadísticas de las matrices MSL-MS2.

Matriz	Grupo	N	Número de subepiloci	%loci con variación	Número de marcadores de condición		
					I	II	III
YM+YS	YM	19	367	62,7%	136	119	112
	YS	14	293	49,7%	103	95	95
YM	bot	7	281	68,8%	110	91	80
	flores	4	190	41,5%	76	62	52
	HJ	4	240	57,0%	114	66	60
	HA	4	230	54,3%	100	66	64
YS	bot	6	233	68,7%	107	67	59
	flores	2	119	29,0%	58	35	26
	HJ	3	190	54,3%	72	50	68
	HA	3	173	45,7%	85	53	35

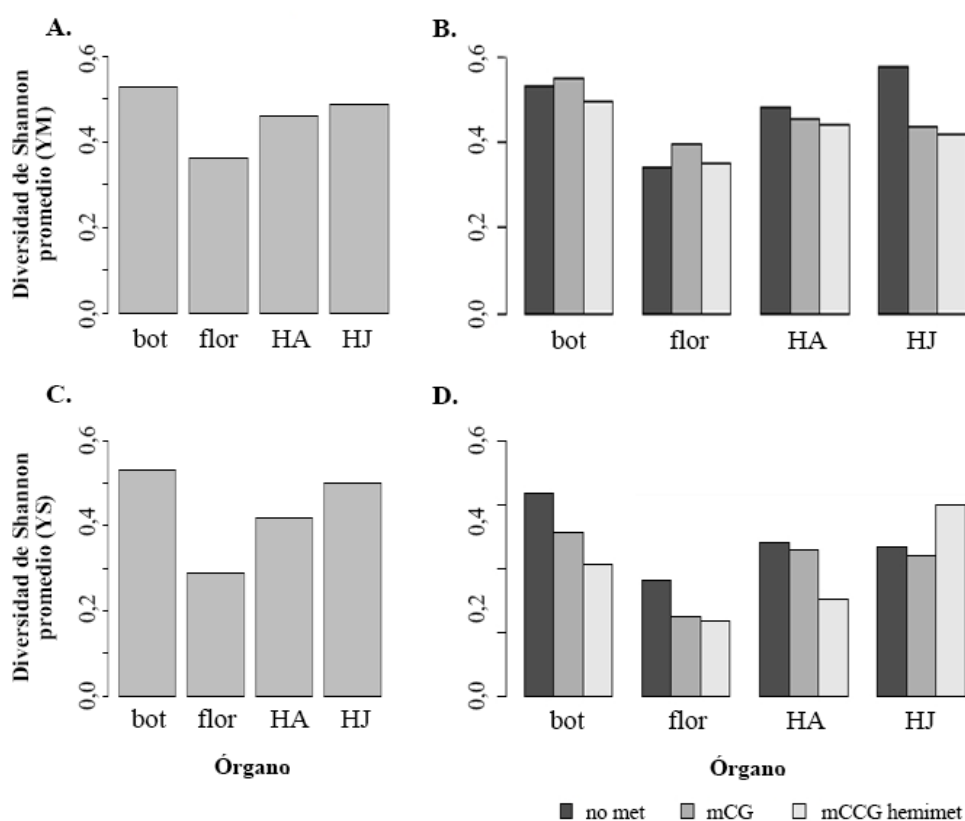
“Subepiloci” refiere a las columnas de la matriz MSL-MS2 resultantes de la transformación aplicada, en donde cada locus se subdivide acorde al número de condiciones que presenta. bot, botones. HJ, hoja joven. HA, hoja adulta.

El análisis de las matrices MSL-MS2 indicó que las muestras de yerba mate albergan mayor diversidad epigenética que las de yerba señorita (**Figura 25A**), y que los índices de diversidad promedio son similares entre los tipos de metilación (**Figura 25B**).



**Figura 25.** Histogramas de diversidad promedio para la matriz YM+YS-MSL-MS2 (interespecífica) **A.** Diversidad para la totalidad de los sitios metilados. **B.** Diversidad estimada por tipo de metilación.

Para ambas especies, los botones y las flores se muestran más contrastantes en cuanto a la diversidad epigenética. Las hojas jóvenes resultaron con mayor diversidad que las hojas adultas, siendo esto más evidente en la yerba señorita (**Figura 26A,C**). Los órganos estudiados de la yerba mate mostraron homogeneidad en términos de diversidad en cuanto a las distintas condiciones de metilación (I-III); aunque las hojas jóvenes exhibieron la mayor diversidad de sitios no metilados (condición I) (**Figura 26B**). En cambio, para la yerba señorita los botones y las flores presentaron una diversidad considerablemente mayor en los sitios no metilados; en los botones la metilación fue más diversa en contextos mCG (condición II) que en los mCCG (condición III). Las hojas jóvenes y adultas contrastaron en la diversidad de los sitios de condición III, siendo considerablemente más alta en las jóvenes y marcadamente menor en las adultas (**Figura 26D**).



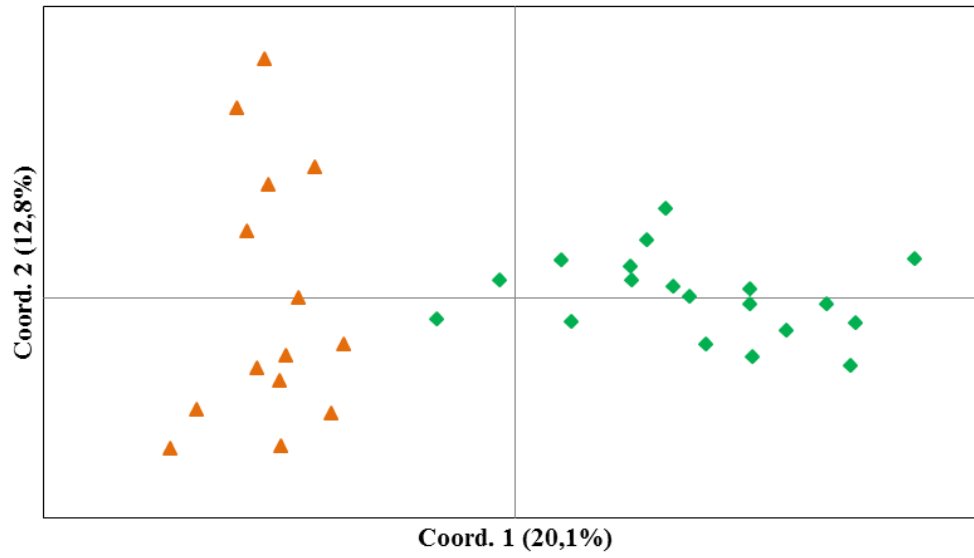
**Figura 26.** Histogramas de los índices de diversidad promedio derivados de la matriz MSL-MS2 y graficados por órgano (A y C) y por tipo de metilación (B y D). A-B. Yerba mate. C-D. Yerba señorita. Bot, botones. HA, hoja adulta. HJ, hoja joven. YM, yerba mate. YS, yerba señorita.

### 2.3. Análisis de agrupamiento

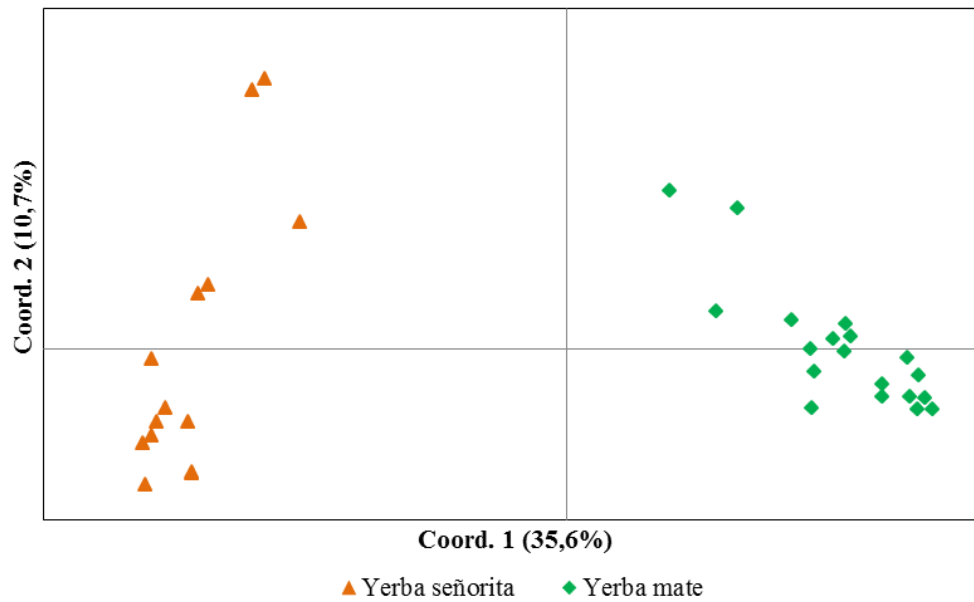
#### 2.3.1. Sitios susceptibles a la metilación, MSL

El AMOVA de las matrices MSL-MS y MSL-MS2 de YM+YS mostraron diferencias significativas entre las especies ( $\Phi_{ST}= 0,263$  con  $p < 0,001$  y  $\Phi_{ST}= 0,299$  con  $p= 0,001$ , respectivamente). En los PCoA se reconocieron dos grupos bien diferenciados según la especie (**Figura 27**). Los agrupamientos de DAPC derivados de MSL-MS no se correspondieron con las especies. Con  $K=2$ , el grupo de muestras de yerba señorita incluyeron 5 de yerba mate (probabilidad de pertenencia,  $P_{bp} \geq 74,5$  %; **Figura AI-2A**); mientras que con  $K=3$  (óptimo) los grupos involucraron muestras de ambas especies ( $P_{bp} \geq 61,9$  %; **Figura AI-3A**). El agrupamiento derivado de la matriz MSL-MS2 con  $K=2$  se corresponde con las especies ( $P_{bp}= 100\%$ ; **Figura AI-3A**); con el óptimo de  $K=3$ , las muestras de yerba señorita se subdividieron en dos grupos, uno incluyendo muestras de yerba mate ( $P_{bp} \geq 83$  %; **Figura AI-3B**).

A. MSL-MS



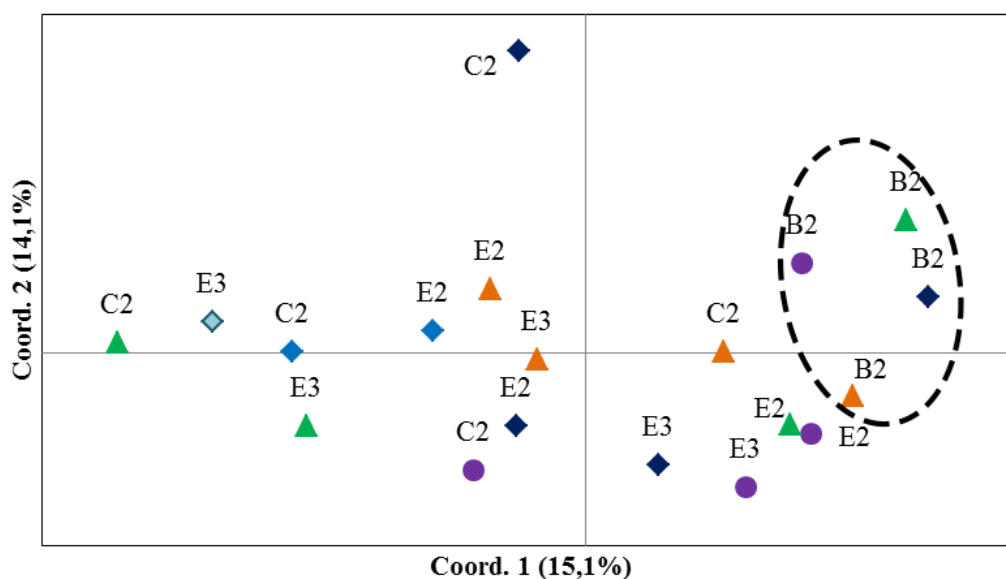
B. MSL-MS2



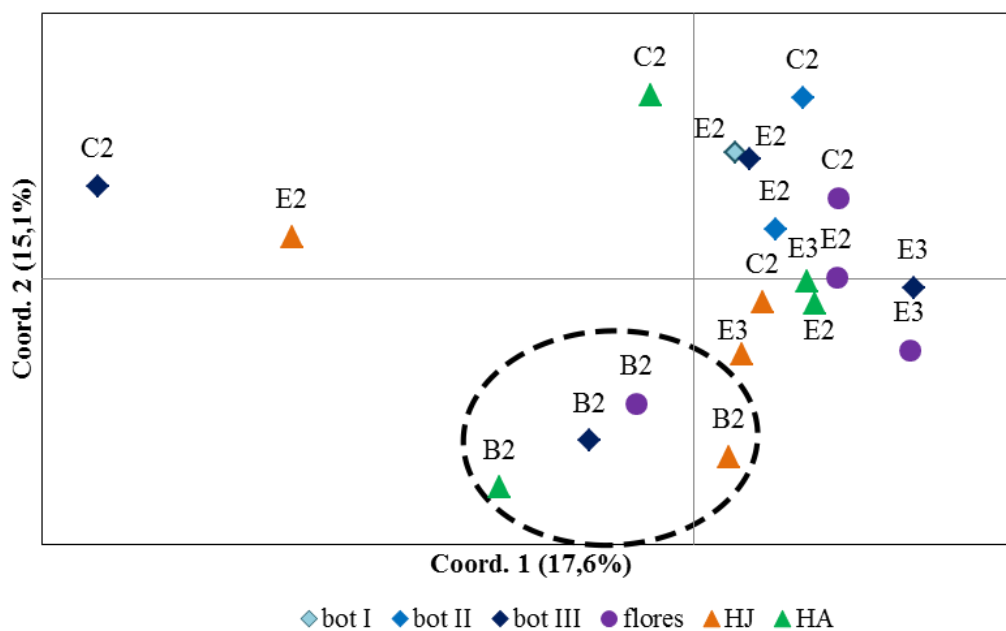
**Figura 27.** Análisis de PCoA interespecífico de los patrones de metilación. **A.** Matriz MSL-MS. **B.** Matriz MSL-MS2.

En los PCoA para ambas matrices de yerba mate (MSL-MS y MSL-MS2) se observó que las muestras del individuo B2♀ conformaron el único agrupamiento evidente (**Figura 28**).

A. MSL-MS



B. MSL-MS2



**Figura 28.** Análisis de PCoA de las muestras de yerba mate. **A.** Matriz YM-MSL-MS (29,2% de la varianza retenida). **B.** Matriz YM-MSL-MS2 (32,7% de la varianza acumulada). Se encierran con líneas punteadas los posibles agrupamientos. botI-III: estadios del desarrollo de los botones florales. HJ, hojas jóvenes. HA, hojas adultas.

Los distintos AMOVA realizados, ya fuera considerando agrupamientos según órgano (distinguiendo o no estadios de botones), tipo de órgano (reproductivo vs. vegetativo) o sexo, no evidenciaron agrupamientos estadísticamente soportados

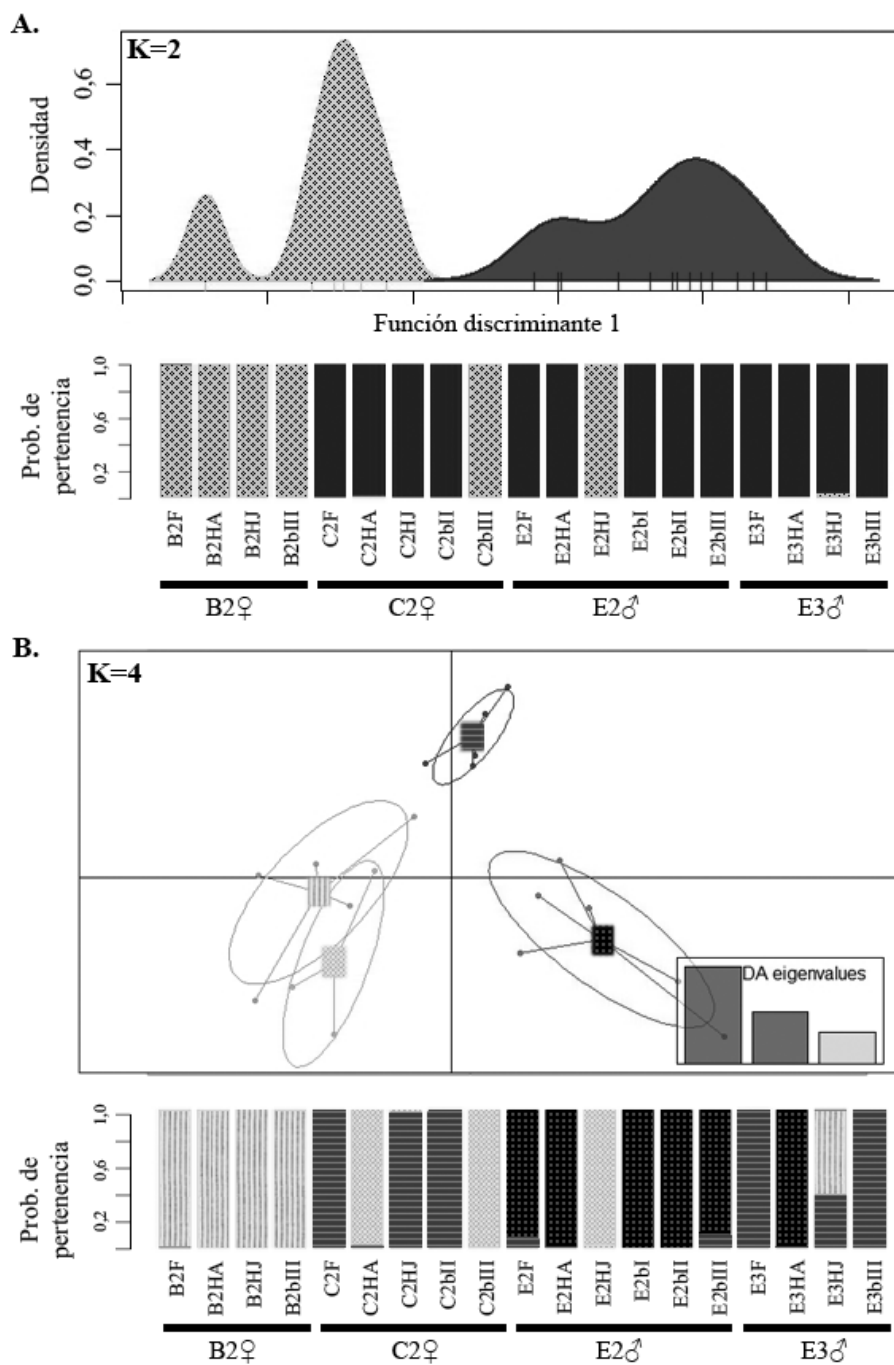
(resultados no mostrados). Únicamente los AMOVA en los que los agrupamientos se realizaron según “individuos” mostraron  $\Phi_{ST}$  globales significativos. Los  $\Phi_{ST}$  de la matriz MSL-MS apoyaron la diferenciación de B2♀; mientras que los de la matriz MSL-MS2 evidenciaron diferenciación entre todos los individuos (**Tabla 19**).

**Tabla 19.**  $\Phi_{ST}$  pareados de las matrices MSL-MS (sobre la diagonal) y MSL-MS2 (debajo de la diagonal) entre muestras de yerba mate discriminadas según individuo.  $\Phi_{ST}$  global<sub>MS</sub> = 0,094 (p= 0,005);  $\Phi_{ST}$  global<sub>MS2</sub> = 0,132 (p= 0,001).

	<b>B2</b>	<b>C2</b>	<b>E2</b>	<b>E3</b>
<b>B2</b>	-	0,160*	0,136*	0,178*
<b>C2</b>	0,162*	-	0,044	0,082
<b>E2</b>	0,160**	0,082*	-	0,002
<b>E3</b>	0,164*	0,122***	0,122**	-

\* p< 0,05; \*\* p< 0,01; \*\*\* p< 0,001.

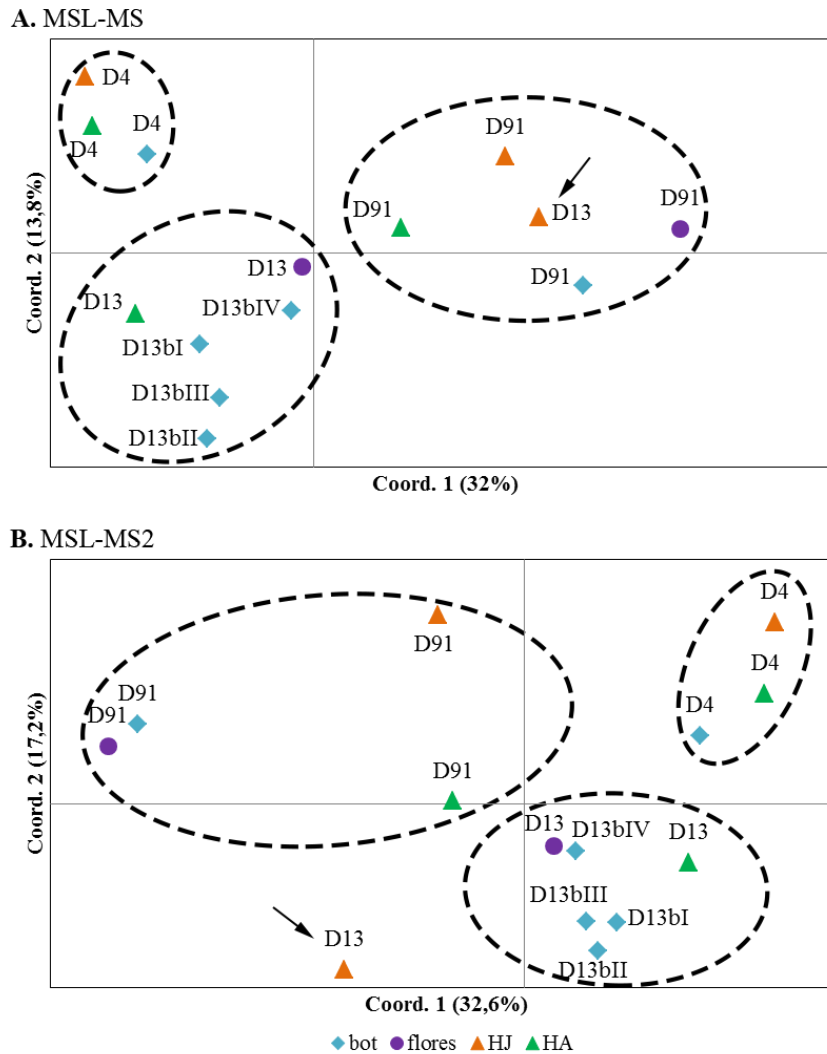
Los DAPC se realizaron con K=2 (óptimo) y K=4, y en ningún caso se obtuvieron grupos acordes a los órganos o sexos. El DAPC de la matriz MSL-MS con K=2 (Pbp  $\geq 97,6$  %) tampoco agrupó las muestras según el individuo (**Figura AI-4A**); mientras que con K=4 se observó un grupo conformado por todas las muestras del individuo B2♀ (junto con otras, y con Pbp  $\geq 97,7$  %), los grupos restantes estuvieron formados por muestras de distintos individuos (Pbp  $\geq 96$  %; **Figura AI-4B**). Para la matriz MSL-MS2, tanto con K=2 como con K=4, solo B2♀ formó un grupo bien definido (Pbp  $\geq 99,6$  % y  $\geq 99,1$  % respectivamente; **Figura 29A,B**).



**Figura 29.** Gráficos DAPC y compoplot de la matriz MSL-MS2 de yerba mate (29,8% de la varianza acumulada). **A.** K=2. **B.** K=4. Los patrones se corresponden entre los DAPC y sus compoplot respectivos.

Los PCoA de yerba señorita también evidenciaron una tendencia al agrupamiento según individuos. La muestra de hoja joven de D13♂ siempre apareció alejada de las restantes del mismo individuo (**Figura 30A,B**), mientras que D91♀ mostró gran dispersión entre sus órganos reproductivos y vegetativos (**Figura 30B**). Únicamente los AMOVA considerando agrupamientos por individuo mostraron resultados

significativos, los  $\Phi_{ST}$  pareados indicaron que todos los individuos se diferenciaron significativamente en sus perfiles de metilación (**Tabla 20**).



**Figura 30.** Análisis de PCoA de las muestras de yerba señorita. **A.** Matriz MSL-MS (45,8% de la varianza retenida). **B.** Matriz MSL-MS2 (49,8% de la varianza acumulada). Con líneas punteadas se indican los posibles agrupamientos según planta donante. Las flechas muestran puntos *outlier*. bot, botones florales. bI-IV: estadios del desarrollo de los botones. HJ, hojas jóvenes. HA, hojas adultas.

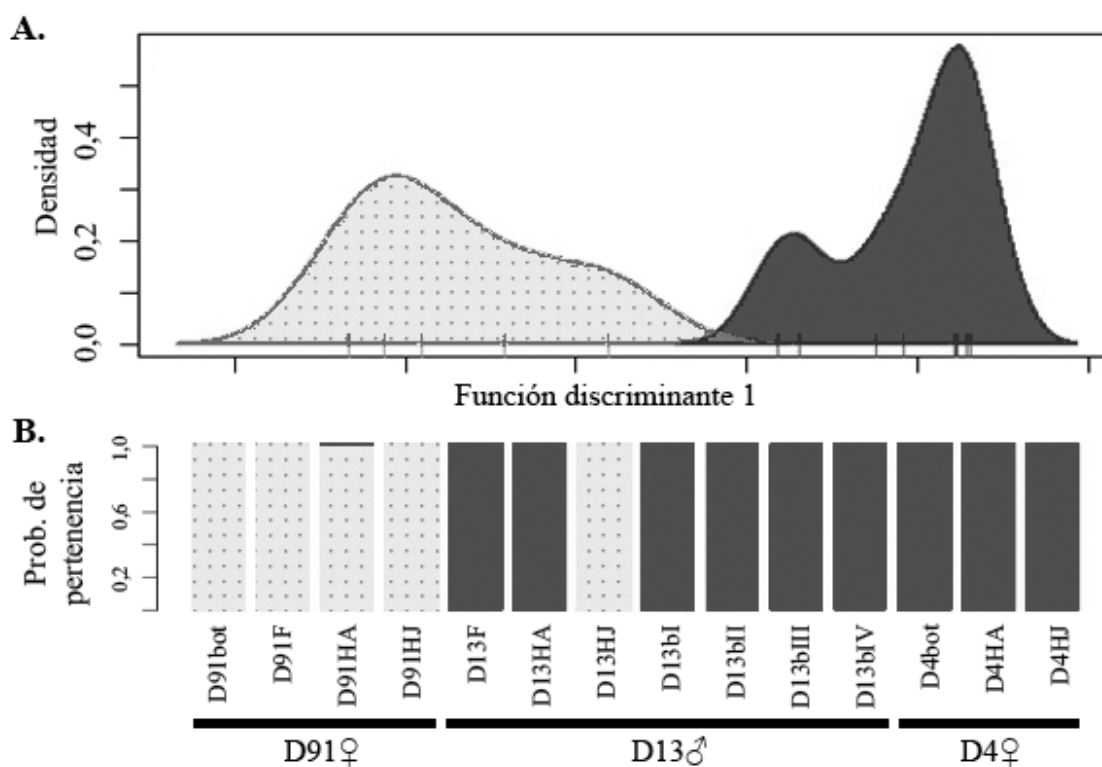


**Tabla 20.**  $\Phi_{ST}$  de a pares de las matrices MSL-MS (sobre la diagonal) y MSL-MS2 (debajo de la diagonal) entre muestras de yerba señorita discriminadas según individuo.  $\Phi_{ST-MS}$  global= 0,330 ( $p < 0,001$ );  $\Phi_{ST-MS2}$  global= 0,225 ( $p = 0,001$ ).

	D4♀	D91♀	D13♂
D4♀	-	0,545*	0,253**
D91♀	0,328*	-	0,268**
D13♂	0,216*	0,186**	-

\*  $p < 0,05$ . \*\*  $p < 0,01$ .

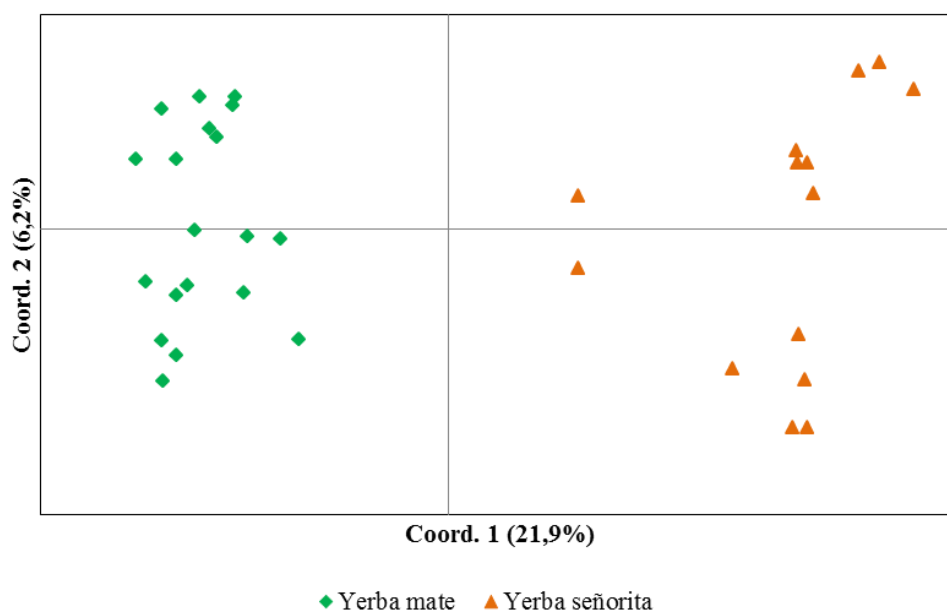
Para ambas matrices, con  $K=2$  (óptimo) las muestras provenientes de D91♀ se separaron del resto ( $Pbp_{MS} \geq 88,8\%$  y  $Pbp_{MS2} \geq 98\%$ ; **Figuras 31** y **AI-5A**). Con  $K=3$ , las muestras tendieron a dividirse según el individuo ( $Pbp_{MS} \geq 62,8\%$  y  $Pbp_{MS2} = 100\%$ ; **Figuras AI-5B** y **AI-6**).



**Figura 31.** Análisis discriminante de la matriz MSL-MS2 de yerba señorita con  $K=2$ . **A.** Gráfico DAPC (varianza retenida= 18,51%). **B.** Gráfico compoplot. Los patrones de los DAPC se corresponden con los de su compoplot respectivo.

2.3.2. Análisis genético NML

El análisis de la matriz NML (es decir, contemplando los sitios no susceptibles a la metilación) permitió la diferenciación significativa de las dos especies (AMOVA:  $\Phi_{ST}=0,309$ ;  $p < 0,001$ ), visualizándose una mayor dispersión en las muestras de yerba señorita (**Figura 32**). Las especies también fueron discriminadas en los DAPC con  $K=2$  (**Figura AI-7**) y con el  $K$  óptimo  $=3$  (**Figura AI-8**), donde las muestras de *I. dumosa* se separaron en dos grupos de acuerdo con los sexos ( $P_{bp} = 100\%$ ).



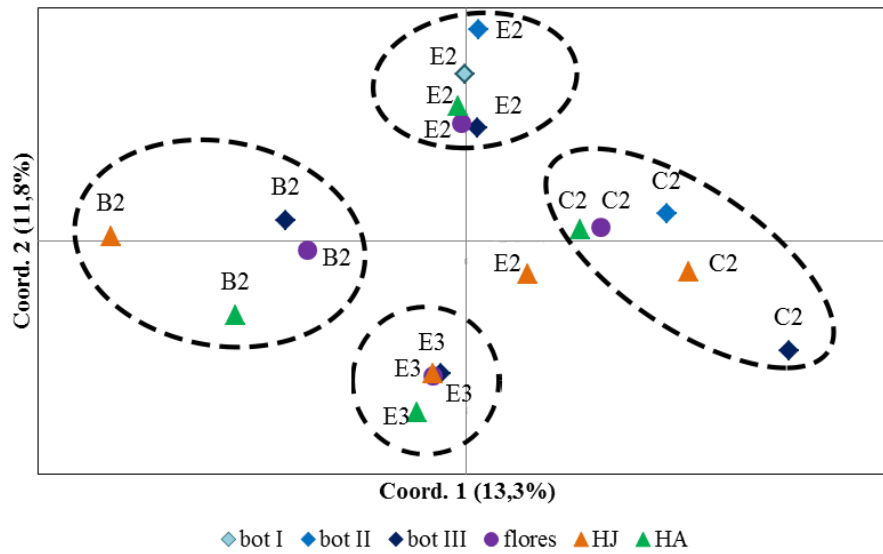
**Figura 32.** Análisis de PCoA interespecífico con datos genéticos (matriz YM+YS-NML).

En yerba mate se detectaron agrupamientos conformados por muestras de un único individuo, significativamente soportados (**Tabla 21**) y no solapantes (**Figura 33**). Cuando se evaluaron los agrupamientos predefinidos (es decir, estadios florales, tipo de órgano o sexos) éstos no resultaron estadísticamente diferenciados (no se muestran).

**Tabla 21.** Valores de  $\Phi_{ST}$  pareado entre individuos de yerba mate a partir de datos MSAP genéticos.  $\Phi_{ST}$  global= 0,185 ( $p < 0,001$ ).

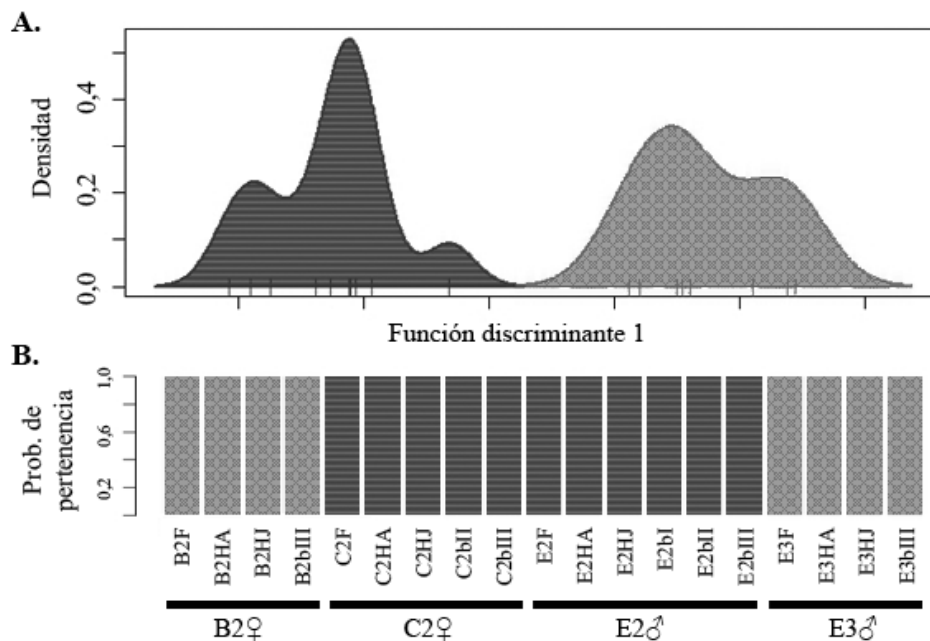
	B2	C2	E2	E3
B2	-			
C2	0,222**	-		
E2	0,161**	0,117**	-	
E3	0,204*	0,177**	0,253**	-

\*  $p < 0,05$ . \*\*  $p < 0,01$ .



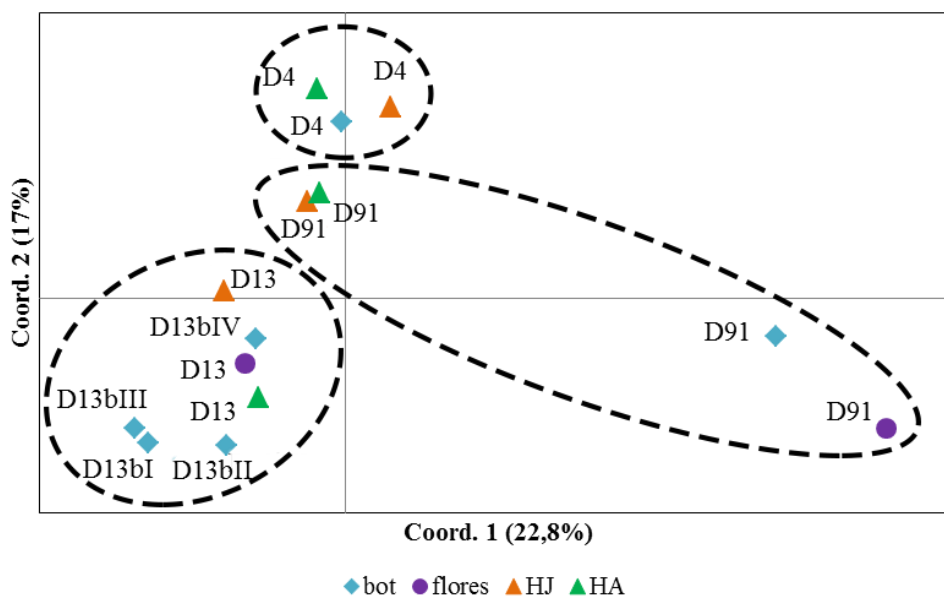
**Figura 33.** Análisis de PCoA de la matriz NML de yerba mate (25,1% de la varianza retenida). Las líneas punteadas muestran los agrupamientos más probables propuestos. botI-III: estadios del desarrollo de los botones florales. HJ, hojas jóvenes. HA, hojas adultas.

Al considerar el óptimo K=2, los dos grupos estuvieron conformados por las muestras de dos individuos (Pbp= 100 %; **Figura 34**). Al forzar la discriminación en K=4, cada grupo correspondió a un individuo (Pbp  $\geq$ 93,9 %; **Figura AI-9A**).



**Figura 34.** Análisis discriminante de la matriz YM-NML con K=2 y una variable clasificadora. **A.** Gráfico DAPC (34,8% de la varianza explicada). **B.** Gráfico compoplot. Los patrones se corresponden entre los distintos gráficos.

Similarmente, las muestras de yerba señorita generaron agrupamientos por individuo (Figura 35), dos de los cuales resultaron diferenciados significativamente (Tabla 22). Los agrupamientos predefinidos no evidenciaron diferencias significativas (no mostrado).



**Figura 35.** Análisis de PCoA de las muestras de yerba señorita. Con líneas punteadas se muestran los posibles agrupamientos según la planta donante. bot, botones florales. HJ, hojas jóvenes. HA, hojas adultas. bI-IV: estadios del desarrollo de los botones.

**Tabla 22.** Valores de  $\Phi_{ST}$  pareado entre individuos de yerba señorita a partir de datos MSAP genéticos.  $\Phi_{ST}$  global = 0,242 ( $p < 0,001$ ).

	D4♀	D91♀	D13♂
D4♀	-		
D91♀	0,122	-	
D13♂	0,350**	0,229**	-

\* Valores significativos con  $p < 0,05$ . \*\* Valores significativos con  $p < 0,01$ .

En el DAPC con  $K=2$  (óptimo) se observó que los órganos reproductivos del individuo D91♀ se apartaron del resto de las muestras ( $P_{bp} = 100\%$ ; Figura AI-10A). Al excluir estas muestras, el análisis con  $K=2$  generó un agrupamiento similar a los resultados del AMOVA pareado ( $P_{bp} \geq 83,3\%$ ; Figura AI-10B). Luego, al forzar  $K=3$

los grupos se correspondieron mayoritariamente con el individuo donante ( $P_{bp} \geq 81,4$  %; **Figura AI-11**).

#### 2.4. Identificación de fragmentos diferencialmente metilados

Se seleccionaron y aislaron 117 bandas que mostraron patrones de metilación diferenciales entre grupos de muestras. Se secuenciaron 28 (**Tabla AI-10**), y se pudo asignar una anotación probable al 25 %, las restantes no arrojaron resultados significativos. Las secuencias identificadas mostraron similitud significativa con entradas de plantas y estarían relacionadas con el metabolismo de carbohidratos, vías de señalización o regulación de la transcripción en respuesta a heridas o en resistencia a patógenos, funciones estructurales, transporte de proteínas, entre otras funciones (**Tabla 23**). No se detectaron secuencias de elementos móviles (datos no mostrados).

**Tabla 23.** Resultados de las comparaciones de los fragmentos secuenciados con las bases de datos públicas.

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado (tipo de metilación)	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
1	346	YS♂ (u,h) vs. YS♀ (f)	LN713260.1; Cromosoma 6 (ADN genómico), <i>Cucumis melo</i> (1/100; 2e-13; 95%; 24%).	-	NS (2,6)	-	
2	290	botI YS♂ (u) vs. f	NS (0,036)	NS (1,6)	NS (8,2)	-	
3	208	YS♂ (u) vs. YS♀ (f)	NS (0,03)	-	NS (0,93)	-	
4	180	YS♂ (i) vs. YM♂ (f,h)	NS (0,002)	-	NS (9,5)	-	
5	362	botones YS♂ (i) vs. f	NS (3e-04)	-	NS (1,4)	-	
6	233	YS♂ (u) vs. f	NS (0,028)	NS (4,1)	NS (0,13)	-	
7	182	YS♂ (u) vs. YS♀ (f)	XM_020670610.1; Beta-glucosidasa 12, ARNm (predicho), <i>Amborella trichopoda</i> (1/100; 3e-13; 76%; 65%).	KHN33321.1; Beta-glucosidasa no cianogénica, <i>Glycine soja</i> (1/100; 5e-13; 69%; 80%; 69%).	XM_016701488.1; tipo furcatin hidrolasa, ARNm (predicho) <i>Capsicum annuum</i> (1/100; 3e-14; 69%; 85%; 74%).	PF00232; Glicósido hidrolasa 1 (3,66e-08)	Metabolismo de carbohidratos, respuestas a hormonas y a estrés salino
8	298	YS♀ (u,i) vs. YS♂ (f)	XM_021430691.1; Homólogo del regulador de transcritos sin sentido 1, ARNm (predicho), <i>Herrania umbratica</i> (1/100; 1e-46; 83%; 66%).	XP_015085052.1; Homólogo del regulador de transcritos sin sentido 1, <i>Solanum pennellii</i> (1/2; 1e-05; 82%; 81%; 44%).	XM_017786253.1; Homólogo del regulador de transcritos sin sentido 1, ARNm (predicho), <i>Vitis vinifera</i> (1/50; 1e-20; 79%; 84%; 94%).	-	NMD ( <i>nonsense-mediated decay</i> ), defensa contra bacterias, señalización, respuesta a heridas
9	220	YS♂ (u) vs. YS♀ (f)	HG975515.1; Cromosoma 3 (ADN genómico), <i>Solanum lycopersicum</i> (1/3; 1e-05; 77%; 35%).	-	NS (3,1)	-	
10*	108	C2♀ (h) vs. f	NS (0,036)	NS (0,006)	NS (1,4)	-	
11	399	YS♀ (i) vs. f	NS (0,18)	NS (0,006)	NS (1,4)	-	

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado (tipo de metilación)	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
12	160	HJ y HA de YM♀ (i) vs. f	NS (0,21)	NS (0,02)	NS (1,8)	-	
13	285	HA de YM (f) vs. u,h	NS (1,5)	-	NS (2,7)	-	
14	128	YS♂ (i) vs. f	NS (0,046)	-	NS (0,34)	-	
15	111	YS♂ (i) vs. f	NS (1,6)	-	NS (2,7)	-	
16	116	HJ de YS♀ (u) vs resto (f)	XM_006476361.2; Transcripto no caracterizado, ARNm (predicho), <i>Citrus sinensis</i> (1/3; 1e-07; 72%, 100%).	XP_017231537.1; Proteína que contiene dominios homólogos a la calponina (predicho), <i>Daucus carota</i> (1/86; 6e-09; 68%; 89%; 98%).	XM_017376050.1; Proteína que contiene dominios homólogos a la calponina, ARNm (predicho), <i>Daucus carota</i> (1/100; 2e-09; 68%; 89%; 99%).	-	Estructura del citoesqueleto, transducción de señales
22	332	bot (f) vs. hojas (i) YM	NS (1,8)	XP_016199504.1; Translocasa de cloroplasto 90, plastídica, <i>Arachis ipaensis</i> (1/6; 3e-05; 43%; 53%; 50%).	XM_016079830.2; Translocasa de cloroplasto 90, plastídica, ARNm (predicho), <i>Arachis duranensis</i> (1/2; 5e-05; 42%; 54%; 51%).	PF11886; Translocasa de cloroplasto 159/132, dominio de anclaje a la membrana (1e-04)	Importación de proteínas al cloroplasto
23	290	bot (u) vs. flores y hojas (f) YM♀	NS (0,036)	-	NS (1,3)	-	
24	175	botI y II (h) vs. botIII (i/f) YM♂	NS (0,069)	-	NS (0,059)	-	
25	170	HA y flores ♂ (u) vs. resto (i) YM	KU365990.1; Proteína de resistencia al oídio tipo 8 (RPW8-a), ARNm, <i>Vitis pseudoreticulata</i> (1/93; 8e-20; 81%; 67%).	ANN89710.1; Proteína de resistencia al oídio tipo 8, <i>Vitis pseudoreticulata</i> (2/100; 3e-15; 71%; 80%; 91%).	KU365990.1; Proteína de resistencia al oídio tipo 8 (RPW8-a), ARNm, <i>Vitis pseudoreticulata</i> (1/100; 1e-15; 71%; 82%; 91%).	-	Resistencia a enfermedades fúngicas
26	152	bot (u) vs. hojas (f) YS♀	NS (0,061)	-	NS (1,4)	-	
27	237	reproductivo (i) vs vegetativo (f)	NS (1,2)	-	NS (8,2)	-	

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado (tipo de metilación)	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
		YS♂					
28	206	botI y II (h) vs resto (f) YM♀	NS (3,6)	1e-06 *	1e-07 *	-	
29	167	bot (u) vs resto (i) YS	NS (0,019)	NS (0,002)	NS (3e-04)	-	
30	217	botI y II (u) vs III (i) YM♂	XP_002278311.3; Fosfatidilinositol 4-quinasa gamma 8, ARNm (predicho), <i>Jatropha curcas</i> (2/69; 2e-17; 75%; 69%).	XP_002278311.3; Fosfatidilinositol 4-quinasa gamma 8 (predicho), <i>Vitis vinifera</i> (2/70; 2e-14; 57%; 74%; 80%).	XM_022998127.1; Fosfatidilinositol 4-quinasa gamma 8, ARNm (predicho), <i>Olea europaea</i> var. <i>sylvestris</i> (1/75; 1e-15; 61%; 75%; 78%).	-	Parte de las vías de señalización celular que involucran al segundo mensajero inositol 1,4,5-trifosfato (InsP3)
31	203	botI y II (u♀; h♂) vs. resto (f) YM	NS (0,082)	NS (6,6)	NS (3,0)	-	
32	173	bot (i) vs resto YS	NS (0,24)	-	NS (0,4)	-	
33 <sup>e</sup>	227	botI y II ♀ (i) vs. resto (f)	<b>base 7 a 61:</b> XM_008344751.2; Proteína de mantenimiento estructural de los cromosomas 4, ARNm (predicho), <i>Malus x domestica</i> (1/31; 1e-08; 87%; 100%).	NS (0,31)	NS (0,57)	-	Estructura de la cromatina
			<b>base 105 a 214:</b> XM_022157489.1; Transcripto no caracterizado, ARNm, <i>Helianthus annuus</i> (1/1; 2e-05; 86%; 100%)	NS (0,042)	NS (0,33)	-	

<sup>a</sup> longitud de la secuencia obtenida.



<sup>b</sup> comparación con bases de datos de nucleótidos: número y anotación del acceso recuperado. Entre paréntesis se indica: la posición del acierto (*hit*) entre los resultados significativos recuperados; valor-E obtenido; porcentaje de identidad; porcentaje de cobertura entre la secuencia incógnita y el acierto.

<sup>c</sup> comparación con bases de datos de proteínas (BlastX) y de secuencias nucleotídicas traducidas (tBlastX) usando la secuencia incógnita traducida: número y anotación del acceso recuperado. Entre paréntesis: posición del *hit*; valor-E; porcentaje de identidad; porcentaje de residuos compartidos positivos; porcentaje de cobertura de la secuencia incógnita por la secuencia recuperada.

<sup>d</sup> en caso de hallarse similitud significativa con un dominio proteico conservado, se muestra su accesión de PFAM, nombre y valor-E obtenido.

<sup>e</sup> esta secuencia resultó ser una quimera, teniendo la secuencia del adaptador *EcoRI* de por medio, por lo que se analizó las partes por separado.

\* no obtuvo aciertos significativos con secuencias de plantas.

NS, no significativo. Entre paréntesis se indica el E-valor obtenido. Los casos en los que no se recuperaron coincidencias significativas con secuencia alguna, se indican con un guión.

bot: botones florales; I, II: estadios del desarrollo de los botones florales. HJ: hojas jóvenes. HA: hojas adultas f: estado de hipermetilación o falta de secuencia diana; i: metilación CG; h: CCG hemimetilado; u: no metilado.

### 3. Evaluación de la variación en la expresión génica diferencial

#### 3.1. Genotipificación cDNA-AFLP

La muestra de botones florales del individuo C2♀ fue excluida de las matrices YM+YS y YM por poseer muchos datos faltantes.

La matriz YM+YS fue generada empleando los patrones de bandas derivados de 15 combinaciones de cebadores y consistió en 1157 *loci* (99 % variables; 0,96 % de datos faltantes; **Tabla 24**). Las matrices YM y YS tuvieron un número similar de *loci* (>700) y de *loci* variables (97 %) (datos faltantes: 1,04 % y 0,67 %, respectivamente). Los índices de diversidad resultaron similares para las tres matrices (**Tabla 24**).

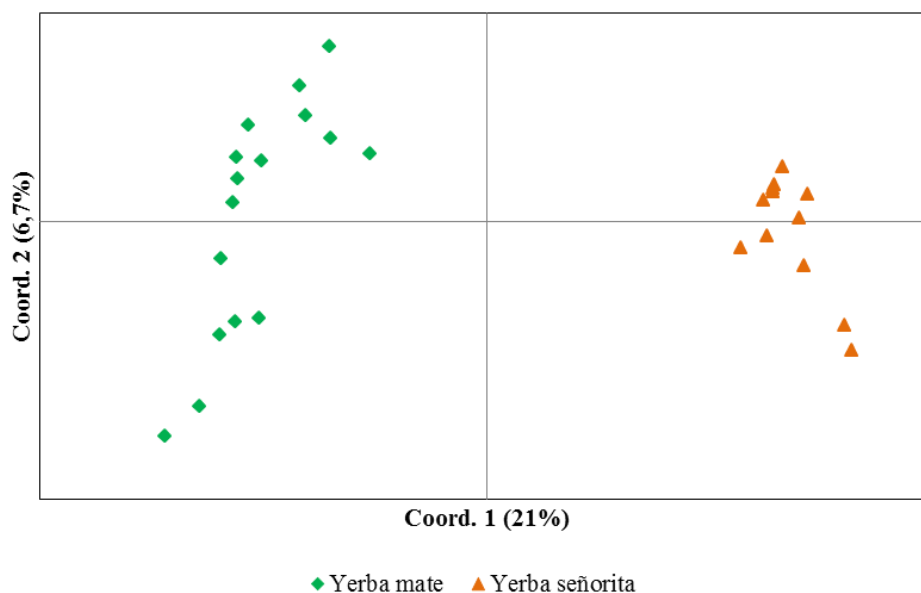
**Tabla 24.** Número de *loci* obtenidos para cada combinación de cebadores e índice de diversidad fenotípica de Shannon (I) promedio, por matriz.

Combinación de cebadores	YM + YS (28 UTOs)	YM (16 UTOs)	YS (12 UTOs)
E-AAC + M-CAC	75	49	33
E-AAG + M-CAA	96	67	60
E-AAG + M-CAG	100	67	63
E-AAG + M-CTA	82	57	54
E-AAG + M-CTG	108	54	78
E-ACA + M-CAT	72	38	49
E-ACC + M-CTC	69	54	37
E-ACG + M-CAA	76	35	54
E-ACG + M-CTT	47	25	28
E-ACT + M-CAC	75	45	40
E-ACT + M-CAT	57	30	36
E-AGC + M-CAG	72	50	32
E-AGC + M-CTA	47	34	30
E-AGG + M-CAA	105	67	67
E-AGG + M-CTT	76	48	45
<b>Total</b>	<b>1157</b>	<b>720</b>	<b>706</b>
<b>Promedio</b>	<b>77,13</b>	<b>48</b>	<b>47,07</b>
<b>I ± DE</b>	<b>0,457 ± 0,260</b>	<b>0,555 ± 0,238</b>	<b>0,605 ± 0,216</b>

DE: desvío estándar.

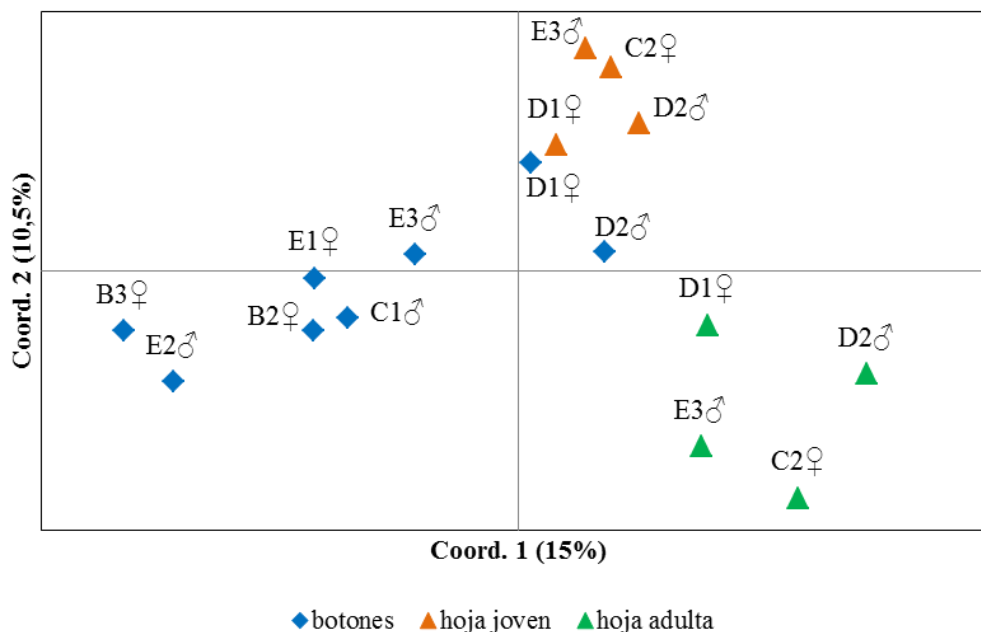
Los análisis a nivel interespecífico (YM+YS) coinciden en indicar que las muestras se agrupan según la especie (AMOVA:  $\Phi_{ST} = 0,295$  con  $p < 0,001$ ; **Figura 36**; Pbp= 100

%, **Figura AI-12**). Con  $K=3$  (valor óptimo) la yerba mate generó dos *clusters*, uno conteniendo únicamente muestras de hoja y el otro las restantes ( $P_{bp} \geq 77,1$  %; **Figura AI-13**).



**Figura 36.** Análisis PCoA de la matriz YM+YS.

Para las muestras de yerba mate se detectó una tendencia al agrupamiento según el tipo de órgano (**Figura 37**), y se hallaron diferencias significativas entre los grupos de órganos y entre todas las comparaciones pareadas (**Tabla 25**).



**Figura 37.** PCoA de la matriz YM.

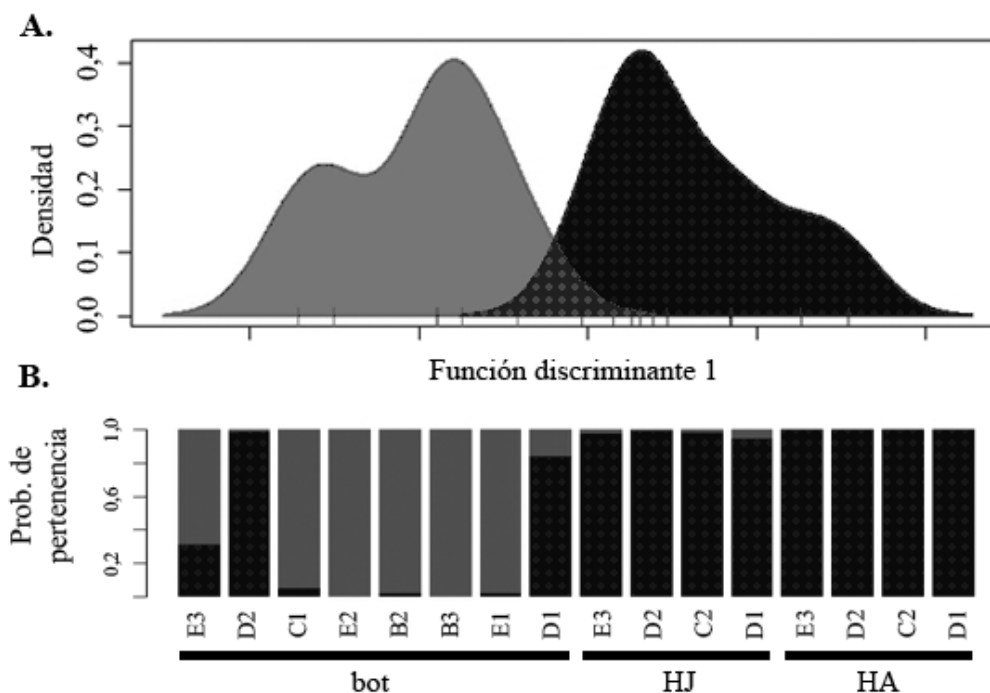
**Tabla 25.** Valores del  $\Phi_{ST}$  de a pares entre muestras de yerba mate (encima de la diagonal;  $\Phi_{ST}$  global= 0,1441 con  $p < 0,0001$ ) y de yerba señorita (debajo de la diagonal;  $\Phi_{ST}$  global= 0,092 con  $p = 0,002$ ), discriminadas según órgano.

	botones	hoja joven	hoja adulta
botones	-	0,111**	0,167**
hoja joven	0,005	-	0,161*
hoja adulta	0,120*	0,142*	-

\*  $p < 0,05$ , \*\*  $p < 0,01$

Con la intención de facilitar la comparación de estos resultados con los epigenéticos, se generó una nueva matriz de yerba mate de 11 UTOs y 595 loci (97 % variables), que permitió desestimar posibles agrupamientos según individuo: las diferencias entre los individuos resultaron no significativas tanto a nivel global ( $\Phi_{ST} = -0,003$ ;  $p = 0,500$ ) como de a pares (dato no mostrado).

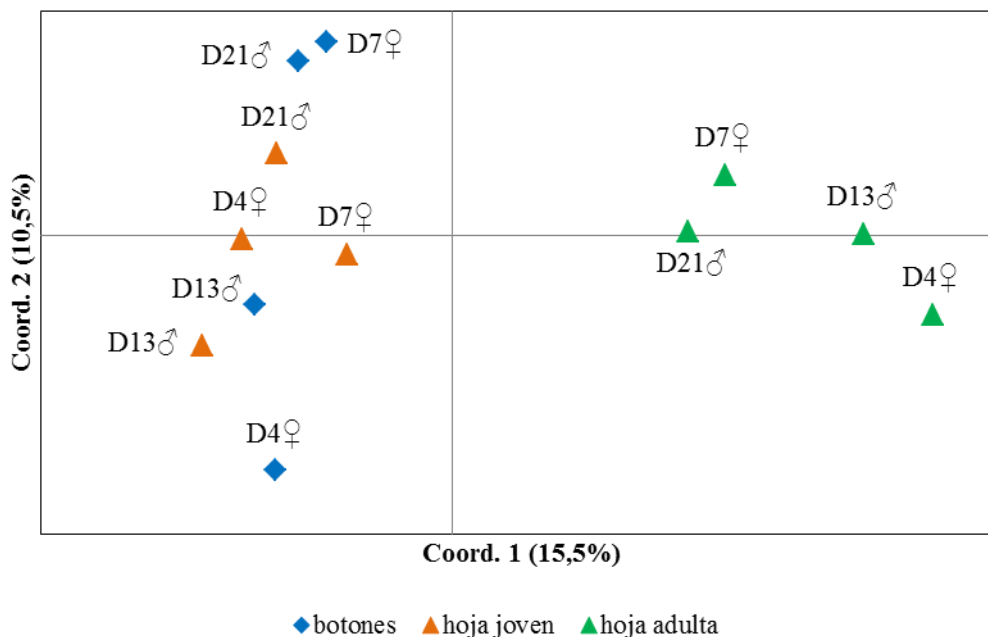
A grandes rasgos, el análisis discriminante con  $K=2$  separó los botones florales de las hojas ( $P_{bp} \geq 69,1$  %; **Figura 38**). Al forzar el agrupamiento según el número de órganos diferentes ( $K=3$ ) se recuperaron sendos grupos, aunque con probabilidad moderada ( $P_{bp} \geq 73$  %; **Figura AI-14**).



**Figura 38.** Análisis discriminante de la matriz YM con K=2. **A.** Gráfico DAPC (14,9% de la varianza explicada). **B.** Gráfico compoplot. Los patrones se corresponden entre los gráficos.

Cuando se analizaron los botones florales separadamente (matriz YM-bot, 16 UTOs x 527 loci –89 % variables–), no se hallaron agrupamientos destacables (**Figuras AI-15 y 16**; AMOVA:  $\Phi_{ST} = -0,001$  con  $p = 0,432$ ).

Para la yerba señorita, las muestras de hoja adulta se diferenciaron significativamente, mientras que los botones florales y las hojas jóvenes se solaparon (**Figura 39**; **Tabla 25**; **Figura AI-17**). Con K=2 (óptimo) todas las muestras tuvieron probabilidad del 100 % de pertenencia; mientras que forzando K=3, ésta fue  $\geq 63,1$  % (**Figura AI-18**). En esta especie tampoco se obtuvieron agrupamientos significativos a nivel individual ( $\Phi_{ST}$  global = -0,012 con  $p = 0,688$ ;  $\Phi_{ST}$  pareados = resultados no significativos, no mostrados).



**Figura 39.** PCoA de la matriz YS.

### 3.2. *Identificación de fragmentos diferencialmente expresados*

Se seleccionaron 148 bandas según su presencia/ausencia diferencial entre los grupos considerados. De éstas, 61 pudieron ser exitosamente aisladas y secuenciadas, obteniéndose 55 secuencias distintas (**Tabla AI-11**).

La comparación vía Blast produjo 23 coincidencias significativas con secuencias nucleotídicas y/o aminoacídicas, de las cuales el 43,5 % mostraron además similitud significativa con dominios proteicos conservados. Las funciones predichas estarían relacionadas con la regulación de la expresión génica a distintos niveles, con la estructura y/o actividad de organelas, con metabolismo, biosíntesis y otros procesos celulares, con resistencia a patógenos, defensa y/o respuesta a estrés, y con la estructura de la pared celular (**Tabla 26**). Ningún fragmento mostró similitud significativa con entradas de retrotransposones (datos no mostrados).

**Tabla 26.** Resultados de las comparaciones de los fragmentos secuenciados con las bases de datos públicas.

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
17	52	bot YM♂ (1) vs. bot YM♀ (0)	NS (0,003)	-	-	-	
18, 19*	93	bot YM y YS♂ (1)	AY727932.1; ARN ribosomal 26S, <i>Ilex repanda</i> (5/100; 2e-36; 99%; 100%).	BAB33421.1; Proteína putativa asociada a la senescencia, <i>Pisum sativum</i> (2/2; 6e-07; 87%; 87%; 96%).	KT179751.1; ARN ribosomal 26S, <i>Antennaria neglecta</i> (1/100; 4e-11; 100%; 100%; 97%).	-	Traducción/ Senescencia
20, 21, 89*	125	bot YM y YS; HA YM♂ (1)	JN705944.1; Factor de transcripción con dominio MADS (AGAMOUS, AG), ARNm, <i>Davidia involucrata</i> (1/100; 2e-39; 90%; 100%).	AEQ75504.1; Factor de transcripción con dominio MADS, <i>Davidia involucrata</i> (1/100; 5e-17; 85%; 95%; 98%).	JN705944.1; Factor de transcripción con dominio MADS (AGAMOUS, AG), ARNm, <i>Davidia involucrata</i> (1/100; 8e-16; 85%; 95%; 100%).	-	Factor de transcripción del desarrollo floral
34	96	YM (1)	XM_011003189.1; Proteína de biogénesis del peroxisoma tipo 19-2, ARNm (predicho), <i>Populus euphratica</i> (2/66; 2e-18; 88%; 84%).	XP_008373765.1; Proteína de biogénesis del peroxisoma tipo 19-2 (predicho), <i>Malus domestica</i> (1/1; 8e-05; 88%; 96%; 81%).	XM_023109390.1; Proteína de biogénesis del peroxisoma tipo 19-2, ARNm (predicho), <i>Cucurbita maxima</i> (1/27; 8e-06; 89%; 100%; 84%).	-	Morfología de los peroxisomas; importación y estabilización post-traduccional de proteínas de la membrana peroxisomal
35	86	HA YS (0)	NS (1,1)	-	NS (3,8)	-	
36	71	HA YM (0)	XM_016039430.1; Proteína serina/treonina quinasa, tipo receptor LRR, ARNm (predicho), <i>Ziziphus jujuba</i> (1/1; 4e-05; 94%; 50%).	-	-	-	Transducción de señales; desarrollo tisular; detección de señales ambientales; respuesta a heridas y patógenos
37	80	bot YS♂ (0)	NS (0,96)	-	-	-	
38	105	HA YM y YS (0)	-	-	-	-	

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
39	62	bot YS♂ (0)	XM_002531792.2; Homólogo de la proteína E3 UFM1-ligasa 1, ARNm (predicho), <i>Ricinus communis</i> (2/2; 3e-05; 86%; 96%).	NS (0,52)	-	-	Modificación post-traducciona por etiquetado de proteínas (similar a la ubiquitinización)
40	191	HA YS (1)	NS (0,076)	NS (8,5)	NS (2,7)	-	
41	111	bot YM♂ (0) vs. bot YM♀ (1)	NS (1,6)	-	NS (9,8)	-	
42	78	bot YM (1)	XM_016704792.1; Componente mitocondrial reductor de amidoxima tipo 2, ARNm (predicho), <i>Capsicum annuum</i> (1/100; 3e-13; 85%; 94%).	XP_010499646.1; Componente mitocondrial reductor de amidoxima tipo 2 (predicho), <i>Camelina sativa</i> (1/46; 1e-05; 83%; 95%; 92%).	CP023758.1; Cromosoma 2 (ADN genómico), <i>Solanum lycopersicum</i> (1/4; 3e-05; 84%; 96%; 97%).	PF03476; Barril beta N-terminal del dominio MOSC de unión al cofactor molibdeno (1,04e-05)	Reducción de profármacos de amidoxima, de una amplia gama de compuestos N-hidroxilados y de nitrito a óxido nítrico (NO)
43	58	bot y HJ YM (1)	XM_020358987.1; Glucomanano 4 – beta-manosiltransferasa tipo 2, ARNm (predicho), <i>Daucus carota</i> (1/86; 7e-12; 90%; 100%).	NS (7e-04)	NS (0,007)	-	Organización de la pared celular
44	144	bot YS♂ (0) vs. bot YS♀ (1)	NS (0,004)	-	NS (7,1)	-	
45	70	bot y HJ de YM y YS (1)	KY828469.1; <i>rrn16</i> (plastídico), <i>Spondias mombin</i> (1/100; 8e-26; 100%; 98%).	KMT07819.1; Proteína hipotética, <i>Beta vulgaris</i> (1/6; 6e-07; 100%; 100%; 94%).	LC342076.1; ADN plastídico, <i>Vitis monticola</i> (1/100; 8e-06; 100%; 100%; 100%).	-	Traducción
46	76	bot YM♂ (0)	NS (0,072)	-	-	-	
47	32	bot YM (1)	NS (0,063)	-	-	-	
48	97	HA YS (1)	NS (1,3)	-	NS (2,8)	-	



Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
49	62	bot YM♂ (0) vs. bot YM♀ (1); HJ YM y YS (1)	XM_022169420.1; Fenilalanina amonio liasa, ARNm (predicho), <i>Helianthus annuus</i> (6/100; 2e-13; 94%; 87%).	NS (6e-04)	-	PLN02457; Fenilalanina amonio liasa (3e-06)	Biosíntesis de diversos productos a partir de L-fenilalanina, regulación de la biosíntesis de lignina y de pigmentos como antocianinas, y mecanismos de defensa contra factores bióticos y abióticos (MacDonald y D’Cunha 2007)
50	40	bot y HJ de YM (1)	NS (3e-04)	-	-	-	
51	81	HA YM (1)	NS (0,28)	-	NS (2,0)	-	
52	200	HA YM y YS (1)	KC753782.1; Flavonol sintasa, ARNm, <i>Ampelopsis grossedentata</i> (1/100; 2e-27; 80%; 71%).	XP_010056462.1; Flavonol sintasa, <i>Camellia sinensis</i> (1/100; 2e-17; 80%; 91%; 69%).	KF157392.1; Leucoantocianidin dioxigenasa, ARNm, <i>Actinidia chinensis</i> (1/10; 6e-10; 68%; 83%; 58%).	PLN02704; PcbC (3,27e-15)	Biosíntesis de flavonoides (metabolitos secundarios)
53	51	bot y HJ YS (1)	NS (0,031)	-	-	-	
54	145	YS (1)	NS (2,3)	-	NS (7,1)	-	
55**	-	HA YS (1)	-	-	-	-	
56	244	bot YM♂ (0) vs. bot YM♀ (1)	KP874095.1; Proteína de unión a clorofila a/b 1 (CAB1), ARNm, <i>Panax ginseng</i> (1/100; 4e-51; 85%; 78%).	AFO67217.1; Proteína de unión a clorofila a/b putativa, <i>Aralia elata</i> (2/100; 7e-30; 90%; 91%; 73%).	KP874095.1; Proteína de unión a clorofila a/b putativa, ARNm, <i>Panax ginseng</i> (1/77; 4e-28; 83%; 88%; 78%).	PLN00025; Dominio de la proteína de unión a clorofila a/b (2,47e-16)	Fotosíntesis
57	234	bot YM y YS + HA YS (1)	XM_022039061.1; Proteína Metalotioneína tipo 3, ARNm (predicho), <i>Carica papaya</i> (2/100; 4e-32; 80%; 70%).	EOY13005.1; Metalotioneína tipo 3, <i>Theobroma cacao</i> (4/11; 2e-05; 70%; 84%; 42%).	NS (0,001)	-	Homeostasis del cobre

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
58	197	HJ YS (1)	XM_019409088.1; 4-cumarato-CoA ligasa tipo 2, ARNm (predicho), <i>Nicotiana attenuata</i> (1/100; 3e-32; 77%; 98%).	AFP49810.1; 4-hidroxicinamoil CoA ligasa 3 (= 4-cumarato-CoA ligasa), <i>Coffea arabica</i> (2/100; 4e-26; 78%; 86%; 98%).	XM_018974791.1; 4-cumarato-CoA ligasa, ARNm (predicho), <i>Juglans regia</i> (1/100; 3e-23; 71%; 80%; 100%).	PLN02246; 4-cumarato-CoA ligasa (6,38e-25)	Metabolismo de compuestos derivados de fenilpropanoides; generación de precursores de compuestos secundarios como lignina, flavonoides, flavonoles, antocianinas, isoflavonoides, furanocumarinas (Yuan et al 2014); defensa contra patógenos
59**	-	bot YS♂ (1) vs. bot YS♀ (0)	-	-	-	-	
60	217	bot YS♂ (0) vs. bot YS♀ (1); HJ YS (1)	XM_021779540.1; Poligalacturonasa, ARNm (predicho), <i>Hevea brasiliensis</i> (1/100; 5e-49; 82%; 95%).	XP_002266600.1; Poligalacturonasa (predicho), <i>Vitis vinifera</i> (3/100; 2e-33; 80%; 91%; 96%).	XM_002266564.4; Poligalacturonasa, ARNm (predicho), <i>Vitis vinifera</i> (1/100; 3e-34; 80%; 91%; 98%).	PLN02793; Probable poligalacturonasa (2,48e-03)	Metabolismo de carbohidratos, organización de la pared celular
61	137	bot♂ YM y YS (1)	KX192073.1; beta-hexosaminidasa 1 (HEXO1), ARNm, <i>Nicotiana benthamiana</i> (1/23; 7e-13; 84%; 55%).	XP_015866524.1; Beta-hexosaminidasa 1 (predicho), <i>Ziziphus jujuba</i> (1/23; 3e-08; 60%; 77%; 98%).	XM_011078592.2; Beta-hexosaminidasa 1, ARNm (predicho), <i>Sesamum indicum</i> (1/32; 1e-06; 61%; 81%; 83%).	-	Metabolismo de carbohidratos
62, 63	142	HA YS (1)	NS (0,64)	-	NS (0,56)	-	
64	123	HJ YM♀ (1)	NS (0,53)	-	NS (5,2)	-	
65	276	YS (1)	NS (1,5)	NS (1,3)	NS (0,12)	-	
66	92	bot YM♂ (0) vs. bot YM♀ (1)	NS (0,35)	-	NS (2,0)	-	
67	319	bot YM♂ (0) vs. bot YM♀ (1); HJ YS♀ (1)	XM_002267251.3; Proteína tirosina fosfatasa (PTP1), ARNm (predicho), <i>Vitis vinifera</i> (1/100; 5e-83; 82%; 100%).	XP_008237328.1; Proteína tirosina fosfatasa PTP1 (predicho), <i>Prunus mume</i> (1/100; 1e-61; 81%; 94%; 99%).	XM_023039048.1; Proteína tirosina fosfatasa tipo PTP1, ARNm (predicho), <i>Olea europaea</i> (1/100; 2e-57; 85%; 92%; 100%).	cd00047; Proteína tirosina fosfatasa (4,36e-16)	Regulación negativa de la defensa, transducción de señales

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
68	240	YS (1)	XM_021789133.1; ARN helicasa ATP-dependiente, factor de procesamiento de pre-ARNm (DEAH2), ARNm (predicho), <i>Hevea brasiliensis</i> (1/100; 1e-45; 84%; 73%).	XP_022770772.1; probable ARN helicasa ATP-dependiente, factor de procesamiento de pre-ARNm (DEAH2), <i>Durio zibethinus</i> (4/100; 1e-15; 97%; 100%, 42%).	XM_012582491.1; probable ARN helicasa ATP-dependiente, factor de procesamiento de pre-ARNm (DEAH2), ARNm (predicho), <i>Gossypium raimondii</i> (1/100; 5e-15; 94%; 100%; 63%).	-	Procesamiento ( <i>splicing</i> ) de pre-ARNm
69	116	YS (1)	NS (0,50)	-	NS (2,8)	-	
70	86	YM + YS (1)	NS (1,1)	-	-	-	
71	75	HJ YM♀ (1)	NS (0,072)	-	-	-	
72	71	bot YM♂ (0) vs. bot YM♀ (1); HJ YM (1)	NS (2,8)	-	-	-	
73	244	YS (1)	XM_018969258.1; Fumarato hidratasa 1, mitocondrial, ARNm (predicho), <i>Juglans regia</i> (1/100; 3e-46; 84%; 75%).	XP_007200983.1; Fumarato hidratasa 1, mitocondrial, <i>Prunus persica</i> (3/100; 9e-27; 90%; 96%; 73%).	XM_006479950.2; Fumarato hidratasa 1, mitocondrial, ARNm (predicho), <i>Citrus sinensis</i> (1/100; 2e-27; 90%; 95%; 93%).	PLN00134; Fumarato hidratasa (7,80e-28)	Metabolismo del fumarato y del malato; ciclo de Krebs; respuesta al estrés oxidativo; respuesta al estrés salino
74, 75*	238	YM (1)	NS (0,008)	XP_002520774.1; Péptido-N4-(N-acetil-β-glucosaminil) asparagina amidasa tipo A (predicho), <i>Ricinus communis</i> (1/57; 9e-13; 53%; 68%; 95%).	XM_002520728.2; Péptido-N4-(N-acetil-β-glucosaminil) asparagina amidasa tipo A, ARNm (predicho), <i>Ricinus communis</i> (1/54; 2e-11; 53%; 68%; 95%).	-	Escisión de glicopéptidos
76	169	bot YM y YS (1)	XM_002283274.4; Peptidasa mitocondrial procesadora de proteínas, subunidad alfa, ARNm (predicho), <i>Vitis vinifera</i> (2/100; 3e-25; 83%; 72%).	XP_019191999.1; Peptidasa mitocondrial procesadora de proteínas, subunidad alfa (predicho), <i>Ipomoea nil</i> (2/100; 7e-12; 79%; 100%; 67%).	AM489156.2; Secuencia genómica, <i>Vitis vinifera</i> (1/100; 2e-14; 60%; 86%; 94%).	-	Eliminación de secuencias señal de los precursores de proteínas mitocondriales; cadena respiratoria mitocondrial

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
77	169	bot YS♂ (0) vs. bot YS♀ (1); HA YS (1)	XM_018773391.1; Probable aminotransferasa TAT2, ARNm (predicho), <i>Nicotiana tomentosiformis</i> (1/100; 5e-29; 81%; 88%).	XP_009795220.1; Probable aminotransferasa TAT2 (predicho), <i>Nicotiana sylvestris</i> (1/100; 8e-21; 84%; 94%; 88%).	XM_018773391.1; Probable aminotransferasa TAT2, ARNm (predicho), <i>Nicotiana tomentosiformis</i> (1/100; 1e-20; 84%; 94%; 96%).	PLN00145; Tirosina / nicotianamina aminotransferasa (1,73e-21)	Catabolismo de tirosina; biosíntesis de vitamina E
78	44	HA YS (1)	NS (3,1)	-	-	-	
79	184	bot YM (1)	NS (0,91)	-	NS (0,56)	-	
80	176	bot YM (1)	NS (0,86)	NS (7,9)	NS (2,7)	-	
81	124	HA YS (0) vs. bot y HJ YS (1); YM (1)	NS (0,013)	-	NS (1,1)	-	
82	125	YM bot y HJ (1) vs. HA YM (0)	NS (0,004)	-	NS (0,57)	-	
83	99	bot YS♂ (0) vs. bot YS♀ (1); HA YS y bot YM (1)	NS (4,8)	-	NS (7,2)	-	
84	214	bot y HA YM (1)	XM_010253966.1; Ubiquitina carboxi-terminal hidrolasa 27, ARNm (predicho), <i>Nelumbo nucifera</i> (1/50; 8e-15; 82%; 44%).	OVA09172.1; Ubiquitina carboxi-terminal hidrolasa, familia 2 <i>Macleaya cordata</i> (1/100; 1e-14; 60%; 72%; 77%).	XM_015727608.1; Ubiquitina carboxi-terminal hidrolasa 27, ARNm (predicho), <i>Ricinus communis</i> (1/100; 3e-12; 53%; 66%; 85%).	cd02662; Peptidasa C19F (3,23e-03)	Desubiquitinación de proteínas; proceso catabólico de proteínas dependiente de ubiquitina
85	196	bot YS♂ (0) vs. bot YS♀ (1)	-	NS (1,6)	NS (1,0)	-	

Sec.	Long. (pb) <sup>a</sup>	Grupo diferenciado	BlastN <sup>b</sup>	BlastX <sup>c</sup>	tBlastX <sup>c</sup>	Dominio <sup>d</sup>	Función
86, 87*	169	YM + YS (1)	XM_009600494.2; Proteína con dominio BTB/POZ, ARNm (predicho), <i>Nicotiana tomentosiformis</i> (2/91; 2e-28; 86%; 69%).	XP_022738339.1; Proteína con dominio BTB/POZ, <i>Durio zibethinus</i> (1/100; 1e-17; 75%; 86%; 90%).	XM_022882605.1; Proteína con dominio BTB/POZ, ARNm (predicho), <i>Durio zibethinus</i> (1/100; 6e-19; 77%; 89%; 92%).	-	Ritmo circadiano; desarrollo del saco embrionario; desarrollo de polen; regulación positiva de la actividad de la telomerasa; regulación de la proteólisis dependiente de ubiquitina; regulación de la transcripción; respuesta al ácido abscísico, a la auxina, a los carbohidratos, al frío, al peróxido de hidrógeno, al ácido jasmónico, al nitrato, al ácido salicílico, al estrés salino y a las heridas; vía de señalización mediada por azúcar
88	144	HA YS♀ (1)	NS (0,19)	-	NS (9,8)	-	

<sup>a</sup> longitud de la secuencia obtenida.

<sup>b</sup> comparación con bases de datos de nucleótidos: número y anotación del acceso recuperado. Entre paréntesis se indica: la posición del acierto (*hit*) entre los resultados significativos recuperados; valor-E obtenido; porcentaje de identidad; porcentaje de cobertura entre la secuencia incógnita y el acierto.

<sup>c</sup> comparación con bases de datos de proteínas (BlastX) y de secuencias nucleotídicas traducidas (tBlastX) usando la secuencia incógnita traducida: número y anotación del acceso recuperado. Entre paréntesis: posición del *hit*; valor-E; porcentaje de identidad; porcentaje de residuos compartidos positivos; porcentaje de cobertura de la secuencia incógnita por la secuencia recuperada.

<sup>d</sup> para los resultados significativos: acesión, nombre y valor-E para la coincidencia entre la incógnita y el acierto.

\* fragmentos distintos que mostraron secuencia idéntica; \*\* secuenciación fallida.

NS, resultado no significativo; entre paréntesis se indica el E-valor obtenido. Los casos en los que no se recuperaron coincidencias significativas con secuencia alguna, se indican con un guión.

bot: botones florales; HJ: hojas jóvenes; HA: hojas adultas

## Discusión

En este trabajo se estudiaron por primera vez marcas epigenéticas en *Ilex*, en particular en la forma de metilación del ADN en los contextos mCG y mCCG, y en distintos órganos. Además, se encaró el estudio comparativo de la expresión génica en los diferentes órganos. Con estas aproximaciones se buscó responder una serie de preguntas biológicas relacionadas con el control genético del desarrollo de órganos vegetativos y reproductivos, y de las vías que conducen a la determinación sexual.

Además de tener interés meramente científico, nuestras preguntas originales apuntaban a lograr un conocimiento más abarcativo de la genética molecular de la yerba mate y de la yerba señorita, que pueda ser aprovechado por los fitomejoradores.

Es importante tener en consideración que las plantas de yerba mate analizadas en este Capítulo se desarrollaron en las condiciones climáticas del Jardín Botánico de Buenos Aires, desde los 6-8 meses de edad (año 2010) hasta el momento del inicio de los muestreos (año 2015). Supuestamente las condiciones edáficas serían las adecuadas, al haber sido implantadas en una parcela rellenada con tierra colorada de Corrientes para mejorar su adaptación ([www.buenosaires.gob.ar/ambienteyspaciopublico/mantenimiento/espaciosverdes/jardinbotanico/colecciones/yerba-mate](http://www.buenosaires.gob.ar/ambienteyspaciopublico/mantenimiento/espaciosverdes/jardinbotanico/colecciones/yerba-mate)). En las diversas visitas realizadas al Jardín Botánico el aspecto externo de las plantas no dio indicios de afecciones notorias. Esto no excluye que estén o hayan estado en condiciones de estrés que pueda afectar los patrones epigenéticos y/o de expresión génica. Aun así, las plantas del Jardín Botánico estuvieron bajo las mismas condiciones (estresantes o no) a lo largo de los muestreos. Por tanto, se considera que, en principio, las condiciones ambientales no afectarían las características analizadas, pero sí podrían crear ciertas distorsiones respecto a muestras de plantas provenientes del ambiente original, dado que ciertas condiciones estresantes podrían inducir cambios en la expresión génica a través de la hipo- o hipermetilación del ADN (Chinnusamy y Zhu 2009). Una consideración aparte debe realizarse para las plantas de yerba señorita, dado que se trabajó con dos grupos de plantas. Unas estuvieron bajo condiciones de crecimiento aún más contrastantes por estar implantadas en macetas individuales y variada exposición a luz solar, características edáficas, etc. De éstas, las plantas femeninas tuvieron dificultad para florecer, y baja productividad reiterada. El segundo grupo de plantas provino de una misma locación de Corrientes, y fueron sometidas a las mismas condiciones ambientales.

Estas dificultades, y el bajo número de muestras con las que se contó, podrían haberse evitado de habernos permitido el acceso al material del Banco de Germoplasma de la EEA-INTA CA, entre 2013 y 2016.

Es conocido que las dos especies estudiadas aquí están muy bien diferenciadas genéticamente (Gottlieb *et al.* 2005, 2011; Gottlieb y Poggio 2014). Los resultados obtenidos en este trabajo concuerdan en ese sentido, e incorporan el conocimiento de la diferenciación a nivel epigenético y de la expresión génica.

En términos generales, los resultados obtenidos estarían indicando que las muestras de plantas de yerba mate albergan mayor variación y diversidad epigenética que las de la yerba señorita, mientras que la diversidad en la expresión génica sería similar entre las dos especies. Es probable que las plantas de yerba mate analizadas sean producto de generaciones de selección artificial, generando un aumento secundario de la frecuencia y variedad de epialelos (Richards 2011); mientras que la yerba señorita se encontraría menos manipulada.

Las dos especies evidenciaron niveles similares de frecuencia de alternativas de metilación (no metilado, mCG y mCCG hemimetilado), y predominancia de sitios hipermetilados o mutados. En angiospermas se ha registrado casos en que el establecimiento, el mantenimiento y las consecuencias de la metilación del ADN serían dependientes de la especie y del contexto nucleotídico particular en el que se encuentra la marca (Niederhuth *et al.* 2016). Los resultados presentados también sugieren frecuencias relativas similares de estas condiciones de metilación entre órganos, con mayor proporción de diferencias genéticas y/o de hipermetilaciones, seguido por marcas mCG, por las condiciones no metiladas y finalmente por las de hemimetilación mCCG. Si bien existen metilaciones en otros contextos, que no son detectadas por el método MSAP, la mayor frecuencia de marcas mCG frente a las mCCG en todas las comparaciones coincide con lo reportado en bibliografía para otras plantas (Henderson y Jacobsen 2007; Jablonka y Raz 2009; Angers *et al.* 2010; Schulz *et al.* 2013; Fulneček y Kovařík 2014; Niederhuth *et al.* 2016).

Un análisis más detallado tomando en cuenta el número de *subepiloci* generados, permite concluir que habría más *loci* conteniendo sitios no metilados, seguido de mCG, y finalmente mCCG. Con este mismo método de cuantificación Schulz *et al.* (2014), estudiando poblaciones de *Viola elatior*, informaron la misma relación que la obtenida aquí, aunque con diferencias más marcadas entre las condiciones. Por su parte, Cara

*et al.* (2013) reportaron en especies de *Solanum* (parentales e híbrido interespecífico) niveles similares de mCG y CCGG no metilado, y muy superiores a los de mCCG hemimetilados.

Además, en las dos especies estudiadas en esta Tesis la diversidad epigenética resultó similar o mayor a la genética, coincidiendo con lo observado en otras plantas (Herrera y Bazaga 2010; Herrera *et al.* 2013; Schulz *et al.* 2013; Avramidou *et al.* 2015b). Aquí se pudo observar que para todos los conjuntos de datos se obtuvo una mayor diversidad bajo el enfoque analítico MS, aunque esto podría ser producto del método de transformación de los datos. Por tal motivo, y por las diferencias en la resolución de los agrupamientos obtenidos, los análisis de diversidad se centraron en el enfoque MS2.

La yerba mate y la señorita mostraron niveles de diversidad similares de las tres condiciones de metilación. Además, los órganos en crecimiento activo (botones y hojas jóvenes) han evidenciado los mayores valores de variación y diversidad epigenética entre plantas, así como mayor número de *subepiloci* totales, mientras que los patrones epigenéticos de las flores abiertas ya diferenciadas resultan más similares entre plantas, aun de distinto sexo. También hemos podido detectar diferencias en los niveles de diversidad de las marcas epigenéticas entre estadios de desarrollo foliar, así, las hojas jóvenes de la yerba mate mostraron la mayor diversidad en sitios no metilados. Lo mismo se observó para los botones y flores de la yerba señorita. La alta diversidad de sitios no metilados podría indicar una mayor variedad entre plantas de *loci* en expresión activa, probablemente influida por el fondo genético. Por otro lado, en la yerba señorita se detectó una marcada diferencia en los niveles de diversidad de los sitios mCCG entre las hojas (mayor en jóvenes y menor en adultas), aunque este resultado podría estar sesgado por el inusual comportamiento de la muestra de hoja joven de D13♂.

Los resultados obtenidos en este trabajo evidencian una mayor variación genética respecto a la epigenética, reflejando una alta diferenciación genética entre los individuos. La bibliografía consultada informa lo contrario, tanto para poblaciones naturales (por ejemplo, *Solanum ruiz-lealii*, Marfil *et al.* 2009; *Viola cazorensis*, Herrera y Bazaga 2010) como para cultivares (por ejemplo, *Allium sativum*, Gimenez *et al.* 2015). Cabe recordar que la existencia de distintos criterios analíticos para los datos MSAP (Schulz *et al.* 2013; Fulneček y Kovařík 2014; Alonso *et al.* 2016) y el uso de esta técnica en diseños experimentales muy variados limita los trabajos estrictamente comparables.



Las plantas de yerba mate del Jardín Botánico mostraron alto nivel de variación y de diversidad ISSR, y se pudo descartar la presencia de clones. Esto resultó sorprendente, teniendo en cuenta la baja variabilidad reportada para poblaciones naturales de Brasil usando marcadores RAPD e isoenzimas (Gauer y Cavalli-Molina 2000; Neumann Wendt 2005), y la limitada variación microsatélite registrada en accesiones de diverso origen geográfico (Cascales *et al.* 2014). En contraposición, pocos trabajos informaron altos niveles de variación. Pereira *et al.* (2013) detectó altos niveles de variabilidad microsatélite en dos poblaciones de Brasil. Luego, los únicos antecedentes en el uso de marcadores ISSR en yerba mate; analizando más de 100 plantas de yerbales comerciales, estimaron un alto valor global de diversidad genética (Paiva *et al.* 2016, 2017), aunque 0,73 veces el obtenido aquí. En otras especies cultivadas, como accesiones comerciales de la planta medicinal *Jatropha curcas* (Grativol *et al.* 2011), cultivares de cereza (*Prunus avium*; Ganopoulos *et al.* 2011) y accesiones de *Curcuma longa* (Singh *et al.* 2012), se estimaron menores valores de diversidad ISSR (0,53-0,77 veces) que el estimado para las plantas del Jardín Botánico.

En este trabajo hemos encontrado correlación positiva y significativa entre la información de sitios susceptibles a la metilación (MSL) y de sitios no susceptibles (NML), para todos los conjuntos de datos. Según Richards (2006), esto podría explicarse por una influencia o asociación del fondo genético sobre la variación epigenética. Dicho de otro modo, los mecanismos que controlan la metilación estarían relacionados con la variación de secuencia alrededor de los motivos CCGG (Cervera *et al.* 2002). Esta correlación también fue hallada en poblaciones de *Solanum* (Cara *et al.* 2013) y de *Viola elatior* (Schulz *et al.* 2014). Recientemente, Niederhuth *et al.* (2016) reportaron que, a nivel intraespecífico, los sitios metilados son altamente estables entre los órganos y a través de las generaciones.

Los análisis epigenéticos revelaron una clara tendencia al agrupamiento de muestras de acuerdo a la planta donante, refutando nuestras hipótesis iniciales sobre la diferenciación entre órganos y/o sexos. En particular para la yerba mate, el individuo B2♀ evidenció perfiles de epi- y genotipificación que lo distinguieron del resto. Asimismo, en la yerba señorita, la información epigenética discriminó las muestras de acuerdo a su planta originante y en cierta medida con diferencias ambientales, teniendo en cuenta su procedencia y las consideraciones realizadas al comienzo de esta Discusión. Para el individuo D91♀ sí se distinguieron los órganos reproductivos respecto de los vegetativos tanto mediante la información epigenética como genética.

Teniendo en mente que en este trabajo se detectó una mayor frecuencia de condición IV en todas las comparaciones, y particularmente en yerba señorita, los botones femeninos (pistilados) se distinguieron de los masculinos (estaminados) por una mayor frecuencia de sitios hipermetilados (silenciados), o una alta diferenciación genética con respecto al otro sexo. En yerba señorita, los botones y flores femeninos evidenciaron, además de los altos niveles de hipermetilación, menores frecuencias de mCG y de sitios no metilados que las contrapartes masculinas. Por tanto, sería plausible que los mecanismos operantes resulten similares a los reportados por Rico *et al.* (2014) para *Quercus ilex*, basados en aumentos significativos en los niveles de hipermetilación y disminución significativa de mCG, en respuesta al estrés. Por otro lado, las muestras de plantas femeninas no se diferencian estadísticamente entre sí a nivel genético, pero sí respecto a la masculina.

Otros autores informaron diferencias significativas en el nivel de metilación de distintos órganos (Cervera *et al.* 2002). En *A. thaliana* el nivel de metilación global del ADN (sin distinción de contextos) es mayor en las hojas maduras que en las plántulas (Finnegan *et al.* 1998), similar a lo detectado en tomate mediante mediciones HPLC (Messeguer *et al.* 1991). Estas diferencias fueron relacionadas con la expresión génica: como en general los niveles elevados de metilación se asocian con la inhibición de la transcripción, es probable que los tejidos maduros solo necesiten un subconjunto de los genes requeridos por los tejidos en crecimiento activo, y en consecuencia estos órganos mostrarían mayores niveles de metilación (Messeguer *et al.* 1991). Xiong *et al.* (1999), detectaron mediante MSAP un mayor nivel de metilación en las hojas de las plántulas de arroz que en las hojas bandera, en contraposición a lo hallado en tomate por Messeguer *et al.* (1991). Sin embargo, los resultados no son estrictamente comparables, ya que las estimas de metilación derivadas del MSAP no pueden considerarse estimas de metilación global, y suelen ser considerablemente más altas que las obtenidas mediante HPLC (Alonso *et al.* 2016). Lu *et al.* (2008), también usando MSAP, compararon en maíz la inflorescencia masculina, hoja bracteal y hoja de la mazorca, hallando diferencias significativas entre éstos, y concluyendo que la complejidad de la metilación cambia durante el crecimiento y desarrollo de la planta.

Del análisis de los resultados obtenidos se desprende que, para ambas especies, la regulación génica que deriva en el crecimiento y diferenciación de los estadios analizados no comprometería patrones generales de metilación, como los estudiados aquí, soportando la estabilidad en la metilación propuesta por Niederhuth *et al.* (2016).

No se descarta que sí involucren otros procesos epigenéticos, como modificaciones de histonas, rutas de ARNi, etc., como proponen Pikaard y Scheid (2014). Alternativamente, cabría suponer que el cambio en el estado de metilación fuera relevante para unos pocos *loci* reguladores de una vía relativamente independiente de la metilación del ADN, en sintonía con la idea de que el efecto de la metilación en la eucromatina podría ser *locus* específica (X Zhang *et al.* 2008).

La información contenida en los patrones de expresión génica diferencial posibilitó evidenciar, en ambas especies, agrupamientos de acuerdo a los órganos analizados, en contraposición a los resultados epigenéticos. Al tratarse de una aproximación más directa al panorama de activación de genes que el análisis de la metilación, es más informativo en cuanto a lo que ocurre a nivel celular en estos órganos, y estaría reflejando más fielmente la actividad génica en procesos de crecimiento, desarrollo y proliferación celular.

En el estudio de la diferenciación de los sexos de plantas dioicas existen reportes del rol de las marcas epigenéticas (por ej., Janoušek *et al.* 1996; Zhang *et al.* 2008; Song *et al.* 2012a; Bräutigam *et al.* 2017), así como también de *loci* expresados diferencialmente entre órganos de sexo opuesto (Zluvova *et al.* 2006; Murase *et al.* 2017). Bajo las condiciones del presente estudio, no fue posible distinguir si los botones o primordios florales provenían de plantas femeninas o masculinas. Esto podría sugerir que las diferencias en regulación y expresión génica entre los sexos ocurrirían en estadios del desarrollo aún más tempranos que los estudiados aquí. Los resultados histológicos presentados por Giberti (2001) sugieren que el arresto de los órganos del sexo opuesto ocurriría muy temprano en el desarrollo del botón floral. En estadios similares a los aquí analizados, Giberti (2001) observó estaminodios presentando anteras con sacos polínicos vacíos y estigmas con óvulos visibles en los botones de plantas femeninas, y sacos polínicos con granos de polen distinguibles en los botones de plantas masculinas. Cabe recordar que, en el género *Ilex*, las flores tienen los órganos del sexo opuesto abortados y estériles, y que se desconoce qué mecanismos desencadenan y/o mantienen este estado. Alternativamente, la diferenciación entre botones pistilados y estaminados estaría dada por unos pocos genes, y por tanto no detectados aquí, como ocurre en las especies dioicas *Silene latifolia* (Zluvova *et al.* 2006), *Diospyros lotus* (Akagi *et al.* 2014) y *Asparagus officinalis* (Murase *et al.* 2017). En *D. lotus* se encontró un ARN pequeño codificado en el cromosoma Y, sin

homólogos en las plantas femeninas, que interfiere los ARNm de un gen feminizante, silenciándolo (Akagi *et al.* 2014). Si bien es probable que la determinación sexual en los distintos grupos de especies dioicas involucre diferentes genes y mecanismos (Käfer *et al.* 2017), no se puede descartar que en las especies de *Ilex* los genes de ambos sexos se expresen (originando los futuros pistilodios y estaminodios), pero que exista un mecanismo de silenciamiento post-transcripcional o post-traduccional que bloquee la función de los productos génicos que llevan al desarrollo de uno de los géneros, como en *D. lotus*.

En la Introducción de este Capítulo se mencionó que durante la gametogénesis de *A. thaliana* (tanto femenina como masculina) ocurre una pérdida de metilación en las células involucradas (Law y Jacobsen 2010). En la yerba mate y la yerba señorita, la tendencia observada a la disminución de mCCG a medida que avanza el desarrollo de los botones florales podría ser análoga a las pérdidas de metilación observadas en la gametogénesis de *A. thaliana*. De todas formas, no debería perderse de vista que, como fue advertido por Giberti y Gurni (2008), la distinción en estadios del desarrollo para los botones florales en el género *Ilex* es muy dificultosa.

A enero de 2018, las bases de datos del NCBI contienen un total de 471 secuencias de yerba mate y 266 de yerba señorita (**Tabla 27**). De las proteínas, 86 son plastídicas y derivan de esta Tesis. Por su parte, Acevedo *et al.* (2013, 2016) también contribuyeron con secuencias de ARN y proteínas de la yerba mate. Entre los valores mencionados no se computa la primera versión del genoma de la yerba mate (secuenciado en el marco del proyecto PRO.MATE.AR 2015) ya que su anotación aún está en proceso. Tampoco fue posible acceder al transcriptoma de esta especie (Debat *et al.* 2014), depositado en la base de datos *Short Read Archive*, del NCBI. Por lo tanto, los fragmentos identificados en nuestro trabajo resultaron de la comparación con secuencias de otras especies de angiospermas.

**Tabla 27.** Secuencias de yerba mate y yerba señorita alojadas en bases de datos públicas de NCBI.

<b>Base de datos</b>	<b><i>I. paraguariensis</i></b>	<b><i>I. dumosa</i></b>
<b>Nucleótidos</b>	170	14
<b>Proteínas</b>	120	89
<b>Genes</b>	134	132
<b>GSS*</b>	42	31
<b>GEO Datasets**</b>	5	0
<b>Total</b>	<b>471</b>	<b>266</b>

\* Base de datos de lecturas únicas no anotadas; \*\* Base de datos de expresión génica curada.

De los fragmentos secuenciados en este trabajo (83 secuencias distintas en total), se pudo identificar el 36 %, distinguiéndose –a grandes rasgos– 5 categorías generales. La primera y más representada (30%) contiene fragmentos relacionados con metabolismo y biosíntesis de compuestos como flavonoides y vitaminas. La segunda (27%) reúne genes participantes en la regulación de la expresión a nivel transcripcional, post-transcripcional, traduccional, o post-traduccional. La tercera categoría (23%) está conformada por proteínas que tienen algún papel en la resistencia a patógenos, en la defensa contra la herbivoría y en la respuesta a algún tipo de estrés. La siguiente incluye funciones en procesos celulares como fotosíntesis, respiración celular, actividad de peroxisomas y de cloroplastos (13%). Finalmente, se hallaron fragmentos con función estructural en la pared celular y citoesqueleto (7%). En ningún caso se encontraron secuencias de elementos móviles de retroelementos del tipo LTR, ampliamente distribuidos en plantas (Llorens *et al.* 2009), a diferencia de lo registrado para estas especies por Gottlieb y Poggio (2014). A continuación se discuten los resultados más promisorios.

La beta-glucosidasa es una exocelulasa que cataliza la hidrólisis de enlaces glicosídicos liberando glucosa (Morant *et al.* 2008). Participa en muchos procesos en las plantas, como en la defensa química, el metabolismo de los alcaloides, la hidrólisis de los oligosacáridos, la regulación de las fitohormonas y la lignificación (Seshadri *et al.* 2009). Resulta de interés la detección en yerba señorita de un fragmento del gen de beta-glucosidasa en estado no metilado (es decir, activo) en las muestras provenientes de la planta masculina, que en las plantas femeninas estaría hipermetilado o con sitio de corte mutado. Esto podría interpretarse, tentativamente, como evidencia de un metabolismo de carbohidratos más activo en las muestras del donante masculino. La

sacarosa se produce en los órganos fotosintéticos (hojas), desde donde es transportada al polen vía floema (Paupière *et al.* 2014), avalando el hallazgo en hojas.

También relacionado al metabolismo de carbohidratos, aquí fue detectado un fragmento del transcripto de la proteína beta-hexominidasa (HEXO1) únicamente en los botones masculinos de ambas especies. Presente en el lumen vacuolar de raíces de *A. thaliana*, esta proteína participa en la conversión de N-glicanos en N-glicanos paucimannosídicos (Liebminger *et al.* 2011). Se ha constatado que la correcta N-glicosilación es indispensable para la función de dos receptores quinasas involucrados en respuestas a infecciones bacterianas en *A. thaliana* (Nekrasov *et al.* 2009; Häweker *et al.* 2010). Por otro lado, en combinación con la detección mencionada de una beta-glucosidasa probablemente activa en botones, flores y hojas de plantas masculinas de yerba señorita, refleja la necesidad energética del desarrollo y maduración del polen.

En botones y en hojas jóvenes de yerba mate, se detectó un fragmento con similitud significativa a una proteína receptora *leucine rich repeat* (LRR), perteneciente a la mayor subfamilia de receptores kinasas transmembrana en plantas, con más de 200 miembros en *Arabidopsis* (Torii 2004). Estos receptores reconocen una amplia gama de ligandos, incluidas pequeñas hormonas y péptidos, y están involucrados en una variedad de procesos de señalización durante el desarrollo y la respuesta inmune, regulando la respuesta a patógenos fúngicos, la autoincompatibilidad, y el crecimiento, proliferación y diferenciación celular, entre otros (Afzal *et al.* 2008; De Smet *et al.* 2009; Nekrasov *et al.* 2009; Häweker *et al.* 2010; Jaillais *et al.* 2011; Racolta *et al.* 2014; Zambounis *et al.* 2016). Queda pendiente identificar fehacientemente la secuencia completa del transcripto, y verificar su expresión diferencial.

La degradación del ARNm mediada por mutación terminadora (NMD, *Nonsense Mediated Decay*) es un mecanismo molecular conservado que destruye ARNm aberrantes (es decir, con codón de terminación de la traducción prematuro, denominados “sin sentido”), y que en diversos organismos regula la expresión normal de numerosos genes (Rayson *et al.* 2012). En las muestras masculinas de la yerba señorita se halló que el gen de un regulador de transcritos sin sentido estaba hipermetilado, y por tanto probablemente silenciado. En cambio, tendría distintos niveles de expresión en las femeninas, en base al estado de metilación. Estas diferencias podrían conducir a cascadas de regulación génica diferencial entre los sexos.

En hojas jóvenes de plantas femeninas de esa misma especie, se detectó un fragmento no metilado similar a una kinesina con un dominio de homología a calponina

(KCH), muy conocido por ser un motivo de unión a actina, y que ha sido asociado tentativamente con el rol de la interacción microtúbulo-microfilamento (Frey *et al.* 2010; Friedberg 2011). En arroz se ha reportado expresión diferencial de kinesinas que presentan este motivo en tejidos jóvenes y en desarrollo (Frey *et al.* 2010). Si bien resta identificar con precisión a qué proteína pertenece el fragmento encontrado, su estado no metilado en hojas jóvenes podría sugerir un rol en la división y el desarrollo celular.

En los botones de la yerba mate se detectó el gen de la translocasa de cloroplastos 90 en estado hipermetilado, mientras que en las hojas dicho fragmento presenta mCG. La proteína GTPasa codificada por este gen forma parte de un complejo translocador que facilita la importación a los cloroplastos de proteínas precursoras codificadas en el genoma nuclear (Hiltbrunner *et al.* 2004). Probablemente este hallazgo se deba a la actividad fotosintética fundamental de las hojas, y en este caso la metilación mCG podría indicar su expresión constitutiva (Bewick *et al.* 2016; Bewick y Schmitz 2017).

Las hojas adultas y las flores masculinas de yerba mate tienen en estado no metilado un gen codificante de una proteína que confiere resistencia a patógenos causantes del oídio (*Erysiphe*); el resto de las muestras, tienen el fragmento en estado mCG. En *A. thaliana* se hallaron diferentes genes de resistencia a esta enfermedad que actúan de manera aditiva (Sáenz-Mata y Jiménez-Bremont 2012) y fueron usados exitosamente para la generación de plantas transgénicas de *N. tabacum* resistentes (Xiao *et al.* 2003). Lo observado en yerba mate podría denotar la expresión basal del gen.

Otro fragmento mostró similitud significativa con el gen de la fosfatidilinositol 4-kinasa gamma 8 (PI4KG8). En *A. thaliana* esta proteína cataliza la transferencia de fosfato a un precursor del mensajero secundario inositol trisfosfato (InsP3), el cual participa, entre otras cosas, en la proliferación celular y en la respuesta a las condiciones de estrés, regulando la liberación de  $\text{Ca}^{2+}$  (Alimohammadi *et al.* 2015). Estas enzimas estarían implicadas en la señalización desencadenada por el ácido salicílico, molécula central en la defensa contra patógenos, y que participa en procesos como la germinación, la floración, la senescencia y la adquisición de termotolerancia (Krinke *et al.* 2007). En yerba mate estaría no metilada en los botones masculinos de estadio temprano y medio frente a los de estadio tardío (mCG). Este cambio en el estado de metilación podría indicar modificaciones en los niveles de su expresión durante el desarrollo floral masculino.

Las mitocondrias de las plantas, además de su rol primordial en la respiración celular y generación de energía, están involucradas en la adaptación celular a estrés abióticos

(Choudhury *et al.* 2017). Estudiando plantas de yerba mate sometidas a estrés hídrico, Acevedo *et al.* (2013) hallaron la sobreexpresión de una flavoproteína relacionada con el Ciclo de Krebs y la cadena respiratoria, como respuesta al aumento del nivel de ácido abscísico causado por el estrés aplicado. Aquí se halló expresión de una peptidasa mitocondrial en los botones florales de ambas especies. De acuerdo con nuestros resultados, cabría suponer una menor expresión de esta peptidasa en hojas jóvenes y adultas. Esta peptidasa cataliza en el citoplasma la desglicosilación de las glicoproteínas desnaturalizadas que presentan enlaces N y ayuda a su degradación mediada por proteosoma; en *A. thaliana* se expresa constitutivamente a niveles bajos a lo largo de todas las etapas de desarrollo de la planta (Diepold *et al.* 2007). La metalotioneína tipo 3, proteína implicada en la homeostasis del cobre aquí hallada en los botones de ambas especies y en las hojas adultas de yerba señorita, también fue detectada en la respuesta al estrés hídrico de la yerba mate, como participante en la defensa contra el estrés oxidativo (Acevedo *et al.* 2016).

La síntesis de aminoácidos aromáticos y de metabolitos secundarios (incluyendo auxinas, salicilatos, alcaloides antimicrobianos, flavonoides y compuestos polifenólicos) forma parte de un sistema de defensa general ante estrés biótico y abiótico (Pourcel *et al.* 2007). Los flavonoides, valorados por los humanos por sus propiedades nutraceuticas (Filip *et al.* 2001; Heck y de Mejía 2007; Bracesco *et al.* 2011), en los tejidos vegetales participan en pigmentación, en la defensa contra los patógenos y en la protección contra la radiación UV, a la vez que actúan como moléculas señal en las interacciones planta-microbio (Pelletier *et al.* 1997). Pueden ser importantes reguladores del crecimiento y desarrollo de las plantas en respuesta a la señalización de auxinas y del etileno, y modular el transporte de auxinas (Lewis *et al.* 2011). En yerba mate se asociaron experimentalmente con un papel protector durante eventos de estrés hídrico (Acevedo *et al.* 2016). En este trabajo se aislaron cuatro fragmentos relacionados con la biosíntesis o el metabolismo de aminoácidos y metabolitos secundarios, dos de ellos se asocian a los flavonoides. Uno, detectado en los botones de yerba mate, mostró similitud con un componente mitocondrial reductor de amidoxima (mARC, *mitochondrial Amidoxime Reducing Component*). Una búsqueda mediante algoritmo BlastX en Uniprot mostró similitud con la enzima sulfurasa del cofactor molibdeno (mejor acierto con *Macleaya cordata*; valor-E =  $1,4 \times 10^{-07}$ ), una proteína citosólica que regula específicamente la actividad de las enzimas aldehído oxidasa y xantina deshidrogenasa. Con el cofactor en estado sulfatado, la enzima



xantina deshidrogenasa metaboliza la xantina (Heidenreich *et al.* 2005). Como fue expuesto en la Introducción general de esta Tesis, las formas metiladas de las xantinas (metilxantinas –teofilina, cafeína y teobromina) tienen interés farmacológico (Fredholm 2011).

Otro transcripto identificado, de expresión activa en las hojas adultas de yerba mate y yerba señorita, es codificante para la enzima flavonol sintasa, que cataliza la biosíntesis de flavonoles (co-pigmentos) a partir de dihidroflavonoles (Prescott *et al.* 2002). En *A. thaliana*, la auxina y el etileno regulan la biosíntesis de flavonol (Lewis *et al.* 2011), y se encontraron evidencias de la existencia de parálogos de flavonol sintasa diferencialmente expresados en distintos órganos, con roles regulatorios en la biosíntesis de flavonoides (Ayabe *et al.* 2010). Resulta llamativo no haber detectado este fragmento en las hojas jóvenes, ya que también son ricas en este tipo de compuestos secundarios. Un tercer fragmento asignado a la categoría de biosíntesis de metabolitos secundarios fue aislado de las hojas jóvenes de yerba señorita, y resultó similar a la enzima 4-cumarato-CoA ligasa, que produce tioésteres de CoA que se utilizan para sintetizar varios compuestos que incluyen antocianinas, flavonoides, isoflavonoides, cumarinas, lignina, suberina y otros fenólicos (Ehlting *et al.* 1999). Un cuarto fragmento evidenció expresión diferencial de fenilalanina amonio liasa (PAL). Esta enzima participa en la biosíntesis de diversos productos a partir de L-fenilalanina, que a su vez derivan en la biosíntesis de lignina y antocianinas, y está involucrada en mecanismos de defensa (MacDonald y D’Cunha 2007). Song *et al.* (2012a), estudiando la metilación y expresión diferencial en la especie dioica *Populus tomentosa*, hallaron que las flores femeninas tenían la PAL metilada y con expresión reprimida, mientras que las masculinas mostraban lo contrario, sugiriendo su participación en la morfogénesis de los órganos florales, en la fertilidad del polen y en la modificación de proteínas clave del desarrollo floral (Song *et al.* 2012a). Recientemente, K Zhang *et al.* (2016) reportaron en arroz que la sobreexpresión de la enzima PAL redundó en un desarrollo anormal de anteras y polen, y consecuentemente semi-esterilidad. En nuestro estudio se halló expresión en botones femeninos de ambas especies, y en los masculinos de yerba señorita, por lo que estos resultados contradicen los de Song *et al.* (2012a) pero no se descarta que esta enzima tenga alguna función importante en la fertilidad de uno o ambos sexos.

Asociado con el desarrollo floral, se ha detectado un fragmento similar al transcripto de la proteína homeótica floral AGAMOUS, un factor de transcripción –con dominio

MADS de interacción con el ADN– que controla la identidad de órganos reproductivos (estambres y carpelos) durante los primeros pasos del desarrollo floral (Ito *et al.* 2004; Ng *et al.* 2009; Ahn *et al.* 2015). Este transcripto fue hallado en los botones de ambos sexos de las dos especies estudiadas y en hojas adultas de yerba mate. van der Linden *et al.* (2002) detectaron la expresión de un gen muy similar al AGAMOUS en tejidos vegetativos de *Malus x domestica*, aunque a menores niveles que en estambres y carpelos, sugiriendo que la proteína podría tener más de una actividad. En *A. thaliana* se describió que ciertas proteínas tipo AGAMOUS son necesarias para bloquear la iniciación de programas reproductivos en los órganos vegetativos (Fernandez *et al.* 2014), lo que podría explicar su hallazgo en hojas de yerba mate.

En los botones y en las hojas jóvenes de las plantas femeninas se halló expresión de la proteína tirosina fosfatasa (PTP1), enzima que, en condiciones de estrés no oxidativo, desfosforila y probablemente inhibe la MAP kinasa 6; esta última modula respuestas a estrés participando en transducción de señales (Bartels *et al.* 2009). Además, PTP1 participaría en la regulación negativa de la defensa, reprimiendo el ácido salicílico y la biosíntesis de camalexina, metabolito secundario que combate patógenos bacterianos y fúngicos (Nafisi *et al.* 2007). Las características de este transcripto lo hacen atractivo para su futuro estudio exhaustivo.

## CONCLUSIONES

La detección de variantes génicas de interés y la comprensión de su regulación constituyen el punto de partida para la aplicación de planes de mejoramiento biotecnológico exitosos. Haciendo uso de la información plastómica aquí generada, proponemos que la tecnología de transformación de cloroplastos debe ser considerada como una herramienta de mejora posible en yerba mate.

De acuerdo con nuestras hipótesis iniciales, en esta Tesis se concluye que:

- ☛ Los plastomas de la yerba mate y la yerba señorita están altamente conservados, compartiendo 99 % de identidad nucleotídica y evidenciando sintenia, incluso con especies filogenéticamente distantes (*P. ginseng*).
- ☛ Las secuencias plastídicas no codificantes albergan los mayores valores de variación nucleotídica.
- ☛ Se seleccionaron 9 IGS y 7 intrones, y se generaron 30 marcadores microsatélite, aportando así marcadores con utilidad en estudios poblaciones, filo- y biogeográficos, y filogenéticos en *Ilex*.
- ☛ La yerba mate y la yerba señorita pueden distinguirse tanto a nivel epigenético como por sus perfiles de expresión, en concordancia con los reportes de diferencias notables en su fenotipo, fisiología y características genéticas.
- ☛ El análisis de la expresión génica evidenció patrones característicos de los órganos estudiados.
- ☛ Entre los fragmentos diferenciales aislados, se cuentan genes o transcritos que codifican para proteínas relacionadas con metabolismo y biosíntesis de compuestos como flavonoides y vitaminas (30 %), regulación de la expresión a distintos niveles (27 %), resistencia a patógenos, en la defensa contra la herbivoría y en la respuesta a algún tipo de estrés (23 %), procesos celulares como fotosíntesis, respiración celular, actividad de peroxisomas y de cloroplastos (13 %), y con función estructural en la pared celular y citoesqueleto (7 %).

## *Conclusiones*

Por otro lado, las siguientes conclusiones se oponen a las hipótesis iniciales:

- La epigenotipificación no evidenció patrones particulares de los órganos estudiados.
- No se hallaron patrones obvios que distingan entre los sexos, bajo ninguna aproximación metodológica.

## **PERSPECTIVAS**

- ✓ Futuras investigaciones sobre el rol de la epigenética en estas especies podrían abordarse con metodologías más exhaustivas, como la secuenciación del epigenoma completo.
- ✓ Tanto la metilación como la expresión diferencial de los genes identificados deben ser confirmadas mediante otras técnicas. Por ejemplo, via *Methylation-Specific PCR* (MS-PCR; Herman *et al.* 1996) acoplado al análisis de la expresión diferencial mediante PCR cuantitativa en tiempo real con ARNm molde (RT-qPCR). Esto permitirá verificar el estado de metilación de los genes detectados en las distintas muestras y asociarlo a los niveles de expresión.
- ✓ Las secuencias de los genes detectados serán utilizadas como sondas para hibridaciones *in situ* histológicas.
- ✓ La caracterización de los fragmentos diferenciales pendientes brindará la oportunidad de hallar nuevas secuencias de interés.
- ✓ El estudio del desarrollo diferencial de los órganos sexuales podría ser abordado, por ejemplo, mediante la caracterización de ARN pequeños de interferencia.
- ✓ Como parte del plan Postdoctoral, se llevará a cabo una caracterización más exhaustiva mediante RT-qPCR de la expresión de regiones génicas relacionadas con la respuesta a patógenos (fúngicos, virales, etc.) y con la producción de metabolitos secundarios, y de aquellas sospechadas de estar involucradas en la determinación del sexo.

## REFERENCIAS BIBLIOGRÁFICAS

- Acevedo RM, Maiale SJ, Pessino SC, Bottini R, Ruiz OA, Sansberro PA. 2013. A succinate dehydrogenase flavoprotein subunit-like transcript is upregulated in *Ilex paraguariensis* leaves in response to water deficit and abscisic acid. *Plant Physiology and Biochemistry* **65**: 48-54.
- Acevedo RM, Ruiz OA, Sansberro PA. 2016. Gene expression changes in response to drought stress in *Ilex paraguariensis* leaves. *Plant OMICS* **9**: 334-343.
- Adam H, Collin M, Richaud F, et al. 2011. Environmental regulation of sex determination in oil palm: Current knowledge and insights from other species. *Annals of Botany* **108**: 1529-1537.
- Afzal AJ, Wood AJ, Lightfoot DA. 2008. Plant receptor-like serine threonine kinases: Roles in signaling and plant defense. *Molecular Plant-Microbe Interactions* **21**: 507-517.
- Ahn MS, Kim YS, Han JY, Yoon ES, Choi YE. 2015. *Panax ginseng* PgMADS1, an AP1/FUL-like MADS-box gene, is activated by hormones and is involved in inflorescence growth. *Plant Cell, Tissue and Organ Culture* **122**: 161-173.
- Akagi T, Henry IM, Tao R, Comai L. 2014. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* **346**: 646-650.
- Alikaridis F. 1987. Natural constituents of *Ilex* species. *Journal of Ethnopharmacology* **20**: 121-144.
- Alimohammadi M, Lahiani MH, Khodakovskaya M V. 2015. Genetic reduction of inositol triphosphate (InsP3) increases tolerance of tomato plants to oxidative stress. *Planta* **242**: 123-135.
- Alonso C, Pérez R, Bazaga P, Medrano M, Herrera CM. 2014. Individual variation in size and fecundity is correlated with differences in global DNA cytosine methylation in the perennial herb *Helleborus foetidus* (Ranunculaceae). *American Journal of Botany* **101**: 1309-1313.
- Alonso C, Pérez R, Bazaga P, Medrano M, Herrera CM. 2016. MSAP markers and global cytosine methylation in plants: A literature survey and comparative analysis for a wild-growing species. *Molecular Ecology Resources* **16**: 80-90.
- Anesini C, Turner S, Cogoi L, Filip R. 2012. Study of the participation of caffeine and polyphenols on the overall antioxidant activity of mate (*Ilex paraguariensis*). *LWT - Food Science and Technology* **45**: 299-304.
- Angers B, Castonguay E, Massicotte R. 2010. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. *Molecular Ecology* **19**: 1283-1295.
- Avramidou E V, Doulis AG, Aravanopoulos FA. 2015a. Determination of epigenetic inheritance, genetic inheritance, and estimation of genome DNA methylation in a full-sib family of *Cupressus sempervirens* L. *Gene* **562**: 180-187.
- Avramidou E V, Ganopoulos I V, Doulis AG, Tsaftaris AS, Aravanopoulos FA. 2015b. Beyond population genetics: natural epigenetic variation in wild cherry (*Prunus avium*). *Tree Genetics & Genomes* **55**: 904-906.
- Ayabe S, Uchiyama H, Aoki T, Akashi T. 2010. Plant phenolics: Phenylpropanoids. En: Mander L, Liu H-W, eds. *Comprehensive Natural Products II*. Elsevier, 929-976.
- Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *Plant Journal* **9**: 745-753.
- Bachem CWB, Oomen RJFJ, Visser RGF. 1998. Transcript imaging with cDNA-AFLP: A step-by-step protocol. *Plant Molecular Biology Reporter* **16**: 157-173.
- Baránek M, Křížan B, Ondrušíková E, Pidra M. 2010. DNA-methylation changes in grapevine somaclones following in vitro culture and thermotherapy. *Plant Cell, Tissue and Organ Culture* **101**: 11-22.
- Barral G, Poggio L, Giberti GC. 1995. Chromosome numbers and DNA content from *Ilex argentina* (Aquifoliaceae). *Boletín de la Sociedad Argentina de Botánica* **30**: 243-248.
- Bartels S, Anderson JC, Gonzalez Besteiro MA, et al. 2009. MAP KINASE PHOSPHATASE1 and PROTEIN TYROSINE PHOSPHATASE1 are repressors of salicylic acid synthesis and SNC1-mediated responses in *Arabidopsis*. *The Plant Cell* **21**: 2884-2897.
- Bernasconi NK, Mroginski LA, Sansberro PA, Rey HY. 1996. Micropropagación de la yerba mate (*Ilex paraguariensis* St. Hil.): Efecto del tipo de explante en el establecimiento de los cultivos *in vitro*. *Phyton* **58**: 23-31.
- Bewick AJ, Ji L, Niederhuth CE, et al. 2016. On the origin and evolutionary consequences of gene body DNA methylation. *Proceedings of the National Academy of Sciences* **113**: 9111-9116.
- Bewick AJ, Schmitz RJ. 2017. Gene body DNA methylation in plants. *Current Opinion in Plant Biology* **36**: 103-110.

## Referencias bibliográficas

- Bird A. 2002.** DNA methylation patterns and epigenetic memory DNA methylation patterns and epigenetic memory. *Genes & Development*: 6-21.
- Bobadilla Landey R, Cenci A, Georget F, et al. 2013.** High genetic and epigenetic stability in *Coffea arabica* plants derived from embryogenic suspensions and secondary embryogenesis as revealed by AFLP, MSAP and the phenotypic variation rate. *PLoS ONE* **8**: e56372.
- Bock R. 2007.** Structure, function, and inheritance of plastid genomes. En: Bock R, ed. *Topics in Current Genetics*. Göteborg, Sweden: Springer-Verlag Berlin Heidelberg, 29-63.
- Borsch T, Quandt D. 2009.** Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* **282**: 169-199.
- Bouyer D, Kramdi A, Kassam M, et al. 2017.** DNA methylation dynamics during early plant life. *Genome Biology* **18**: 179.
- Bracesco N, Sanchez AG, Contreras V, Menini T, Gugliucci A. 2011.** Recent advances on *Ilex paraguariensis* research: Minireview. *Journal of Ethnopharmacology* **136**: 378-384.
- Bräutigam K, Soolanayakanahally R, Champigny M, et al. 2017.** Sexual epigenetics: Gender-specific methylation of a gene in the sex determining region of *Populus balsamifera*. *Scientific Reports* **7**: 45388.
- Bremer B, Bremer K, Heidari N, et al. 2002.** Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molecular Phylogenetics and Evolution* **24**: 274-301.
- Brudno M, Do CB, Cooper GM, et al. 2003.** LAGAN and Multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA. *Genome Research* **13**: 721-731.
- Buduba C. 2011.** Mutagénesis y antimutagénesis en extractos acuosos, clorofórmicos y acetónicos de *Ilex paraguariensis* var. *paraguariensis* e *Ilex dumosa* var. *dumosa*. Tesis de Maestría, Universidad Nacional de La Plata, Argentina.
- Burggren WW. 2016.** Epigenetic inheritance and its role in evolutionary biology: re-evaluation and new perspectives. *Biology* **5**: 24.
- Burggren WW, Crews D. 2014.** Epigenetics in comparative biology: why we should pay attention. *Integrative and Comparative Biology* **54**: 7-20.
- Burtnik OJ. 2003.** *Manual del pequeño yerbatero correntino*. Corrientes, Argentina: INTA AER Santo Tomé.
- Burtnik OJ. 2006.** *Yerba mate : Manual de producción*. Corrientes, Argentina: INTA – Agencia de Extensión Rural Santo Tomé.
- Cai Z. 2010.** Comparative analyses of land plant plastid genomes. PhD Thesis, University of Texas, EEUU.
- Cara N, Marfil CF, Masuelli RW. 2013.** Epigenetic patterns newly established after interspecific hybridization in natural populations of *Solanum*. *Ecology and Evolution* **3**: 3764-3779.
- Cardozo Junior EL, Morand C. 2016.** Interest of mate (*Ilex paraguariensis* A. St.-Hil.) as a new natural functional food to preserve human cardiovascular health - A review. *Journal of Functional Foods* **21**: 440-454.
- Cascales J, Bracco M, Garberoglio MJ, Poggio L, Gottlieb AM. 2017.** Integral phylogenomic approach over *Ilex* L. species from southern South America. *Life* **7**: 47.
- Cascales J, Bracco M, Poggio L, Gottlieb AM. 2014.** Genetic diversity of wild germplasm of «yerba mate» (*Ilex paraguariensis* St. Hil.) from Uruguay. *Genetica* **142**: 563-573.
- Cervera M-T, Ruiz-García L, Martínez-Zapater M. 2002.** Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics* **268**: 543-552.
- Chen Y-L, Hwang T-L, Yu H-P, et al. 2016.** *Ilex kaushue* and its bioactive component 3,5-dicaffeoylquinic acid protected mice from lipopolysaccharide-induced acute lung injury. *Scientific reports* **6**: 34243.
- Chinnusamy V, Zhu JK. 2009.** Epigenetic regulation of stress responses in plants. *Current Opinion in Plant Biology* **12**: 133-139.
- Choudhury FK, Rivero RM, Blumwald E, Mittler R. 2017.** Reactive oxygen species, abiotic stress and stress combination. *The Plant Journal* **90**: 856-867.
- Chumley TW, Palmer JD, Mower JP, et al. 2006.** The complete chloroplast genome sequence of *Pelargonium × hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Molecular Biology and Evolution* **23**: 2175-2190.
- Coulleri JP. 2010.** Gene flow and local adaptation: antagonistic forces shape populations of *Ilex dumosa* (Aquifoliaceae). *Boletín de la Sociedad Argentina de Botánica* **45**: 333-342.
- Cuénoud P, Del Pero Martinez MA, Loizeau PA, Spichiger R, Andrews S, Manen JF. 2000.** Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). *Annals of Botany* **85**: 111-122.

## Referencias bibliográficas

- Dangi R, Tamhankar S, Choudhary RK, Rao S. 2016.** Molecular phylogenetics and systematics of *Trigonella* L. (Fabaceae) based on nuclear ribosomal ITS and chloroplast *trnL* intron sequences. *Genetic Resources and Crop Evolution* **63**: 79-96.
- Daniell H, Kumar S, Dufourmantel N. 2005.** Breakthrough in chloroplast genetic engineering of agronomically important crops. *Trends in Biotechnology* **23**: 238-245.
- Daniell H, Lin C-S, Yu M, Chang W-J. 2016.** Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology* **17**: 134.
- Day T, Bonduriansky R. 2011.** A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *The American Naturalist* **178**: E18-E36.
- Day A, Madesis P. 2007.** DNA replication, recombination, and repair in plastids. En: Bock R, ed. *Topics in Current Genetics*. Göteborg, Sweden: Springer-Verlag Berlin Heidelberg, 65-119.
- Debat HJ, Grabile M, Aguilera PM, et al. 2014.** Exploring the genes of yerba mate (*Ilex paraguariensis* A. St.-Hil.) by NGS and de novo transcriptome assembly. *PLoS ONE* **9**: e109835.
- Diekmann K, Hodkinson TR, Fricke E, Barth S. 2008.** An optimized chloroplast DNA extraction protocol for grasses (Poaceae) proves suitable for whole plastid genome sequencing and SNP detection. *PLoS ONE* **3**: e2813.
- Diepold A, Li G, Lennarz WJ, Nürnberger T, Brunner F. 2007.** The *Arabidopsis* *AtPNG1* gene encodes a peptide: N-glycanase. *The Plant Journal* **52**: 94-104.
- Dong W, Liu J, Yu J, Wang L, Zhou S. 2012.** Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE* **7**: e35071.
- Dorken ME, Pannell JR. 2008.** Density- dependent regulation of the sex ratio in an annual plant. *The American Naturalist* **171**: 824-830.
- Drescher A, Ruf S, Calsa T, Carrer H, Bock R. 2000.** The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant Journal* **22**: 97-104.
- Du FK, Lang T, Lu S, Wang Y, Li J, Yin K. 2015.** An improved method for chloroplast genome sequencing in non-model forest tree species. *Tree Genetics & Genomes* **11**: 114.
- Edgar RC. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792-1797.
- Ehltling J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E. 1999.** Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *The Plant Journal* **19**: 9-20.
- Excoffier L, Smouse PE, Quattro JM. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.
- Felsenfeld G. 2014.** A brief history of epigenetics. *Cold Spring Harbor Perspectives in Biology* **6**: a018200.
- Fernandez DE, Wang C-T, Zheng Y, et al. 2014.** The MADS-domain factors AGAMOUS-LIKE15 and AGAMOUS-LIKE18, along with SHORT VEGETATIVE PHASE and AGAMOUS-LIKE24, are necessary to block floral gene expression during the vegetative phase. *Plant Physiology* **165**: 1591-1603.
- Ferreira AG, Kaspary R, Ferreira HB, Rosa LM. 1983.** Proporção de sexo e polinização em *Ilex paraguariensis* St. Hill. *Brasil Forestal* **53**: 29-33.
- Filip R. 2009.** Mate (*Ilex paraguariensis*). En: Imperato F, ed. *Recent advances in Phytochemistry Research*. Signpost, Kerala, 113.
- Filip R. 2011.** Yerba mate: de bebida tradicional a alimento funcional. Perspectivas para el siglo XXI. En: *5to Congreso Sudamericano de la Yerba Mate*. Posadas, Misiones, 26.
- Filip R, López P, Coussio J, Ferraro G. 1998.** Mate substitutes or adulterants: study of xanthine content. *Phytotherapy Research* **12**: 129-131.
- Filip R, López P, Giberti GC, Coussio J, Ferraro G. 2001.** Phenolic compounds in seven South American *Ilex* species. *Fitoterapia* **72**: 774-778.
- Filip R, Lotito SB, Ferraro G, Fraga CG. 2000.** Antioxidant activity of *Ilex paraguariensis* and related species. *Nutrition Research* **20**: 1437-1446.
- Finnegan EJ, Genger RK, Peacock WJ, Dennis ES. 1998.** Dna methylation in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 223-247.
- Flatscher R, Frajman B, Schönswetter P, Paun O. 2012.** Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation? *Genetics Research International*: 1-9.
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004.** VISTA: Computational tools for comparative genomics. *Nucleic Acids Research* **32**: W273-279.



## Referencias bibliográficas

- Fredholm BB. 2011.** *Methylxanthines - Handbook of Experimental Pharmacology 200* (F Hofmann, Ed.). Springer Heidelberg Dordrecht London New York.
- Frey JE, Frey B, Forcioli D. 2005.** Quantitative assessment of heteroplasmy levels in *Senecio vulgaris* chloroplast DNA. *Genetica* **123**: 255-261.
- Frey N, Klotz J, Nick P. 2010.** A kinesin with calponin-homology domain is involved in premitotic nuclear migration. *Journal of Experimental Botany* **61**: 3423-3437.
- Friedberg F. 2011.** Single and multiple CH (calponin homology) domain containing multidomain proteins in *Arabidopsis* and *Saccharomyces*: An inventory. *Molecular Biology Reports* **38**: 213-218.
- Fulneček J, Kovařík A. 2014.** How to interpret Methylation Sensitive Amplified Polymorphism (MSAP) profiles? *BMC Genetics* **15**: 2.
- Galle FC. 1997.** *Hollies: The genus Ilex*. Portland, Oregon, USA: Timber Press.
- Ganopoulos I V, Kazantzis K, Chatzicharisis I, Karayiannis I, Tsaftaris AS. 2011.** Genetic diversity, structure and fruit trait associations in Greek sweet cherry cultivars using microsatellite based (SSR/ISSR) and morpho-physiological markers. *Euphytica* **181**: 237-251.
- Gauer L, Cavalli-Molina S. 2000.** Genetic variation in natural populations of maté (*Ilex paraguariensis* A. St.-Hil., Aquifoliaceae) using RAPD markers. *Heredity* **84**: 647-656.
- Giberti GC. 1979.** Las especies argentinas del género *Ilex* L. (Aquifoliaceae). *Darwiniana* **22**: 217-240.
- Giberti GC. 1989.** Los parientes silvestres de la yerba mate y el problema de su adulteración. *Dominguezia* **7**: 1-22.
- Giberti GC. 1994a.** Aquifoliaceae. En: *Flora Fanerogámica de Argentina*. Córdoba, Argentina: Programa PROFLOA (CONICET).
- Giberti GC. 1994b.** Aquifoliaceae. En: Spichiger R, Ramella L, eds. *Flora del Paraguay*. Missouri Botanical Garden, Switzerland: Editions des Conservatoire et Jardin botaniques de la Ville de Geneva, 24.
- Giberti GC. 2001.** Diferentes aspectos del género *Ilex* (Aquifoliaceae). Corología, arquitectura floral, posición sistemática. Tesis de Doctorado. Universidad de Buenos Aires, Argentina.
- Giberti GC. 2008.** Aquifoliaceae. En: Zuloaga F, Morrone O, Belgrano M, eds. *Catálogo de las Plantas vasculares del Cono Sur (Argentina, Sur de Brasil, Chile, Paraguay y Uruguay)*, Vol. 2. *Dicotyledoneae: Acanthaceae–Fabaceae* (Abarema – Schizolobium). Monographs in Systematic Botany from the Missouri Botanical Garden, 1143-1146.
- Giberti GC. 2011a.** La «yerba mate» (*Ilex paraguariensis*, Aquifoliaceae) en tempranos escritos rioplatenses de bonpland y su real distribución geográfica en sudamerica austral. *Bonplandia* **20**: 203-212.
- Giberti GC. 2011b.** South American *Ilex* raw materials for regional infusions: Advanced arboreal crop systems versus wild forest extractive production. En: Filip R, ed. *Multidisciplinary Approaches on Food Science and Nutrition for the XXI Century*. Kerala, India, 17-27.
- Giberti GC, Gurni AA. 2008.** Anatomía floral comparada de once especies sudamericanas de *Ilex* L. (Aquifoliaceae) relacionadas con la yerba mate. *Dominguezia* **24**: 77-94.
- Gimenez MD, Yañez-Santos AM, Paz RC, et al. 2015.** Assessment of genetic and epigenetic changes in virus-free garlic (*Allium sativum* L.) plants obtained by meristem culture followed by *in vitro* propagation. *Plant Cell Reports* **35**: 129-141.
- Goldberg AD, Allis CD, Bernstein E. 2007.** Epigenetics: a landscape takes shape. *Cell* **128**: 635-638.
- Gorelick R. 2003.** Evolution of dioecy and sex chromosomes via methylation driving Muller's ratchet. *Biological Journal of the Linnean Society* **80**: 353-368.
- Gottlieb AM, Giberti GC, Poggio L. 2005.** Molecular analyses of the genus *Ilex* (Aquifoliaceae) in southern South America, evidence from AFLP and its sequence data. *American Journal of Botany* **92**: 352-369.
- Gottlieb AM, Giberti GC, Poggio L. 2011.** Evaluación del germoplasma de *Ilex paraguariensis* e *Ilex dumosa* (Aquifoliaceae). *Boletín de la Sociedad Argentina de Botánica* **46**: 113-123.
- Gottlieb AM, Poggio L. 2010.** Genomic screening in dioecious «yerba mate» tree (*Ilex paraguariensis* A. St. Hill., Aquifoliaceae) through representational difference analysis. *Genetica* **138**: 567-578.
- Gottlieb AM, Poggio L. 2014.** Quantitative and qualitative genomic characterization of cultivated *Ilex* L. species. *Plant Genetic Resources: Characterisation and Utilisation*: 1-11.
- Gower JC. 1966.** Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**: 325-338.
- Grant SR. 1999.** Genetics of gender dimorphism in higher plants. En: Geber MA, Dawson TE, Dephl LF, eds. *Gender and sexual dimorphism in flowering plants*. New York, USA: Springer, Berlin Heidelberg, 247-274.
- Grativol C, Da Fonseca Lira-Medeiros C, Hemerly AS, Ferreira PCG. 2011.** High efficiency and reliability of inter-simple sequence repeats (ISSR) markers for evaluation of genetic diversity in Brazilian cultivated *Jatropha*

## Referencias bibliográficas

- curcas L. accessions. *Molecular Biology Reports* **38**: 4245-4256.
- Greizerstein EJ, Giberti GC, Poggio L. 2004.** Cytogenetic studies of southern South American *Ilex* species. *Caryologia* **57**: 19-23.
- Grela IA. 2004.** Geografía florística de las especies arbóreas de Uruguay: propuesta para la delimitación de dendrofloras. Tesis de Maestría, Universidad de la República, Uruguay.
- Grewal SIS, Jia S. 2007.** Heterochromatin revisited. *Nature Reviews Genetics* **8**: 35-46.
- Hao D-C, Gu X, Xiao P-G, Liang Z, Xu L, Peng Y. 2013.** Research progress in the phytochemistry and biology of *Ilex* pharmaceutical resources. *Acta Pharmaceutica Sinica B* **3**: 8-19.
- Hao D-C, Xiao P-G, Peng Y, Dong J, Liu W. 2012.** Evaluation of the chloroplast barcoding markers by mean and smallest interspecific distances. *Pakistan Journal of Botany* **44**: 1271-1274.
- Häweker H, Rips S, Koiwa H, et al. 2010.** Pattern recognition receptors require *N*-glycosylation to mediate plant immunity. *The Journal of Biological Chemistry* **285**: 4629-4636.
- Heck CI, de Mejía EG. 2007.** Yerba mate tea (*Ilex paraguariensis*): A comprehensive review on chemistry, health implications, and technological considerations. *Journal of Food Science* **72**: R138-151.
- Heidenreich T, Wollers S, Mendel RR, Bittner F. 2005.** Characterization of the NifS-like domain of ABA3 from *Arabidopsis thaliana* provides insight into the mechanism of molybdenum cofactor sulfuration. *Journal of Biological Chemistry* **280**: 4213-4218.
- Henderson IR, Jacobsen SE. 2007.** Epigenetic inheritance in plants. *Nature* **447**: 418-424.
- Herrera CM, Bazaga P. 2010.** Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytologist* **187**: 867-876.
- Herrera CM, Medrano M, Bazaga P. 2013.** Epigenetic differentiation persists after male gametogenesis in natural populations of the perennial herb *Helleborus foetidus* (Ranunculaceae). *PLoS ONE* **8**: e70730.
- Herrera CM, Medrano M, Bazaga P. 2014.** Variation in DNA methylation transmissibility, genetic heterogeneity and fecundity-related traits in natural populations of the perennial herb *Helleborus foetidus*. *Molecular Ecology* **23**: 1085-1095.
- Hiltbrunner A, Grünig K, Alvarez-Huerta M, Infanger S, Bauer J, Kessler F. 2004.** AtToc90, a new GTP-binding component of the *Arabidopsis* chloroplast protein import machinery. *Plant Molecular Biology* **54**: 427-440.
- Holliday R. 2006.** Epigenetics: A historical overview. *Epigenetics* **1**: 76-80.
- Ikeda Y, Nishimura T. 2015.** The role of DNA methylation in transposable element silencing and genomic imprinting. En: Pontes O, Jin H, eds. *Nuclear Functions in Plant Transcription, Signaling and Development*. New York: Springer New York, 13-29.
- Ito T, Wellmer F, Yu H, et al. 2004.** The homeotic protein AGAMOUS controls microsporogenesis by regulation of *SPOROCYTELESS*. *Nature* **430**: 356-360.
- Jablonka E, Raz G. 2009.** Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology* **84**: 131-176.
- Jaillais Y, Belkhadir Y, Balsemao-Pires E, Dangl JL, Chory J. 2011.** Extracellular leucine-rich repeats as a platform for receptor/coreceptor complex formation. *Proceedings of the National Academy of Sciences* **108**: 8503-8507.
- Janoušek B, Široký J, Vyskot B. 1996.** Epigenetic control of sexual phenotype in a dioecious plant, *Melandrium album*. *Molecular and General Genetics* **250**: 483-490.
- Jansen RK, Cai Z, Raubeson LA, et al. 2007.** Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences* **104**: 19369-19374.
- Jansen RK, Ruhlman TA. 2012.** Plastid genomes of seed plants. En: Bock R, Knoop V, eds. *Genomics of Chloroplasts and Mitochondria. Advances in Photosynthesis and Respiration*. Dordrecht, Netherlands: Springer Science+Business Media B.V., 103-126.
- Jombart T. 2008.** adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403-1405.
- Jombart T, Devillard S, Balloux F. 2010.** Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* **11**: 94.
- Käfer J, Marais GAB, Pannell JR. 2017.** On the rarity of dioecy in flowering plants. *Molecular Ecology* **26**: 1225-1241.
- Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772-780.
- Kejnovsky E, Vyskot B. 2010.** *Silene latifolia*: The classical model to study heteromorphic sex chromosomes.

## Referencias bibliográficas

- Cytogenetic and Genome Research* **129**: 250-262.
- Keller HA. 2007.** Origen mítico de la yerba mate. *Ilex paraguariensis* A. st. Hil (Aquifoliaceae), una versión mbya guaraní. *Suplemento antropológico* **42**: 335-344.
- Keller HA. 2008.** *La fotosíntesis de la cultura: estudios etnobiológicos en comunidades guaraníes de Misiones, Argentina*. Asunción, Paraguay: Centro de Estudios Antropológicos de la Universidad Católica.
- Keller HA. 2013a.** Ka'aguachu: la selva en un sólo árbol. Una contribución de la mitología ava chiripa a la toponimia de la región guaraníca. *Estudios Socioterritoriales* **13**: 101-123.
- Keller HA. 2013b.** Árboles y arbustos en mitos sobre el origen y el fin del mundo de los guaraníes meridionales: elucidación de algunas expresiones fitonímicas. *Bonplandia* **22**: 149-158.
- Keller HA, Giberti GC. 2011.** Primer registro para la flora Argentina de *Ilex affinis* (Aquifoliaceae), sustituto de la «yerba mate». *Boletín de la Sociedad Argentina de Botánica* **46**: 187-194.
- Kim K-J, Lee H-L. 2004.** Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Research* **11**: 247-261.
- Kim HK, Saifullah, Khan S, et al. 2010.** Metabolic classification of South American *Ilex* species by NMR-based metabolomics. *Phytochemistry* **71**: 773-784.
- Koop HU, Herz S, Golds TJ, Nickelsen J. 2007.** The genetic transformation of plastids. En: Bock R, ed. *Topics in Current Genetics*. Springer-Verlag Berlin Heidelberg, 457-510.
- Krinke O, Ruelland E, Valentova O, et al. 2007.** Phosphatidylinositol 4-kinase activation is an early response to salicylic acid in *Arabidopsis* suspension cells. *Plant Physiology* **144**: 1347-1359.
- Kumar G, Rattan UK, Singh AK. 2016.** Chilling-mediated DNA methylation changes during dormancy and its release reveal the importance of epigenetic regulation during winter dormancy in Apple (*Malus x domestica* Borkh.). *PLoS ONE* **11**: e0149934.
- Kurtz S. 2001.** REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29**: 4633-4642.
- Lam VKY, Gomez MS, Graham SW. 2015.** The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae) is colinear with its green relatives and is under strong purifying selection. *Genome Biology and Evolution* **7**: 2220-2236.
- Latzel V, Allan E, Bortolini Silveira A, Colot V, Fischer M, Bossdorf O. 2013.** Epigenetic diversity increases the productivity and stability of plant populations. *Nature Communications* **4**: 1-7.
- Law JA, Jacobsen SE. 2010.** Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* **11**: 204-220.
- Lee NS, Xeu SH, Park JO, Roh MS. 2006.** Molecular evidence for hybridization of *Ilex wandoensis* (Aquifoliaceae) by RAPD analysis. *Journal of Plant Biology* **49**: 491-497.
- Lewis WH, Kennelly EJ, Bass GN, Wedner HJ, Elvin-Lewis MP, Fast WD. 1991.** Ritualistic use of the holly *Ilex guayusa* by Amazonian Jivaro Indians. *Journal of Ethnopharmacology* **33**: 25-30.
- Lewis DR, Ramirez M V, Miller ND, et al. 2011.** Auxin and ethylene induce flavonol accumulation through distinct transcriptional networks. *Plant Physiology* **156**: 144-164.
- Li R, Ma PF, Wen J, Yi TS. 2013.** Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. *PLoS ONE* **8**: e78568.
- Li Y, Shan X, Liu X, Hu L, Guo W, Liu B. 2008.** Utility of the methylation-sensitive amplified polymorphism (MSAP) marker for detection of DNA methylation polymorphism and epigenetic population structure in a wild barley species (*Hordeum brevisubulatum*). *Ecological Research* **23**: 927-930.
- Li L, Xu LJ, Ma GZ, Dong YM, Peng Y, Xiao P-G. 2013.** The large-leaved Kudingcha (*Ilex latifolia* Thunb and *Ilex kudingcha* C.J. Tseng): A traditional Chinese tea with plentiful secondary metabolites and potential biological activities. *Journal of Natural Medicines* **67**: 425-437.
- Liebming E, Veit C, Pabst M, et al. 2011.**  $\beta$ -N-acetylhexosaminidases HEXO1 and HEXO3 are responsible for the formation of paucimannosidic N-glycans in *Arabidopsis thaliana*. *Journal of Biological Chemistry* **286**: 10793-10802.
- van der Linden CG, Vosman B, Smulders MJM. 2002.** Cloning and characterization of four apple MADS box genes isolated from vegetative tissue. *Journal of Experimental Botany* **53**: 1025-1036.
- Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. 2010.** Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS ONE* **5**: e10326.
- Liu C, Shi L, Zhu Y, et al. 2012.** CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* **13**: 715.

## Referencias bibliográficas

- Llorens C, Futami R, Covelli L, et al. 2011.** The Gypsy Database (GyDB) of mobile genetic elements: Release 2.0. *Nucleic Acids Research* **39**: D70-D74.
- Llorens C, Muñoz-Pomer A, Bernad L, Botella H, Moya A. 2009.** Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. *Biology Direct* **4**: 41.
- Loesener T. 1901.** Monographia Aquifoliacearum, Part 1. *Nova Acta Academiae Caesareae Leopoldino-Carolinae Germanicae Naturae Curiosorum* **78**: 1-589.
- Loesener T. 1908.** Monographia Aquifoliacearum, Part 2. *Nova Acta Academiae Caesareae Leopoldino-Carolinae Germanicae Naturae Curiosorum* **89**: 1-313.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013.** OrganellarGenomeDRAW--a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic acids research* **41**: W575-81.
- Loizeau PA, Barriera G, Manen JF, Broennimann O. 2005.** Towards an understanding of the distribution of *Ilex* L. (Aquifoliaceae) on a World-wide scale. *Biologische Skrifter* **55**: 501-520.
- Loizeau PA, Savolainen V, Andrews S, Spichiger R. 2016.** Aquifoliaceae. En: Kadereit JW, Bittrich V, eds. *Flowering plants. Eudicots, The families and genera of vascular plants 14*. Basel, Switzerland: Springer International Publishing, 31-36.
- Lu Y, Rong T, Cao M. 2008.** Analysis of DNA methylation in different maize tissues. *Journal of Genetics and Genomics* **35**: 41-48.
- Luna C, Acevedo R, Collavino MM, González AM, Mroginski LA, Sansberro PA. 2013.** Endophytic bacteria from *Ilex paraguariensis* shoot cultures: Localization, characterization, and response to isothiazolone biocides. *In Vitro Cellular and Developmental Biology - Plant* **49**: 326-332.
- Luna C, González AM, Mroginski LA, Sansberro PA. 2017.** Anatomical and histological features of *Ilex paraguariensis* leaves under different *in vitro* shoot culture systems. *Plant Cell, Tissue and Organ Culture* **129**: 457-467.
- Luo G, He C. 2017.** DNA N6 -methyladenine in metazoans: functional epigenetic mark or bystander? *Nature Structural & Molecular Biology* **24**: 503-506.
- Luo J, Pan X, Peng T, et al. 2016.** DNA methylation patterns of banana leaves in response to *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *Journal of Integrative Agriculture* **15**: 2736-2744.
- MacDonald MJ, D'Cunha GB. 2007.** A modern view of phenylalanine ammonia lyase. *Biochemistry and Cell Biology* **85**: 273-282.
- Maier RM, Schmitz-Linneweber C. 2004.** Plastid genomes. En: Daniell H, Chase C, eds. *Molecular Biology and Biotechnology of Plant Organelles*. Springer, 115-150.
- Maiocchi MG, Del Vitto LA, Petenatti ME, Marchevsky EJ, Pellerano RG, Petenatti EM. 2016.** Multielemental composition and nutritional value of «dumosa» (*Ilex dumosa*), «yerba mate» (*I. paraguariensis*) and their commercial mixture in different forms of use. *Revista de la Facultad de Ciencias Agrarias* **48**: 145-159.
- Manen JF. 2004.** Are both sympatric species *Ilex perado* and *Ilex canariensis* secretly hybridizing? Indication from nuclear markers collected in Tenerife. *BMC Evolutionary Biology* **4**: 46.
- Manen JF, Barriera G, Loizeau PA, Naciri Y. 2010.** The history of extant *Ilex* species (Aquifoliaceae): Evidence of hybridization within a Miocene radiation. *Molecular Phylogenetics and Evolution* **57**: 961-977.
- Manen JF, Boulter MC, Naciri-Graven Y. 2002.** The complex history of the genus *Ilex* L. (Aquifoliaceae): evidence from the comparison of plastid and nuclear DNA sequences and from fossil data. *Plant Systematics and Evolution* **235**: 79-98.
- Marfil CF, Camadro EL, Masuelli RW. 2009.** Phenotypic instability and epigenetic variability in a diploid potato of hybrid origin, *Solanum ruiz-lealii*. *BMC Plant Biology* **9**: 21.
- Martin HA. 1977.** The history of *Ilex* (Aquifoliaceae) with special reference to Australia: Evidence from pollen. *Australian Journal of Botany* **25**: 655-673.
- Martins WS, Lucas DCS, Neves KFS, Bertioli DJ. 2009.** WebSat - A web software for microsatellite marker development. *Bioinformatics* **3**: 282-283.
- Marx F, Janssens MJJ, Urfer P, Scherer R. 2003.** Caffeine and theobromine composition of mate (*Ilex paraguariensis*) leaves in five plantations of Misiones, Argentina. *Plant Foods for Human Nutrition* **58**: 1-8.
- Masuelli RW, Marfil CF. 2010.** Epigenética y evolución. En: Levitus G, Echenique V, Rubinstein C, Hopp E, Mroginski L, eds. *Biotechnology y Mejoramiento vegetal II*. Ediciones INTA, Argentina.
- Masuelli RW, Marfil CF. 2011.** Variabilidad epigenética en plantas y evolución. *Journal of Basic & Applied Genetics* **22**: 1-8.

## Referencias bibliográficas

- Mayol MR. 1997.** La experiencia argentina en sistemas de poda de yerba mate. En: *1er congreso sudamericano de la yerba mate. II reunión técnica del cono sur sobre el cultivo de yerba mate*. Curitiba (PR) Brasil, 337-350.
- Mayol MR, Molina SP. 2011.** Sustratos en la propagación de estacas adultas de yerba mate. En: *5to Congreso Sudamericano de la Yerba Mate*. Posadas, Misiones, 42.
- Mayr E. 1980.** *The evolutionary synthesis: perspectives on the unification of biology*. Cambridge, Massachusetts; London, England: Harvard Univ. Press.
- McFadden GI, Van Dooren GG. 2004.** Evolution: Red algal genome affirms a common origin of all plastids. *Current Biology* **14**: 514-516.
- McKeown PC, Spillane C. 2014.** Landscaping plant epigenetics. En: *Methods in Molecular Biology*. 1-24.
- Medrano M, Herrera CM, Bazaga P. 2014.** Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology* **23**: 4926-4938.
- Messeguer R, Ganal MW, Steffens JC, Tanksley SD. 1991.** Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. *Plant Molecular Biology* **16**: 753-770.
- Middleton CP, Senerchia N, Stein N, et al. 2014.** Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the triticeae tribe. *PLoS ONE* **9**: e85761.
- Mohanty A, Martín JP, Aguinagalde I. 2002.** Population genetic analysis of European *Prunus spinosa* (Rosaceae) using chloroplast DNA markers. *American Journal of Botany* **89**: 1223-1228.
- Moliterni VMC, Cattivelli L, Ranalli P, Mandolino G. 2004.** The sexual differentiation of *Cannabis sativa* L.: A morphological and molecular study. *Euphytica* **140**: 95-106.
- Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010.** Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences* **107**: 4623-4628.
- Morant AV, Jørgensen K, Jørgensen C, et al. 2008.**  $\beta$ -Glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**: 1795-1813.
- Mroginski LA, Rouvier SM, Fabisik JC, et al. 1999.** Effect of medium composition and light supply on *in vitro* shoot proliferation in *Ilex paraguariensis* (Aquifoliaceae). *Journal of Plant Nutrition* **22**: 359-368.
- Murase K, Shigenobu S, Fujii S, et al. 2017.** MYB transcription factor gene involved in sex determination in *Asparagus officinalis*. *Genes to Cells* **22**: 115-123.
- Myers EW, Sutton GG, Delcher AL, et al. 2000.** A whole-genome assembly of *Drosophila*. *Science* **287**: 2196-2204.
- Nafisi M, Goregaoker S, Botanga CJ, et al. 2007.** *Arabidopsis* cytochrome P450 monooxygenase 71A13 catalyzes the conversion of indole-3-acetaldoxime in camalexin synthesis. *The Plant Cell* **19**: 2039-2052.
- Nahar L, Russell WR, Middleton M, Shoeb M, Sarker SD. 2005.** Antioxidant phenylacetic acid derivatives from the seeds of *Ilex aquifolium*. *Acta pharmaceutica (Zagreb, Croatia)* **55**: 187-193.
- Nekrasov V, Li J, Batoux M, et al. 2009.** Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *The EMBO Journal* **28**: 3428-3438.
- Neumann Wendt S. 2005.** Genética de populações em *Ilex paraguariensis* St. Hil.
- Ng KH, Yu H, Ito T. 2009.** AGAMOUS controls *GIANT KILLER*, a multifunctional chromatin modifier in reproductive organ patterning and differentiation. *PLoS Biology* **7**: e1000251.
- Niederhuth CE, Bewick AJ, Ji L, et al. 2016.** Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17**: 194.
- Niklas CO. 1987.** Estudios embriológicos y citológicos en la yerba mate *Ilex paraguariensis* (Aquifoliaceae). *Bonplandia* **6**: 45-56.
- Noirot M, Barre P, Duperray C, Louarn J, Hamon S. 2003.** Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA: Consequences on genome size evaluation in coffee tree. *Annals of Botany* **92**: 259-264.
- Paiva DI, Cascales J, Gauchat ME, Scherer RA, Gottlieb AM. 2016.** Caracterización molecular de materiales seleccionados de yerba mate (*Ilex paraguariensis* St. Hil.). En: *Journal of Basic & Applied Genetics* **27**(Suppl.): 242.
- Paiva DI, Cascales J, Gauchat ME, Scherer RA, Gottlieb AM. 2017.** Genotipificación ISSR de materiales argentinos de yerba mate (*Ilex paraguariensis* St. Hil.). En: *Anais VII Congresso Sul-Americano da erva-mate; III Simpósio Internacional de Erva-Mate; I Feira de Tecnologia na Indústria Ervateira*. Erechim, Brasil, 50.
- Pannell JR, Eppley SM, Dorken ME, Berjano R. 2014.** Regional variation in sex ratios and sex allocation in androdioecious *Mercurialis annua*. *Journal of Evolutionary Biology* **27**: 1467-1477.
- Paupière MJ, van Heusden AW, Bovy AG. 2014.** The metabolic basis of pollen thermo-tolerance: Perspectives for breeding. *Metabolites* **4**: 889-920.

## Referencias bibliográficas

- Peakall R, Smouse PE. 2012.** GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**: 2537-2539.
- Pelletier K, Murrell JR, Shirley BW. 1997.** Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis*. *Plant Physiology* **113**: 1437-1445.
- Pereira MF, Ciampi AY, Inglis PW, Souza VA, Azevedo VCR. 2013.** Shotgun sequencing for microsatellite identification in *Ilex paraguariensis* (Aquifoliaceae). *Applications in Plant Sciences* **1**: 1200245.
- Pérez-Figueroa A. 2013.** msap: A tool for the statistical analysis of methylation-sensitive amplified polymorphism data. *Molecular Ecology Resources* **13**: 522-527.
- Péros J-P, Berger G, Portemont A, Boursiquot J-M, Lacombe T. 2011.** Genetic variation and biogeography of the disjunct *Vitis* subg. *Vitis* (Vitaceae). *Journal of Biogeography* **38**: 471-486.
- Pikaard CS, Scheid OM. 2014.** Epigenetic regulation in plants (CD Allis, M-L Caparros, T Jenuwein, y D Reinberg, Eds.). *Cold Spring Harbor Perspectives in Biology* **6**: a019315.
- Pourcel L, Routaboul JM, Cheynier V, Lepiniec L, Debeaujon I. 2007.** Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends in Plant Science* **12**: 29-36.
- Prat Kricun SD. 1995.** 2º Curso de capacitación en producción de yerba mate. Estación Experimental Agropecuaria Cerro Azul, INTA.
- Prat Kricun SD. 2001.** Cruzamientos dirigidos en *Ilex paraguariensis* St. Hil. EEA Cerro Azul. *Informe Técnico N°73*.
- Prat Kricun SD. 2009.** Antiguo cultivo, nueva infusión: inscriben dos variedades de *Ilex dumosa*. <http://www.agrositio.com/vertext/vertext.asp?id=106378andse=94>.
- Prat Kricun SD. 2011.** Apuntes sobre el mejoramiento genético de la yerba mate. *Compartiendo Tecnología*: 46-51.
- Prat Kricun SD, Belingheri LD. 2003.** Programa de Mejoramiento genético de la Yerba Mate en el INTA. En: 3º Congreso Sul-Americano da Erva Mate. Chapecó, SC, Brasil, 88.
- Prescott AG, Stamford NPJ, Wheeler G, Firmin JL. 2002.** *In vitro* properties of a recombinant flavonol synthase from *Arabidopsis thaliana*. *Phytochemistry* **60**: 589-593.
- Prince LM. 2015.** Plastid primers for angiosperm phylogenetics and phylogeography. *Applications in Plant Sciences* **3**: 1400085.
- Provan J, Biss PM, McMeel D, Mathews S. 2004.** Universal primers for the amplification of chloroplast microsatellites in grasses (Poaceae). *Molecular Ecology Notes* **4**: 262-264.
- Provan J, Powell W, Hollingsworth PM. 2001.** Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution* **16**: 142-147.
- Racolta A, Bryan AC, Tax FE. 2014.** The receptor-like kinases GSO1 and GSO2 together regulate root growth in *Arabidopsis* through control of cell division and cell fate specification. *Developmental Dynamics* **243**: 257-278.
- R Core Team. 2016.** R: a language and environment for statistical computing. URL: <https://www.r-project.org/>
- Radice M, Vidari G. 2007.** Caracterización fitoquímica de la especie *Ilex guayusa* Loes. y elaboración de un prototipo de fitofármaco de interés comercial. *La Granja. Revista de Ciencias de la Vida* **6**: 3-11.
- Rafiee Seyed MA. 2012.** Determinación de arbutina en las especies de *Ilex*. Tesis de Licenciatura, Universidad de Belgrano, Argentina.
- Rakocevic M, Medrado MJS, Martim SF, Assad ED. 2009.** Sexual dimorphism and seasonal changes of leaf gas exchange in the dioecious tree *Ilex paraguariensis* grown in two contrasted cultivation types. *Annals of Applied Biology* **154**: 291-301.
- Rakocevic M, Janssens M, Scherer R. 2012.** Light responses and gender issues in the domestication process of yerba-mate, a subtropical evergreen. En: Bezerra AD, Ferreira TS, eds. *Evergreens*. Nova Science Publishers, Inc, 63-95.
- Rambaut A. 2016.** FigTree v1.4.3 software. Institute of Evolutionary Biology, University of Edinburgh.
- Rau V. 2009.** La yerba mate en Misiones (Argentina). Estructura y significados de una producción localizada. *Agroalimentaria* **28**: 49-58.
- Raubeson LA, Jansen RK. 2005.** Chloroplast genomes of plants. En: Henry RJ, ed. *Plant diversity and evolution: genotypic and phenotypic variation in higher plants*. Oxfordshire, UK: CABI, 45-68.
- Raubeson LA, Peery R, Chumley TW, et al. 2007.** Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* **8**: 174.
- Rayson S, Arciga-Reyes L, Wootton L, et al. 2012.** A role for nonsense-mediated mRNA decay in plants: Pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. *PLoS ONE* **7**: e31917.
- Realini MF, González GE, Font F, Picca PI, Poggio L, Gottlieb AM. 2015.** Phylogenetic relationships in *Opuntia* (Cactaceae, Opuntioideae) from southern South America. *Plant Systematics and Evolution* **301**: 1123-1134.

## Referencias bibliográficas

- Realini MF, Wulff AF, Gonzalez GE, Poggio L, Gottlieb AM. 2014.** Estudios citogenéticos en plantas femeninas y masculinas de yerba mate. En: *VI Congreso Sudamericano Yerba Mate*. Montevideo, Uruguay, 137.
- Rebollo R, Horard B, Hubert B, Vieira C. 2010.** Jumping genes and epigenetics: towards new species. *Gene* **454**: 1-7.
- Rendell S, Ennos RA. 2003.** Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology* **12**: 2681-2688.
- Rey HY, Sansberro PA, Collavino MM, Daviña JR, González AM, Mroginski LA. 2002.** Colchicine, trifluralin, and oryzalin promoted development of somatic embryos in *Ilex paraguariensis* (Aquifoliaceae). *Euphytica* **123**: 49-56.
- Reyna-Lopez GE, Simpson J, Ruiz-Herrera J. 1997.** Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Molecular and General Genetics* **253**: 703-710.
- Richards EJ. 2006.** Inherited epigenetic variation — revisiting soft inheritance. *Nature Reviews Genetics* **7**: 395-401.
- Richards EJ. 2011.** Natural epigenetic variation in plant species: A view from the field. *Current Opinion in Plant Biology* **14**: 204-209.
- Richards CL, Schrey AW, Pigliucci M. 2012.** Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecology Letters* **15**: 1016-1025.
- Rico L, Ogaya R, Barbeta A, Peñuelas J. 2014.** Changes in DNA methylation fingerprint of *Quercus ilex* trees in response to experimental field drought simulating projected climate change. *Plant Biology* **16**: 419-427.
- Riggs AD, Porter TN. 1996.** Overview of epigenetic mechanisms. En: Russo VEA, Martienssen R, Riggs AD, eds. *Epigenetic Mechanisms of Gene Regulation*. New York, USA: Cold Spring Harbor Laboratory Press, 29-45.
- Roberts DW. 2016.** Package ‘labdsv’. URL: <https://cran.r-project.org/web/packages/labdsv>
- Rodríguez López CM, Wetten AC, Wilkinson MJ. 2010.** Progressive erosion of genetic and epigenetic variation in callus-derived cocoa (*Theobroma cacao*) plants. *New Phytologist* **186**: 856-868.
- Ross S, Arriaga ME, Pechi E. 2017.** Establecimiento *in vitro* de yerba mate (*Ilex paraguariensis* A. St.-Hil.) nativa de Uruguay. *Agrociencia Uruguay* **21**: 15-23.
- Ruhlman TA, Jansen RK. 2014.** The plastid genomes of flowering plants. En: *Methods in Molecular Biology*, 3-38.
- Sabir JSM, Arasappan D, Bahieldin A, et al. 2014.** Whole mitochondrial and plastid genome SNP analysis of nine date palm cultivars reveals plastid heteroplasmy and close phylogenetic relationships among cultivars. *PLoS ONE* **9**: e94158.
- Sáenz-Mata J, Jiménez-Bremont JF. 2012.** HR4 gene is induced in the *Arabidopsis-Trichoderma atroviride* beneficial interaction. *International Journal of Molecular Sciences* **13**: 9110-9128.
- Sansberro PA, Rey HY, Bernardis A, Luna C, Collavino MM, Mroginski LA. 2000.** Plant regeneration of *Ilex paraguariensis* (Aquifoliaceae) by *in vitro* culture of nodal segments. *Biocell* **24**: 53-63.
- Sansberro PA, Rey HY, Mroginski LA. 2001a.** *In vitro* culture of zygotic embryos of *Ilex* species. *HortScience* **36**: 351-352.
- Sansberro PA, Rey HY, Mroginski LA, Collavino MM. 1998.** *In vitro* culture of rudimentary embryos of *Ilex paraguariensis*: Responses to exogenous cytokinins. *Journal of Plant Growth Regulation* **17**: 101-105.
- Sansberro PA, Rey HY, Mroginski LA, Collavino MM. 1999.** *In vitro* plant regeneration of *Ilex paraguariensis* (Aquifoliaceae). *In Vitro Cellular & Developmental Biology - Plant* **35**: 401-402.
- Sansberro PA, Rey HY, Mroginski LA, Krivenki MA. 2001b.** Plant regeneration from *Ilex* spp. (Aquifoliaceae) *in vitro*. *Biocell* **25**: 139-146.
- Schattner P, Brooks AN, Lowe TM. 2005.** The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research* **33**: 569-572.
- Scherer R, Janssens MJJ, Marx F, Urfer P, Schneider E. 2006.** Saponin content and quality-related traits of mass-selected yerba maté (*Ilex paraguariensis* A. St.-Hil.) trees. *Journal of Herbs, Spices & Medicinal Plants* **12**: 73-85.
- Scherer R, Urfer P, Mayol MR, Belingheri LD, Marx F, Janssens MJJ. 2002.** Inheritance studies of caffeine and theobromine content of Mate (*Ilex paraguariensis*) in Misiones, Argentina. *Euphytica* **126**: 203-210.
- Schlüter PM, Harris SA. 2006.** Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes* **6**: 569-572.
- Schrey AW, Alvarez M, Foust CM, et al. 2013.** Ecological epigenetics: Beyond MS-AFLP. *Integrative and Comparative Biology* **53**: 340-350.
- Schulz B, Eckstein RL, Durka W. 2013.** Scoring and analysis of methylation-sensitive amplification polymorphisms for epigenetic population studies. *Molecular Ecology Resources* **13**: 642-653.

## Referencias bibliográficas

- Schulz B, Eckstein RL, Durka W. 2014. Epigenetic variation reflects dynamic habitat conditions in a rare floodplain herb. *Molecular Ecology* **23**: 3523-3537.
- Selbach-Schnadelbach A, Smith Cavalli S, Manen JF, Coelho GC, De Souza-Chies TT. 2009. New information for *Ilex* phylogenetics based on the plastid *psbA-trnH* intergenic spacer (Aquifoliaceae). *Botanical Journal of the Linnean Society* **159**: 182-193.
- Seshadri S, Akiyama T, Opassiri R, Kuaprasert B, Cairns JK. 2009. Structural and enzymatic characterization of Os3BGlu6, a rice  $\beta$ -glucosidase hydrolyzing hydrophobic glycosides and (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 2)-linked disaccharides. *Plant Physiology* **151**: 47-58.
- Setoguchi H, Watanabe I. 2000. Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *American Journal of Botany* **87**: 793-810.
- Shannon CE, Weaver W. 1949. *The mathematical theory of communication*. Urbana, Illinois, EEUU: University of Illinois Press.
- Shaw J, Lickey EB, Beck JT, et al. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**: 142-166.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* **94**: 275-288.
- Shaw J, Shafer HL, Rayne Leonard O, Kovach MJ, Schorr M, Morris AB. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *American Journal of Botany* **101**: 1987-2004.
- Shi C, Hu N, Huang H, Gao J, Zhao YJ, Gao LZ. 2012. An improved chloroplast DNA extraction procedure for whole plastid genome sequencing. *PLoS ONE* **7**: e31468.
- Shi L, Li N, Wang S, et al. 2016. Molecular evidence for the hybrid origin of *Ilex dabieshanensis* (Aquifoliaceae). *PLoS ONE* **11**: e0147825.
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO Journal* **5**: 2043-2049.
- Sides CB, Enquist BJ, Ebersole JJ, Smith MN, Henderson AN, Sloat LL. 2014. Revisiting Darwin's hypothesis: Does greater intraspecific variability increase species ecological breadth? *American Journal of Botany* **101**: 56-62.
- Singh S, Panda MK, Nayak S. 2012. Evaluation of genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. *Industrial Crops and Products* **37**: 284-291.
- Skinner MK. 2014. Endocrine disruptor induction of epigenetic transgenerational inheritance of disease. *Molecular and Cellular Endocrinology* **398**: 4-12.
- Skinner MK. 2015. Environmental epigenetics and a unified theory of the molecular aspects of evolution: A neo-Lamarckian concept that facilitates neo-Darwinian evolution. *Genome Biology and Evolution* **7**: 1296-1302.
- Skinner MK, Manikkam M, Guerrero-Bosagna C. 2010. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends in Endocrinology and Metabolism* **21**: 214-222.
- De Smet I, Voß U, Jürgens G, Beeckman T. 2009. Receptor-like kinases shape the plant. *Nature Cell Biology* **11**: 1166-1173.
- Sneath PHA, Sokal RR. 1973. *Numerical taxonomy: The principles and practice of numerical classification* (WH Freeman, Ed.). San Francisco, California.
- Soares Lopes MR, Martinez ST, Chaves VC, da Silva Ramos Rocha A, do Amarante L. 2007. Determinação por HPLC de cafeína e teobromina em folhas jovens e velhas de *Ilex paraguariensis*. *Revista Brasileira de Biociências* **5**: 954-956.
- Song Y, Ma K, Bo W, Zhang Z, Zhang D. 2012a. Sex-specific DNA methylation and gene expression in andromonoecious poplar. *Plant Cell Reports* **31**: 1393-1405.
- Song Y, Wang Z, Bo W, Ren Y, Zhang Z, Zhang D. 2012b. Transcriptional profiling by cDNA-AFLP analysis showed differential transcript abundance in response to water stress in *Populus hopeiensis*. *BMC Genomics* **13**: 286.
- Stark LA. 2010. Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE-Life Sciences Education* **9**: 6-9.
- Stevens PF. 2017. *Angiosperm Phylogeny Website. Version 14*. <http://www.mobot.org/MOBOT/research/APweb/>.
- Stumm GB, Vedder H, Schlegel J. 1997. A simple method for isolation of PCR fragments from silver-stained polyacrylamide gels by scratching with a fine needle. *Technical Tips Online* **2**: 116-117.
- Talamali A, Bajji M, Le Thomas A, Kinet JM, Dutuit P. 2003. Flower architecture and sex determination: How does



## Referencias bibliográficas

- Atriplex halimus* play with floral morphogenesis and sex genes? *New Phytologist* **157**: 105-113.
- Talamali A, Dutuit P, Le Thomas A, Gorenflot R. 2001.** Polygamy in *Atriplex halimus* L. (Chenopodiaceae). *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie* **324**: 107-113.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.
- Tangphatsornruang S, Sangsrakru D, Chanprasert J, et al. 2010.** The chloroplast genome sequence of mungbean (*Vigna radiata*) determined by high-throughput pyrosequencing: structural organization and phylogenetic relationships. *DNA Research* **17**: 11-22.
- Tarragó J, Sansberro PA, Filip R, et al. 2005.** Effect of leaf retention and flavonoids on rooting of *Ilex paraguariensis* cuttings. *Scientia Horticulturae* **103**: 479-488.
- Thea AE. 2016.** Actualización sobre los compuestos nutricionales y bioactivos de la yerba mate (*I. paraguariensis*). En: *III Jornada de Divulgación Científica sobre Yerba Mate y Salud*. Córdoba, Argentina, 28.
- Timme RE, Kuehl J V, Boore JL, Jansen RK. 2007.** A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: Identification of divergent regions and categorization of shared repeats. *American Journal of Botany* **94**: 302-312.
- Torii KU. 2004.** Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *234*: 1-46.
- Untergasser A, Cutcutache I, Koressaar T, et al. 2012.** Primer3-new capabilities and interfaces. *Nucleic Acids Research* **40**: e115.
- Varotto S, Forestan C. 2010.** Understanding epialleles formation in response to environmental cues and their heritability in plants. *Maydica* **55**: 275-279.
- Verma D, Daniell H. 2007.** Chloroplast vector systems for biotechnology applications. *Plant Physiology* **145**: 1129-1143.
- Vos P, Hogers R, Bleeker M, et al. 1995.** A new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407-4414.
- Vyskot B, Hobza R. 2004.** Gender in plants: sex chromosomes are emerging from the fog. *TRENDS in Genetics* **20**: 432-438.
- Vyskot B, Hobza R. 2015.** The genomics of plant sex chromosomes. *Plant Science* **236**: 126-135.
- Waddington C. 1953.** Epigenetics and evolution. En: Brown R, Danielli JF, eds. *Evolution: Society of Experimental Biology Symposium*. Cambridge: Cambridge University Press, 186-199.
- Wakasugi T, Sugita M, Tsudzuki T, Sugiura M. 1998.** Updated gene map of tobacco chloroplast DNA. *Plant Molecular Biology Reporter* **16**: 231-241.
- Weising K, Gardner RC. 1999.** A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* **42**: 9-19.
- Winge H. 1997.** Conservação genética do erva-mate no Brasil. En: *II Reunião Técnica do Cono Sul sobre a cultura do Erva Mate*. 209-226.
- Xiao S, Charoenwattana P, Holcombe L, Turner JG. 2003.** The *Arabidopsis* genes *RPW8.1* and *RPW8.2* confer induced resistance to powdery mildew diseases in tobacco. *Molecular Plant-Microbe Interactions* **16**: 289-294.
- Xiong LZ, Xu CG, Saghai Maroof MA, Zhang Q. 1999.** Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Molecular and General Genetics* **261**: 439-446.
- Yang J-B, Yang SX, Li HT, Yang J, Li DZ. 2013.** Comparative chloroplast genomes of *Camellia* species. *PLoS ONE* **8**: e73053.
- Yang M, Zhang X, Liu G, et al. 2010.** The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *PLoS ONE* **5**: e12762.
- Yao X, Liu Y-Y, Tan Y-H, Song Y, Corlett RT. 2016a.** The complete chloroplast genome sequence of *Helwingia himalaica* (Helwingiaceae, Aquifoliales) and a chloroplast phylogenomic analysis of the Campanulidae. *PeerJ* **4**: e2734.
- Yao X, Tan Y-H, Liu Y-Y, Song Y, Yang J-B, Corlett RT. 2016b.** Chloroplast genome structure in *Ilex* (Aquifoliaceae). *Scientific Reports* **6**: 28559.
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012.** Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* **13**: 134.
- Yu G, Duan J, Yan H, Song B, He Z. 2011.** cDNA-AFLP analysis of gene expression differences between the flower bud and sprout-shoot apical meristem of *Angelica sinensis* (Oliv.) Diels. *Genetics and Molecular Biology* **34**: 274-

## Referencias bibliográficas

279.

- Zambounis A, Psomopoulos FE, Ganopoulos I, et al. 2016.** *In silico* analysis of the LRR receptor-like serine threonine kinases subfamily in *Morus notabilis*. *Plant OMICS* **9**: 319-326.
- Zhang X, Shiu S, Cal A, Borevitz JO. 2008.** Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genetics* **4**: e1000032.
- Zhang K, Song Q, Wei Q, et al. 2016.** Down-regulation of *OsSPX1* caused semi-male sterility, resulting in reduction of grain yield in rice. *Plant Biotechnology Journal* **14**: 1661-1672.
- Zhang W, Wang X, Yu Q, Ming R, Jiang J. 2008.** DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. *Genome Research* **18**: 1938-1943.
- Zhou CP, Qi YP, You X, et al. 2013.** Leaf cDNA-AFLP analysis of two citrus species differing in manganese tolerance in response to long-term manganese-toxicity. *BMC Genomics* **14**: 621.
- Zietkiewicz E, Rafalski A, Labuda D. 1994.** Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* **20**: 176–183.
- Zluvova J, Nicolas M, Berger A, Negrutiu I, Monéger F. 2006.** Premature arrest of the male flower meristem precedes sexual dimorphism in the dioecious plant *Silene latifolia*. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 18854-18859.

## APÉNDICES

## Apéndice I – Tablas y figuras suplementarias

**Tabla AI-1.** Genes y marcos de lectura abiertos codificados en los plastomas de plantas superiores. Adaptado de (Bock 2007).

Gen	Producto génico	Funciones y observaciones
<b>Componentes fotosintéticos</b>		
<i>psaA, B</i>	Subunidades A y B del FSI	Subunidades del centro de reacción, esenciales para la función del FSI
<i>psaC</i>	Subunidad C del FSI	Subunidad esencial de unión al cofactor
<i>psaI, J</i>	Subunidades I y J del FSI	Subunidades pequeñas, no esenciales para la función del FSI
<i>ycf3, 4</i>	Proteínas Ycf3 e Ycf4	Factores esenciales de ensamblaje del FSI
<i>psbA, D</i>	Proteínas D1 y D2 del FSII	Centros de reacción, esenciales para la función del FSII
<i>psbB, C</i>	Subunidades CP47 y CP43 del FSII	Proteínas de la antena interna, esenciales para la función del FSII
<i>psbE, F</i>	Subunidad $\alpha$ y $\beta$ del citocromo $b_{559}$	Esenciales para el ensamblado, estabilidad y función del FSII, protección del FSII contra la fotoinhibición, oxidación oscura de plastoquinol
<i>psbH-N, T, Z</i>	Subunidades H-N, T y Z del FSII	Subunidades pequeñas, con funciones en el ensamblado del FSII, estabilización y fotoprotección, transferencia de electrones, reparación de centros de reacción dañados
<i>petA, B</i>	Subunidades f y $b_6$ del cyt $b_6f$	Subunidades centrales del complejo cyt $b_6f$ , esenciales para su función
<i>petD, G, L, N</i>	Subunidades IV, G, L y N del cyt $b_6f$	Esenciales para la función y estabilización del cyt $b_6f$
<i>atpA</i>	Subunidad $\alpha$ de la ATP sintasa	CF <sub>1</sub> , sitio de unión a nucleótidos
<i>atpB</i>	Subunidad $\beta$ de la ATP sintasa	CF <sub>1</sub> , sitio catalítico
<i>atpE</i>	Subunidad $\epsilon$ de la ATP sintasa	CF <sub>1</sub> , regulación de la activación de CF <sub>1</sub> CF <sub>0</sub> , necesaria para el pasaje de protones
<i>atpF</i>	Subunidad b de la ATP sintasa	CF <sub>0</sub> , unión de CF <sub>1</sub>
<i>atpH, I</i>	Subunidades c y a de la ATP sintasa	CF <sub>0</sub> , translocación de protones
<i>ndhA-K</i>	Subunidades A-K de la NAD(P)H deshidrogenasa	Clororespiración, transferencia cíclica de electrones, ausente en plastomas de gimnospermas
<b>Componentes del sistema genético</b>		
<i>rbcL</i>	Subunidad grande de la Rubisco	Fijación del CO <sub>2</sub>
<i>rpoA, B, C1, C2</i>	Subunidades $\alpha$ , $\beta$ , $\beta'$ y $\beta''$ de la ARN polimerasa	Transcripción, ARN polimerasa plastídica tipo <i>E. coli</i> (PEP)
<i>matK</i>	Maturasa de intrones	Factor de <i>splicing</i> para intrones de tipo II
<i>infA</i>	Factor iniciador de la traducción 1	Traducción, pseudogen inactivo o gen perdido (y transferido al núcleo) en varios linajes
<i>rrn16, 23, 5, 4.5</i>	ARN ribosomal 16S, 23S, 5S, 4.5S	Traducción, subunidad menor del ribosoma
<i>trnA-UGC</i>	ARNt-Alanina(UGC)	Traducción
<i>trnC-GCA</i>	ARNt-Cisteína(GCA)	
<i>trnD-GUC</i>	ARNt-Aspartato(GUC)	
<i>trnE-UUC</i>	ARNt-Glutamato(UUC)	
<i>trnF-GAA</i>	ARNt-Fenilalanina(GAA)	Traducción
<i>trnG-GCC / UCC</i>	ARNt-Glicina(GCC) / (UCC)	
<i>trnH-GUG</i>	ARNt-Histidina(GUG)	
<i>trnI-CAU / GAU</i>	ARNt-Isoleucina(CAU) / (GAU)	
<i>trnK-UUU</i>	ARNt-Lisina(UUU)	

Gen	Producto génico	Funciones y observaciones
<i>trnL-CAA / UAA / UAG</i>	ARNt-Leucina(CAA) / (UAA) / (UAG)	
<i>trnM-CAU</i>	ARNt-Metionina(CAU)	
<i>trn<sup>f</sup>MCAU</i>	ARNt-N-Formilmetionina(CAU)	Iniciación de la traducción
<i>trnN-GUU</i>	ARNt-Asparagina(GUU)	Traducción
<i>trnP-UGG</i>	ARNt-Prolina(UGG)	
<i>trnQ-UUG</i>	ARNt-Glutamina(UUG)	
<i>trnR-ACG / UCU</i>	ARNt-Arginina(ACG) / (UCU)	
<i>trnS-GCU / GGA / UGA</i>	ARNt-Serina(GCU) / (GGA) / (UGA)	
<i>trnT-GGU / UGU</i>	ARNt-Treonina(GGU) / (UGU)	
<i>trnV-GAC / UAC</i>	ARNt-Valina(GAC) / (UAC)	
<i>trnW-CCA</i>	ARNt-Triptofano(CCA)	
<i>trnY-GUA</i>	ARNt-Tirosina(GUA)	
<i>rps2-4, 7, 8, 11, 12, 14-16, 18, 19</i>	Proteínas ribosomales S2, S3, S4, S7, S8, S11, S12, S14, S15, S16, S18, S19	
<i>rpl2, 14, 16, 20, 22, 23, 32, 33, 36</i>	Proteínas ribosomales L2, L14, L16, L20, L22, L23, L32, L33, L36	Traducción, subunidad mayor del ribosoma
<b>Otros</b>		
<i>clpP</i>	Subunidad catalítica de la proteasa Clp	Degradación de proteínas dependiente de ATP, esencial para la supervivencia celular
<i>accD</i>	Subunidad carboxilasa del acetyl-CoA	Biosíntesis de ácidos grasos, esencial para la supervivencia celular
<i>yef5 / ccsA</i>	Subunidad A del complejo del sistema II para biogénesis de citocromos tipo c	Necesario para la fijación del grupo hemo a los citocromos plastídicos de tipo c
<i>yef10 / cemA</i>	Proteína de la membrana interna	Presuntamente involucrado en la captación de carbono inorgánico
<b>ORFs conservados de función desconocida</b>		
<i>yef1</i>	Proteína putativa Ycf1	Gen esencial, función desconocida
<i>yef2</i>	Proteína putativa Ycf2	Gen esencial, función desconocida, contiene un dominio putativo de unión a nucleótidos
<i>yef15</i>	Desconocido	ORF con un significado funcional poco claro
<i>sprA</i>	ARN pequeño	Función desconocida

FSI: fotosistema I; FSII: fotosistema II. LHCII: Complejo de captura externa de luz. Cyt b<sub>6</sub>f: citocromo b<sub>6</sub>f. CF<sub>0</sub>: complejo principal de la ATP sintasa, que atraviesa la membrana tilacoides; CF<sub>1</sub>: complejo principal de la ATP sintasa, que sobresale de la membrana tilacoides. PEP: “*plastid-encoded plastid RNA polymerase*”. Ycf: “*hypothetical chloroplast reading frame*”, marco de lectura abierto en el plastoma. ORF: “*open reading frame*”, marco de lectura abierto.

**Tabla AI-2.** Marcadores cpSSR utilizados para realizar el genotipado preliminar en *I. dumosa* e *I. paraguariensis*.

Región	Nombre cpSSR	Motivo de repetición	Cebadores	
			Directo 5' - 3'	Reverso 5' - 3'
5' de <i>trnS</i>	<i>ccmp2</i>	(A) <sub>11</sub>	GATCCCGGACGTAATCCTG	ATCGTACCGAGGGTTCGAAT
Intrón <i>trnG-GCC</i>	<i>ccmp3</i>	(T) <sub>11</sub>	CAGACCAAAAGCTGACATAG	GTTTCATTCGGCTCCTTTAT
Intrón <i>atpF</i>	<i>ccmp4</i>	(T) <sub>13</sub>	AATGCTGAATCGAYGACCTA	CCAAAATATTBGGAGGACTCT
Intrón 2 <i>ycf3</i>	<i>ccmp6</i>	(T) <sub>5</sub> C(T) <sub>17</sub>	CGATGCATATGTAGAAAGCC	CATTACGTGCGACTATCTCC
IGS <i>atpB-rbcL</i>	<i>ccmp7</i>	(A) <sub>13</sub>	CAACATATACTACTGTCAAG	ACATCATTATTGTATACTCTTTC
IGS <i>rpl2-rps19</i>	<i>ccmp10</i>	(A) <sub>11</sub>	TTTTTTTTTAGTGAACGTGTCA	TTCGTCGDCGTAGTAAATAG

Y= C o T; B= C, G o T; D= A, G o T

**Tabla AI-3.** Plastomas utilizados como referencia para realizar los ensamblados y para seleccionar las secuencias de cloroplastos.

Nombre	Accesión
<i>Camellia cuspidata</i> voucher Yang3159	NC_022459.1
<i>C. danzaiensis</i> voucher Yang3147	NC_022460.1
<i>C. impressinervis</i> voucher Yang1080	NC_022461.1
<i>C. pitardii</i> voucher Yang3148	NC_022462.1
<i>C. sinensis</i>	NC_020019.1
<i>C. sinensis</i> cultivar Longjing43	KF562708.1
<i>C. taliensis</i> voucher Yang3157	NC_022264.1
<i>C. taliensis</i> voucher Yang3158	KF156836.1
<i>C. yunnanensis</i> voucher Yang1090	NC_022463.1
<i>Coffea arabica</i>	NC_008535.1
<i>Theobroma cacao</i>	NC_014676.2

**Tabla AI-4.** Números de accesoión de GenBank de las especies adicionales utilizadas para evaluar la utilidad filogenética de los intrones.

Intrón	<i>I. cornuta</i>	<i>I. kaushue</i>	<i>I. pentagona</i>	Otras especies
<i>atpF</i>	GQ997302	nd	nd	nd
<i>clpP</i>	GQ997307	nd	nd	nd
<i>ndhA</i>	GQ997310	FR849975	FR849977	nd
<i>ndhB</i>	GQ997311	nd	nd	nd
<i>petB</i>	GQ997322	nd	nd	nd
<i>petD</i>	GQ997323	nd	nd	AM396557 <i>I. aquifolium</i>
<i>rpl16</i>	GQ997349	nd	nd	KC881529 <i>I. opaca</i>
<i>rpl2</i>	GQ997350	nd	nd	nd
<i>rpoC1</i>	GQ997359	nd	nd	nd
<i>rps12</i>	GQ997362	nd	nd	nd
<i>rps16</i>	GQ997365	nd	nd	nd
<i>trnA(UGC)</i>	HQ664579	nd	nd	nd
<i>trnG(UCC)</i>	FR849992	FR849996	FR849997	nd
<i>trnI(GAU)</i>	HQ664579	nd	nd	nd
<i>trnK(UUU)</i>	GQ997309	nd	nd	AF542607.2 <i>I. aquifolium</i> ; AJ429376 <i>Ilex</i> sp.
<i>trnL(UAA)</i>	FJ394743	nd	nd-	AJ492570.2 <i>I. argentina</i> ; AJ492575 <i>I. brasiliensis</i> ; AJ492576 <i>I. brevicuspis</i> ; AJ492590 <i>I. crenata</i> ; AJ492577 <i>I. integerrima</i> ; AJ492603 <i>I. leucoclada</i> ; AJ492578.2 <i>I. microdonta</i> ; AJ492587 <i>I. mutchagara</i> ; AJ492574 <i>I. pseudobuxus</i> ; AJ492579.2 <i>I. theezans</i> ; AJ275346 <i>I. yunnanensis</i> .
<i>trnV(UAC)</i>	nd	nd	nd	AJ429722 <i>Ilex</i> sp.
<i>ycf3</i>	GQ997379	nd	nd	nd

nd, sin datos de secuencia en GenBank. Las especies asiáticas están subrayadas.

**Tabla AI-5.** Números de accesoión de Genbank de las especies adicionales utilizadas para evaluar la utilidad filogenética de los IGS. Las especies asiáticas se muestran subrayadas.

IGS <sup>a</sup>	<u><i>I. cornuta</i></u>	<u><i>I. kaushue</i></u>	<u><i>I. pentagona</i></u>	<u><i>I. latifolia</i></u> *	Otras especies
<i>accD-psaI</i>	FR849953	FR849954	FR849955	nd	nd
<i>atpB-rbcL</i>	AF471624	nd	KF255714	AF471655	AF471597 <i>I. amara</i> ; AF471600 <i>I. brasiliensis</i> ; AF471649 <i>I. integerrima</i> ; AF471688 <i>I. pseudobuxus</i> ; AF471709 <i>I. theezans</i>
<i>atpH-atpI</i>	FR849969	FR849967	FR849968	nd	nd
<i>ndhC-trnV</i>	FR849958	FR849957	FR849960; FR849956	nd	nd
<i>ndhF-rpl32</i>	nd	FR849979	FR849981	FR849980	nd
<i>ndhJ-trnF</i>	FR849962	FR849961 <sup>b</sup>	FR849964	nd	nd
<i>petA-psbJ</i>	FR849950	FR849951	FR849952	nd	nd
<i>psbE-petL</i>	FR849946	FR849947	FR849945	nd	nd
<i>rpl32-trnL</i>	FR849989	FR849986	FR849991	FR849984	nd
<i>trnK-3' rps16</i>	FR849971	FR849970	FR849973	nd	nd
<i>trnL-trnF</i>	FJ394743	FR849961 <sup>b</sup>	KF255729	AJ492604	FJ394730 <i>I. amara</i> ; AJ492570.2 <i>I. argentina</i> ; AJ492575 <i>I. brasiliensis</i> ; AJ492576 <i>I. brevicuspis</i> ; AJ492577 <i>I. integerrima</i> ; AJ492578.2 <i>I. microdonta</i> ; AJ492574 <i>I. pseudobuxus</i> ; AJ492579.2 <i>I. theezans</i>
<i>trnS-trnG</i>	FR849992	FR849994	FR849997	nd	nd
<i>trnH-psbA</i>	JN044938	JN044946	JN044953	HG765001	EU359340 <i>I. affinis</i> ; EU359360 <i>I. amara</i> ; EU359339 <i>I. argentina</i> ; EU359357 <i>I. brasiliensis</i> ; EU359341 <u><i>I. pedunculosa</i></u> ; EU359352 <i>I. theezans</i>
<i>trnT-psbD</i>	FR849965	FR849966	nd	nd	nd

\* Secuencias de *I. latifolia* disponibles previo a la publicación de su plastoma completo.

<sup>a</sup> regiones ubicadas en la LSC, con excepción de *ndhF-rpl32* y *rpl32-trnL*, que pertenecen a la SSC.

<sup>b</sup> este fragmento contiene tanto el IGS *ndhJ-trnF* como *trnL-trnF*.

nd, sin datos en Genbank.

**Tabla AI-6.** Estadísticas de los ensamblados realizados en INDEAR, por individuo, por especie y en total. Id: *I. dumosa*; Ip: *I. paraguariensis*.

Muestra	Datos crudos (lecturas)			Ensamblado ( <i>contigs</i> )							
	Total <sup>a</sup>	Cloroplastos		Total <sup>a</sup>		Cloroplastos					
	n°	n°	%	n°	Tamaño (pb) <sup>b</sup>	n°	%	Tamaño (pb) <sup>b</sup>	%	Mayor (pb) <sup>c</sup>	N50 <sup>d</sup>
<b>Id222</b>	31.602	307	0,971	1.522	904.190	30	1,971	46.116	5,100	2.934	1.633
<b>Id55</b>	25.449	558	2,193	1.202	733.959	40	3,328	94.458	12,870	9.815	3.643
<b>IdCRV7</b>	28.271	492	1,740	1.335	833.147	29	2,172	88.394	10,610	11.113	4.013
<b>Id227</b>	12.311	97	0,788	492	262.515	2	0,407	3.050	1,162	1.558	1.558
<b>Ip217</b>	20.480	2.488	12,148	758	532.910	6	0,792	131.137	24,608	73.107	73.107
<b>Ip224</b>	35.533	1.806	5,083	1.498	857.104	13	0,868	130.549	15,231	39.488	13.146
<b>Ip30</b>	19.357	1.579	8,157	59	161.970	16	27,119	103.318	63,788	35.542	8.018
<b>IpJB</b>	36.461	1.169	3,206	1.803	1.119.961	21	1,165	127.511	11,385	23.048	11.243
<i>Total Id + Ip</i>	209.464	8.496	4,056	8.669	5.405.756	157	1,811	724.533	13,403	-	-
<i>Total Id</i>	97.633	1.454	1,489	4.551	2.733.811	101	2,219	232.018	8,487	-	-
<i>Total Ip</i>	111.831	7.042	6,297	4.118	2.671.945	56	1,360	492.515	18,433	-	-
<i>Promedio Id + Ip</i>	26.183	1.062	4,056	1.084	675.720	19.63	1,811	90.567	13,403	24.576	14.545
<i>Total Id</i>	24.408	363.5	1,489	1.138	683.453	25.25	2,219	58.005	8,487	6.355	2.712
<i>Total Ip</i>	27.958	1.760.5	6,297	1.030	667.986	14.00	1,360	123.129	18,433	42.796	26.379

<sup>a</sup> Incluyendo secuencias no plastídicas.

<sup>b</sup> Suma de los tamaños de los *contigs*.

<sup>c</sup> Tamaño del mayor *contig* ensamblado.

<sup>d</sup> Estadístico bioinformático calculado como la longitud media de los *contigs* más largos que, sumados, contribuyen al 50% del total de bases secuenciadas del plastoma.

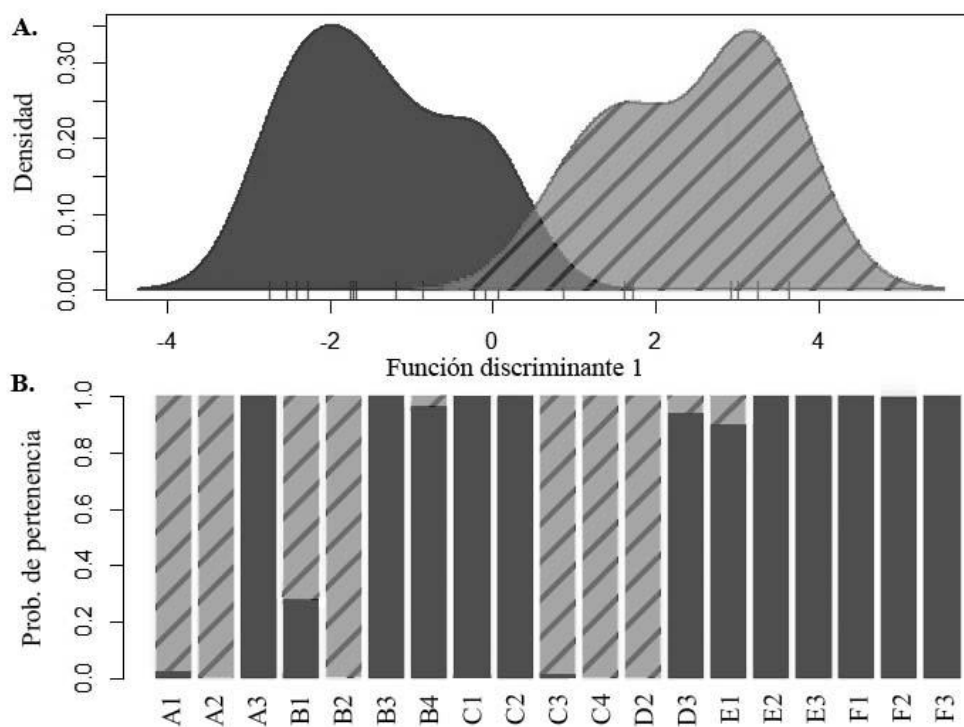


**Tabla AI-7.** Secuencias de los oligonucleótidos adaptadores y cebadores utilizados en MSAP. Las bases selectivas, indicadas como “XXX”, se detallan en la **Tabla 11** del Capítulo 2.

Oligonucleótido	Secuencia (5' - 3')
Adaptador <i>EcoRI</i> Superior	CTCGTAGACTGCGTACC
<i>EcoRI</i> Inferior	AATTGGTACGCAGTCTAC
<i>HpaII/MspI</i> Superior	GATCATGAGTCCTGCT
<i>HpaII/MspI</i> Inferior	CGAGCAGGACTCATGA
Cebador preselectivo <i>EcoRI</i> +A	GACTGCGTACCAATTCA
<i>HpaII/MspI</i> +A	ATGAGTCCTGCTCGGA
<i>HpaII/MspI</i> +T	ATGAGTCCTGCTCGGT
Cebador selectivo <i>EcoRI</i> +3	GACTGCGTACCAATTCXXX
<i>HpaII/MspI</i> +3	ATGAGTCCTGCTCGGXXX

**Tabla AI-8.** Secuencias de adaptadores y cebadores utilizados en cDNA-AFLP. Las bases selectivas se detallan en la **Tabla 13** del Capítulo 2.

Adaptador/cebador	Secuencia (5' - 3')
Adaptador <i>EcoRI</i> Superior	CTCGTAGACTGCGTACC
Adaptador <i>EcoRI</i> Inferior	AATTGGTACGCAGTCTAC
Adaptador <i>MseI</i> Superior	GACGATGAGTCCTGAG
Adaptador <i>MseI</i> Inferior	TACTCAGGACTCAT
Cebador preselectivo <i>EcoRI</i> +A	GACTGCGTACCAATTCA
Cebador preselectivo <i>MseI</i> +C	GATGAGTCCTGAGTAAC
Cebador selectivo <i>EcoRI</i> +XXX	GACTGCGTACCAATTCAXXX
Cebador selectivo <i>MseI</i> +XXX	GATGAGTCCTGAGTAAXXX

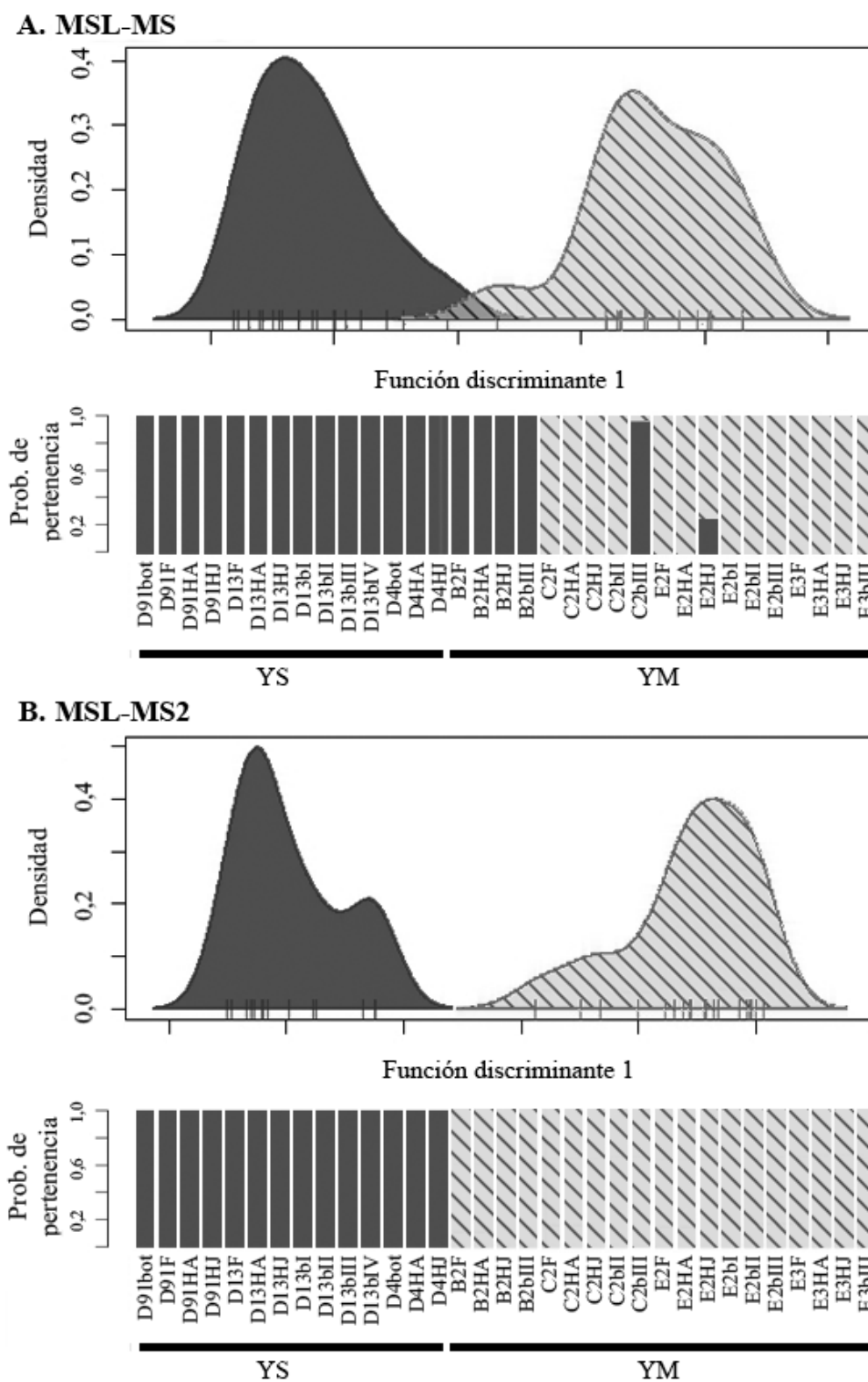


**Figura AI-1.** Análisis discriminante de los componentes principales de las muestras de yerba mate, analizadas con ISSR. **A.** Gráfico DAPC (24,7% de la varianza retenida). **B.** Gráfico compoplot (probabilidades de pertenencia al grupo asignado).

**Tabla AI-9.** Coeficientes de correlación *r* entre las matrices MSL-MS, MSL-MS2 y NML, a nivel inter e intraespecífico.

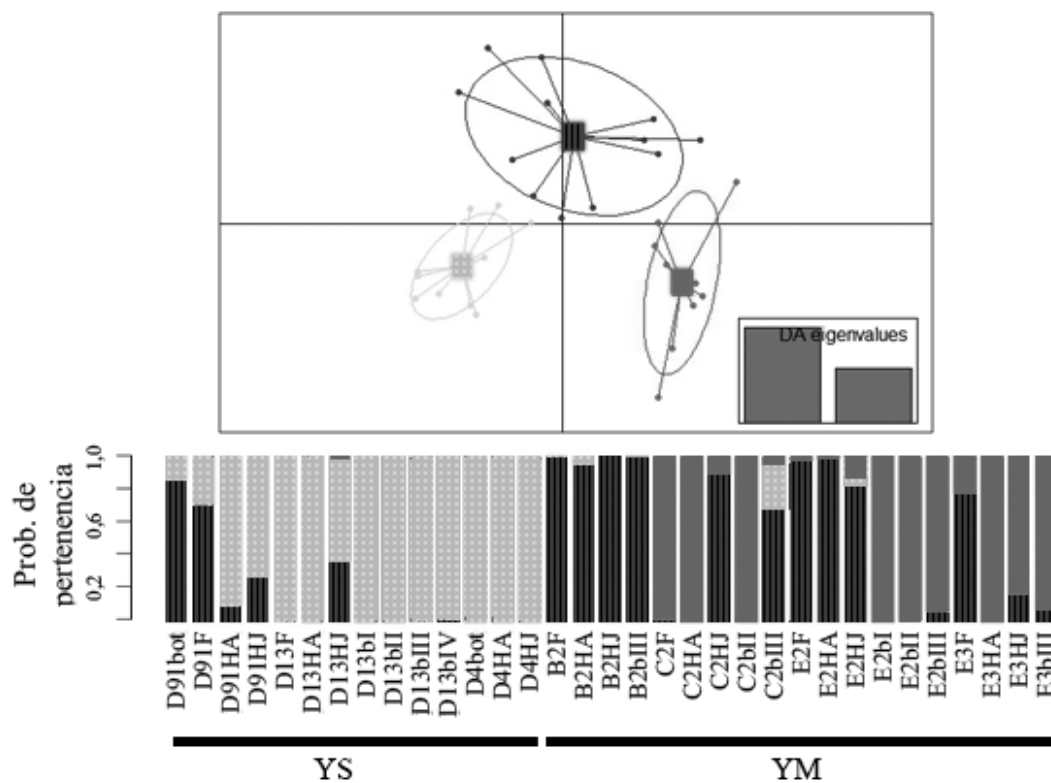
Matriz	Test de Mantel ( <i>r</i> )		
	MSL-MS vs. NML	MSL-MS2 vs. NML	MSL-MS vs. MSL-MS2
YM+YS	0,476**	0,874**	0,667**
YM	0,385*	0,692**	0,676**
YS	0,299*	0,664**	0,643**

\*  $p < 0,01$ ; \*\*  $p = 0,001$

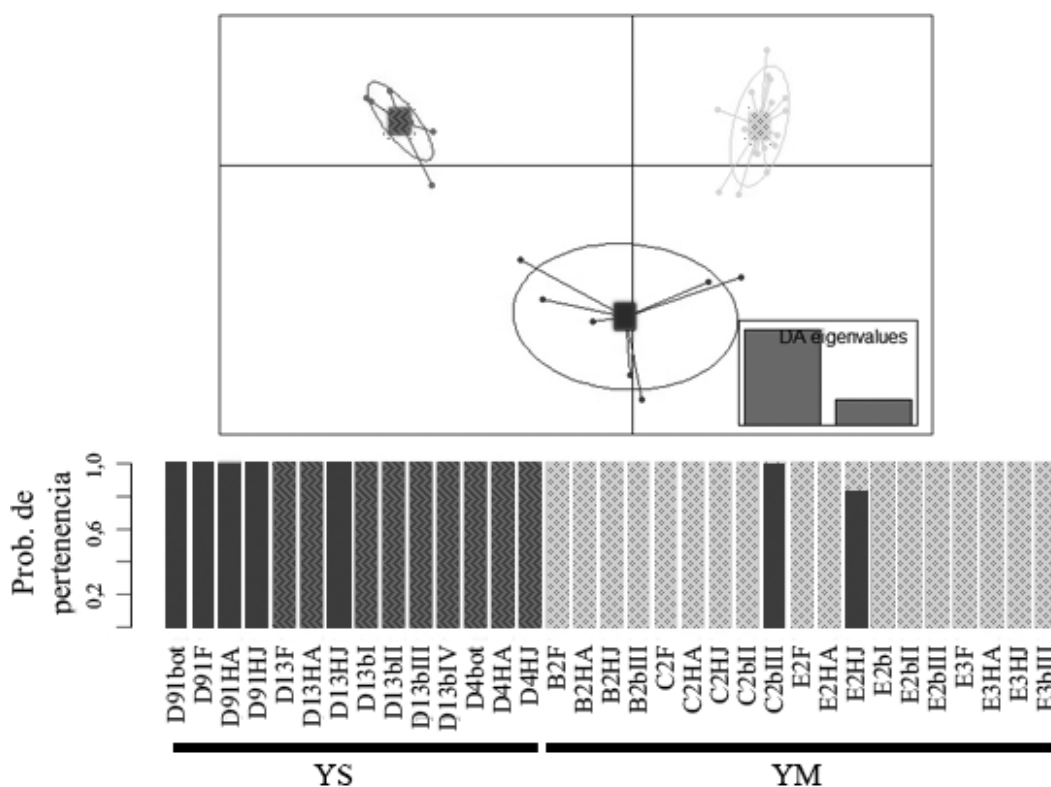


**Figura AI-2.** Análisis DAPC de los datos epigenéticos de YM+YS con K=2, y sus gráficos compplot respectivos. **A.** MSL-MS (24,83% de la varianza acumulada). **B.** MSL-MS2 (21,7% de la varianza). Los patrones se corresponden entre gráficos.

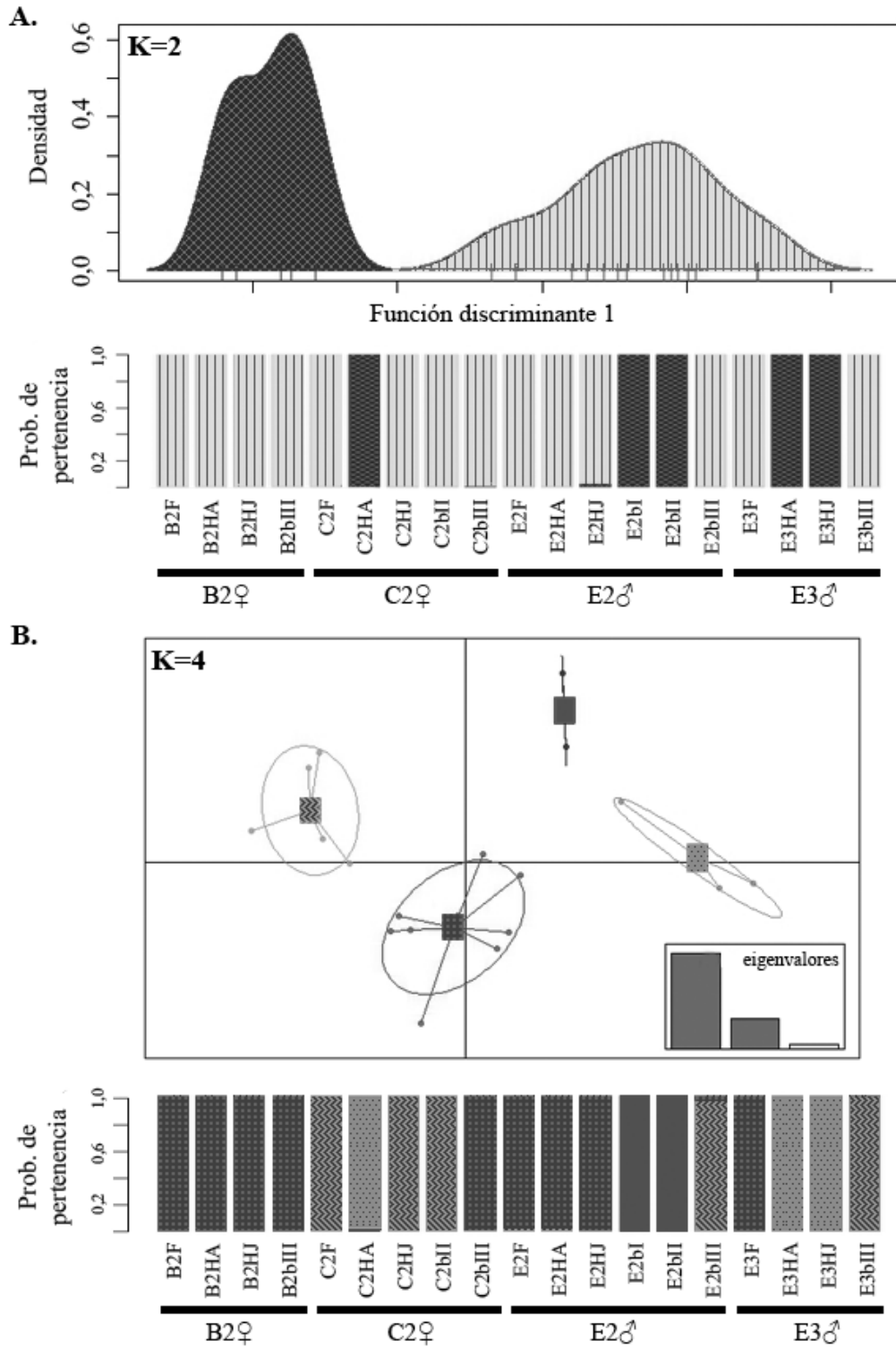
**A. MSL-MS**



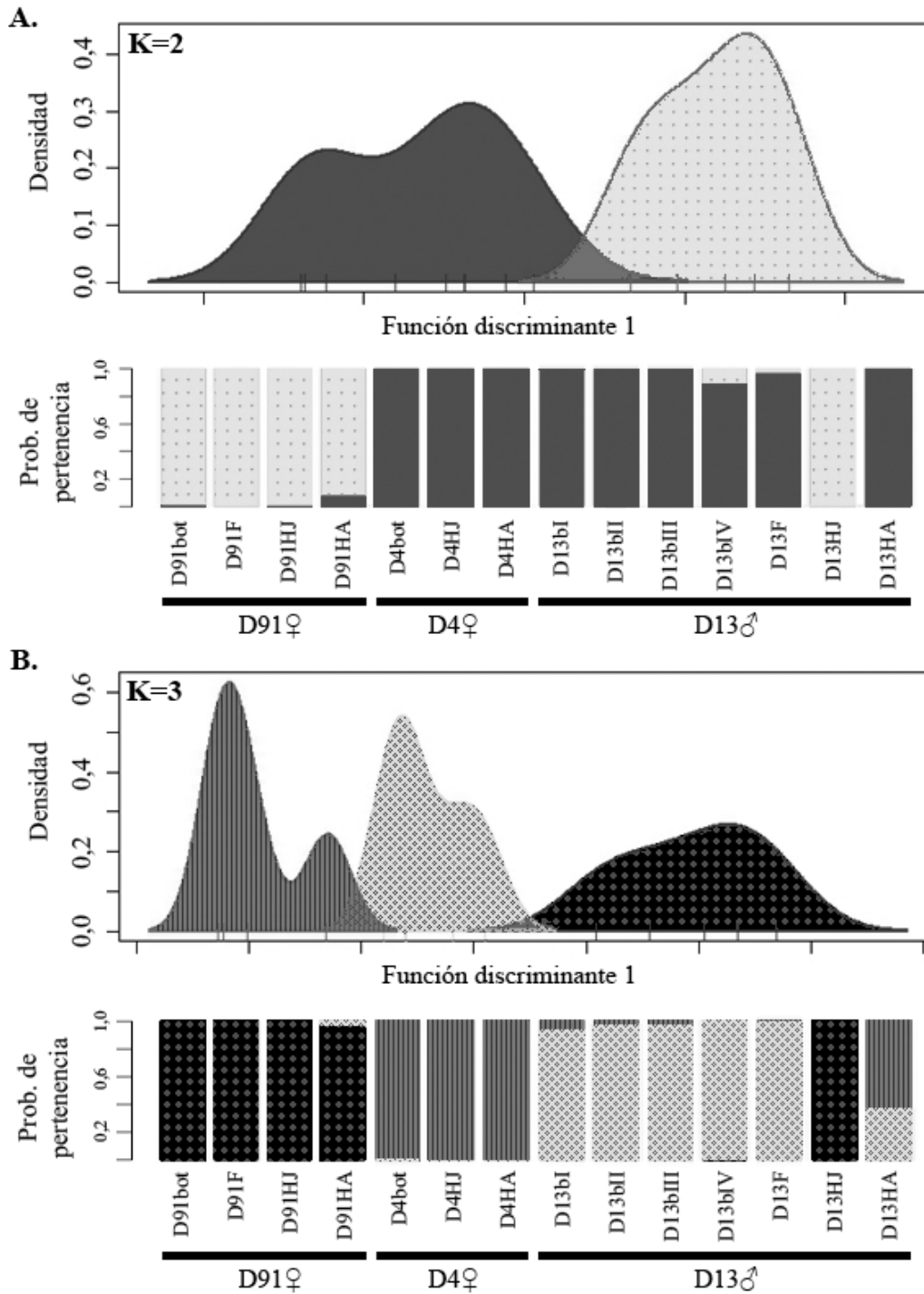
**B. MSL-MS2**



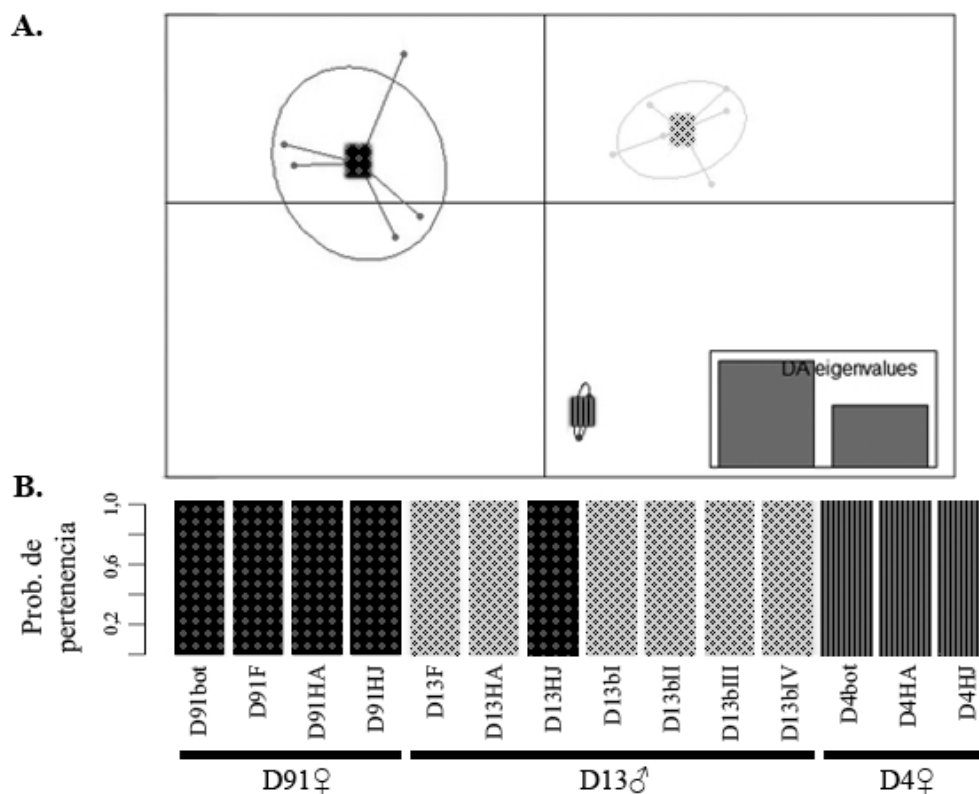
**Figura AI-3.** Análisis discriminante de las matrices epigenéticas interespecíficas (con K=3) y gráficos compoplot. **A.** MSL-MS (17,61% de la varianza acumulada). **B.** MSL-MS2 (33,8% de la varianza). El patrón del centroide de cada grupo se corresponde con el del compoplot.



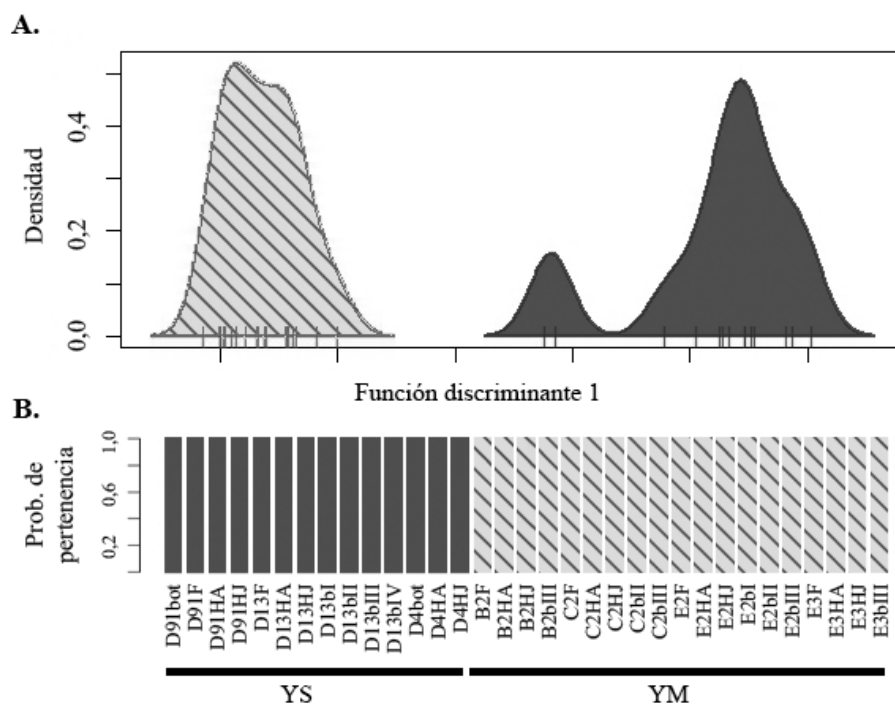
**Figura AI-4.** Análisis DAPC de la matriz MSL-MS de yerba mate y probabilidad de pertenencia de cada muestra al grupo asignado. **A.** K=2 (31,8% de la varianza acumulada). **B.** K=4 (31,8% de la varianza). Los patrones son coherentes entre los gráficos DAPC y su compplot correspondiente.



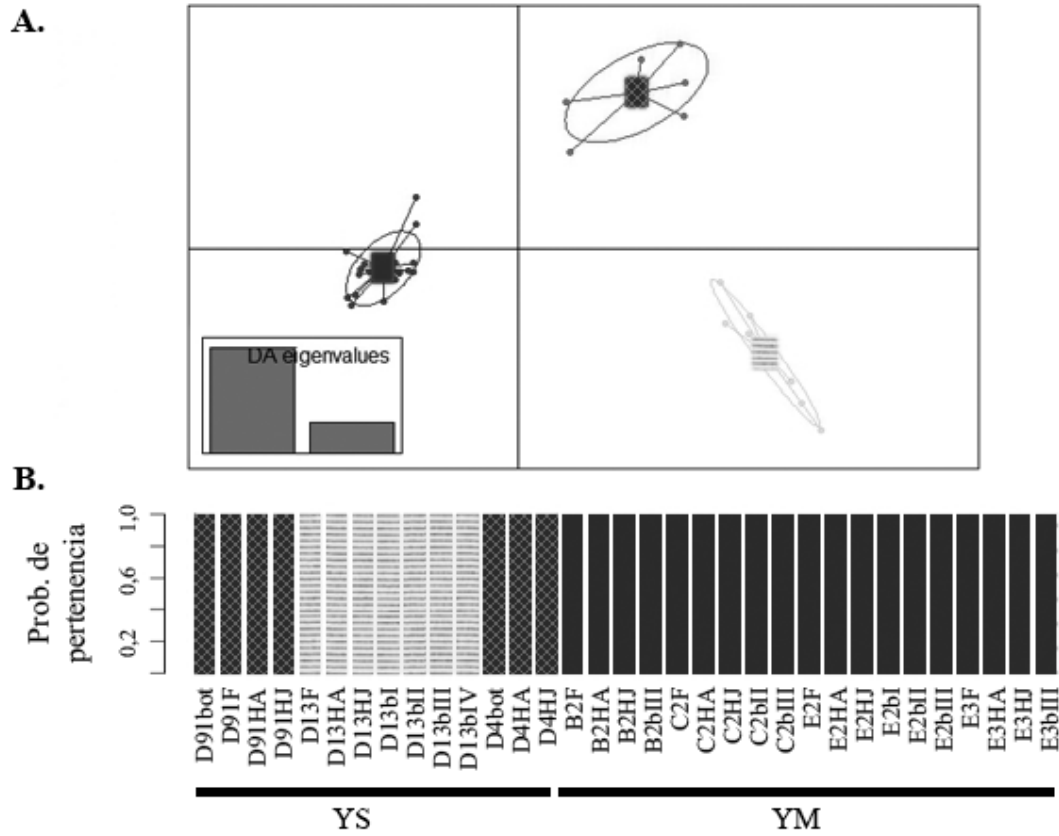
**Figura AI-5.** Análisis DAPC de la matriz MSL-MS de yerba señorita y gráfico compoplot asociado (varianza retenida= 18,9%). **A.** K=2. **B.** K=3. Los patrones de los DAPC se corresponden con los de su compoplot respectivo.



**Figura AI-6.** Análisis discriminante de la matriz MSL-MS2 de yerba señorita con K=3. **A.** Gráfico DAPC (varianza acumulada= 44,16%). **B.** Gráfico compoplot. Los patrones de los DAPC se corresponden con los de su compoplot respectivo.

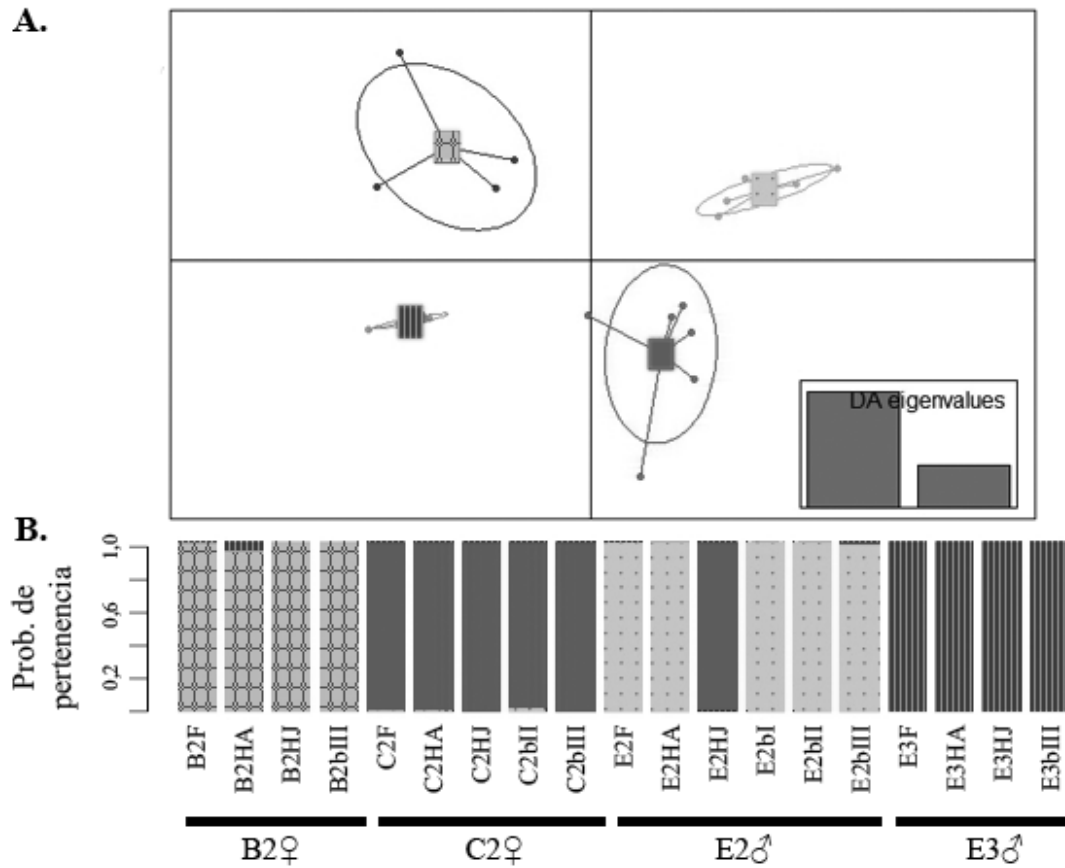


**Figura AI-7.** Análisis discriminante de los componentes principales de la matriz YM+YS-NML. **A.** Gráfico DAPC con K=2 (20,94% de la varianza). **B.** Gráfico compoplot. Los patrones de los DAPC se corresponden con los de su compoplot respectivo.

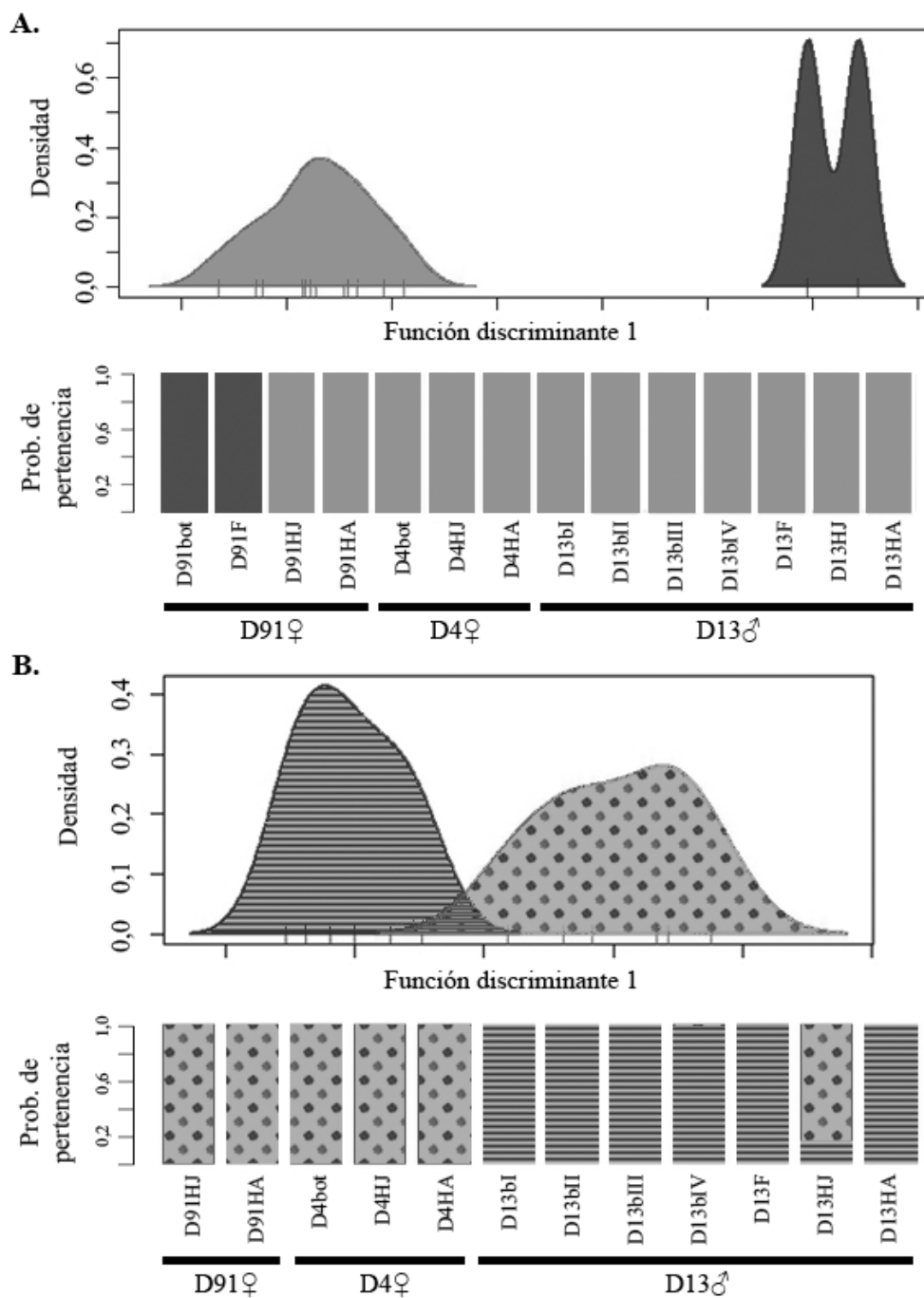


**Figura AI-8.** Análisis discriminante de la matriz NML interespecífica con  $K=3$ . **A.** Gráfico DAPC (32,98% de la varianza). **B.** Gráfico compoplot. En el centroide de cada grupo se indica su patrón en el compoplot.

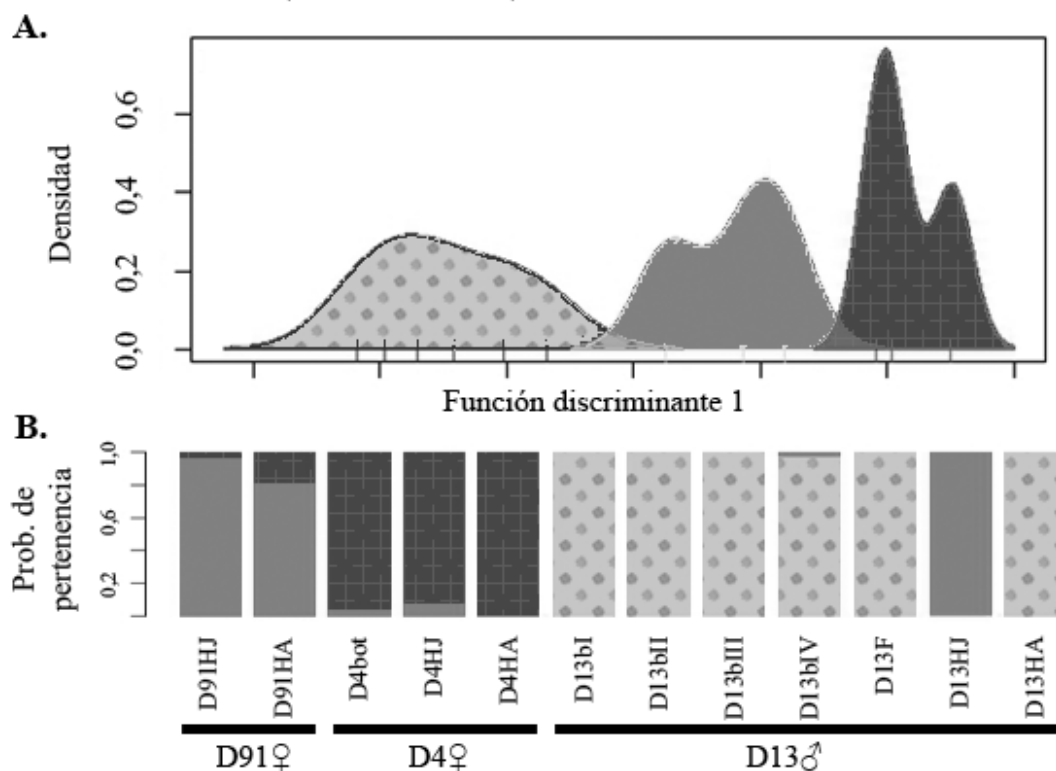




**Figura AI-9.** Análisis discriminante en base a YM-NML considerando el número de individuos donantes (K=4). **A.** Gráfico DAPC (24,8% de la varianza acumulada). **B.** Gráfico compoplot. Los patrones se corresponden entre los DAPC y su compoplot correspondiente.



**Figura AI-10.** Gráficos DAPC y compplot en base a los NML de las muestras de yerba señorita (K=2). **A.** Incluyendo los tejidos sexuales de D91 (21,8% de la varianza retenida); **B.** descartándolos (24,9% de la varianza). Los tonos de grises y los patrones se corresponden entre los gráficos DAPC y los compplot asociados.



**Figura AI-11.** Análisis discriminante en base a YS-NML considerando el número de individuos donantes (K=3). **A.** Gráfico DAPC (24,9% de la varianza acumulada). **B.** Gráfico compoplot. Los patrones se corresponden entre los DAPC y su compoplot correspondiente.

**Tabla AI-10.** Características de las bandas MSAP seleccionadas para secuenciar.

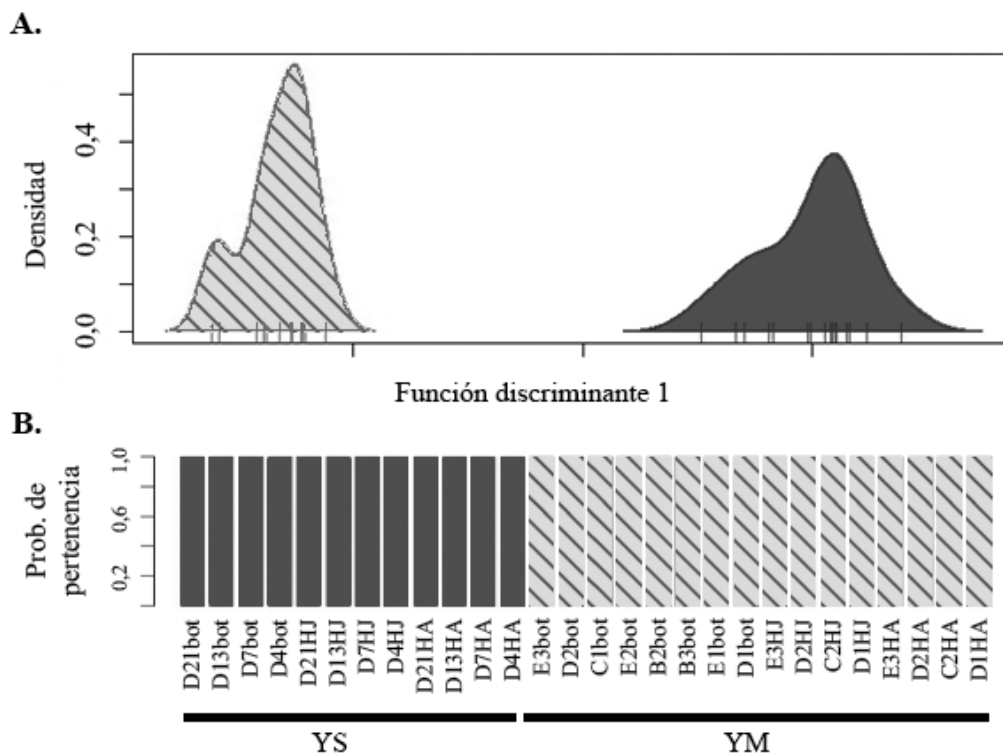
Fragmento	Muestra <sup>a</sup>	Cebadores	Tamaño observado en los gels (pb) <sup>b</sup>
1	D13 - bot I - H	E-AAC + H/M-ATG	450
2	D13 - flores - H	E-AAC + H/M-ATG	350
3	D13 - bot I - H	E-AAC + H/M-ATG	270
4	E2 - HA - H	E-AAG + H/M-TTC	240
5	D13 - bot I - M	E-AAG + H/M-ATT	400-450
6	D13 - bot I - M	E-AAG + H/M-AAC	300
7	D13 - bot I - M	E-AAG + H/M-AAC	250
8	D91 - bot - M	E-ACT + H/M-AAG	400
9	D13 - bot I - H	E-ACT + H/M-AAG	290
10	C2 - bot II - H	E-ACT + H/M-AAG	180
11	D91 - bot - M	E-AGC + H/M-TAC	500
12	B2 - HJ - M	E-AGC + H/M-TAC	230
13	E2 - bot I - H	E-AGC + H/M-ATT	300 - 400
14	D13 - flores - M	E-AGG + H/M-TAG	200
15	D13 - bot I - M	E-AGG + H/M-TAG	200
16	D4 - HJ - H	E-AGG + H/M-AAG	150
22	E3 - HJ - M	E-AAC + H/M-ATG	350

23	E3 - HJ - M	E-AAC + H/M-ATG	350
24	E3 - HJ - M	E-AAC + H/M-ATG	240
25	E3 - HJ - M	E-AAC + H/M-ATG	240
26	D13 - bot I - M	E-AAC + H/M-ATG	235
27	D13 - bot III - M	E-ACA + H/M-ATG	300
28	C2 - bot I - H	E-ACA + H/M-ATG	250
29	D4 - bot - M	E-AAG + H/M-ATT	220
30	E2 - bot III - M	E-AGC + H/M-ATT	200-250
31	C2 - bot II - H	E-AAG + H/M-AAC	300
32	D13 - bot III - M	E-AAG + H/M-AAC	250-300
33	C2 - bot II - M	E-ACT + H/M-AAG	300-350

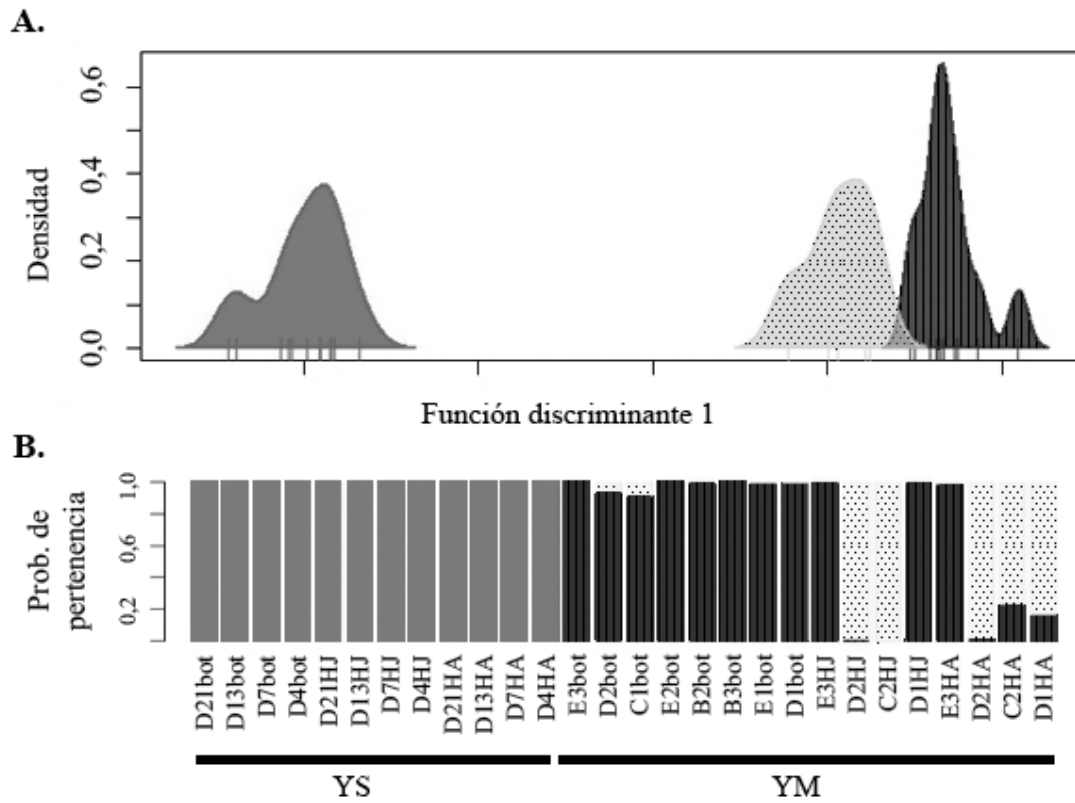
<sup>a</sup> datos de la muestra de la que se aisló la banda para secuenciar: nombre del individuo, órgano muestreado, y enzima de restricción empleada.

<sup>b</sup> tamaño aproximado en pares de bases.

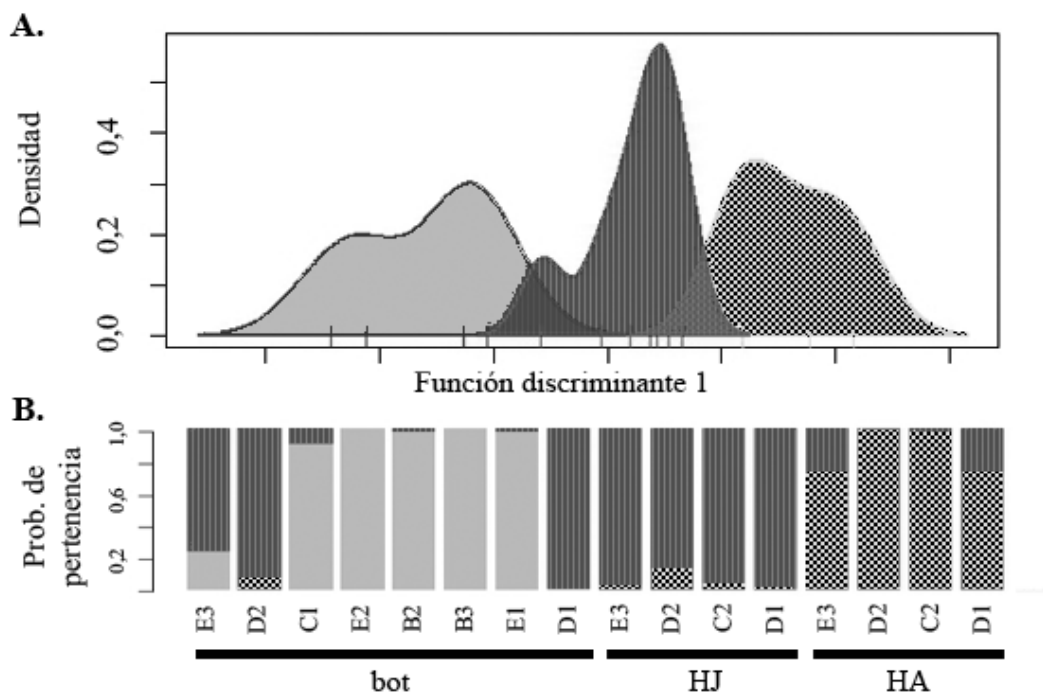
bot: botones florales; I, II: estadios del desarrollo de los botones florales.



**Figura AI-12.** Análisis discriminante de la matriz YM+YS con K=2. **A.** Gráfico DAPC (20,87% de la varianza retenida). **B.** Gráfico compoplot. Los patrones se corresponden entre los distintos gráficos.



**Figura AI-13.** Análisis discriminante con K=3 de la matriz YM+YS. **A.** Gráfico DAPC con 20,87% de la varianza retenida. **B.** Gráfico compoplot. Los patrones se corresponden entre los gráficos.



**Figura AI-14.** Análisis DAPC de la yerba mate con K=3. **A.** Gráfico DAPC (14,9% de la varianza explicada). **B.** Gráfico compoplot. Los patrones se corresponden entre los gráficos DAPC y su respectivo compoplot.

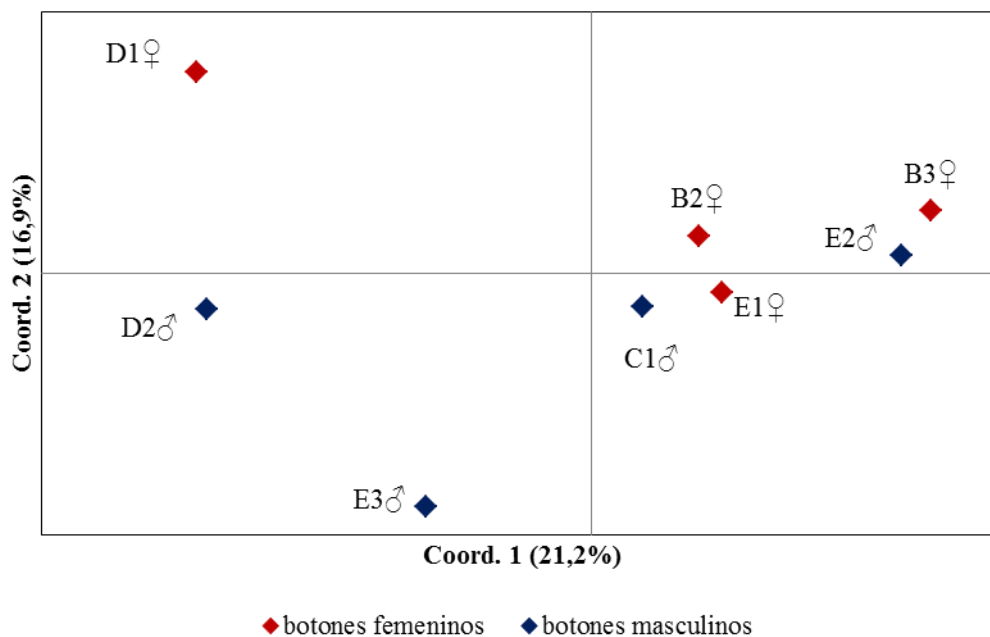


Figura AI-15. PCoA de la matriz YM-bot.

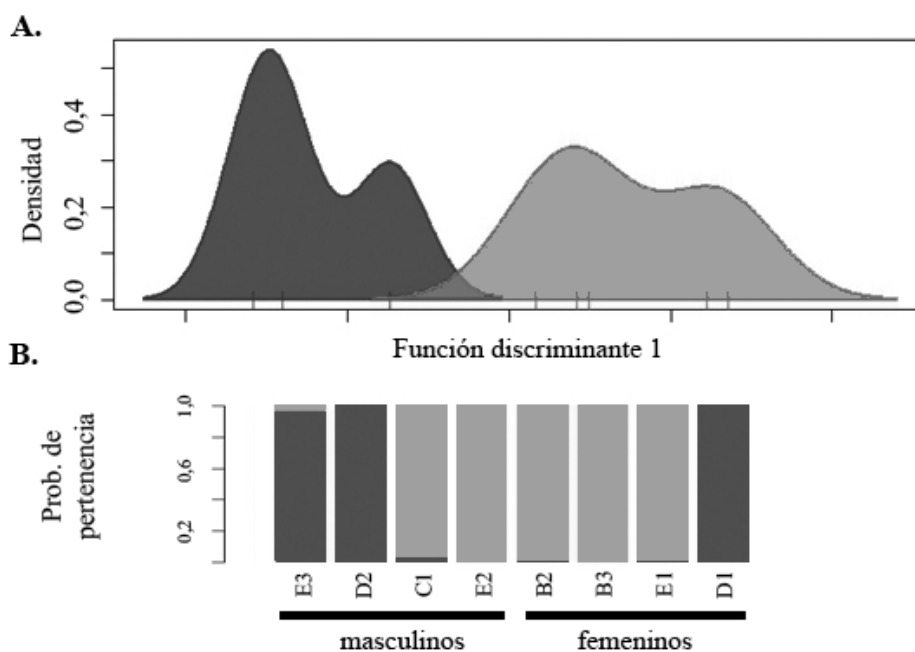
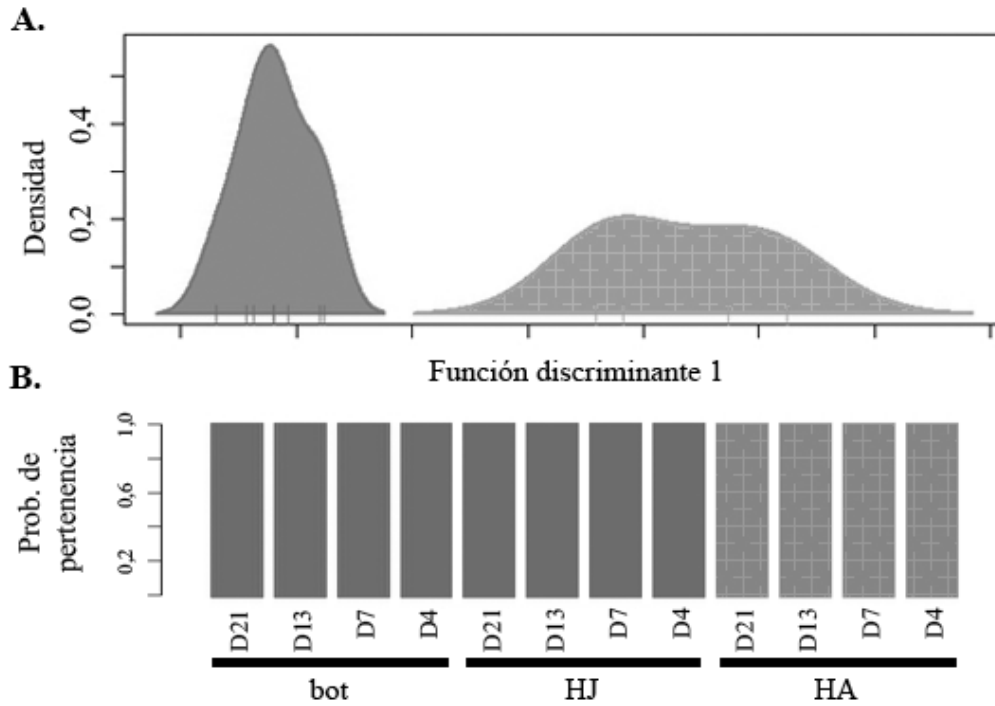
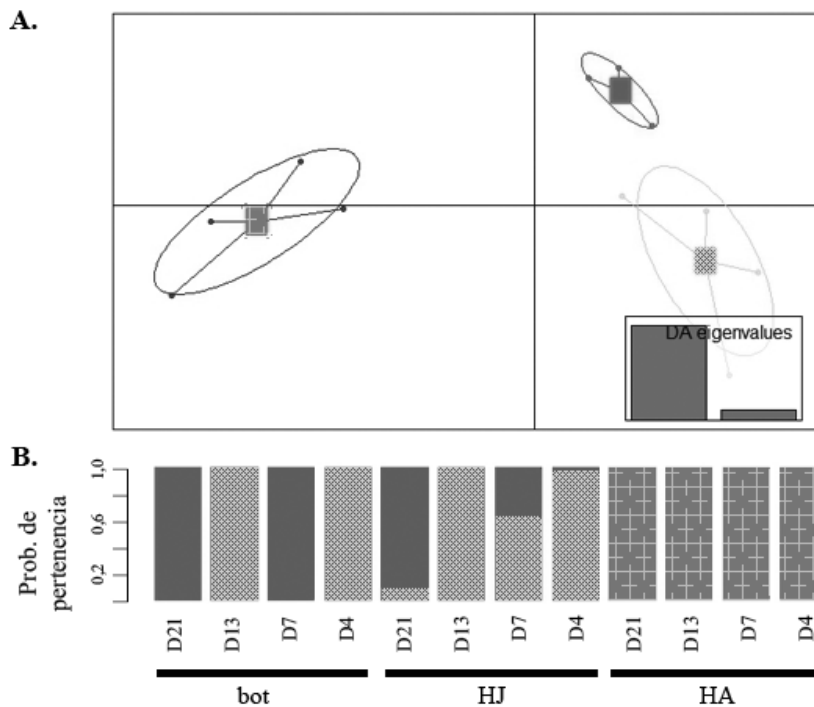


Figura AI-16. Análisis discriminante de la matriz YM-bot. **A.** Gráfico DAPC (20,89% de la varianza explicada). **B.** Gráfico compoplot. Las tonalidades de gris representan los mismos grupos entre gráficos.



**Figura AI-17.** Análisis discriminante de la matriz YS (K=2). **A.** Gráfico DAPC (15,5% de la varianza retenida). **B.** Gráfico compoplot. Los patrones se corresponden entre los distintos gráficos.



**Figura AI-18.** Análisis DAPC de yerba señorita con K=3. **A.** Gráfico DAPC (26,1% de la varianza retenida). **B.** Gráfico compoplot. Los patrones se corresponden entre los gráficos DAPC y su respectivo compoplot.

**Tabla AI-11.** Características de las bandas cDNA-AFLP seleccionadas para secuenciar.

Secuencia N°	Individuo – órgano <sup>a</sup>	Cebadores	Tamaño observado en los geles (pb) <sup>b</sup>
17	E3 bot	E-ACA + M-CAT	110
18	D13 bot	E-ACC + M-CTC	150
19	E3 bot	E-ACC + M-CTC	150
20*	E3 HA	E-ACG + M-CTT	195
21*	D1 bot	E-ACG + M-CTT	195
34	D1 bot	E-AAG + M-CAG	150
35	D4 bot	E-AAG + M-CAG	150
36	D1 bot	E-AAG + M-CAG	135
37	D4 bot	E-AAG + M-CAG	120
38	D21 HJ	E-AAG + M-CTG	175
39	D7 bot	E-AAG + M-CTG	120
40	D13 HA	E-ACA + M-CAT	240
41	D1 bot	E-ACA + M-CAT	180
42	D1 bot	E-ACA + M-CAT	135
43	E3 bot	E-ACA + M-CAT	120
44	D4 bot	E-ACC + M-CTC	205
45	D4 bot	E-ACC + M-CTC	190
46	D1 HA	E-ACG + M-CAA	160
47	B3 bot	E-ACG + M-CAA	90
48	D13 HA	E-ACT + M-CAC	160
49	B2 bot	E-ACT + M-CAC	125
50	C2 HJ	E-ACT + M-CAC	100
51	D2 HA	E-ACT + M-CAT	150
52	E3 HA	E-AGC + M-CTA	270
53	D4 bot	E-AGG + M-CAA	110
54	D7 bot	E-AAG + M-CTG	210
55	D21 HA	E-AAG + M-CTG	190
56	D1 HJ	E-ACC + M-CTC	300
57	D4 bot	E-ACC + M-CTC	295
58	D21 HJ	E-ACC + M-CTC	265
59	D4 HA	E-ACG + M-CTT	150
60	D7 bot	E-ACT + M-CAT	250
61	D7 bot	E-ACT + M-CAT	210
62	D7 HA	E-AGC + M-CAG	190
63	D13 HA	E-AGC + M-CAG	190
64	C2 HJ	E-AGG + M-CAA	170
65	D4 HA	E-AAC + M-CAC	300
66	E1 bot	E-AAC + M-CAC	185
67	B2 bot	E-AAG + M-CAA	400
68	D7 HA	E-AAG + M-CAA	300
69	D7 HJ	E-AAG + M-CAA	170



Secuencia N°	Individuo – órgano <sup>a</sup>	Cebadores	Tamaño observado en los geles (pb) <sup>b</sup>
70	D13 HJ	E-AAG + M-CAA	150
71	C2 HJ	E-AAG + M-CAA	145
72	D1 bot	E-AAG + M-CAA	145
73	D21 HA	E-AAG + M-CTA	300
74	D4 bot	E-AAG + M-CTA	300
75	B3 bot	E-AAG + M-CTA	245
76	B2 bot	E-AAG + M-CTA	245
77	D7 bot	E-AAG + M-CTA	245
78	D13 HA	E-AAG + M-CTA	215
79	E2 bot	E-AAG + M-CTA	215
80	E2 bot	E-AAG + M-CTA	215
81	E3 bot	E-AAG + M-CTA	205
82	E3 bot	E-AAG + M-CTA	190
83	D4 bot	E-AAG + M-CTA	170
84	E3 HA	E-AGG + M-CTT	285
85	D4 bot	E-AGG + M-CTT	270
86	D1 HJ	E-AGG + M-CTT	240
87	C1 bot	E-AGG + M-CTT	230
88	D7 HA	E-AGG + M-CTT	210
89	E2 bot	E-AGG + M-CTT	200

<sup>a</sup> datos de la muestra de la que se cortó la banda para secuenciar: nombre del individuo y órgano muestreado.

<sup>b</sup> tamaño aproximado de cada banda en pares de bases.

\* en un re-análisis se consideraron pertenecientes al mismo *locus*.

bot, botones florales. HJ, hoja joven. HA, hoja adulta.

**Apéndice II – Soluciones y protocolos****Soluciones utilizadas****Buffer TAE 50X**

Tris	60,5 g
EDTA 0,5 M pH 8	25 ml
Ácido acético glacial	14,3 ml
H <sub>2</sub> O bidestilada	hasta 250 ml

*Buffer TAE 1X:* Tris 0,04 M; ácido acético glacial 0,114% (v/v); EDTA 1 mM pH 8.

**Buffer TBE 5X**

Tris	54 g
Ácido bórico	27,5 g
EDTA 0,5 M pH 8	20 ml
H <sub>2</sub> O bidestilada	hasta 1000 ml

*Buffer TBE 1X:* Tris 0,09 M; ácido bórico 0,09 M; EDTA 2 mM pH 8.

**Buffer TE 10mM pH 8**

Tris HCl 1M	1 ml
EDTA 0,5 M	0,2 ml
H <sub>2</sub> O bidestilada	Hasta 100 ml

**Buffer de siembra en geles de agarosa**

Glicerol	5 ml
<i>Buffer TAE 50X</i>	200 µl
Azul de bromofenol, solución saturada	1 ml
Xilene cianol, solución al 10%	1 ml
H <sub>2</sub> O bidestilada	2,8 ml

**Buffer de siembra en geles de poliacrilamida**

Formamida	950 µl
Sc acuosa saturada de Azul de bromofenol (50mg/ml)	20 µl
Sc acuosa saturada de Xilene cianol (50mg/ml)	20 µl
EDTA 0,5 M pH 8	20 µl

**Solución stock acrilamida-bisacrilamida (29:1)**

Acrilamida	145 g
Bis-acrilamida	5 g
H <sub>2</sub> O bidestilada	hasta 500 ml
Filtrar y conservar a 4 °C.	

Composición de los geles de poliacrilamida desnaturalizantes 5% (p/v), urea 8M

Buffer TBE 5X	12 ml
Solución acrilamida-bisacrilamida al 30%	10 ml
Urea	28,8 g
TEMED (diamina de N,N,N,N'-tetrametiletileno)	60 µl
Persulfato de amonio 25% (p/v)	36 µl
H2O bidestilada	hasta 60 ml

Composición de los geles de poliacrilamida desnaturalizantes 6% (p/v), urea 8M

Buffer TBE 5X	12 ml
Solución acrilamida-bisacrilamida al 30%	12 ml
Urea	28,8 g
TEMED	60 µl
Persulfato de amonio 25% (p/v)	36 µl
H2O bidestilada	hasta 60 ml

Buffer C

NaCl	0,87 g
Tris-HCl 1M pH 8	15 ml
EDTA 0,5 M pH 8	15 ml
H2O bidestilada	hasta 150 ml

ProtocolosExtracción de cloroplastos

- ✓ Se utilizaron los equipos y *buffers* pre-enfriados: los tubos Falcon y los morteros se enfriaron a -20°C, y las centrifugaciones se realizaron a 2-4°C, con rotores pre-enfriados.
- ✓ Para obtener un mayor rendimiento, se mantuvieron las muestras en oscuridad completa antes y durante la preparación. Los tubos se cubrieron con papel aluminio para evitar la exposición a la luz.

Día 1

1. Preparar *buffer* CIB 1X usando agua miliQ estéril. Mezclar bien. Por cada 10 g de hojas, preparar 70 ml de CIB 1X.
2. Preparar solución CIB 1X + BSA (70 ml por muestra, mezclando 68,6 ml de *buffer* CIB 1X con 1,4 ml BSA 50mg/ml). Separar 10 ml, y al resto agregarle 0,06 g de PVP-40 (0,1% p/v) (solución "CIB 1X + BSA + PVP")
3. Lavar 10 g de hojas exhaustivamente con agua destilada. Remover el exceso de agua con papel. Quitar la nervadura central y cortar las hojas en pequeños pedazos (1-3 cm) usando una tijera o *cutter* limpio.
4. Agregar 60 ml de CIB 1X + BSA + PVP (0,1% p/v) (6 ml/g de hojas, proporción 6:1, recomendada para tabaco por el fabricante) y procesar las hojas con 2 a 4 golpes de licuadora (5 segundos) para lograr un grueso macerado con un mínimo de espuma. Alternativamente: procesar el material en seco con picadora o mortero enfriado en hielo, y luego agregar CIB 1X + BSA + PVP.

5. Filtrar el macerado hacia 2 tubos de 50 ml comunes sumergidos en hielo (aprox. 30 ml c/u). Dividir el filtrado entre 4 tubos de 50 ml tal que el volumen no exceda los 2/3 del tubo (aprox. 35 ml).
6. Centrifugar por 3 min a 200 x g. Precipita un *pellet* blanco (paredes y células enteras). Si no hay precipitado, aumentar velocidad hasta 400 x g.
7. Transferir el sobrenadante a tubos de 50 ml limpios y fríos. Centrifugar por 10 min a 2500 x g para sedimentar los cloroplastos en un *pellet* verde. Pasar el sobrenadante a nuevo tubo, y volver a centrifugar 10 min a 3000 x g.
8. Descartar el sobrenadante y romper suavemente el *pellet* de los 4 tubos golpeando con el dedo. Resuspender cada uno en 1 ml de CIB 1X + BSA mezclando suavemente por pipeteo. Evitar la formación de espuma. Juntar los *pellets* resuspendidos en un único tubo.
9. Purificación de cloroplastos intactos: centrifugar los 4 ml de suspensión de cloroplastos encima de un gradiente Percoll 40/80%.

**Preparación del gradiente Percoll 40/80%** (para 4 ml de suspensión de cloroplastos):

1. Preparar 2,5 ml de Percoll 80% en un tubo Falcon de 14 ml, mezclando 2 ml de Percoll con 500 µl de CIB 1X + BSA.
2. Preparar 5 ml de Percoll 40%, mezclando 2 ml de Percoll con 3 ml de CIB 1X + BSA. Colocar cuidadosamente la capa de Percoll 40% sobre la de 80%.
3. Aplicar cuidadosamente la suspensión de cloroplastos encima del gradiente Percoll y centrifugar sin freno por 15 min a 3200 x g. Los cloroplastos rotos forman una banda superior, y los intactos forman una banda en la interfase entre las capas de Percoll.
4. Recoger la banda de la interfase con una pipeta Pasteur de vidrio, colocarla en un tubo de 14 ml, suspender en 3 vol de CIB 1X sin BSA, y centrifugar a 1700 x g por 1 min. Mantener frío.
5. Resuspender el *pellet* de cloroplastos en 500 µl de CIB 1X sin BSA. La suspensión debe mantenerse en oscuridad, en hielo, hasta su uso.

**Extracción de ADN** (Adaptado de Diekmann *et al.* 2008 y Shi *et al.* 2012)

1. Agregar 2 ml de *buffer* C, 375 µl de SDS 20%, 5 µl de β-Mercaptoetanol y 7,5 µl de proteinasa K (10 mg/ml) al *pellet* de cloroplastos purificados, e incubar en baño térmico a 55°C con agitación suave ON (o al menos 4 hs).

---

*Día 2*

2. Poner los tubos de centrifugación en hielo por 5 min, agregar 450 µl de KAc 5 M (pH=5,2) y llevar a freezer (-20°C) por 30 min.
  3. Dividir en tubos de 2 ml para centrifugar 15 min a 8000 rpm, y transferir el sobrenadante a un tubo limpio.
  4. Agregar 1 vol de fenol (pH=8), mezclar bien, y centrifugar 20 min a 10000 rpm. Transferir la fase acuosa (superior) con cuidado a un nuevo tubo, con tip de punta cortada.
  5. Agregar un volumen de cloroformo:isoamílico (24:1) (remueve el fenol), centrifugar 20 min a 10000 rpm, transferir la fase superior a un nuevo tubo.
  6. Repetir el paso 5.
-

7. Precipitar el ADN con 2 vol de etanol 100% frío, llevar a freezer (-20°C) ON, o 2 hs a -80°C.
- 

## Día 3

8. Centrifugar a 8000 rpm por 20 min.
  9. Lavar el pellet 2 veces con etanol 70%, centrifugar 5 min a 13000 rpm, secar, y resuspender en un total de 200 µl de Tris-HCl 10 mM (dividir ese volumen por la cantidad de tubos, luego juntar en un único tubo).
  10. Tratar con 1 µl de RNAsa 20 mg/ml por cada 250 µl (200 µl → 0,8 µl), incubar 1 hora en baño a 37°C.
  11. Llevar a 500 µl con Tris-HCl 10mM para eliminar la RNAsa con 1 volumen de cloroformo:isoamílico 24:1. Centrifugar 5 min a máx velocidad, pasar el sobrenadante a nuevo tubo y precipitar con 2 vol de EtOH 100%. Llevar a freezer (-20°C) ON.
- 

## Día 4

12. Centrifugar 10 min a máx velocidad. Descartar sobrenadante.
  13. Resuspender en 55 µl de Tris-HCl 10 mM.
  14. Cuantificar.
- 

Protocolo MSAP**A. Restricción con *EcoRI***

1. En un tubo de 0,5 ml común pipetear 500 ng de ADN en volumen máximo = 22,8 µl.

**Por tubo**

<b>Mix</b>	<i>Buffer</i> REact3 10X	3 µl
	<i>EcoRI</i> 10 u/µl	0,5 µl (5 unidades/tubo)
	H <sub>2</sub> O dd estéril	3,7 µl

**V<sub>f</sub> = 30 µl**

- ✓ Vortexear la mix y dar un spin
- ✓ Repartir **7,2 µl por tubo** y vortexear.
  - a) Incubar a 37°C por 2 hs.
  - b) Inactivar a 80°C por 15 min.
  - c) Colocar en hielo y dar spin, luego al hielo nuevamente.
  - d) Chequear la digestión: sembrar 5 µl en gel de agarosa 1,5%.

**B. Limpieza antes de la segunda digestión:**

- a) Agregar 175 µl de Tris-HCl 10 mM y 200 µl de cloroformo:isoamílico 24:1.
- b) Centrifugar 15 min a máxima velocidad.
- c) Pasar fase superior a nuevo tubo de 1,5 ml.
- d) Precipitar con 2 vol de EtOH 100% frío, ON o 2 hs a -80°C.
- e) Centrifugar 15 min a máxima velocidad y secar el pellet.
- f) Resuspender en 50 µl de Tris-HCl 10 mM, incubar 1 h a 60°C.
- g) Dividir el volumen en 2 tubos de 0,5 ml.

**C. Restricción con *MspI* o *HpaII***

En cada tubo: 25 µl de ADN digerido con *EcoRI*.

**Por tubo**

<b>Mix M</b>	<i>Buffer Cut Smart</i> 10X	3 µl
	<i>MspI</i> 20 u/µl	0,125 µl (2,5 unidades/tubo)
	H <sub>2</sub> O estéril	1,875 µl

**V<sub>f</sub> = 30 µl**

**Por tubo**

<b>Mix H</b>	<i>Buffer Cut Smart</i> 10X	3 µl
	<i>HpaII</i> 10 u/µl	0,25 µl (2,5 unidades/tubo)
	H <sub>2</sub> O estéril	1,75 µl

**V<sub>f</sub> = 30 µl**

- ✓ Vortexear la mix y dar un spin.
- ✓ Repartir 5 µl por tubo y vortexear.
- a) Incubar a 37°C por 1 h 30 min.
- b) Inactivar a 80°C por 20 min.

**D. Ligado**

Dividir en dos tubos los ADN digeridos, almacenar uno a -20°C. Al otro (15 µl de ADN digerido) agregarle la mix de ligado.

**Por tubo**

<b>Mix</b>	<i>Buffer</i> ligasa 5X	5 µl
	Adaptador <i>EcoRI</i> (5 pmol/ µl)	0,5 µl
	Adaptador <i>MspI/HpaII</i> (50 pmol/ µl)	0,5 µl
	H <sub>2</sub> O	3,5 µl
	T4 ligasa (1u/µl)	0,5 µl (0,5 unidad)

**V<sub>f</sub> = 25 µl**

- ✓ Vortexear la mix y dar un spin.
- ✓ Repartir 10 µl a cada tubo y vortexear.
- c) Incubar 2 hs a 37°C.
- d) Incubar a 4°C ON.

**E. Preamplificación**

- ✓ Diluir 1/100. La porción no utilizada de la mezcla puede guardarse a -20°C.

En un tubo de 0,5 ml de pared delgada, pipetear 5 µl de ADN ligado 1/100:

**Por tubo**

<b>Mix</b>	<i>Buffer</i> 10X	5 µl	
	MgCl <sub>2</sub> 50 mM	1,5 µl	[final]=1,5 mM
	dNTPs [1,25 mM c/u]	8 µl	[final]=0,2 mM
	Cebador <i>EcoRI</i> + A [30 ng/µl]	3 µl	final=90 ng
	Cebador <i>MspI/HpaII</i> + T/A [30 ng/µl]	3 µl	final=90 ng
	H <sub>2</sub> O	24,3 µl	
	Taq [5u/µl]	0,2 µl	

**V<sub>f</sub> = 50 µl**

- ✓ Vortexear la mix y dar un spin.
- ✓ Programa de PCR: 20 ciclos de 94°C por 30 seg, 56°C por 1 min, 72°C por 1 min

- a) Diluir 1/5: transferir 10 µl de la mezcla a un tubo de 0,5 ml, agregar 40 µl de Tris-HCl 10 mM y mezclar bien. Las mezclas diluidas y no diluidas pueden guardarse a -20°C.
- b) Chequear en gel de agarosa 1,5%.

**F. Amplificación selectiva**

Pipetear en cada tubo 5 µl de ADN preamplificado 1/5.

**Por tubo**

<b>Mix</b>	<i>Buffer</i> 10X	2 µl	
	MgCl <sub>2</sub> 50 mM	0,6 µl	[final]=1,5 mM
	dNTPs [1,25 mM c/u]	4 µl	[final]=0,25 mM
	Cebador <i>EcoRI</i> + 3 [30 ng/µl]	1,2 µl	final=36 ng, 1,8 ng/µl
	Cebador <i>MseI</i> + 3 [30 ng/µl]	1,2 µl	final=36 ng, 1,8 ng/µl
	H <sub>2</sub> O	5,85 µl	
	Taq [5u/µl]	0,15 µl	

**V<sub>f</sub>= 20 µl**

- ✓ Vortexear la mix y dar un spin, repartir 15 µl por tubo.
- ✓ Programa de PCR:
  - 13 ciclos de:
    - 94°C, 30 seg
    - 65°C, 30 seg (con *touchdown* de 0,7°C)
    - 72°C, 30 seg
  - 23 ciclos de:
    - 94°C, 30 seg
    - 56°C, 30 seg
    - 72°C, 1 min
  - Extensión final: 72°C, 10 min

Protocolo cDNA-AFLP

**A. Restricción**

Digestión con *EcoRI*-HF y *MseI*, en un tubo de 0,5 ml común, conteniendo 300 ng de ADNc en un volumen máximo de 26,55 µl.

**Por tubo**

<b>Mix</b>	<i>Buffer CutSmart</i> 10X	3 µl
	<i>EcoRI</i> -HF 20 u/µl	0,15 µl (3 unidades/tubo)
	<i>MseI</i> 10 u/µl	0,3 µl (3 unidades/tubo)
	H <sub>2</sub> O	Hasta <b>30 µl</b>

- ✓ Vortexear la mix y dar un spin
- ✓ Repartir y vortexear cada tubo.
- e) Incubar a 37°C por 2 hs.
- f) Inactivar a 65°C por 20 min.
- g) Colocar en hielo y dar un spin, luego hielo nuevamente.

**B. Ligado**

Dividir en dos tubos los ADN digeridos, almacenar uno a -20°C. Al otro (15 µl de ADN digerido) agregarle la mix de ligado.

**Por tubo**

<b>Mix</b>	<i>Buffer</i> ligasa 5X	5 $\mu$ l
	Adaptador <i>EcoRI</i>	0,5 $\mu$ l (5 pmol)
	Adaptador <i>MseI</i>	0,5 $\mu$ l (50 pmol)
	H <sub>2</sub> O	3,5 $\mu$ l
	T4 ligasa [1u/ $\mu$ l]	0,5 $\mu$ l

**V<sub>f</sub> = 25  $\mu$ l**

- ✓ Vortexear la mix y dar un spin.
- ✓ Repartir 10  $\mu$ l a cada tubo y vortexear.
- e) Incubar 2 hs a temperatura ambiente.
- f) Incubar a 4°C ON.

**C. Preamplificación y amplificaciones selectivas**

- ✓ Proceder de la misma manera que para MSAP, utilizando cebadores *EcoRI* + A y *MseI* + C (preamplificación) y *EcoRI* + 3 en combinación con *MseI* + 3 (amplificaciones selectivas).