



**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO**

---

---

**INSTITUTO DE CIENCIAS AGROPECUARIAS**

**ESTUDIO MOLECULAR Y MORFOLÓGICO DE CONÍFERAS DEL  
ESTADO DE HIDALGO**

**T E S I S**

**Que para obtener el grado de**

**DOCTOR EN CIENCIAS AGROPECUARIAS**

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

**SERGIO HERNÁNDEZ LEÓN**

**DIRECTOR DE TESIS: DR. OTILIO A. ACEVEDO SANDOVAL**

**CODIRECTOR DE TESIS: DAMON P. LITTLE**

**Tulancingo de Bravo, Hidalgo. Noviembre 2018**



## COORDINACIÓN DE POSGRADO DEL ICAP

Actas de la reunión del Comité de Tesis de Doctorado en Ciencias Agropecuarias

### Apertura:

La reunión ordinaria para evaluar los avances de la tesis intitulada: **"Estudio molecular y morfológico de coníferas del estado de Hidalgo"**, que desarrolla el estudiante M. en C. Sergio Hernández León.

### Asistentes:

Dr. Otilio A. Acevedo Sandoval

Dra. Mariana Saucedo García

Dr. Rodrigo Rodríguez Laguna

Dr. Ramón Razo Zárate

### Revisión de Trabajo de Tesis

### Observaciones:

El comité revisó con antelación el trabajo de tesis en extenso propuesto por el estudiante, comunicando al estudiante M. en C. Sergio Hernández León, oportunamente las correcciones, adiciones y/o modificaciones que debería considerar para mejorar su trabajo y poder continuar con el proceso de obtención de grado. El estudiante atendió de forma conveniente las sugerencias del comité.

### Acuerdos

En esta fecha, se comunica atentamente que el comité conformado por los profesores firmantes, otorgamos nuestra autorización para que el estudiante imprima su trabajo final de tesis, y continúe con los trámites necesarios para la obtención del grado de doctorado respectivo.

ATENTAMENTE

"AMOR, ORDEN Y PROGRESO"

Tulancingo de Bravo, Hidalgo a 27 de noviembre del 2018.

Dr. Otilio A. Acevedo Sandoval

Dra. Mariana Saucedo García

Dr. Rodrigo Rodríguez Laguna

Dr. Ramón Razo Zárate

  
Mariana Saucedo García  
  
  
COORDINACIÓN DE INVESTIGACIÓN Y POSGRADO  
INSTITUTO DE CIENCIAS AGROPECUARIAS



## **Agradecimientos**

Agradezco al Doctorado en Ciencias Agropecuarias de la Universidad Autónoma del Estado de Hidalgo, por permitirme cursar mis estudios de Doctorado bajo un nivel académico de excelencia.

La realización de este trabajo fue posible gracias al apoyo financiero otorgado por el Consejo Nacional de Ciencia y Tecnología a través de la Red Temática Código de Barras de la Vida (Continuidad de Redes Temáticas) No. CONACyT: 271108.

A los integrantes de mi comité tutorial Dr. Otilio A. Acevedo Sandoval, PhD. Damon P. Little, Dra. Mariana Saucedo García, Dr. Rodrigo Rodríguez Laguna y Dr. Ramón Razo Zárate por su tiempo, apoyo y guía para concluir satisfactoriamente con mis estudios de doctorado.

A los integrantes de mi comité de evaluación Dr. Alfonso Suárez Islas, Dr. Oscar Arce Cervantes, Dr. Ramón Razo Zárate y Dr. Rodrigo Rodríguez Laguna por sus invaluable comentarios para mejorar versiones anteriores de este escrito.

A mi estudiante y amigo Cristhian Espinoza Pelcastre con quien tuve la oportunidad de desarrollar el tercer capítulo de esta investigación.

## **Agradecimientos a título personal**

A mis profesores Dr. Juan Capulín Grande, Dr. Alfredo Madariaga Navarrete, Dr. Rodrigo Rodríguez Laguna, Dr. Ramón Razo Zárate, Dr. Martín Amador Meza Nieto y Dra. Norma Güemes Vera.

A la Dra. Norma Güemes Vera y Dr. Javier Piloni Martini Por su apoyo en los procesos académicos y administrativos durante mi estancia en el Doctorado en Ciencias Agropecuarias.

A los integrantes del cuerpo académico de Agrobiotecnología: Dr. Benito Flores Chávez, Dra. Josefa Espitia López, Dra. Mariana Saucedo García, Dr. Oscar Arce Cervantes, Dr. Paul Misael Garza López por ser un gran equipo de apoyo personal y académico.

Amigos, alumnos, colaboradores y compañeros del Programa Educativo de Ing. Forestal. Al Dr. José González Ávalos y la M. en C. Elizabeth Pérez Soto eternamente agradecido por hacer posible mi incorporación a esta gran institución.

A los directivos del Instituto de Ciencias Agropecuarias: Director Dr. Miguel A. Miguez Escorcia, Secretario Académico M. en C. M. Jesús Franco Fernández por las facilidades otorgadas durante el tiempo de mis estudios de doctorado.

Al Dr. David S. Gernandt y familia siempre agradecido por su amistad y apoyo en el campo académico y de investigación.

Por su amistad y acompañamiento durante el trabajo de campo y laboratorio a Alejandro López Reyes, Cristhian Espinoza Pelcastre, Edith Jiménez Muñoz, Elena Rosales Islas, Jesús Antonio Bello, Laura Margarita Márquez Valdelamar, Lidia Irene Cabrera Martínez, Perla Barrón Hernández, Ramón Escarcega Vargas, Shamed Popoca Paredes y Patricia Rosas Escobar.

A esta gran institución y Alma mater que es la Universidad Autónoma del Estado de Hidalgo que me ha permitido continuar con mi desarrollo personal y profesional.

## **Dedicatoria**

**Este logro lo dedico a mi familia con quienes es un privilegio compartir cada día de esta vida, mi mamá Lucia León López, mis hermanos Julián, Verónica, Vicente, Yazmín y mis sobrinos Carlos, Erick, Fernanda, Mateo, Mayte y Renata, gracias por siempre poder contar con ustedes, por su amor y apoyo incondicional.**

## Índice general

Relación de cuadros .....	10
Relación de figuras .....	11
Resumen general.....	14
General abstract .....	15
Capítulo I.....	16
1.1 Introducción general.....	16
1.2 Objetivo general.....	20
1.2.1 Objetivos particulares .....	20
1.3 Revisión de la literatura.....	21
1.3.1 Características de las especies en categorías de extinción y categorías de riesgo .....	21
1.3.2 Los códigos de barra genéticos .....	23
1.3.3 Secuenciación Sanger .....	23
1.3.4 Taxonomía imperfecta .....	24
1.3.5 Hibridación.....	24
1.3.6 Polimorfismo ancestral compartido.....	25
1.3.7 Tasas bajas de mutación.....	25
1.3.8 Universalidad .....	26
1.3.9 Calidad de la secuencia.....	26
1.3.10 Complejidad lingüística de la secuencia .....	27
1.3.11 Discriminación de especies .....	28
1.3.12 Métodos basados en árboles, filogenéticos o de agrupamiento jerárquico .....	29
1.3.13 Métodos diagnósticos basados en caracteres .....	30
1.3.13 Métodos de similitud .....	30
1.3.14 Métodos híbridos .....	31
1.3.15 El gap del código de barras genético .....	31
1.3.16 Antecedentes de los códigos de barras genéticos en plantas.....	32
1.4 Literatura citada .....	38
Capítulo II .....	48
Desempeño de los códigos de barras genéticos centrales en plantas a escala local: Identificación de las coníferas del estado de Hidalgo, México .....	48
Resumen .....	48
Abstract .....	49

2.1	Introducción.....	51
2.2	Método.....	53
2.2.1	Extracción de DNA, PCR y alineamiento de secuencias.....	63
2.2.2	Análisis de datos.....	63
2.3	Resultados.....	67
2.3.1	Caracterización de las secuencias.....	67
2.3.2	Calidad de la secuencia.....	68
2.3.3	Complejidad lingüística.....	68
2.3.4	Gap del código de barras genético.....	69
2.3.5	Discriminación de especies basada en caracteres.....	70
2.3.6	Identificación de especies basada en árboles filogenéticos.....	73
2.4	Discusión.....	77
2.4.1	Caracterización de las secuencias.....	77
2.4.2	Calidad de la secuencia, complejidad lingüística y gap del código de barras genético.....	79
2.4.3	Identificación de especies basada en caracteres.....	81
2.4.4	Identificación de especies basada en árboles filogenéticos.....	82
2.5	Conclusiones.....	93
2.6	Recomendaciones.....	94
2.7	Literatura citada.....	95
	Capítulo III.....	107
	Clave de identificación taxonómica con caracteres anatómicos foliares para especies de <i>Pinus</i> L. en Hidalgo.....	107
	Resumen.....	107
	Abstract.....	108
3.1	Introducción.....	109
3.2	Materiales y métodos.....	110
3.2.1	Trabajo de campo.....	111
3.2.2	Cortes foliares y tinción.....	111
3.2.3	Codificación de caracteres.....	114
3.2.3	Análisis de datos.....	115
3.3	Resultados y discusión.....	115
3.3.1	Trabajo de campo.....	115
3.3.2	Codificación de caracteres.....	116

3.3.3 Análisis de agrupamiento.....	124
3.3.4 Clave de identificación .....	126
3.4 Conclusión.....	126
3.5 Literatura citada .....	127
Capítulo IV .....	131
Conclusiones generales y perspectivas .....	131

## **Relación de cuadros**

**Cuadro 2.1.** Coníferas nativas del estado de Hidalgo, México. Esta lista de especies actualizada es resultado de trabajo de campo y reporte previos.

**Cuadro 2.2.** Familias, géneros y especies de coníferas del estado de Hidalgo analizados en este estudio. Para las coordenadas geográficas se utilizó el Sistema de Coordenadas Geográficas WGS84. Los especímenes se depositaron en el Herbario Nacional de México (MEXU).

**Cuadro 2.3.** Características de las secuencias de los códigos de barras genéticos *matK* y *rbcL*.

**Cuadro 2.4.** Variación diagnóstica de los códigos de barras *matK* y *rbcL* analizados con BRONx.

**Cuadro 3.1.** Datos detallados de colecta de especies de *Pinus* del estado de Hidalgo.

**Cuadro 3.2.** Caracteres y estados de carácter utilizados para la descripción anatómica foliar de las especies de *Pinus* del estado de Hidalgo.

**Cuadro 3.3.** Matriz de datos morfoanatómicos de acículas para las especies de *Pinus* naturalmente distribuidas en el estado de Hidalgo.

## Relación de figuras

**Figura 2.1.** Localidades muestreadas. Distribución geográfica de las especies de coníferas muestreadas (círculos azules) en diversas localidades del estado de Hidalgo, México. Se utilizó el Sistema de Coordenadas Geográficas WGS84.

**Figura 2.2.** Calidad de la secuencia ( $B_{30}$ ) y Complejidad lingüística (LC) para especímenes con secuencias completas (*matK* y *rbcL*). Cada círculo representa una secuencia individual de *matK* (m; negro) y *rbcL* (r; gris). Los recuadros negros representan el promedio por marcador. Las barras de error representan tres desviaciones estándar.

**Figura 2.3.** Comparación de la Máxima distancia intraespecífica y la Mínima distancia interespecífica. Los círculos opacos representan secuencias diagnósticas para *matK* (negros) y *rbcL* (gris). Los círculos semitransparentes representan secuencias no diagnósticas. La línea gris representa igual variación intra e interespecífica, por encima de la cual se localizan las secuencias correspondientes con especies que presentan el gap del código de barras genético.

**Figura 2.4.** Proporción de especies que pueden ser distinguidas y número de especies combinadas. Para *matK* (m; negro), *rbcL* (r; gris claro) y la combinación de marcadores (mr; gris). Los cuadrados en el centro indican promedios de discriminación de especies. Las barras de error indican intervalos de confianza al 95%.

**Figura 2.5.** Relaciones filogenéticas resultantes del análisis de la región *matK*. Consenso estricto de 13 árboles más parsimoniosos (L=475; IC= 0.845; IR=0.984). Los valores de soporte Jackknife mayores al 50%, se muestran próximos a los nodos. Las especies que pueden ser identificadas con el criterio del clado menos inclusivo se muestran en negritas.

**Figura 2.6.** Relaciones filogenéticas resultantes del análisis de la región *rbcL*. Consenso estricto de 3 árboles más parsimoniosos (L=133; IC= 0.797; IR=0.978). Los valores de soporte Jackknife mayores al 50%, se muestran próximos a los nodos. Las especies que pueden ser identificadas con el criterio del clado menos inclusivo se muestran en negritas.

**Figura 2.7.** Relaciones filogenéticas resultantes del análisis de la región (*matK + rbcL*). Consenso estricto de 14 árboles más parsimoniosos (L = 610; IC = 0.826; IR = 0.979). Los valores de soporte Jackknife mayores al 50%, se muestran próximos a los nodos. Las especies que pueden ser identificadas con el criterio del clado menos inclusivo se muestran en negritas.

**Figura 3.1.** Secciones transversales de acículas de especies de *Pinus* del estado de Hidalgo. **A** *Pinus ayacahuite*. **B** *P. cembroides*. **C** *P. devoniana*. **D** *P. greggii*. **E** *P. hartwegii*. **F** *P. leiophylla*. **G** *P. montezumae*. **H** *P. oocarpa*. **I** *P. patula*. **J** *P. pinceana*. **K** *P. pseudostrobus*. **L** *P. teocote*. \* indica la posición de los estomas. Abreviaciones: **Cr**, Canal resinífero; **End**, Endodermis; **Epi**, Epidermis; **Escl**, Esclerenquima; **Flo**, Floema; **Hip**, Hipodermis; **Mes**, Mesófilo; **Tt**, Tejido de transfusión; **Xil**, Xilema.

**Figura 3.2.** Secciones transversales de acículas de especies de *Pinus* del estado de Hidalgo. **A** *Pinus devoniana*. **B, E, I** *P. teocote*. **C-D** *P. greggii*. **F, H** *P. cembroides*. **G** *P. oocarpa*. Abreviaciones de los estados de caracteres morfoanatómicos: **Cre**, Canal resinífero externo; **Cri**, Canal resinífero interno; **Crm**, Canal resinífero medio; **Crs**, Canal resinífero septal; **End**, Endodermis; **Epi**, Epidermis; **Flo**, Floema; **Hip**, Hipodermis; **Hv**, Haz vascular; **Mes**, Mesófilo; **Tt**, Tejido de transfusión; **Xil**, Xilema.

**Figura 3.3.** Dendrograma obtenido por análisis de agrupamiento UPGMA, con el uso del coeficiente de similitud de Jaccard en 12 especies de *Pinus* distribuidas en el estado

de Hidalgo. Coeficiente de correlación cofenético = 0.847018494508054. El número en cada terminal corresponde con el número del ejemplar por especie.

## **Resumen general**

En este estudio se analizan las coníferas con distribución natural en el estado de Hidalgo, México, mediante técnicas moleculares y morfológicas con la finalidad de contribuir a su conocimiento, manejo, aprovechamiento y conservación. El código de barras genético constituye una herramienta fundamental para la identificación de especies. En la primera parte del estudio se caracterizaron y evaluaron las regiones centrales del código de barras genético en plantas, *matK* y *rbcL*, para evaluar su utilidad en la identificación de las 25 especies de coníferas en el estado; incluyendo diez bajo diversas categorías de riesgo. Para las regiones del código de barras genético se obtuvo la calidad de las secuencias, su complejidad lingüística y la presencia del gap del código de barras genético. Se compararon dos métodos para la identificación de especies. En la segunda parte del estudio se estableció una clave dicotómica basada en la morfoanatomía foliar para la identificación de las 12 especies de pinos distribuidas de manera natural en el estado de Hidalgo. Se obtuvieron cortes transversales de las acículas, a partir de los cuales se codificó la matriz de caracteres, la cual fue sometida a un análisis de similitud UPGMA, se elaboró una clave de identificación taxonómica, con la cual se logran diferenciar 10 de las 12 especies, incluyendo *Pinus greggii* una especie bajo categoría de riesgo. El análisis UPGMA resultó en dos agrupaciones principales, en la primera se recuperaron dos subgrupos correspondientes con las especies del subgénero *Strobus* y del subgénero *Pinus*, en la segunda se recuperaron únicamente especies del subgénero *Pinus*. Se concluye que la clave de identificación elaborada es robusta para la identificación de especies de pinos a escala local utilizando exclusivamente caracteres anatómicos foliares.

**Palabras clave:** Acículas, anatomía foliar, clave dicotómica, código de barras genético, IUCN, *matK*, *rbcL*.

## General abstract

Forests cover more than 31% of the land surface on earth and provide many goods and services worldwide. In this study we analyzed conifers naturally distributed in the state of Hidalgo, through morphological and molecular techniques to contribute to their knowledge, management and conservation. DNA barcoding constitutes a fundamental tool for species identification. In the first part of this study we characterize and evaluate the plant core barcoding regions, *matK* and *rbcL*, to identify the 25 conifer species from the state of Hidalgo, Mexico, including ten species under various threat categories. Sequence quality, linguistic complexity and the presence of the barcode gap was estimated. Two methods were compared for successful species identification. In the second part of this study we established an identification key based on leaf anatomy for the identification of 12 pine species occurring in the state of Hidalgo. Permanent slides were prepared from transverse leaf sections from which a character matrix was codified from leaf anatomy features and analyzed with the UPGMA method. The matrix was used to construct an identification key. The key allows for the identification of 10 out of the 12 species, including *Pinus greggii* a species under risk. UPGMA analysis generated two principal clusters, the first group included two interior clusters corresponding with subgenus *Strobus* species and subgenus *Pinus* species, the second principal cluster included species from subgenus *Pinus* only. This study contributes with a robust identification key for pine species at a local scale using exclusively secondary leaf anatomy characters.

**Key words:** Dichotomous key, DNA barcoding, IUCN, leaf anatomy, needles, *matK*, *rbcL*.

# Capítulo I

## 1.1 Introducción general

Los bosques cubren más del 31% de la superficie terrestre y ayudan a preservar la biodiversidad global, son la base de la industria forestal, proveen diversas fuentes de energía, agua, comida y recreación además proveen bienes y servicios diversos a nivel mundial (Keenan et al., 2015; Secretariat of the Convention on Biological Diversity, 2009). Los bosques están formados por diversas especies arbóreas incluidas las coníferas que además de ser ornamentales y excelentes para forestación, son una de las mayores fuentes de madera a nivel mundial, altamente valorada para la fabricación de muebles, pulpa para papel y otros diversos productos. *Abies*, *Pinus* y *Pseudotsuga* se encuentran entre los géneros más aprovechados en la industria forestal a nivel global (Armenise, Simeone, Piredda, y Schirone, 2012).

Las coníferas (Pinophyta) están clasificadas en seis familias (Araucariaceae, Cupressaceae, Pinaceae, Podocarpaceae, Sciadopityaceae y Taxaceae), 70 géneros y aproximadamente 670 especies que constituyen el 0.3% de la diversidad de las plantas terrestres (Eckenwalder, 2009; Gernandt y Vázquez-Lobo, 2014; Gernandt, Willyard, Syring, y Liston, 2011; Gernandt, Geada López, Ortiz García, y Liston, 2005).

Las coníferas (Pinophyta) incluyen árboles o arbustos leñosos con hojas simples, conos polínicos simples y conos ovulados compuestos o reducidos (Christenhusz et al., 2011). Generalmente exhiben una alta diversidad genética (niveles altos de heterocigocidad y diversidad intraespecífica) promovida por grandes tamaños poblacionales, sistemas reproductivos entrecruzados, largas esperanzas de vida, dispersión a grandes distancias de polen y algunas veces semillas que permiten un flujo genético eficiente, sobre todo mediante la polinización por viento la cual mantiene la

similitud alélica a lo largo de grandes distancias geográficas y bajos niveles de diferenciación genética entre poblaciones, altas tasas de mutación y promedio bajo de desequilibrio de ligamento (Gernandt et al., 2011; Rémy J Petit y Vendramin, 2007; Rémy J. Petit et al., 2005). En conjunto, estos rasgos provocan en gran medida una evolución independiente de los genes y al mismo tiempo favorecen la selección natural eficiente (en comparación con la deriva genética) en especies con grandes tamaños poblacionales conduciendo al establecimiento de clinas de genotipos localmente adaptados (Prunier, Verta, y Mackay, 2016).

La distribución de las coníferas en el mundo ocurre en casi todos los principales tipos de vegetación con la excepción del bosque tropical de tierras bajas, el salar, pantano de sal y vegetación acuática y subacuática (Farjon, 2008). En México, las coníferas son dominantes en bosques de *Abies*, *Callitropsis* (*Cupressus*), *Picea*, *Pinus* y *Pseudotsuga*, matorrales o bosques de *Juniperus* y *Pinus* además de estar presentes en otros tipos de vegetación tales como matorrales xerófilos, bosques mesófilos de montaña, bosques de pino-encino y bosque de galería de *Taxodium huegelii* (Rzedowski, 1978).

Las coníferas son particularmente diversas en el hemisferio norte donde se encuentra cerca del 70% de las especies, son cosmopolitas, incluyen especies ampliamente distribuidas así como endémicas de distribución restringida (Farjon, 2010), México tiene gran riqueza de coníferas, con aproximadamente 94 especies (14%) de las reconocidas a nivel mundial. Los estados de Hidalgo, Puebla, Oaxaca y Veracruz cuentan con la mayor diversidad de géneros de coníferas en México (Gernandt y Pérez-De La Rosa, 2014).

El estado de Hidalgo se localiza entre los paralelos 19° 36'; 21° 24' latitud norte y los meridianos 97° 58'; 99° 53' de longitud oeste. Con aproximadamente 20, 905 km<sup>2</sup>,

representa el 1.1% del territorio mexicano (INEGI, 2014). El estado de Hidalgo se encuentra dentro de la Zona de Transición Mexicana, su territorio incluye cuatro provincias biogeográficas: Altiplano Mexicano, la costa del Golfo de México, la Sierra Madre Oriental y el Eje Volcánico Transmexicano (Morrone, 2005; Rzedowski, 1978). La Sierra Madre Oriental junto con la Sierra Madre Occidental concentran la mayor diversidad de coníferas en México (Gernandt y Pérez-de La Rosa, 2014).

En el estado de Hidalgo se distribuyen de manera natural cuatro familias de coníferas, (Cupressaceae, Pinaceae, Podocarpaceae y Taxaceae); ocho géneros (*Abies*, *Callitropsis*, *Juniperus*, *Pinus*, *Podocarpus*, *Pseudotsuga*, *Taxodium* y *Taxus*), así como 25 especies. La mayoría de estas especies tiene afinidad florística septentrional excepto *Podocarpus*, que tiene una afinidad florística meridional (Halffter, Llorente-Bousquets, y Morrone, 2008). La mayoría de los recursos forestales maderables del estado de Hidalgo son obtenidos a partir de especies de coníferas, particularmente del género *Pinus* (SEMARNAT, 2010).

Diez especies de coníferas en el estado aparecen en listados de especies en categorías de riesgo por lo que su conservación debe ser una prioridad. Siete especies aparecen en el listado del gobierno mexicano de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT, 2010), en la categoría de sujetas a protección especial (Pr; *Callitropsis lusitanica* (Mill.) D. P. Little, *Juniperus monticola* Martínez, *Podocarpus matudae* Lundell y *Pseudotsuga menziesii* var. *glauca* (Bleissn.) Franco) o como en riesgo de extinción (P; *Abies guatemalensis* Rehder, *Pinus pinceana* Gordon y *Taxus globosa* Schltdl.). Por otro lado, cinco especies se encuentran en el listado rojo de la Unión Internacional para la Conservación de la Naturaleza (IUCN) en la categoría de Vulnerable (*Abies hidalgensis*, *Juniperus angosturana* y *Pinus greggii*) o como En peligro de extinción (*A. guatemalensis* y *T. globosa*; IUCN, 2018). Ocho de las especies

que se distribuyen de manera natural en el estado de Hidalgo son endémicas de México: *Abies hidalgensis*, *A. religiosa*, *Callitropsis benthamii*, *Juniperus angosturana*, *Pinus greggii*, *P. patula*, *P. pinceana*, y *P. teocote*. La primera es una especie poco conocida que posiblemente es endémica del estado de Hidalgo (IUCN, 2018).

A pesar de la importancia del grupo hasta el momento no se cuenta con un listado actualizado de especies de coníferas para el estado de Hidalgo, por lo que es de suma importancia una revisión taxonómica a nivel de estado, basada en la revisión de la literatura, ejemplares recolectados en campo y determinados mediante claves de identificación taxonómica. Aunado a lo anterior la implementación de herramientas alternativas de identificación tales como los códigos de barras genéticos o las claves de identificación basadas en caracteres foliares morfoanatómicos es crucial para asegurar la conservación, el aprovechamiento y el manejo sostenible de este importante recurso forestal.

## **1.2 Objetivo general**

Estudiar las especies de coníferas con distribución natural en el estado de Hidalgo mediante análisis moleculares y morfológicos con la finalidad de contribuir al conocimiento, manejo, aprovechamiento y conservación de este grupo de suma importancia como recurso natural forestal.

### **1.2.1 Objetivos particulares**

- Establecer un listado actualizado de las especies de coníferas en el estado de Hidalgo mediante su estudio morfológico.
- Evaluar los códigos de barras genéticos (*matK* y *rbcL*) para la identificación molecular de especies a nivel regional, comparando dos métodos.
- Establecer una clave dicotómica para la identificación de los pinos con distribución natural en el estado de Hidalgo, con base en la anatomía de las acículas.

## 1.3 Revisión de la literatura

### 1.3.1 Características de las especies en categorías de extinción y categorías de riesgo

En el listado de la NOM-059-SEMARNAT- 2010 *Callitropsis lusitanica*, *Juniperus monticola*, *Podocarpus matudae*, y *Pseudotsuga menziesii* var. *glauca* se encuentran en la categoría de sujetas a protección especial (Pr) la cual incluye a "aquellas especies que podrían llegar a encontrarse amenazadas por factores que inciden negativamente en su viabilidad, por lo que se determina la necesidad de propiciar su recuperación y conservación o la recuperación y conservación de poblaciones de especies asociadas". Por otro lado *Abies guatemalensis*, *Pinus pinceana* y *Taxus globosa*, se encuentran en la categoría de en riesgo de extinción (P) "aquellas especies cuyas áreas de distribución o tamaño de sus poblaciones en el Territorio Nacional han disminuido drásticamente poniendo en riesgo su viabilidad biológica en todo su hábitat natural, debido a factores tales como la destrucción o modificación drástica del hábitat, aprovechamiento no sustentable, enfermedades o depredación, entre otros" (SEMARNAT, 2010).

Las categorías de la IUCN son más específicas por lo que a continuación se presenta el extracto de la información por especie.

*Abies guatemalensis*. Status: En peligro de extinción (EN). Especie con poblaciones con tendencia a disminuir. Esta especie se encuentra enfrentando un muy alto riesgo de extinción en la naturaleza, exhibe una presunta reducción del tamaño poblacional ya sea observada, estimada o inferida de  $\geq 50\%$  en los últimos 10 años o tres generaciones o lo que sea mayor, donde las causas de la reducción no son claramente reversibles, entendibles y suspendibles, con base en el rango geográfico, la observación directa, un índice de abundancia acorde al taxon, niveles de explotación

actual o potencial, los efectos de taxa introducidos, hibridación, patógenos, contaminantes, competidores o parásitos. El área de ocupación estimada es menor de 500 km<sup>2</sup>, los cálculos indican una fragmentación severa o es sabido que la especie se encuentra en no mas de cinco localidades. Se observa, inferiere o proyecta un declive continuo de la extensión de presencia, área de acupación, área, extensión y/o calidad del hábitat, número de localidades o poblaciones o número de individuos maduros (IUCN, 2018).

*Abies hidalgensis*. Status: Vulnerable (VU). Las tendencias poblacionales se desconocen. La especie enfrenta un alto riesgo de extinción en la naturaleza. La población es muy pequeña con un área de ocupación restringida (típicamente menos de 20 km<sup>2</sup>) o con pocas localidades (cinco a menos), de tal manera que es susceptible a los efectos de actividades humanas o eventos estocásticos dentro de un periodo de tiempo muy corto con un futuro incierto, por lo que la especie puede llegar a estar Críticamente en Vías de Extinción o incluso Extinta en un corto periodo de tiempo (IUCN, 2018).

*Pinus greggii*. Status: Vulnerable (VU). La población presenta una tendencia a disminuir. Se considera que esta especie esta enfrentando alto riesgo de extinción en la naturaleza. Se estima que el área de ocupación es menor de 2000 km<sup>2</sup>, severamente fragmentada o la existencia de la especie se conoce en no mas de 10 localidades. Con una disminución continua del área de ocupación, extensión y/o calidad del hábitat o el número de individuos maduros, ya sea observada, inferida o proyectada (IUCN, 2018).

*Taxus globosa*. Status: En peligro de extinción (EN). La población presenta una tendencia a disminuir. Se considera que esta especie se encuentra enfrentando un muy alto riesgo de extinción en el medio silvestre. Se observa, estima, infiere o supone una reducción del tamaño poblacional de  $\geq 50\%$  durante los últimos 10 años o tres generaciones, lo que sea mayor, donde las causas de la reducción son claramente

reversibles, entendiibles e interrumpibles, con base en la observación directa, un índice de abundancia apropiada para el taxon, un declive del área de ocupación, grado de ocurrencia y/o calidad del hábitat, niveles de explotación actual o potencial y los efectos de taxa introducidos, hibridación, patógenos, contaminantes, competidores o parásitos (IUCN, 2018).

### **1.3.2 Los códigos de barra genéticos**

En términos generales los códigos de barras genéticos permitirán la identificación taxonómica de cualquier muestra biológica que contenga DNA (aun cuando este se encuentre altamente fragmentado), amplificado mediante PCR y secuenciado mediante tecnología Sanger (Big Dye Terminator) (Hebert *et al.*, 2003; Hollingsworth, 2007).

### **1.3.3 Secuenciación Sanger**

La secuenciación Sanger (Big Dye Terminator) utiliza un cebador para sintetizar una hebra de DNA complementaria a una cadena molde de hebra sencilla empleando 2',3'-didesoxinucleótidos-5'-Trifosfato (ddNTPs), que pueden inhibir la actividad de la DNA polimerasa I actuando como terminadores de cadena, debido a que cuando un ddNTPs es incorporado en la cadena en crecimiento el proceso termina, cada uno de los cuatro ddNTPs está marcado fluorescentemente con un pigmento correspondiente a una de las cuatro bases nitrogenadas (Sanger, Nicklen, y Coulson, 1977; Smith et al., 1986).

El proceso anterior resulta en una mezcla de cadenas de DNA de diferentes tamaños, las cuales difieren únicamente por un nucleótido y terminando con un ddNTP que se encuentra marcado con un pigmento específico en su extremo 3'. Mediante electroforesis capilar, los productos de secuenciación de DNA entran en el capilar como resultado de inyección electrocinética, posteriormente una carga de alto voltaje aplicada

al amortiguador de reacción de secuenciación impulsa los fragmentos cargados negativamente hacia el interior de los capilares, los productos de secuenciación son separados por tamaño con base en su carga total. La secuencia de DNA se decodifica a partir de los patrones de color correspondientes a la secuencia de nucleótidos (Sanger et al., 1977; Smith et al., 1986).

El éxito de identificación a nivel de especies mediante los códigos de barras genéticos puede ser más difícil en grupos como coníferas debido a que presentan 1) taxonomía imperfecta, 2) hibridación con subsecuente introgresión, 3) polimorfismo ancestral compartido y 4) tasas de mutación bajas (P M Hollingsworth, Graham, y Little, 2011; D. P. Little, 2011); por lo que son necesarios más estudios de caso, que permitan transformar estas generalizaciones en predicciones cuantitativas (Hollingsworth *et al.*, 2011).

#### **1.3.4 Taxonomía imperfecta**

Grupos con incertidumbre taxonómica, causada por límites de especies pobremente definidos, circunscripciones estrechas o amplias de especies o presencia de especies crípticas, causando sobre o subestimación del número de especies (P M Hollingsworth et al., 2011; Little y Stevenson, 2007; Syring, Farrell, Businský, Cronn, y Liston, 2007).

#### **1.3.5 Hibridación con subsecuente introgresión**

Los grupos con hibridación y subsecuente introgresión, son especialmente frecuentes en regiones de simpatria, con híbridos que exhiben cierto grado de fertilidad. La habilidad de hibridar bajo condiciones naturales o artificiales en géneros como *Pinus* ha sido ampliamente documentada, así como su potencial de especiación por hibridación (Rémy J Petit y Excoffier, 2009; Wang, Tsumura, Yoshimaru, Nagasaka, y Szmidt, 1999). La variación genética compartida entre especies cercanamente relacionadas puede ser causada por introgresión después de contacto secundario (Zhou et al., 2017). La

introgresión es la transferencia de genes entre especies principalmente por retrocruza (Anderson, 1949; Arnold, 1992; Rieseberg and Carney, 1998) en Twyford y Ennos (2012).

### **1.3.6 Polimorfismo ancestral compartido**

La retención del prolimorfismo ancestral puede ser causado por reparto de linajes incompleto (Zhou et al., 2017). El reparto de linajes o sorteo de linajes es el proceso por el cual el polimorfismo ancestral se reparte y posteriormente se fija entre los linajes descendientes mediante deriva genética o selección (Broughton y Harrison, 2003), por otro lado, el reparto de linajes incompleto se refiere a la permanencia y retención de polimorfismo ancestral a través de múltiples eventos de especiación. Cuando los tiempos de los eventos de divergencia entre especies son menores a los intervalos de ramificación de linajes al interior de las especies, el polimorfismo ancestral puede ser acarreado a través de eventos sucesivos de divergencia (Syring, Farrell, Businský, Cronn, & Liston, 2007). La permanencia del reparto de polimorfismo ancestral incompleto es favorecida en grupos de radiación rápida o reciente, o en grupos que históricamente presentan grandes tamaños poblacionales (P M Hollingsworth, Graham, & Little, 2011).

### **1.3.7 Tasas de mutación bajas**

La escasa variabilidad molecular interespecífica, en grupos de reciente radiación, limita las aplicaciones potenciales del código de barras genético a un subconjunto particular de todas las especies conocidas hasta el momento.

Entre los parámetros evaluados en los códigos de barras genéticos se encuentran la universalidad (secuenciación rápida de alta calidad, confiable y de bajo costo), la calidad de la secuencia, la complejidad lingüística (una medida de la repetitividad de la secuencia), la discriminación de especies (variación para distinguir entre especies), y la

presencia/ausencia del gap del código de barras genético (mayor variación interespecífica que variación intraespecífica; Armenise et al., 2012; CBOL Plant Working Group et al., 2009; Fazekas et al., 2008; Ford et al., 2009; Maia et al., 2012; Pesole, Attimonelli, y Saccone, 1995).

### **1.3.8 Universalidad**

La universalidad se mide por el potencial de los cebadores para amplificar regiones blanco mediante la reacción en cadena de la polimerasa (PCR) y la generación de secuencias de alta calidad. La universalidad depende de regiones adyacentes conservadas que permitan la amplificación incluso en taxones altamente divergentes (P. M. Hollingsworth, 2007; Little, 2011; Maia et al., 2012). Los locus más variables también son los más difíciles de amplificar entre diversos taxones pero al mismo tiempo son los que permiten las mayores tasas de discriminación de especies. La especificidad de los cebadores consiste en la amplificación exclusiva de las regiones blanco (Ford et al., 2009).

### **1.3.9 Calidad de la secuencia**

La calidad de la secuencia está determinada por el log10 de la probabilidad de error multiplicada por -10,  $-10 \times \log_{10} p$ ; un valor de calidad de 10 (Q10) indica que existe una oportunidad en 10 de que el programa de secuenciación asigne de manera incorrecta una base determinada; un valor de calidad de Q20 significa una probabilidad de una asignación incorrecta en 100. Para estudios de códigos de barras genéticos es altamente deseable tener valores de calidad de Q30 o mayores (Hanner, 2009; Little, 2010a).

De acuerdo con Little (2010) la calidad total de una secuencia dada ( $S_R$ ) puede ser estimada para un umbral definido de calidad ( $q$ ) utilizando la siguiente función:

$$SR = \sum_{i=1}^j \begin{cases} 0, Ri < q \\ 1, Ri \geq q \end{cases}$$

Donde  $R_i$  es el estimado de la calidad para la posición  $i$  de la secuencia ( $R$ ),  $j$  es la longitud de la secuencia, la calidad total de la secuencia puede ser estimada mediante la suma de la calidad de las lecturas de la secuencia en la secuencia consenso ( $S_R$ ).

La siguiente función se utiliza para normalizar por longitud y cobertura de la secuencia:

$$Bq = \frac{\sum_{R=1}^k SR}{cx}$$

Donde  $k$  es el número total de lecturas de secuencia en la secuencia consenso,  $c$  es la longitud observada en la secuencia consenso y  $x$  es la cobertura esperada. La calidad de la secuencia determina la confiabilidad y la robustez de los datos y los resultados.

### **1.3.10 Complejidad lingüística de la secuencia**

La complejidad lingüística es una medida de la repetitividad de la secuencia. Asumiendo que el material genético puede considerarse como un lenguaje escrito, que utiliza un alfabeto de cuatro elementos ACGT, se pueden hacer palabras, denominadas cadenas, motivos y patrones (Pesole, Attimonelli, y Saccone, 1995).

Las secuencias nucleotídicas comparten algunas características con los lenguajes escritos, como el hecho de que son combinaciones de caracteres unidimensionales con un grado considerable de repetición, y son secuencias en las que se pueden identificar propiedades no aleatorias. La aproximación lingüística para el análisis de las secuencias nucleotídicas es una herramienta poderosa en diversas situaciones tales como la identificación de motivos con un rol funcional, el establecimiento de correlaciones funcionales entre secuencias y la relación filogenética existente entre textos genéticos (Pesole et al., 1995). Las regiones genómicas con microsatélites múltiples tienen valores menores de calidad y mayor complejidad lingüística (M. L. Hollingsworth et al., 2009). La complejidad lingüística de una secuencia nucleotídica ( $C$ ) puede describirse mediante la siguiente fórmula empírica:

$$C = \prod_{j=1}^{L-1} U_j$$

Donde

$$U_j = \left\{ \sum_{i=1}^{4^j} \frac{X_{ij}}{L - j + 1} \right\} \max \left( 1, \frac{L}{4^j - j + 1} \right)$$

Donde  $L$  es la longitud de la secuencia nucleotídica,  $X_{ij} = 1$  si la secuencia  $i$  de longitud  $j$  se encuentra al menos una vez en la secuencia que está siendo examinada, o 0 si está ausente. Entre mayor sea el número de secuencias de cebadores diferentes de longitud 1, 2, 3, ...,  $L-1$  en la secuencia, mayor será la complejidad global de la secuencia. Entre mayor sea la complejidad global de la secuencia existirá mayor heterogeneidad que refleja una menor restricción estructural y funcional sobre la región analizada, exhibiendo mayor variación y por lo tanto permitiendo un mayor éxito de identificación.

### 1.3.11 Discriminación de especies

Para la discriminación de especies los marcadores de los códigos de barras genéticos, deben exhibir suficiente variación para contar con sustituciones únicas que sirvan como marcadores a nivel de especie y que sirvan para separar incluso pares de especies hermanas (Chase et al., 2005); por otro lado deben mostrar consistencia de secuenciación suficiente para asegurar que la variación intraespecífica no afecte la asignación de especies (Ford et al., 2009).

Los métodos de identificación mediante los códigos de barras genéticos recientemente han incorporado, métodos basados en caracteres de similitud especial (sinapomorfias), en lugar de la similitud total. La similitud especial puede ser utilizada en forma de árboles filogenéticos o implícitamente como descriptores taxonómicos

jerárquicos (Farris, 1979; D Little, 2011). Sin embargo, se requiere del desarrollo de un marco teórico y metodológico sólido para traducir secuencias de DNA a información taxonómica de especies (D Little, 2011).

Los métodos de identificación de especies mediante los códigos de barras genéticos pueden estar basados en 1) árboles y métodos filogenéticos (e.g. parsimonia, vecino más cercano, máxima verosimilitud), 2) métodos diagnósticos libres de alineamiento (e.g. DNA-BAR, DOME ID y ATIM), 3) similitud (distancias genéticas pareadas, BLAST, BLAT, FASTA, BRONX y mega BLAST) o 4) combinaciones de los tres tipos de métodos anteriores (e.g. Similitud y árboles (BLAST y parsimonia, BLAST y vecino más cercano), métodos diagnósticos y de similitud) (Little y Stevenson, 2007; Tanabe y Toju, 2013).

#### **1.3.12 Métodos basados en árboles, filogenéticos o de agrupamiento jerárquico**

El clado menos inclusivo es un método de agrupamiento jerárquico, basado en métodos filogenéticos tales como el vecino más cercano, parsimonia o máxima verosimilitud, para implementar la identificación, una secuencia blanco se alinea con un conjunto de secuencias de referencia, posteriormente se construye una jerarquía para determinar su identidad. El clado menos inclusivo donde se encuentre la secuencia blanco se toma como la identificación del espécimen (Little y Stevenson, 2007).

Los métodos de agrupamiento jerárquico pueden generar identificaciones ambiguas en al menos tres situaciones. Primero cuando una secuencia blanco se recupera como grupo hermano de un grupo conocido, segundo cuando una secuencia se recupera en una posición entre un grupo monofilético (para o polifilético) y otro grupo; y tercero cuando se obtiene un árbol no resuelto. Además de identificaciones ambiguas, la búsqueda de árboles es un problema NP-complejo, por otro lado los métodos

filogenéticos pueden fallar como resultado de inconsistencias en el alineamiento (Little y Stevenson, 2007; Little, 2011).

### **1.3.13 Métodos diagnósticos basados en caracteres**

Estos métodos constituyen una adaptación de la aproximación morfológica convencional al código de barras genético, están basados en caracteres y no se incorporan las relaciones filogenéticas. Se pueden utilizar caracteres no homólogos que usualmente son más informativos que los caracteres estrictamente homólogos (Little y Stevenson, 2007).

Utilizando los métodos diagnósticos, se puede hacer la diagnosis de las especies (e.g. una única secuencia o una lista de motivos por región). Los especímenes pueden ser identificados mediante caracteres diagnósticos, buscando en secuencias registradas en una base de referencia. Una de las principales ventajas de este tipo de métodos es que no se requiere de un alineamiento de secuencias, pero la falta de variabilidad puede causar que no todas las especies tengan caracteres diagnósticos únicos, sin embargo, pueden ser diferenciadas por combinaciones únicas de los caracteres diagnósticos más comunes.

### **1.3.13 Métodos de similitud**

Los métodos de similitud usan un umbral de divergencia estricto, por encima del cual la secuencia blanco es considerada de una especie diferente y por debajo del cual una secuencia es considerada como proveniente de un individuo conspecífico. La similitud puede ser estimada para la porción de la secuencia blanco alineada con una secuencia de referencia, por lo que la secuencia de referencia con la menor distancia pareada se toma como la identificación (Little y Stevenson, 2007).

Los métodos de similitud pueden generar identificaciones de secuencias inconsistentes o incluso incorrectas, en situaciones reales. Una secuencia de referencia incompleta puede generar una identificación incorrecta debido a los parámetros utilizados para construir el alineamiento pareado. Todos los métodos de similitud pueden producir identificaciones ambiguas cuando una secuencia blanco tiene un empalme idéntico con más de una secuencia de la base de datos de referencia (Little y Stevenson, 2007). Los métodos basados en similitud no están diseñados para tomar en cuenta las relaciones filogenéticas jerárquicas, además al ser corregidos para evitar sesgos biológicos o de muestreo, pueden aunque no intencionalmente opacar la mínima cantidad de variación en las secuencias de especies estrechamente relacionadas (Little, 2011).

#### **1.3.14 Métodos híbridos**

Un ejemplo es el uso de métodos de similitud en combinación con métodos basados en árboles, en los que se seleccionan las secuencias de la base de datos de referencia que más se parezcan a la secuencia blanco (e.g. 100), para posteriormente construir un árbol de distancias y poder determinar la identificación final (Little y Stevenson, 2007).

#### **1.3.15 El gap del código de barras genético**

Representa la variación intraespecífica relativa comparada con la variación interespecífica relativa (Meyer y Paulay, 2005; Pesole et al., 1995). La identificación efectiva de especies mediante los códigos de barras genéticos depende de que la variación interespecífica exceda por un orden de magnitud la variación intraespecífica; creando un notorio gap del código de barras genético y monofilia recíproca entre las especies, que permita la identificación de un espécimen desconocido con una tasa de error mínima (Hebert, Cywinska, Ball, y DeWaard, 2003; Wiemers y Fiedler, 2007). La falta de un gap del código de barras genético evidenciado mediante la parafilia de la

especie y tasas de identificación bajas, puede ser explicada a causa de límites de especie poco claros (Maia et al., 2012).

### **1.3.16 Antecedentes de los códigos de barras genéticos en plantas**

Encontrar un código de barras genético efectivo para plantas ha sido difícil; en comparación con animales, en los que la región del citocromo oxidasa 1 del genoma mitocondrial ha permitido la identificación taxonómica de diversas especies de aves, peces e insectos (Hebert et al., 2003; P. M. Hollingsworth, 2011; Kerr et al., 2007; Ward, Zemlak, Innes, Last, y Hebert, 2005).

Para la implementación de los códigos de barras genéticos en plantas, muchas combinaciones de marcadores han sido evaluadas, cada una difiriendo en su habilidad de amplificación o identificación de especies (e.g. P. M. Hollingsworth, 2011; Roxanne Steele y Chris Pires, 2011).

En 2009 el grupo de trabajo de plantas del Consorcio del Código de Barras de la Vida (CBOL), propusieron dos regiones codificantes del genoma del plástido como los códigos de barras genéticos centrales (*matK* y *rbcL*) suplementadas por otras regiones en caso de ser necesario (P. M. Hollingsworth, 2011).

La región *rbcL* (Subunidad mayor de la ribulosa-1,5-bisfosfato carboxilasa/oxigenasa) es una región codificante de copia única del genoma del plastidio, con longitud de 1,428 pares de bases (pb), y tasa de evolución conservativa (Chase et al., 1993). La región *matK* (maturasa K) es una región codificante de copia única del genoma del plastidio, con una longitud de 1,500 pb, funciona como una maturasa con un papel esencial en el empalme de RNA. La maturasa K exhibe tasas mayores de evolución que *rbcL* (Barthet y Hilu, 2007; Sugita, Shinozaki, y Sugiura, 1985). Las regiones *matK* y *rbcL* han sido ampliamente utilizadas para estudios filogenéticos y más recientemente para la identificación de especies, por lo que es

importante evaluar su desempeño cuando se cambia de un contexto filogenético a uno de identificación de especies (Chase et al., 2005).

Los marcadores del plastidio son heredados vía paterna en la mayoría de las familias de coníferas excepto en *Cephalotaxaceae* y *Cunninghamia* de la familia *Cupressaceae* (Lu, Peng, Cheng, Hong, y Chiang, 2001; Mogensen, 1996; Rémy J Petit y Vendramin, 2007; Whittle y Johnston, 2002; Worth, Yokogawa, y Isagi, 2014). Sin embargo, evidencia reciente sugiere que existe herencia materna del DNA del plastidio en la cruce de *P. mugo* x *P. sylvestris*, mientras que en la cruce opuesta (*P. sylvestris* x *P. mugo*) la herencia es vía paterna (Kormutak et al., 2018). Los marcadores del plastidio tienen mayor probabilidad de ser congruentes con los límites de especies morfológicos como resultado de sus patrones de dispersión, en comparación con los marcadores mitocondriales que en coníferas son heredados por vía materna y presentan menor dispersión (É. Aguirre-Planter et al., 2012; Du, Petit, y Liu, 2009; Mogensen, 1996; R J Petit, Kremer, y Wagner, 1993; Rémy J Petit y Excoffier, 2009).

Entre las ventajas del uso de marcadores del genoma del plastidio, se encuentra la haploidia, un tamaño relativamente pequeño y conservado, una estructura estable y el buen conocimiento de la función y posición de sus regiones, para las cuales existen cebadores disponibles que permiten su fácil amplificación, el plastidio que de forma general se hereda de forma uniparental (sin embargo ver Kormutack et al., 2018), se encuentra en copias numerosas por célula lo cual permite la amplificación de DNA a partir incluso de material altamente degradado. En comparación con el genoma mitocondrial presenta tasas de sustitución mayores y recombinación intramolecular menor, siendo una de sus principales limitantes su tasa de sustitución relativamente baja (D. Gernandt et al., 2011; Mower, Touzet, Gummow, Delph, y Palmer, 2007; Olmstead y Palmer, 1994).

La región *trnH-psbA* es más variable que *matK* o *rbcL*, sin embargo dicha variación no siempre es suficiente para discriminar entre especies (e.g. *Pinus*, subsección *Ponderosae*). Entre los factores que han evitado que *trnH-psbA* sea usado como código de barras genético se encuentran: los rearrreglos cromosómicos (inversiones, deleciones y microsatélites) comunes al interior de las especies; la presencia de poliestructuras largas que causan dificultades de secuenciación y su longitud relativamente corta (Dong et al., 2015; Hernández-León, Gernandt, Pérez de la Rosa, y Jardón-Barbolla, 2013). En *Pinus* sección *Trifoliae* se cuenta con evidencia de haplotipos compartidos entre subsecciones para esta región y falta de universalidad de los cebadores que resulta en la incapacidad de obtener secuencias bidireccionales en grupos particulares (Hernández-León et al., 2013). En un estudio de las coníferas de Italia se documentó una inversión de 34-pb para una secuencia de *trnH-psbA* exclusiva de un individuo de *Picea abies*, además de variación diferencial entre familias para *trnH-psbA* siendo mayor en Pinaceae que en la familia Cupressaceae (Armenise et al., 2012).

La región *ycfI* no había sido propuesta como un potencial código de barras genético para plantas debido a su longitud y a la falta de universalidad de los cebadores particularmente para gimnospermas (Dong et al., 2015; Hernández-León et al., 2013). Dong (et al., 2015) reportan 90.91% de éxito de amplificación de *ycfI* para gimnospermas y un mejor desempeño general cuando se compara con *matK*, *rbcL* y *trnH-psbA*. Hernández-León (et. al. 2013) encontraron en *Pinus* sección *Trifoliae* que entre varias regiones del genoma del cloroplasto, *ycfI* fue la única en generar más haplotipos que especies reconocidas, por lo que el uso de cebadores específicos por familia sería una aproximación de mucha utilidad para la identificación de especímenes mediante esta región. El uso de cebadores específicos (la no universalidad de los

cebadores) no debería ser una limitante, cuando el fin último de los códigos de barras genéticos es la identificación taxonómica de los especímenes (Dong et al., 2015).

Marcadores nucleares como el nrITS (Espaciador transcrito interno) han sido ampliamente utilizados para estudios filogenéticos, documentándose múltiples casos de evolución concertada incompleta. La evolución concertada es una forma de evolución de familias multigénicas en la que los diversos genes de una familia tienden a homogeneizarse y se asume que evolucionan de manera dependiente en conjunto como una sola unidad (Liao, 1999), pero en algunos casos la evolución concertada se da de manera incompleta entre los repetidos, por lo que ocasionalmente se pueden detectar en un mismo individuo copias parálogas, incluyendo pseudogenes, recombinados y múltiples copias funcionales de ITS que causan variación intraindividual por lo que no son viables para la identificación efectiva de especies mediante los códigos de barras genéticos (Bailey, Carr, Harris, y Hughes, 2003; Dong et al., 2015; Xu, Zeng, Gao, Jin, y Zhang, 2017).

La región nuclear ITS ocasionalmente exhibe evolución concertada incompleta (slow concerted evolution), variación intragenómica, presencia de pseudogenes y complejidad experimental en especies de origen híbrido (Dong et al., 2015; D S Gernandt, Liston, y Piñero, 2001), existe evidencia que indica que en gimnospermas ITS no presenta un buen desempeño como código de barras genético (D.-Z. Li et al., 2011). ITS (ITS1 e ITS2) como parte del DNA nuclear es uno de los marcadores más secuenciados en plantas debido a su rápida evolución concertada al interior y entre las subunidades componentes, su tasa de evolución rápida, tamaño reducido y disponibilidad de cebadores universales (D.-Z. Li et al., 2011); ITS1 comparado con ITS2 tiene un mejor desempeño como código de barras genético en plantas, en términos de eficiencia, universalidad e identificación de especímenes, además de que genera un

producto de amplificación de menor tamaño y presenta un menor contenido de GC (Wang et al., 2015)

Otra región nuclear que tiene el potencial de incrementar la tasa de identificación de especímenes es *NEEDLY* intron 2, una región de fácil recuperación en familias como Podocarpaceae (Knopf, Schulz, Little, Stützel, y Stevenson, 2012; D P Little, 2004).

La identificación basada en los códigos de barras genéticos continua siendo una herramienta atractiva en términos de exactitud, precisión, velocidad, costo y funcionalidad (Armenise, Simeone, Piredda, & Schirone, 2012).

Una biblioteca de referencia será de gran valor para un amplio espectro de usuarios de los códigos de barras genéticos, particularmente para aquellos que no están familiarizados con la morfología y anatomía de coníferas (Hollingsworth, 2007), será la base de iniciativas sobre inventarios de biodiversidad, monitoreo del comercio de especies, entre otras que son de particular importancia para especies amenazadas en vía de extinción. La conservación y manejo sostenible de los recursos forestales deben estar basados en datos sólidos bien documentados (Little, Knopf, y Schulz, 2013). De manera adicional las coníferas del estado de Hidalgo constituyen un modelo para evaluar y predecir el éxito de identificación mediante los códigos de barras genéticos centrales en otros grupos de plantas con rasgos de historia de vida similares y a una escala local.

En el capítulo II de este estudio se establece un listado actualizado de las especies de coníferas del estado de Hidalgo y se evalúa su identificación a nivel de familia, género y especie mediante los códigos de barras genéticos (*matK* y *rbcL*), estimando la calidad de la secuencia, la complejidad lingüística y la presencia del gap del código de barras. Se establece la biblioteca de referencia para la identificación taxonómica a nivel de familia, género y especie, comparando el desempeño de dos

métodos: BRONX (Barcode Recognition Obtained with Nucleotide eXposés; Little, 2012) y el clado menos inclusivo (Little y Stevenson, 2007). En el capítulo III se establece una clave dicotómica para la identificación de los pinos con distribución natural en el estado de Hidalgo, con base en un análisis UPGMA de caracteres morfoanatómicos foliares.

#### 1.4 Literatura citada

- Aguirre-Planter, Érika et al. 2012. “Phylogeny, Diversification Rates and Species Boundaries of Mesoamerican Firs (*Abies*, Pinaceae) in a Genus-Wide Context.” *Molecular Phylogenetics and Evolution* 62(1): 263–74.
- Armenise, Laura, Marco C Simeone, Roberta Piredda, and Bartolomeo Schirone. 2012. “Validation of DNA Barcoding as an Efficient Tool for Taxon Identification and Detection of Species Diversity in Italian Conifers.” *European Journal of Forest Research* 131(5): 1337–53.
- Bailey, C. Donovan, Timothy G. Carr, Stephen A. Harris, and Colin E. Hughes. 2003. “Characterization of Angiosperm nrDNA Polymorphism, Paralogy, and Pseudogenes.” *Molecular Phylogenetics and Evolution* 29(3): 435–55.
- Barthet, Michelle M., and Khidir W. Hilu. 2007. “Expression of *matK*: Functional and Evolutionary Implications.” *American Journal of Botany* 94(8): 1402–12.
- Broughton, R. E., & Harrison, R. G. (2003). Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. *Genetics*, 163(4), 1389-1401.
- CBOL Plant Working Group et al. 2009. “A DNA Barcode for Land Plants.” *Proceedings of the National Academy of Sciences of the United States of America* 106(31): 12794–97. <http://www.pnas.org/content/106/31/12794.full>.
- Chase, Mark W et al. 1993. “Phylogenetics of Seed Plants: An Analysis of Nucleotide Sequences from the Plastid Gene *rbcL*.” *Annals of the Missouri Botanical Garden*: 528–80.

- . 2005. “Land Plants and DNA Barcodes: Short-Term and Long-Term Goals.” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360(1462): 1889–95.
- Christenhusz, Maarten J M et al. 2011. “A New Classification and Linear Sequence of Extant Gymnosperms.” *Phytotaxa* 19: 55–70.
- Dong, W et al. 2015. “ycf1, the Most Promising Plastid DNA Barcode of Land Plants.” *Sci Rep* 5: 8348. <http://www.ncbi.nlm.nih.gov/pubmed/25672218>  
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4325322/pdf/srep08348.pdf>.
- Du, Fang K., Remy J. Petit, and Jian Quan Liu. 2009. “More Introgression with Less Gene Flow: Chloroplast vs. Mitochondrial DNA in the *Picea Asperata* Complex in China, and Comparison with Other Conifers.” *Molecular Ecology* 18(7): 1396–1407.
- Eckenwalder, JE. 2009. *Conifers of the World: The Complete Reference*. Timber Press.
- Farjon, A. 2008. *A Natural History of Conifers*. Timber Press.
- . 2010. *A Handbook of the World's Conifers*. Leiden: Brill.
- Farris, J S. 1979. “The Information Content of the Phylogenetic System.” *Systematic Zoology* 28: 483–519. <http://www.jstor.org/stable/pdfplus/2412562.pdf>.
- Fazekas, Aron J. et al. 2008. “Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well.” *PLoS ONE* 3(7).

- Ford, Caroline S. et al. 2009. "Selection of Candidate Coding DNA Barcoding Regions for Use on Land Plants." *Botanical Journal of the Linnean Society* 159(1): 1–11.
- Gernandt, D S, a Liston, and D Piñero. 2001. "Variation in the nrDNA ITS of Pinus Subsection Cembroides: Implications for Molecular Systematic Studies of Pine Species Complexes." *Molecular phylogenetics and evolution* 21(3): 449–67.
- Gernandt, David S., Gretel Geada López, Sol Ortiz García, and Aaron Liston. 2005. "Phylogeny and Classification of Pinus." *Taxon* 54(1): 29–42.
- Gernandt, David S., and Jorge A. Pérez-De La Rosa. 2014. "Biodiversidad de Pinophyta (coníferas) En México." *Revista Mexicana de Biodiversidad* 85(SUPPL.).
- Gernandt, DS, and A Vázquez-Lobo. 2014. "Conifers: The Most Diverse Group of Naked Seed Plants." In *The Tree of Life Evolution and Classification of Living Organisms*, eds. P Vargas and R Zardoya. USA: Sinauer Associates, Inc. Publishers, 713.
- Gernandt, DS, A Willyard, JV Syring, and A Liston. 2011. "The Conifers (Pinophyta)." In *Genetics, Genomics and Breeding of Conifers*, eds. C Plomion, J Bousquet, and C Kole. USA: Science Publishers, 1–39.
- Halffter, Gonzalo, Jorge Llorente-Bousquets, and Juan J. Morrone. 2008. "La Perspectiva Biogeográfica Histórica." In *Capital Natural de México*, , 67–86.
- Hanner, By Robert. 2005. "Proposed Standards for BARCODE Records in INSDC (BRIs)." *Global Biodiversity* (December): 1–3.

- Hebert, P D N, A Cywinska, S L Ball, and J R DeWaard. 2003. “Biological Identifications through DNA Barcodes.” *Proceedings. Biological sciences / The Royal Society* 270(1512): 313–21.
- Hernández-León, Sergio, David S Gernandt, Jorge A Pérez de la Rosa, and Lev Jardón-Barbolla. 2013. “Phylogenetic Relationships and Species Delimitation in *Pinus* Section *Trifoliae* Inferred from Plastid DNA.” *PLoS ONE* 8(7): 1–14.
- Hollingsworth, Michelle L. et al. 2009. “Selecting Barcoding Loci for Plants: Evaluation of Seven Candidate Loci with Species-Level Sampling in Three Divergent Groups of Land Plants.” *Molecular Ecology Resources* 9(2): 439–57.
- Hollingsworth, P M, S W Graham, and D P Little. 2011. “Choosing and Using a Plant DNA Barcode.” *Plos One* 6(5): e19254.  
<http://www.ncbi.nlm.nih.gov/pubmed/21637336>.
- Hollingsworth, P. M. 2007. “DNA Barcoding : Potential Users.” *Policy* 3(2): 44–47.
- . 2011. “Refining the DNA Barcode for Land Plants.” *Proceedings of the National Academy of Sciences* 108(49): 19451–52.
- INEGI. 2014. *Anuario Estadístico Y Geográfico de Hidalgo 2013*. México: Instituto Nacional de Estadística y Geografía.
- IUCN. 2014. “The IUCN Red List of Threatened Species.” *Versión 2014.2*.  
<http://www.iucnredlist.org> (August 26, 2014).

- Keenan, Rodney J. et al. 2015. "Dynamics of Global Forest Area: Results from the FAO Global Forest Resources Assessment 2015." *Forest Ecology and Management* 352: 9–20.
- Kerr, Kevin C R et al. 2007. "Comprehensive DNA Barcode Coverage of North American Birds." *Molecular Ecology Notes* 7(4): 535–43.
- Knopf, Patrick et al. 2012. "Relationships within Podocarpaceae Based on DNA Sequence, Anatomical, Morphological, and Biogeographical Data." *Cladistics* 28(3): 271–99.
- Kormutak, Andrej et al. 2018. "Maternal Inheritance of Chloroplast DNA in *Pinus Mugo* Turra: A Case Study of *Pinus mugo* × *Pinus sylvestris* Crossing." *Plant Systematics and Evolution* 304(1): 71–76.
- Li, D.-Z. et al. 2011. "From the Cover: Comparative Analysis of a Large Dataset Indicates That Internal Transcribed Spacer (ITS) Should Be Incorporated into the Core Barcode for Seed Plants." *Proceedings of the National Academy of Sciences* 108(49): 19641–46.
- Liao, D. 1999. "Concerted Evolution: Molecular Mechanism and Biological Implications." *American Journal of Human Genetics* 64(1): 24–30.
- Little, D P. 2004. "Documentation of Hybridization between Californian Cypresses: *Cupressus Macnabiana* X *Sargentii*." *Systematic Botany* 29: 825–33. <Go to ISI>://WOS:000225245000003.

- Little, D. P. 2010. "A Unified Index of Sequence Quality and Contig Overlap for DNA Barcoding." *Bioinformatics (Oxford, England)* 26(21): 2780–81.  
<http://www.ncbi.nlm.nih.gov/pubmed/20837570>.
- . 2011. "DNA Barcode Sequence Identification Incorporating Taxonomic Hierarchy and within Taxon Variability." *PLoS ONE* 6(8).
- . 2012. "BRONX: Barcode Recognition Obtained with Nucleotide eXpose's." <http://www.nybg.org/files/scientists/dlittle/BRONX2.html>.
- Little, D. P., P. Knopf, and C. Schulz. 2013. "DNA Barcode Identification of Podocarpaceae-The Second Largest Conifer Family." *PLoS ONE* 8(11).
- Little, D. P., and D. W. Stevenson. 2007. "A Comparison of Algorithms for the Identification of Specimens Using DNA Barcodes: Examples from Gymnosperms." *Cladistics* 23(1): 1–21.
- Lu, S Y et al. 2001. "Chloroplast DNA Phylogeography of *Cunninghamia Konishii* (Cupressaceae), an Endemic Conifer of Taiwan." *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 44(5): 797–807.
- Maia, Vitor Hugo et al. 2012. "DNA Barcoding Bromeliaceae: Achievements and Pitfalls." *PLoS ONE* 7(1).
- Meyer, Christopher P., and Gustav Paulay. 2005. "DNA Barcoding: Error Rates Based on Comprehensive Sampling." *PLoS Biology* 3(12): 1–10.

- Mogensen, H. Lloyd. 1996. "The Hows and Whys of Cytoplasmic Inheritance in Seed Plants." *American Journal of Botany* 83(3): 383–404.
- Morrone, Juan J. 2005. "Hacia Una Síntesis Biogeográfica de México." *Revista Mexicana de Biodiversidad* 76(002): 207–52.
- Mower, Jeffrey P et al. 2007. "Extensive Variation in Synonymous Substitution Rates in Mitochondrial Genes of Seed Plants." *BMC evolutionary biology* 7: 135. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1973135&tool=pmcentrez&rendertype=abstract>.
- Olmstead, Richard G., and Jeffrey D. Palmer. 1994. "Chloroplast DNA Systematics: A Review of Methods and Data Analysis." *American Journal of Botany* 81(9): 1205. <http://links.jstor.org/sici?sici=0002-9122%28199409%2981%3A9%3C1205%3ACDSARO%3E2.0.CO%3B2-U&origin=crossref>.
- Pesole, G, M Attimonelli, and C Saccone. 1995. "Linguistic Analysis of Nucleotide Sequences: Algorithms for Pattern Recognition and Analysis of Codon Strategy." *Methods in enzymology* 266: 281–94.
- Petit, R J, A Kremer, and D B Wagner. 1993. "Finite Island Model for Organelle and Nuclear Genes in Plants." *Heredity* 71(6): 630–41.
- Petit, Rémy J, and Laurent Excoffier. 2009. "Gene Flow and Species Delimitation." *Trends in Ecology & Evolution* 24(7): 386–93. <http://www.ncbi.nlm.nih.gov/pubmed/19409650>.

- Petit, Rémy J, and Giovanni G Vendramin. 2007. “Plant Phylogeography Based on Organelle Genes: An Introduction.” In *Phylogeography of Southern European Refugia – Evolutionary Perspectives on the Origins and Conservation of European Biodiversity*, , 23–97.
- Petit, Rémy J. et al. 2005. “Comparative Organization of Chloroplast, Mitochondrial and Nuclear Diversity in Plant Populations.” *Molecular Ecology* 14(3): 689–701.
- Prunier, Julien, Jukka Pekka Verta, and John J. Mackay. 2016. “Conifer Genomics and Adaptation: At the Crossroads of Genetic Diversity and Genome Function.” *New Phytologist* 209(1): 44–62.
- Rzedowski, J. 1978. *Vegetación de México*. México: Limusa.
- Sanger, F, S Nicklen, and a R Coulson. 1977. “DNA Sequencing with Chain-Terminating Inhibitors.” *Proceedings of the National Academy of Sciences of the United States of America* 74(12): 5463–67.  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=431765&tool=pmcentrez&rendertype=abstract>.
- SEMARNAT. 2010. *Norma Oficial Mexicana NOM-059-SEMARNAT-2010*.
- Smith, Lloyd M. et al. 1986. “Fluorescence Detection in Automated DNA Sequence Analysis.” *Nature* 321(6071): 674–79.  
<http://www.nature.com/nature/journal/v321/n6071/abs/321674a0.html>  
<http://www.nature.com/nature/journal/v321/n6071/pdf/321674a0.pdf>.

- Steele, P. R., and J. C. Pires. 2011. "Biodiversity Assessment: State-of-the-Art Techniques in Phylogenomics and Species Identification." *American Journal of Botany* 98(3): 415–25.
- Sugita, M, K Shinozaki, and M Sugiura. 1985. "Tobacco Chloroplast tRNA(UUU) Gene Contains a 2.5-Kilobase-Pair Intron: An Open Reading Frame and a Conserved Boundary Sequence in the Intron." *Proceedings of the National Academy of Sciences of the United States of America* 82(11): 3557–61.
- Syring, John et al. 2007. "Widespread Genealogical Nonmonophyly in Species of Pinus Subgenus Strobilus." *Systematic biology* 56(2): 163–81.
- Tanabe, Akifumi S., and Hirokazu Toju. 2013. "Two New Computational Methods for Universal DNA Barcoding: A Benchmark Using Barcode Sequences of Bacteria, Archaea, Animals, Fungi, and Land Plants." *PLoS ONE* 8(10).
- Twyford, A. D., and R. A. Ennos. 2012. "Next-Generation Hybridization and Introgression." *Heredity* 108(3): 179–89.
- Wang, Xiao Ru et al. 1999. "Phylogenetic Relationships of Eurasian Pines (Pinus, Pinaceae) Based on Chloroplast RbcL, matK, rpl20-rps18 Spacer, and trnV Intron Sequences." *American Journal of Botany* 86(12): 1742–53.
- Wang, Xin Cun et al. 2015. "ITS1: A DNA Barcode Better than ITS2 in Eukaryotes?" *Molecular Ecology Resources* 15(3): 573–86.
- Ward, Robert D et al. 2005. "DNA Barcoding Australia's Fish Species." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360(1462): 1847–57.

- Whittle, Carrie-Ann, and Mark O Johnston. 2002. "Male-Driven Evolution of Mitochondrial and Chloroplastial DNA Sequences in Plants." *Molecular Biology and Evolution* 19(6): 938–49.
- Wiemers, Martin, and Konrad Fiedler. 2007. "Does the DNA Barcoding Gap Exist? - a Case Study in Blue Butterflies (Lepidoptera: Lycaenidae)." *Frontiers in zoology* 4(September 2004): 8.  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1838910&tool=pmcentrez&rendertype=abstract>.
- Worth, James R P, Masashi Yokogawa, and Yuji Isagi. 2014. "Outcrossing Rates and Organelle Inheritance Estimated from Two Natural Populations of the Japanese Endemic Conifer *Sciadopitys Verticillata*." *Journal of Plant Research* 127(5): 617–26. <http://dx.doi.org/10.1007/s10265-014-0646-y>.
- Xu, Bo et al. 2017. "ITS Non-Concerted Evolution and Rampant Hybridization in the Legume Genus *Lespedeza* (Fabaceae)." *Scientific Reports* 7.
- Zhou, Y. et al. 2017. "Importance of Incomplete Lineage Sorting and Introgression in the Origin of Shared Genetic Variation between Two Closely Related Pines with Overlapping Distributions." *Heredity* 118(3): 211–20.

## Capítulo II

### Desempeño de los códigos de barras genéticos centrales en plantas a escala local:

#### Identificación de las coníferas del estado de Hidalgo, México

##### Resumen

El código de barras genético constituye una herramienta fundamental para la identificación de especies, especialmente para regiones con alta diversidad que cuentan con pocos estudios florísticos. En este estudio se caracterizan y evalúan las regiones centrales del código de barras genético en plantas, *matK* y *rbcL*, para identificar las 25 especies de coníferas con distribución natural en el estado de Hidalgo, México; incluyendo diez especies bajo diversas categorías de riesgo. Para las regiones del código de barras genético se obtuvo la calidad de las secuencias, su complejidad lingüística y la presencia del gap del código de barras genético. Se compararon dos métodos para la identificación de especies: BRONX y el clado menos inclusivo. Se generaron 89 secuencias para *matK* y 93 para *rbcL*. La región *matK* presentó la mayor diversidad haplotípica y la diversidad nucleotídica (II), incluyendo 6 indels (inserciones o deleciones). La combinación de las dos regiones generó 21 haplotipos únicos. En *rbcL* se observó mayor calidad de secuencias, así como mayor complejidad lingüística. Todas las especies diagnosticables exhibieron el gap del código de barras genético. Todos los especímenes se asignaron de manera no ambigua a familia y género, sin importar el marcador o método utilizado. La región *matK* analizada con BRONX exhibió el mayor éxito de identificación (44%). A pesar del bajo éxito de identificación, es posible establecer proyectos de manejo, conservación y monitoreo, aun en ausencia de caracteres morfológicos taxonómicamente diagnósticos, para al menos 50% de las especies en categorías de riesgo. La baja divergencia entre especies cercanamente

relacionadas puede ser resultado de tasas de evolución molecular bajas en las regiones centrales de los códigos de barras genéticos, hibridación o sorteo de linajes incompleto. En grupos con rasgos de historia de vida comparables a coníferas, se esperan éxitos de identificación similares en regiones con condiciones de tamaño y diversidad taxonómica equiparables con las del estado de Hidalgo. La reducción del área geográfica no necesariamente resulta en mayor éxito de identificación, especialmente para centros de diversificación o regiones con alta diversidad de especies.

**Key words:** BRONX, identificación de especímenes basada en caracteres, código de barras genético, IUCN, clado menos inclusivo, complejidad lingüística, *matK*, *rbcL*, calidad de la secuencia, identificación de especímenes basada en árboles.

### **Abstract**

DNA barcoding constitutes a fundamental tool for species identification, especially for highly diverse regions with few floristic studies. Here, we characterize and evaluate the plant core barcoding regions, *matK* and *rbcL*, to identify the 25 conifer species from the state of Hidalgo, Mexico, including 10 species under various threat categories. Sequence quality, linguistic complexity and the presence of the barcode gap was estimated. Two methods were compared for successful species identification: BRONX and the least inclusive clade. We generated 89 sequences for *matK* and 93 for *rbcL*. *matK* exhibited the highest haplotype diversity and nucleotide diversity (II), including six indels. The combination of both barcode regions resulted in 21 unique haplotypes. Higher sequence quality and linguistic complexity was observed in *rbcL*. Every diagnosable species had the barcode gap. All specimens were assigned unambiguously to family and genus, regardless of the marker or method employed. The analysis of

*matK* through BRONx produced the highest species identification success (44%). Despite the low specimen identification, it will be possible to establish local management, conservation and monitoring projects for at least half of the threatened species, in the absence of taxonomically diagnostic morphological characters. The low divergence between closely related species can result from the slow rate of molecular evolution of the core barcoding markers or from hybridization or incomplete lineage sorting. Similar identification success is expected for groups with comparable life history traits as conifers in regions with similar size and taxonomic diversity as the state of Hidalgo. A reduction of the geographic area will not necessarily translate into higher identification success, especially for high-diversity regions and centers of diversification.

**Key words:** BRONX, character-based specimen identification, DNA barcoding, IUCN, least inclusive clade, linguistic complexity, *matK*, *rbcL*, sequence quality, tree-based specimen identification.

## 2.1 Introducción

Las coníferas se identifican primeramente mediante claves taxonómicas basadas en morfología (estructuras externas) y anatomía (estructuras internas); además de diversas características tales como hábito de crecimiento, forma de la copa, tallo, características de la corteza y la madera, patrones de ramificación, ramillas y brotes, posición de hojas, apariencia, así como órganos reproductivos (conos de polen y semillas) y rasgos del polen, ocasionalmente se utiliza información sobre su distribución geográfica, los rasgos de su historia de vida y características ecológicas, el uso de dichas claves requiere cierto grado de experiencia para la interpretación de caracteres morfológicos en individuos maduros (Roxanne Steele y Chris Pires, 2011). Los conos ovulados son particularmente importantes para la identificación de especies en géneros tales como *Pinus* (Eckenwalder, 2009; D. Gernandt, Willyard, Syring, y Liston, 2011; David S. Gernandt y Pérez-De La Rosa, 2014). Las especies al interior de los géneros idealmente se identifican mediante caracteres vegetativos (hojas, tallos y raíces), debido a que los caracteres de las estructuras reproductivas generalmente son temporales o se producen en bajas cantidades cada año después de la madurez sexual (Eckenwalder, 2009).

La identificación de especies es fundamental para la implementación de proyectos de conservación, monitoreo y manejo de recursos forestales. Los inventarios de biodiversidad son el primer paso necesario para establecer políticas de conservación de especies amenazadas, medir el impacto ambiental de proyectos de desarrollo económico, así como para la exploración de nuevos productos naturales (Cardinale et al., 2012). La identificación de especímenes es difícil en ausencia de caracteres morfológicos diagnósticos particularmente de estructuras reproductivas (e.g., semillas o plántulas). En muchos casos es casi imposible identificar material tal como fragmentos de órganos vegetales, cuando los especímenes han sido parcial o totalmente procesados

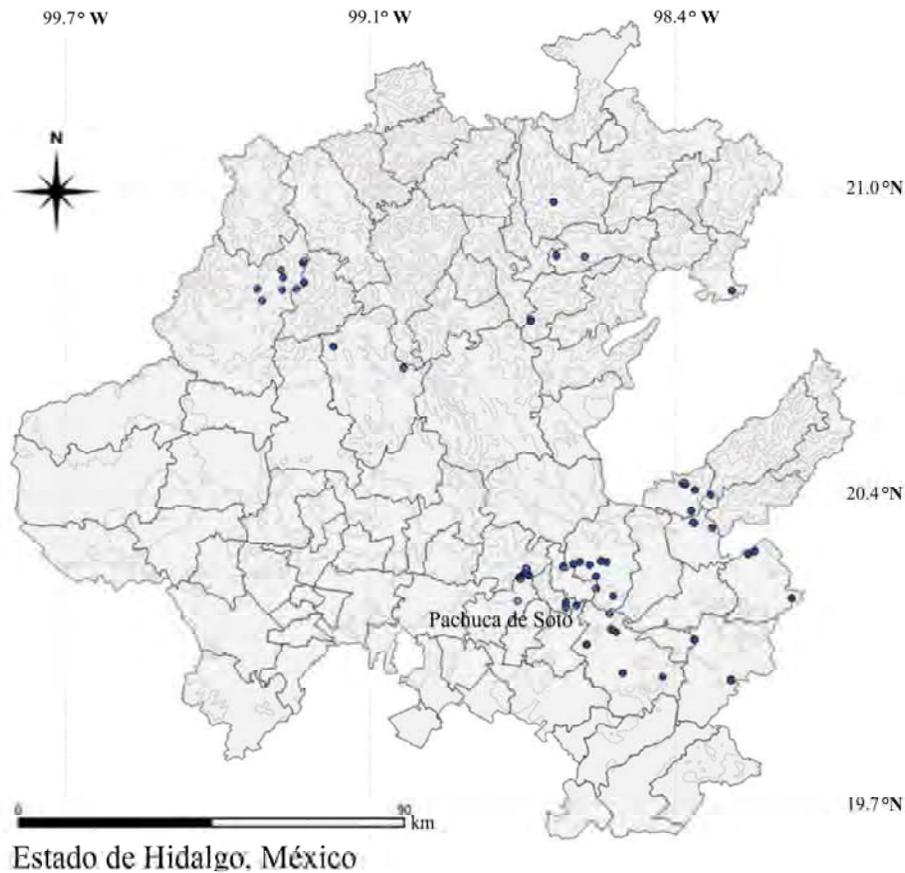
[e.g., en heces, mezclas de plantas en suplementos herbales o té (Hollingsworth, 2007; Steele y Pires, 2011)]. En situaciones como las anteriores los códigos de barras genéticos constituyen una herramienta de diagnóstico alternativa, para lograr la identificación taxonómica de las muestras (Hebert et al., 2003; Maia et al., 2012). Los códigos de barras genéticos aspiran a asignar especímenes a nivel de especie mediante una o pocas regiones genómicas cortas y estandarizadas (Hebert et al., 2003; Hollingsworth, 2011).

La identificación de especies de coníferas es sumamente importante debido a su abundancia, además de que aproximadamente 25% de la diversidad de coníferas enfrenta actualmente peligro de extinción (Armenise et al., 2012; FAO, 2010; Farjon y Page, 1999).

En este estudio se establece la biblioteca de referencia para la identificación de las coníferas del estado de Hidalgo mediante los códigos de barras genéticos centrales en plantas (*matK* y *rbcL*; CBOL, 2009), evaluando el éxito de identificación de especies mediante dos métodos -El clado menos inclusivo (Little y Stevenson, 2007) y BRONX (Little, 2012). Los objetivos específicos de esta investigación fueron: 1) Establecer un listado actualizado de especies de coníferas en el estado de Hidalgo mediante el análisis morfológico de especímenes en campo y en herbario, 2) Establecer una biblioteca de referencia de las regiones centrales del código de barras genético (*matK* y *rbcL*) para la identificación taxonómica de las coníferas que se distribuyen de manera natural en el estado, 3) Evaluar en las secuencias generadas la calidad de la secuencia, la complejidad lingüística y la presencia del gap del código de barras genético y 4) Comparar el desempeño de dos métodos de análisis para la identificación de especímenes a nivel de familia, género y especie: El clado menos inclusivo y BRONX.

## 2.2 Método

Se muestrearon de uno a ocho individuos reproductivamente maduros de las 25 especies de coníferas del estado de Hidalgo (Cuadro 2.1, 2.2; Figura 2.1). Los ejemplares fueron depositados en el Herbario Nacional de México (MEXU).



**Figura 2.1.** Localidades muestreadas. Distribución geográfica de las especies de coníferas muestreadas (círculos azules) en diversas localidades del estado de Hidalgo, México. Para el mapeo de las georreferencias se utilizó el Sistema de Coordenadas Geográficas WGS84.

**Cuadro 2.1.** Coníferas nativas del estado de Hidalgo, México. Esta lista de especies actualizada es resultado de trabajo de campo y reportes previos.

<b>Familia</b>	<b>Género</b>	<b>Especie</b>	
Cupressaceae	<i>Callitropsis</i>	<i>Callitropsis lusitanica</i> (Mill.) D. P. Little	
		<i>C. benthamii</i> (Endl.) D. P. Little	
	<i>Juniperus</i>	<i>Juniperus deppeana</i> Steud.	
		<i>J. flaccida</i> Schltdl.	
		<i>J. angosturana</i> R. P. Adams	
		<i>J. monticola</i> Martínez	
	<i>Taxodium</i>	<i>Taxodium huegelii</i> C. Lawson	
Pinaceae	<i>Abies</i>	<i>Abies guatemalensis</i> Rehder	
		<i>A. hidalgensis</i> Debreczy, I. Rácz & Guizar	
		<i>A. religiosa</i> (Kunth) Schltdl. & Cham.	
	<i>Pinus</i>	<i>Pinus ayacahuite</i> Ehrenb. ex Schltdl.	
		<i>P. cembroides</i> Zucc.	
		<i>P. devoniana</i> Lindl.	
		<i>P. greggii</i> Engelm. ex Parl.	
		<i>P. hartwegii</i> Lindl.	
		<i>P. leiophylla</i> Schiede ex Schltdl. & Cham.	
		<i>P. montezumae</i> Lamb.	
		<i>P. oocarpa</i> Schiede	
		<i>P. patula</i> Schltdl. & Cham.	
		<i>P. pinceana</i> Gordon	
		<i>P. pseudostrobus</i> Lindl.	
		<i>P. teocote</i> Schltdl. & Cham.	
		<i>Pseudotsuga</i>	<i>Pseudotsuga menziesii</i> var. <i>glauca</i> (Bleissn.) Franco

Podocarpaceae *Podocarpus* *Podocarpus matudae* Lundell

Taxaceae *Taxus* *Taxus globosa* Schltdl.

---

**Cuadro 2.2.** Familias, géneros y especies de coníferas del estado de Hidalgo incluidos en este estudio. Para las coordenadas geográficas se utilizó el Sistema de Coordenadas Geográficas WGS84. Los especímenes se depositaron en el Herbario Nacional de México (MEXU) y las secuencias fueron depositadas en GenBank.

<b>Taxon</b>	<b>Localidad</b>	<b>Latitud</b>	<b>Longitud</b>	<b>Especimen voucher</b>	<b>Número de GenBank <i>matK/ rbcL</i></b>
<b>Cupressaceae</b>					
<i>Callitropsis lusitanica</i>	Omitlán	20.198333 N	98.635833 W	DSG722	MH763605/ MH778503
<i>C. lusitanica</i>	Omitlán	20.198333 N	98.635833 W	DSG723	MH763606/ MH778504
<i>C. lusitanica</i>	Omitlán	20.198333 N	98.635833 W	DSG724	MH763607/ MH778505
<i>C. lusitanica</i>	Xochicoatlán	20.71843 N	98.69822 W	SHL147	MH547036/ MH734088
<i>C. lusitanica</i>	Zimapán	20.86888 N	99.21278 W	SHL119	MH577801/ MH734089
<i>C. benthamii</i>	Acaxochitlán	20.232380 N	98.209330 W	SHL40	MH577802/ MH734090
<i>C. benthamii</i>	Agua Blanca	20.357100 N	98.306017 W	SHL82	MH577803/ MH734091
<i>C. benthamii</i>	Pachuca de Soto	20.181944 N	98.723056 W	DSG717	MH763608/ MH778506
<i>Juniperus deppeana</i>	Mineral del Monte	20.108333 N	98.628611 W	DSG417	MH763609/ MH778507
<i>J. deppeana</i>	Singuilucan	19.959300 N	98.61165 W	SHL148	MH577804/ MH718315
<i>J. deppeana</i>	Singuilucan	20.027618 N	98.581537 W	SHL50	MH577805/ MH718316

<i>J. deppeana</i>	Singuilucan	20.056300 N	98.516217 W	SHL19	MH577806/ ---
<i>J. flaccida</i>	El Cardonal	20.68068 N	99.1358 W	SHL127	MH577807/ MH718317
<i>J. flaccida</i>	Huasca de Ocampo	20.211106 N	98.551337 W	SHL34	MH577808/ MH734043
<i>J. flaccida</i>	Huasca de Ocampo	20.211387 N	98.551661 W	SHL33	MH577809/ MH734044
<i>J. flaccida</i>	Zimapán	20.783883 N	99.304750 W	SHL113	MH577810/ MH734045
<i>J. angosturana</i>	Zimapán	20.809883 N	99.314633 W	SHL126	MH577811/ MH734092
<i>J. angosturana</i>	Zimapán	20.809967 N	99.314667 W	SHL124	MH577812/ MH734093
<i>J. monticola</i>	Pachuca de Soto	20.179444 N	98.711111 W	DSG719	MH782656/ MH778508
<i>Taxodium huegelii</i>	El Cardonal	20.68383 N	99.13542 W	SHL129	MH587626/ ---
<i>T. huegelii</i>	Huasca de Ocampo	20.203189 N	98.576575 W	SHL26	MH587627/ MH734046
<i>T. huegelii</i>	Huasca de Ocampo	20.210620 N	98.538186 W	SHL35	MH587628/ MH734047
<i>T. huegelii</i>	Ixmiquilpan	20.4805 N	99.22192 W	SHL112	--- / MH734048
<i>T. huegelii</i>	Santiago Tulantepec	20.037995 N	98.342659 W	SHL24-1	MH587629/ MH734049
<b>Pinaceae</b>					
<i>Pinus ayacahuite</i>	Agua Blanca de Iturbide	20.367778 N	98.342222 W	DSG1039	MH763610/ MH778509

<i>P. ayacahuite</i>	Mineral del Chico	20.196111 N	98.716944 W	DSG413	AY497257/ AY497221
<i>P. ayacahuite</i>	Zimapán	20.86652 N	99.21527 W	SHL121	MH587630/ MH734050
<i>P. cembroides</i>	Cardonal	20.683889 N	99.145833 W	DSG444	MH763611/ MH778510
<i>P. cembroides</i>	Zimapán	20.851530 N	99.263050 W	SHL36	MH587631/ MH734051
<i>P. devoniana</i>	Huasca de Ocampo	20.205000 N	98.612778 W	DSG677	MH763612/ MH778511
<i>P. devoniana</i>	Huasca de Ocampo	20.205556 N	98.612500 W	DSG521	MH763613/ MH778512
<i>P. devoniana</i>	Huasca de Ocampo	20.210556 N	98.597500 W	DSG449	MH763614/ MH778513
<i>P. greggii</i>	Calnali	20.87853 N	98.58818 W	SHL140	MH587632/ MH734052
<i>P. greggii</i>	Calnali	20.883733 N	98.649467 W	SHL142	MH587633/ MH734053
<i>P. greggii</i>	Calnali	20.89645 N	98.62782 W	SHL141	--- / MH734054
<i>P. greggii</i>	Molango	20.74158 N	98.70667 W	SHL136	--- / MH734055
<i>P. greggii</i>	Molango	20.741667 N	98.706667 W	DSG524	MH763615/ MH778514
<i>P. greggii</i>	Tlanchinol	20.98395 N	98.6413 W	SHL137	MH587634/ MH734056
<i>P. greggii</i>	Tlanchinol	21.001667 N	98.655556 W	DSG506	MH778516/ MH778517
<i>P. greggii</i>	Zimapán	20.835556 N	99.258611 W	DSG426	AY497282 / MH778515
<i>P. hartwegii</i>	Singuilucan	19.965278 N	98.501111 W	DSG616	MH778518/ MH782638

<i>P. hartwegii</i>	Singuilucan	19.97075 N	98.471900 W	SHL149	MH587635/ MH734057
<i>P. hartwegii</i>	Zimapán	20.824444 N	99.211667 W	DSG613	MH778519/ MH782639
<i>P. leiophylla</i>	Cuautepec de Hinojosa	19.950129 N	98.260438 W	SHL22	MH587636/ MH734058
<i>P. leiophylla</i>	Singuilucan	19.972120 N	98.43863 W	SHL153	--- / MH734059
<i>P. leiophylla</i>	Zimapán	20.810000 N	99.229722 W	DSG611	MH778520/ MH782640
<i>P. leiophylla</i>	Unknown	0.000000 N	0.000000 W	DSG619	MH778521/ MH782642
<i>P. montezumae</i>	Real del Monte	20.111944 N	98.605556 W	DSG416	AY497269/ MH782641
<i>P. montezumae</i>	Singuilucan	19.955800 N	98.412800 W	SHL155	MH587637/ MH734060
<i>P. montezumae</i>	Singuilucan	19.977920 N	98.449900 W	SHL150	MH791438 / MH734061
<i>P. oocarpa</i>	Calnali	20.882222 N	98.587500 W	DSG711	KY963969/ KY963969
<i>P. oocarpa</i>	Calnali	20.882230 N	98.58745 W	SHL138	MH712864/ MH734062
<i>P. oocarpa</i>	Xochicoatlán *	20.738370 N	98.70768 W	SHL146	MH712865/ MH734063
<i>P. oocarpa</i>	Xochicoatlán *	20.738683 N	98.707383 W	SHL145	MH712866/ MH734064
<i>P. patula</i>	Acaxochitlán	20.129444 N	98.126944 W	DSG1038	MH778522/ MH782644
<i>P. patula</i>	Huasca de Ocampo	20.133056 N	98.523333 W	DSG408	MH778523/ MH782645
<i>P. patula</i>	Huasca de Ocampo	20.151074 N	98.560618 W	SHL30	MH712867/ MH734065

<i>P. patula</i>	Molango	20.743020 N	98.70855 W	SHL144	MH712868/ MH734066
<i>P. patula</i>	Zacualtipán	20.591270 N	98.62047 W	SHL135	--- / MH734067
<i>P. patula</i>	Zimapán	20.866530 N	99.22188 W	SHL122	--- /MH734068
<i>P. pinceana</i>	El Cardonal	20.635200 N	98.990800 W	SHL53	MH791439 / MH734069
<i>P. pinceana</i>	Eloxochitlán	20.635833 N	98.990278 W	DSG440	MH778524/ MH782646
<i>P. pinceana</i>	Eloxochitlán	20.637222 N	98.989444 W	DSG441	MH778525/ MH782647
<i>P. pseudostrobus</i>	Apulco,Metepec	20.295278 N	98.343333 W	DSG411	--- / MH782648
<i>P. pseudostrobus</i>	Apulco,Metepec	20.296389 N	98.345556 W	DSG1040	--- / MH782649
<i>P. pseudostrobus</i>	Cuautepec de Hinojosa	19.950630 N	98.261535 W	SHL25	MH734040/ MH791436
<i>P. pseudostrobus</i>	Metepec	20.282347 N	98.313911 W	SHL130	MH712869/ MH734070
<i>P. pseudostrobus</i>	Zimapán	20.866480 N	99.229030 W	SHL123	MH718318/ MH734071
<i>P. pseudostrobus</i>	Zimapán	0.000000 N	0.000000 W	DSG631	MH778526/ MH782650
<i>P. teocote</i>	Agua Blanca de Iturbide	20.321944 N	98.349722 W	DSG410	MH778527/ MH782651
<i>P. teocote</i>	Cuautepec de Hinojosa	19.948221 N	98.260256 W	SHL20	MH718319/ MH734072
<i>P. teocote</i>	Huasca de Ocampo	20.178629 N	98.561708 W	SHL29	MH718320/ MH734073
<i>P. teocote</i>	Metepec	20.281770 N	98.315650 W	SHL131	--- / MH734074

<i>P. teocote</i>	Singuilucan	19.972150 N	98.438530 W	SHL154	--- / MH734075
<i>P. teocote</i>	Zacualtipán	20.591100 N	98.620330 W	SHL134	--- / MH734076
<i>Pseudotsuga menziesii</i>	Metepéc *	20.283607 N	98.303515 W	SHL133	MH734041/ MH734077
<i>P. menziesii</i>	Mineral del Chico	20.109444 N	98.628056 W	DSG418	MH778528/ ---
<i>P. menziesii</i>	Mineral del Chico	20.171710 N	98.730550 W	SHL75	MH718321/ MH734078
<i>P. menziesii</i>	Mineral del Monte	20.185000 N	98.719167 W	DSG774	MH778529/ MH782652
<i>Abies guatemalensis</i>	Zimapán	0.000000 N	0.000000 W	DSG626	MH778530/ MH782653
<i>A. guatemalensis</i>	Zimapán	0.000000 N	0.000000 W	DSG628	MH778531/ MH782654
<i>A. guatemalensis</i>	Zimapán	20.869722 N	99.213333 W	DSG625	MH778532/ MH782655
<i>A. guatemalensis</i>	Zimapán	20.871617 N	99.212700 W	SHL117	MH718322/ MH734079
<i>A. guatemalensis</i>	Zimapán	20.871620 N	99.212270 W	SHL116	--- / MH734080
<i>A. guatemalensis</i>	Zimapán	20.86889 N	99.212778 W	DSG624	EU269026/ ---
<i>A. hidalgensis</i>	Acaxochitlán	20.227117 N	98.222583 W	SHL86	MH718323/ MH734081
<i>A. religiosa</i>	Cuautepec de Hinojosa	0.000000 N	0.000000 W	ERI03	MH791440/ MH791437
<i>A. religiosa</i>	Singuilucan	19.977720 N	98.449500 W	SHL151	--- / MH734082
<i>A. religiosa</i>	Singuilucan	20.062800 N	98.527717 W	SHL18	MH718324/ MH734083

**Taxaceae**

<i>Taxus globosa</i>	Mineral del Chico	20.185500 N	98.716600 W	DSG772	MH778534/ ---
<i>T. globosa</i>	Mineral del Chico	20.193611 N	98.716111 W	DSG414	--- / MH778502
<i>T. globosa</i>	Mineral del Chico	20.216722 N	98.734389 W	APO02	MH734042/ MH734084
<i>T. globosa</i>	Singuilucan	20.095267 N	98.532650 W	SHL79	MH718325/ MH734085

**Podocarpaceae**

<i>Podocarpus matudae</i>	Agua Blanca de Iturbide	20.381967 N	98.363583 W	SHL106	MH718326/ MH734086
<i>P. matudae</i>	Agua Blanca de Iturbide	20.382823 N	98.363740 W	DSG1219	MH778533/ ---
<i>P. matudae</i>	Agua Blanca de Iturbide	20.383170 N	98.369280 W	SHL45	MH718327/ MH734087

---

\* Nuevas poblaciones registradas. --- = Secuencia no recuperada.

### 2.2.1 Extracción de DNA, PCR y alineamiento de secuencias

El DNA total fue extraído a partir de acículas con CTAB utilizando un protocolo modificado (Doyle y Doyle, 1987). Las secuencias para *matK* y *rbcL* fueron generadas siguiendo protocolos de PCR y secuenciación previamente publicados (David S. Gernandt, Hernández-León, Salgado-Hernández, y Pérez de La Rosa, 2009). Los cebadores utilizados para amplificar *matK* fueron *matK2F-R* (5'-CGT ACT TTT ATG TTT ACA GGC TAA-3'; Wang, Tsumura, Yoshimaru, Nagasaka, y Szmidt, 1999) en combinación con *equisetum-F* (5'-ATA CCC CAT TTT ATT CAT CC-3'; RBG Kew), debido a problemas con la amplificación para algunas muestras se utilizaron cebadores adicionales para amplificar la región *matK*: *GymF1A* (5'-ATY GYR CTT TTA TGT TTA CAR GC-3') en combinación con *GymR1A* (5'-TCA YCC GGA RAT TTT GGT TCG-3'; Y. Li, Gao, Poudel, Li, y Forrest, 2011); para amplificar *rbcL* se utilizaron los cebadores *rbcLa-F* (5'-ATG TCA CCA CAA ACA GAG ACT AAA GC-3'; Levin et al., 2003) y *rbcLajf634-R* (5'-GAA ACG GTC TCT CCA ACG CAT-3'; Fazekas et al., 2008).

Las secuencias fueron ensambladas, editadas y recortadas para excluir sitios de los cebadores mediante Sequencher ver. 4.8 (Gene Codes, Inc., Ann Arbor, Michigan) y BioEdit ver. 7.2.3 (Hall, 1999). Las secuencias fueron verificadas para identificar codones de terminación y desfases del marco de lectura utilizando TranslatorX (Abascal, Zardoya, y Telford, 2010).

### 2.2.2 Análisis de datos

La asignación de los especímenes a nivel de género fue confirmada mediante la comparación de las secuencias con la base de secuencias de GenBank mediante BLAST 2.2.26 (Altschul, Gish, Miller, Myers, y Lipman, 1990). Los aciertos con un

valor  $E$  de  $10^{-20.0}$  o menores se conservaron y se analizaron. El valor  $E$  o valor esperado es un parámetro que describe el número de aciertos que uno esperaría observar solo por azar cuando se realiza una búsqueda en una base de datos de tamaño determinado; el valor  $E$  disminuye exponencialmente conforme aumenta el puntaje (Score,  $S$ ) del empalme. Un valor de  $E$  igual a 1 puede ser interpretado como que se espera observar un empalme en una base de datos determinada solamente por azar. Entre más cercano sea el valor  $E$  a cero, el empalme será más "significativo", sin embargo se debe tener en cuenta que dicho valor es dependiente de la longitud de la secuencia de consulta, por lo que secuencias más cortas tendrán mayor probabilidad de estar presentes en la base de datos únicamente por azar, con valores bajos de  $E$  (NCBI Resource Coordinators, 2017).

Se realizaron alineamientos para las secuencias de las dos regiones en MAFFT 7 con el método de refinamiento iterativo FFT-NS-i (Kato, Asimeno, y Toh, 2009; Kato, Misawa, Kuma, y Miyata, 2002), los indels se codificaron usando 'codificación simple de indels' como se encuentra implementado en 2matrix (Salinas y Little, 2014; Simmons y Ochoterena, 2000), se obtuvieron árboles óptimos para las matrices resultantes en TNT 1.5 para verificar que no existieran contaminantes adicionales (Goloboff y Catalano, 2016).

Las secuencias fueron depositadas en las bases de datos Barcode of Life Datasystems (BOLD) y GenBank. La variación de las secuencias se estimó mediante DnaSP (Librado y Rozas, 2009).

La calidad de la secuencia fue determinada a partir de archivos CAF (Common Assembly Format), usando el Barcode quality index (Little, 2010a). El valor de límite ( $q$ ) se estableció como 30 y la cobertura esperada ( $x$ ) como dos. La cobertura es el

número de veces que se ha secuenciado una región (profundidad de secuenciación; Illimina, 2018). La longitud de la secuencia consenso (c) se estableció como la longitud observada (Little, 2010b).

La complejidad lingüística se calculó mediante COMPLEX 6.1.0 (Rice, Longden, y Bleasby, 2000) con tamaño de la ventana establecida como 100, el tamaño del paso establecido como una base, tamaño mínimo del patrón igual a tres pares de bases y tamaño máximo del patrón igual a seis bases. Se buscaron secuencias con homopolímeros con una longitud igual o mayor a ocho pares de bases.

La correlación entre la calidad de la secuencia y la complejidad lingüística y entre la calidad de la secuencia y la frecuencia de homopolímeros fueron estimados mediante el coeficiente de correlación de Spearman (Spearman, 1904). La diferencia estadística entre la calidad de la secuencia, complejidad lingüística y la frecuencia del gap del código de barras genético, se estimó por marcador utilizando la distribución binominal y una prueba de *t* de dos colas (Welch, 1947). El método BM se utilizó para realizar una corrección por pruebas múltiples en R (Benjamini y Hochberg, 1995). Las distancias pareadas en las regiones *matK*, *rbcL* y su combinación se determinaron para estimar el gap del código de barras genético (Meyer y Paulay, 2005). Las distancias genéticas pareadas se calcularon mediante el uso de *distances.pl* (Little, 2011), el gap del código de barras genético se calculó comparando la máxima distancia genética pareada intraespecífica con la mínima interespecífica para cada especie (Meier, Zhang, y Ali, 2008).

La identificación de especímenes basada en caracteres fue evaluada mediante BRONX (Little, 2011, 2012), la base de datos de referencia se construyó a partir de todas las secuencias para cada marcador y combinación de marcadores, no se incluyeron

aquellas secuencias obtenidas de GenBank. Las secuencias fueron comparadas con la base de referencia, las especies fueron consideradas distintas si todas las secuencias de consulta de una especie determinada retornaban solo secuencias pertenecientes a dicha especie.

Para la identificación de especímenes basada en árboles las secuencias por marcador así como su combinación se alinearon mediante MAFFT 7 (Kato et al., 2009, 2002). *Ginkgo biloba* (GenBank número de acceso: JN867578.1, KP099648.1) fue utilizado para enraizar los árboles. Los indels fueron codificados mediante 'codificación simple de indels' (Simmons y Ochoterena, 2000), los marcadores se concatenaron usando 2matrix (Salinas y Little, 2014). Se utilizó TNT 1.5 (Goloboff y Catalano, 2016) para realizar las búsquedas de los árboles óptimos mediante una búsqueda heurística, con 1000 réplicas de adición de secuencias al azar, conservando dos árboles por réplica, el intercambio de ramas SPR (Subtree pruning and regrafting) fue seguido de intercambio de ramas TBR (Tree Bisection and Reconnection) y BB (Branch and Bound).

Se utilizó TNT 1.5 (Goloboff y Catalano, 2016) para realizar la búsqueda de árboles óptimos mediante una búsqueda heurística, usando 1000 réplicas de adición de secuencias al azar, conservando dos árboles por réplica, el intercambio de ramas por SPR fue seguido por TBR (BB) y 200 iteraciones de matraca (Nixon, 1999), perturbando el 8% de los caracteres por iteración (4% con mayor peso, 4% con menor peso; Little et. al. 2013). La discriminación de especies basada en árboles fue estimada aplicando el método del clado menos inclusivo, mediante el cual el clado que incluye la secuencia de consulta constituye la identificación asignada (Little y Stevenson, 2007). La secuencia de *matK* para *T. globosa* (accesión GenBank MH778534) se fusionó con la secuencia de *rbcL* de *T. globosa* (accesión GenBank MH778502) con la finalidad de

contar con más de una secuencia para dicha especie en el análisis concatenado, ambos individuos provienen de la misma población.

## 2.3 Resultados

En el cuadro 1 se presenta un listado actualizado de las especies de coníferas que se distribuyen de manera natural en el estado de Hidalgo. Durante el trabajo de campo se registraron nuevas poblaciones para *Pinus oocarpa* y *Pseudotsuga menziesii* (Cuadro 2).

### 2.3.1 Caracterización de las secuencias

La caracterización de las secuencias se presenta en el cuadro 3. Las secuencias de *matK* y *rbcL* incluyeron de 1 a 8 individuos por cada una de las 25 especies estudiadas. Se generaron 89 secuencias para *matK* y 93 para *rbcL*. El alineamiento completo (*matK+rbcL*) incluyó 79 secuencias con una longitud de 1, 321 pb. Se detectaron seis indels en la región *matK*, dos de seis pb y cuatro de tres pb. En *rbcL* los indels estuvieron ausentes. Se registraron 15 haplotipos para *rbcL*, 20 para *matK* y 21 para su combinación que corresponden a las 16 especies de la familia Pinaceae, siete a las siete especies de la familia Cupressaceae, dos a la única especie de la familia Taxaceae y uno corresponde a la única especie de Podocarpaceae.

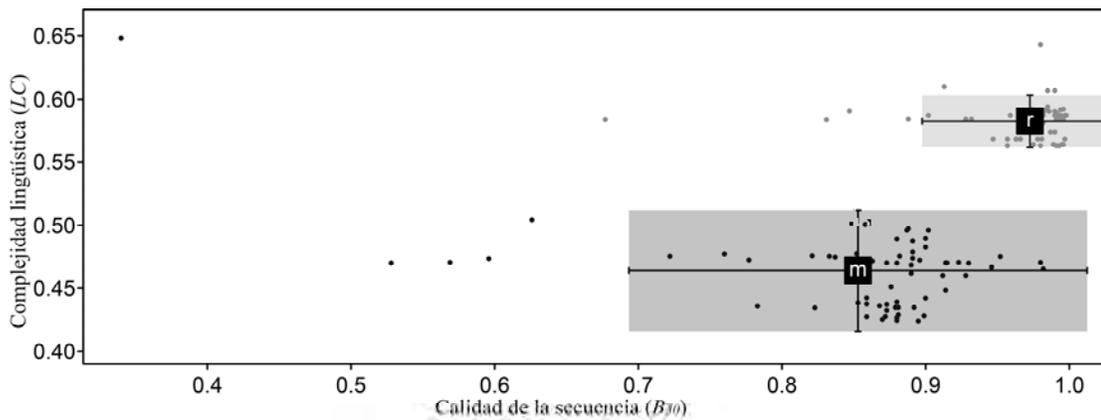
**Cuadro 2.3.** Características del alineamiento de secuencias para los códigos de barras genéticos.

Región	Longitud (pb)	Sv	SIP	DH	II	Indels
<i>matK</i>	797–806	737	737	0.910	0.344	6
<i>rbcL</i>	515	94	93	0.861	0.050	0

pb Pares de bases, Sv Sitios variables, SIP Sitios Informativos bajo Parsimonia, DH Diversidad haplotípica, II Diversidad nucleotídica.

### 2.3.2 Calidad de la secuencia

La calidad de la secuencia para la región *matK* fue de 0.342-0.984 (media= 0.844, rango intercuartil= 0.881-0.9), la calidad de la secuencia para la región *rbcL* fue de 0.466-1.0 (media=0.954, rango intercuartil = 0.961-0.995; Figura 2.2). La diferencia en la calidad de la secuencia entre regiones fue estadísticamente significativa ( $p= 0.05$ ;  $rbcL>matK$ ).



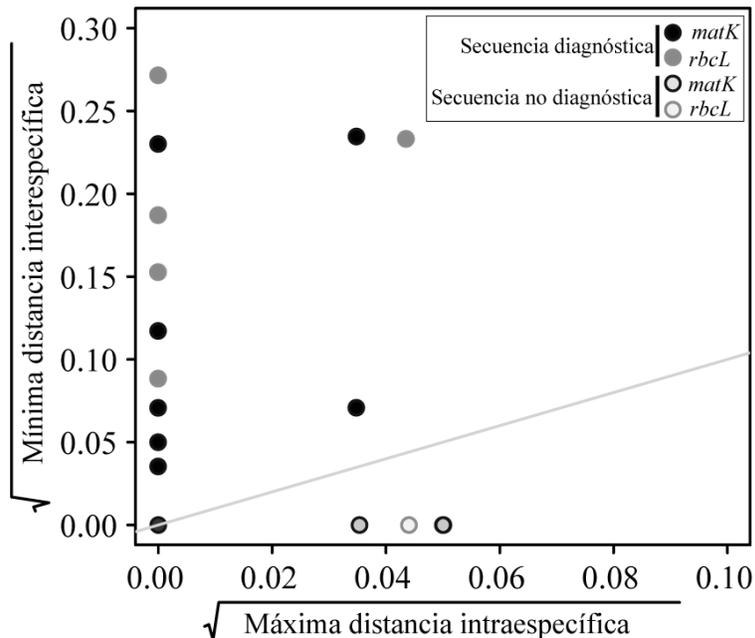
**Figura 2.2.** Calidad de la secuencia ( $B_{30}$ ) contra la Complejidad lingüística (LC) para las secuencias completas. Los círculos representan cada una de las secuencias individuales para *matK* (m; color negro) y *rbcL* (r; color gris claro). Los recuadros en color negro representan los promedios por marcador. Las barras de error indican tres desviaciones estándar.

### 2.3.3 Complejidad lingüística

La complejidad lingüística de la región *matK* fue de 0.424 a 0.648 (media = 0.463, rango intercuartil = 0.438-0.475), y para *rbcL* de 0.563 a 0.643 (media = 0.585; rango intercuartil = 0.581-0.589) (Figura 2.2). La diferencia en complejidad lingüística entre regiones fue estadísticamente significativa ( $p = 0.05$ ;  $rbcL>matK$ ). No se detectaron homopolímeros de ocho o mas pares de bases. No existió correlación entre la calidad de la secuencia y la complejidad lingüística para ninguna de las dos regiones ( $p=0.05$ ).

### 2.3.4 Gap del código de barras genético

La distancia interespecífica mínima para *rbcL* fue de 0.008 (*Pinus ayacahuite*, *P. cembroides* y *P. pinceana*) a 0.074 (*Podocarpus matudae*). Para la región *matK* esta fue de 0.001 (*P. greggii* y *P. montezumae*) a 0.198 (*P. matudae*). El gap del código de barras genético se reportó en la región *matK* para 11 (44%) de las 25 especies, 7 (28%) en *rbcL* y 11 (44%) para su combinación. Todas las especies diagnosticables presentaron el gap del código de barras genético (Figura 2.3). La diferencia en la frecuencia del gap del código de barras genético entre regiones fue estadísticamente significativa ( $p = 0.05$ ;  $matK > rbcL$ ). Para ninguna de las regiones se encontró una relación entre la presencia del gap del código de barras genético y el número de individuos muestreados por especies.

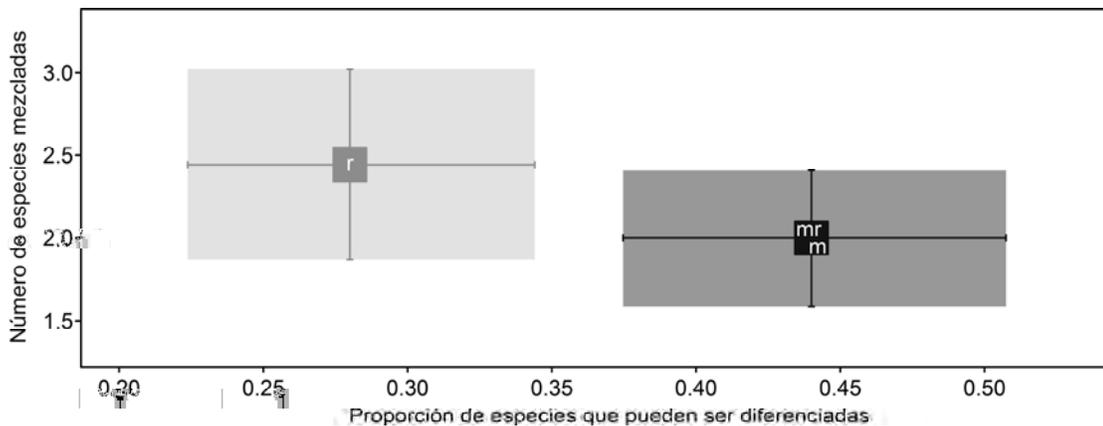


**Figura 2.3.** Comparación de la Máxima distancia intraespecífica y la Mínima distancia interespecífica. La línea gris representa igual variación intra e interespecífica, por encima de la cual se localizan las secuencias correspondientes con especies que presentan el gap del código de barras genético.

### 2.3.5 Discriminación de especies basada en caracteres

El análisis de las regiones *matK* y *rbcL*, en combinación o por separado resultaron en la asignación no ambigua de las secuencias de consulta a una de las cuatro familias y uno de los ocho géneros utilizando el algoritmo BRONX. La región *matK* permitió la identificación de los especímenes de 11 (44%) de las 25 especies totales, así como 6 (60%) de las diez especies en categoría de riesgo (Cuadro 4), *rbcL* resultó en la identificación no ambigua de siete (28%) de las 25 especies totales y cuatro (40%) de las diez especies en categoría de riesgo. El análisis del alineamiento concatenado (*matK* y *rbcL*), permitió la identificación no ambigua de 11 (44%) de las 25 especies totales, así como seis (60%) de las diez especies en categorías de riesgo (Cuadro 4).

No se observaron diferencias significativas entre las regiones con respecto a la identificación de especies ( $p = 0.05$ ). La media de especies mezcladas para la región *matK* fue de dos con un rango de entre 1-3, mientras que para la región *rbcL* la media fue de 2.53 con un rango de entre 1-4. Para la combinación de ambas regiones la media de especies mezcladas fue de dos con un rango de entre 1-3 (Figura 2.4, Cuadro 2.4).



**Figura 2.4.** Proporción de especies que pueden ser diferenciadas contra el número de especies mezcladas, para *matK* (m; color negro), *rbcL* (r; color gris claro) y combinación de marcadores (mr; color gris). Los recuadros en el centro indican el promedio de discriminación de especies. Las barras de error indican intervalos de confianza del 95%.

**Cuadro 2.4.** Variación diagnóstica de las regiones de los códigos de barras genéticos centrales en plantas analizadas mediante BRONX para las muestras con secuencias completas.

Especie	n	Éxito de discriminación			Código de barras genético multilocus	Especies mezcladas con mayor frecuencia (ordenadas por frecuencia)
		matK	rbcL	combinación		
<i>Abies guatemalensis</i> *	4	-	-	-	0	<i>A. hidalgensis</i> , <i>A. religiosa</i>
<i>A. hidalgensis</i> *	1	-	-	-	0	<i>A. guatemalensis</i> , <i>A. religiosa</i>
<i>A. religiosa</i>	2	-	-	-	0	<i>A. guatemalensis</i> , <i>A. hidalgensis</i>
<i>Callitropsis benthamii</i>	3	-	-	-	0	<i>C. lusitanica</i>
<i>C. lusitanica</i> *	5	-	-	-	0	<i>C. benthamii</i>
<i>Juniperus deppeana</i>	3	+	-	+	1	<i>J. angosturana</i>
<i>J. flaccida</i>	4	-	-	-	0	<i>J. monticola</i>
<i>J. angosturana</i> *	2	+	-	+	1	<i>J. deppeana</i>
<i>J. monticola</i> *	1	-	-	-	0	<i>J. flaccida</i>
<i>Pinus ayacahuite</i>	4	+	+	+	1	---
<i>P. cembroides</i>	2	+	+	+	1	---
<i>P. devoniana</i>	3	-	-	-	0	<i>P. hartwegii</i> , <i>P. pseudostrobus</i> , <i>P. montezumae</i>
<i>P. greggii</i> *	6	+	-	+	1	<i>P. leiophylla</i> , <i>P. oocarpa</i> , <i>P. patula</i> , <i>P. teocote</i>
<i>P. hartwegii</i>	3	-	-	-	0	<i>P. devoniana</i> , <i>P. pseudostrobus</i> , <i>P. montezumae</i>
<i>P. leiophylla</i>	3	-	-	-	0	<i>P. oocarpa</i> , <i>P. patula</i> , <i>P. teocote</i> , <i>P. greggii</i>
<i>P. montezumae</i>	3	+	-	+	1	<i>P. devoniana</i> , <i>P. hartwegii</i> , <i>P. pseudostrobus</i>
<i>P. oocarpa</i>	4	-	-	-	0	<i>P. leiophylla</i> , <i>P. patula</i> , <i>P. teocote</i> , <i>P. greggii</i>
<i>P. patula</i>	4	-	-	-	0	<i>P. leiophylla</i> , <i>P. oocarpa</i> , <i>P. teocote</i> , <i>P. greggii</i>
<i>P. pinceana</i> *	3	+	+	+	1	---
<i>P. pseudostrobus</i>	4	-	-	-	0	<i>P. devoniana</i> , <i>P. hartwegii</i> , <i>P. montezumae</i>
<i>P. teocote</i>	3	-	-	-	0	<i>P. leiophylla</i> , <i>P. oocarpa</i> , <i>P. patula</i> , <i>P. greggii</i>

<i>Podocarpus matudae</i> *	2	+	+	+	1	---
<i>Pseudotsuga menziesii</i> *	3	+	+	+	1	---
<i>Taxodium huegelii</i>	3	+	+	+	1	---
<i>Taxus globosa</i> *	3	+	+	+	1	---

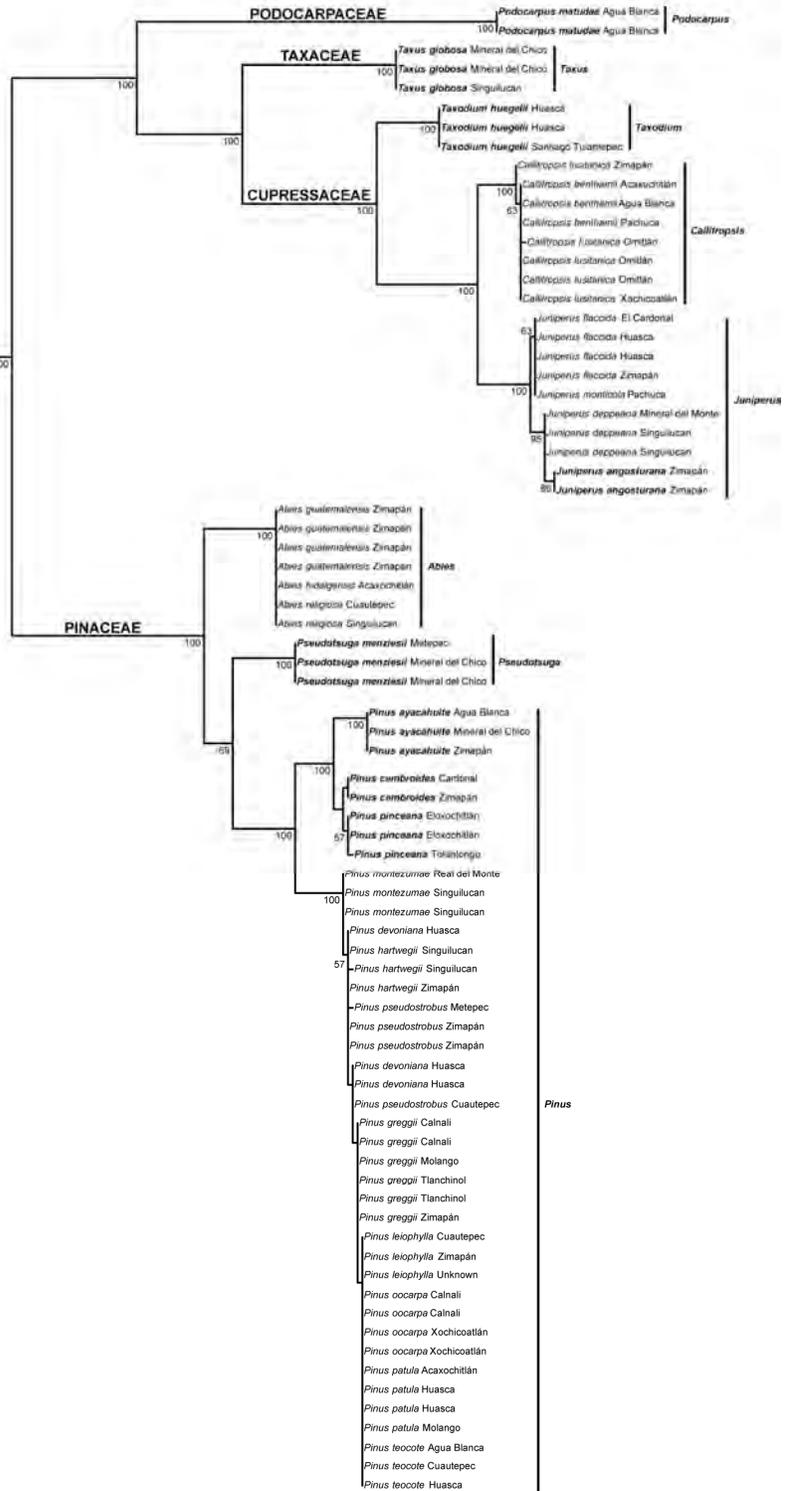
\* Especies en categorías de riesgo en el listado de la IUCN y/ o la NOM-059-SEMARNAT-2010. Éxito de discriminación + = identificación exitosa; - = identificación no exitosa. Gap del código de barras genético multilocus 0 = ausente; 1 = presente.

### 2.3.6 Identificación de especies basada en árboles filogenéticos

El análisis de parsimonia para la región *matK* resultó en 13 árboles igualmente parsimoniosos, el árbol de consenso estricto exhibió 29 nodos resueltos (Figura 2.5). Aplicando el método del clado menos inclusivo los especímenes fueron asignados de manera no ambigua a una de las cuatro familias y uno de los ocho géneros con altos valores de soporte jackknife (62-100%), además ocho (32%) de las 25 especies pueden ser asignadas a especie así como cinco (50%) de las 10 especies en categorías de riesgo: *Juniperus angosturana*, *Pinus pinceana*, *Podocarpus matudae*, *Pseudotsuga menziesii* y *Taxus globosa* (Figura 2.5).

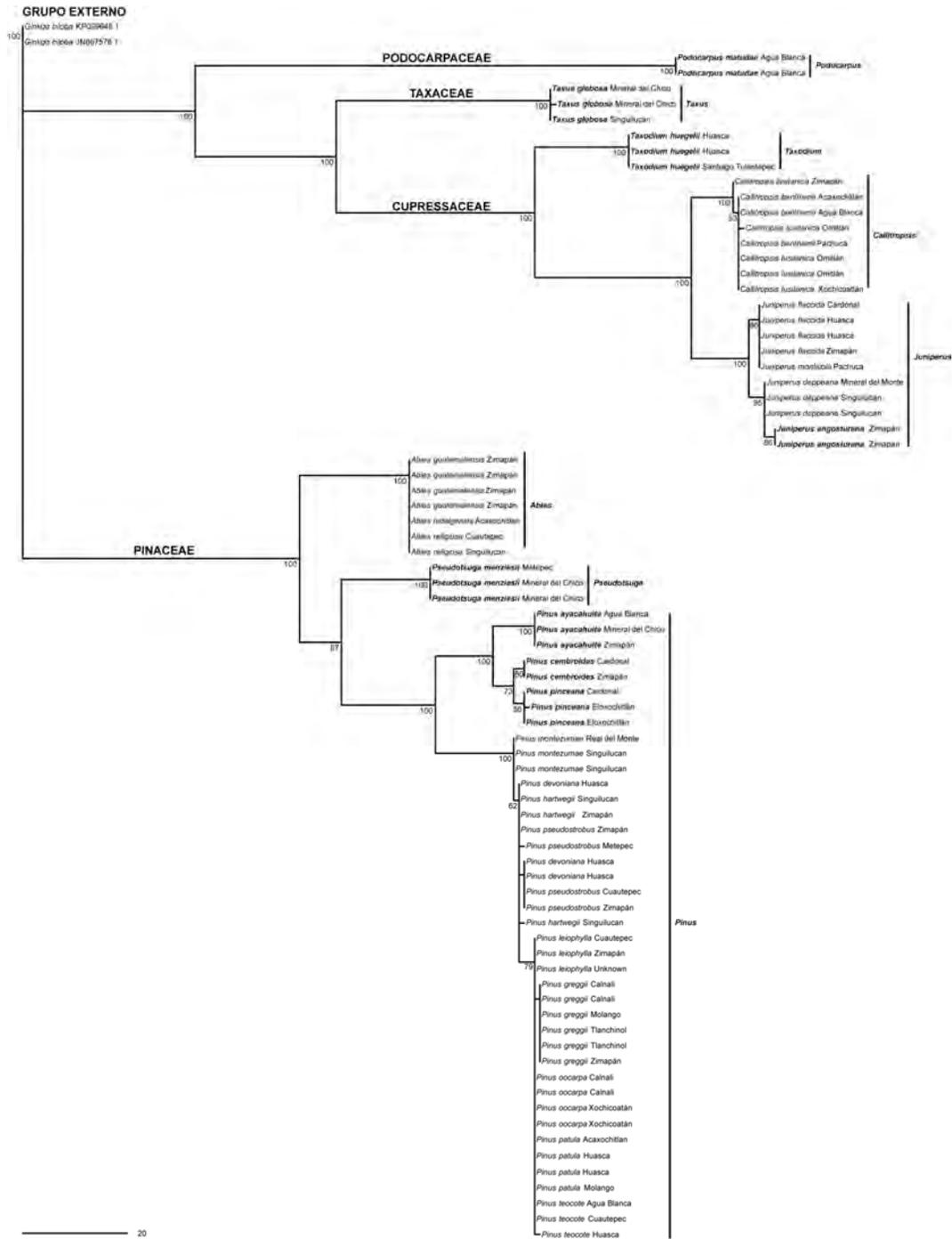
El análisis de parsimonia para la región *rbcL* resultó en tres árboles igualmente parsimoniosos, el árbol de consenso estricto tuvo 21 nodos resueltos (Figura 2.6). Aplicando el método del clado menos inclusivo, los especímenes pueden ser asignados de manera no ambigua a una de las cuatro familias y uno de los ocho géneros con valores altos de soporte jackknife (59-100%), además es posible la asignación de los especímenes de manera no ambigua a siete (28%) de las 25 especies estudiadas y 4 (40%) de las 10 especies en categorías de riesgo: *Pinus pinceana*, *Podocarpus matudae*, *Pseudotsuga menziesii* y *Taxus globosa* (Figura 2.6).

El análisis del alineamiento concatenado (*matK* y *rbcL*) mediante parsimonia generó 14 árboles igualmente parsimoniosos, el árbol de consenso estricto exhibió 28 nodos resueltos (Figura 2.7). Mediante el método del clado menos inclusivo, cada espécimen puede ser asignado de manera no ambigua a familia y género con valores altos de soporte jackknife



**Figura 2.5.** Relaciones filogenéticas resultantes del análisis de la región *matK*. Consenso estricto de 13 árboles más parsimoniosos (L=475; IC= 0.845; IR=0.984). Los valores de soporte Jackknife mayores al 50%, se muestran próximos a los nodos. Las especies que pueden ser identificadas con el criterio del clado menos inclusivo se muestran en negritas.





**Figura 2.7.** Relaciones filogenéticas resultantes del análisis de la región (*matK* + *rbcL*). Consenso estricto de 14 árboles más parsimoniosos (L = 610; IC = 0.826; IR = 0.979). Los valores de soporte Jackknife mayores al 50%, se muestran próximos a los nodos. Las especies que pueden ser identificadas con el criterio del clado menos inclusivo se muestran en negritas.

(86-100%), así como a ocho (32%) de las 25 especies y 6 (60%) de las 10 especies en categorías de riesgo: *Juniperus angosturana*, *Pinus greggii*, *P. pinceana*, *Podocarpus matudae*, *Pseudotsuga menziesii* y *Taxus globosa*. Los ejemplares de *Pinus greggii* son recuperados como un grupo monofilético pero con valores de soporte jackknife menores al 50%, en un análisis de Neighbor Joining los especímenes de *P. greggii* se recuperan como un grupo parafilético por lo que aplicando el método del clado menos inclusivo en el sentido estricto no se consideraría como una identificación exitosa (Figura 2.7).

Independientemente de la región o su combinación, el método del clado menos inclusivo permite la identificación de las siguientes especies que se encuentran en categorías de riesgo: *Pinus pinceana*, *Podocarpus matudae*, *Pseudotsuga menziesii*, y *Taxus globosa* (Figs 2.5-2.7).

## **2.4 Discusión**

### **2.4.1 Caracterización de las secuencias**

La región *matK* permite la asignación de especímenes al mismo número de especies que su combinación con la región *rbcL*, por lo que en términos de tiempo y dinero es más conveniente utilizar únicamente la región *matK* para la identificación de especies de coníferas en el estado de Hidalgo. Por otro lado, *rbcL* presentó las mayores tasas de amplificación mediante la PCR y exhibió la mayor calidad de la secuencia con la única excepción de *Podocarpus matudae*, secuencia que presentó mayor calidad para la región *matK*.

Doce especies cercanamente relacionadas del género *Pinus* no presentaron variación en las regiones de los códigos de barras genéticos analizadas—haciendo imposible su identificación mediante las regiones de los códigos de barras genéticos *matK* y *rbcL*.

Algunas regiones adicionales del genoma del plastidio tales como *trnH-psbA* y *ycf1*, pudieran incrementar el éxito de identificación a nivel de especie (Armenise et al., 2012; Dong et al., 2015; Hernández-León, Gernandt, Pérez de la Rosa, y Jardón-Barbolla, 2013; Knopf, Schulz, Little, Stützel, y Stevenson, 2012), sin embargo debido a que otra fuente potencial de error es la presencia de haplotipos compartidos resultado de introgresión o sorteo de linajes incompleto, (Ledig, 1998; Little y Stevenson, 2007; Little, 2011; Syring et al., 2007), se deben incluir fuentes independientes de información (Hernández-León et al., 2013), tales como datos de secuencias del genoma nuclear o mitocondrial, e.g. nrITS1, nrITS2, o *NEEDLY* intron 2 (Chiou, Yen, Fang, Chen, y Lin, 2007; Hernández-León et al., 2013; Knopf, Schulz, Little, Stützel, y Stevenson, 2012; Ledig, 1998; D.-Z. Li et al., 2011; Maia et al., 2012), la independencia de estas regiones es particularmente importante para grupos en los que se ha documentado la presencia de haplotipos compartidos entre especies (e.g. *Pinus*), en dichos casos los datos no pueden distinguir definitivamente entre el mecanismo causal subyacente de reciente introgresión contra polimorfismo ancestral seguido por sorteo de linajes incompleto (Little, Knopf, y Schulz, 2013). Por otro lado existe evidencia de que el uso de múltiples loci no mejora de manera importante la habilidad de los códigos de barras genéticos para discriminar a nivel de especie (CBOL Plant Working Group et al., 2009; Simeone, Piredda, Papini, Vessella, y Schirone, 2013). El uso de cebadores o fuentes de caracteres adicionales no debe ser sujeto de discusión cuando el objetivo final es la identificación de especies y no el encontrar un marcador universal con la variación precisa que permita la identificación a nivel de especie con técnicas exclusivamente moleculares, lo cual cada vez se hace menos viable de acuerdo con la evidencia disponible (Tanabe y Toju, 2013).

Se registraron 21 haplotipos multilocus para las 25 especies totales, por lo que los haplotipos no estuvieron distribuidos de manera uniforme entre las especies. Se registró variación intraespecífica para *matK* en *C. lusitanica* DSG724 (Norte de Omitlán), *C. lusitanica* SHL119 (Zimapán), *Pinus hartwegii* DSG616 (Singuilucan), *P. pseudostrabus* SHL130 (Cerro Metepec), *P. pinceana* DSG441 (Eloxochitlán), *P. menziesii* DSG418 (Peñas Cargadas). Para el caso de la región *rbcL* se observó variación intraespecífica en *P. teocote* SHL29 (Huasca de Ocampo) y *T. globosa* SHL79 (Singuilucan; Cuadro 2, Figs 2.5-2.7).

La variación intraespecífica y la incongruencia entre la identificación de especímenes con caracteres morfoanatómicos y la circunscripción de especies mediante el código de barras genético requiere de mayor investigación particularmente en grupos de especies tales como el género *Pinus* caracterizados por bajos niveles de variación intraespecífica en las regiones *matK* y *rbcL*, debido a que esta variación puede ser indicio de la presencia de especies crípticas, introgresión o sorteo de linajes incompleto (Little et al., 2013).

#### **2.4.2 Calidad de la secuencia, complejidad lingüística y gap del código de barras genético**

Las secuencias generadas en este estudio (*matK* y *rbcL*) fueron de alta calidad de acuerdo con los estándares de los códigos de barras genéticos, debido a que en promedio el 90% de las posiciones nucleotídicas presentaron un valor de calidad mayor a QV30 (Hanner, 2009). La calidad de la secuencia para *rbcL* fue consistentemente mayor que la de *matK*, en comparación con estudios previos en la familia Podocarpaceae 97% de las posiciones nucleotídicas fueron de alta calidad para la región *matK* (Little et al., 2013); mientras que para la familia Arecaceae 93% de las posiciones

núcleotídicas fueron de alta calidad en secuencias para la región *rbcL* (Jeanson, Labat, y Little, 2011).

Para la complejidad lingüística en la familia Podocarpaceae, se han reportado valores de media para *rbcL* de 0.584 (rango intercuartil = 0.577–0.590) y de 0.443 (rango intercuartil = 0.437–0.449) para la región *matK* (Little et al., 2013), lo cual es consistente con los resultados obtenidos en este estudio, donde se observó menor complejidad lingüística para la región *matK*, probablemente asociada con mayores restricciones funcionales y estructurales.

Un muestreo insuficiente tanto a nivel intraespecífico e interespecífico puede crear un artefacto de gap del código de barras genético (Meyer y Paulay, 2005). La probabilidad de que los porcentajes del gap del código de barras disminuya a escala regional es menor en comparación con una escala global. El gap del código de barras genético se registró en 44% de las especies a partir del análisis de la región *matK*, 28% para la región *rbcL* y 44% para la combinación de ambas regiones. En plantas terrestre se ha logrado el 72% de éxito de identificación de especies en presencia del gap del código de barras genético (CBOL Plant Working Group et al., 2009). En especies de la familia Podocarpaceae se encontró el gap del código de barras genético en 39.1% de las especies a partir del análisis de la región *matK* y en 34% de las especies para la región *rbcL* (Little et al., 2013). En contraste con reportes anteriores la diferencia de la presencia del gap del código de barras genético entre regiones fue estadísticamente significativa ( $p = 0.05$ ;  $matK > rbcL$ ; Little et al., 2013). La presencia del gap del código de barras genético no presentó correlación con el número de ejemplares por especie como ha sido previamente reportado para especies de la familia Podocarpaceae (Little et al., 2013).

### 2.4.3 Identificación de especies basada en caracteres

Los especímenes fueron asignados de manera no ambigua a una de las cuatro familias y uno de los ocho géneros mediante BRONX, independientemente del análisis separado o combinado de las regiones *matK* y *rbcL*. Este es el primer reporte disponible del desempeño de BRONX a escala local, sin embargo algunos métodos equivalentes (usando grupos no delimitados geográficamente y no necesariamente las mismas porciones de los códigos de barras genéticos) mediante el análisis de *matK* y *rbcL* lograron un 46.3% de éxito de identificación para especies de la familia Podocarpaceae (Little et al., 2013), 71.42% en *Parnassia* (Parnassiaceae; Yang, Wang, Möller, Gao, y Wu, 2012), 68.75% en *Primula* sect. *Proliferae* de China (Primulaceae; Yan, Hao, Hu, y Ge, 2011), 59.45% para *Euphorbia* de Madagascar (Euphorbiaceae; Aubriot, Lowry, Cruaud, Couloux, y Haevermans, 2013), 51.85% para Caryoteae (Arecaceae; Jeanson et al., 2011) y 35.71% para *Actaea* (Ranunculaceae; Baker, Stevenson, y Little, 2012), en esta investigación se obtuvo el 44% de éxito de identificación de especies. El análisis de la región *matK*, mediante el algoritmo BRONX permite identificar de manera adicional (respecto de *rbcL*) cuatro especies adicionales (*Juniperus angosturana*, *J. deppeana*, *Pinus greggii* y *P. montezumae*), dos de las cuales se encuentran en categoría de riesgo (Cuadro 4).

Es muy importante destacar que analizando las regiones *matK* y *rbcL* mediante el algoritmo BRONx, se logran identificar especímenes correspondientes a 6 de las 10 especies (60%) en categorías de riesgo, lo cual implica que por ejemplo para *Pinus pinceana* una especie que además de estar en peligro de extinción es endémica de México, es posible el establecimiento de planes de manejo, conservación y monitoreo efectivos a nivel local, basados en la identificación mediante los códigos de barras genéticos.

Las especies en categoría de riesgo que no pueden ser identificadas mediante los códigos de barras centrales (*Abies guatemalensis*, *Abies hidalgensis*, *Callitropsis benthamii* y *Callitropsis lusitanica*), morfológicamente también representan un desafío y su estatus específico aun es controversial y requiere de clarificación. *Abies guatemalensis* y *Abies hidalgensis* no pueden ser diferenciadas de *Abies religiosa* mediante los códigos de barras genéticos centrales, sin embargo dependiendo de la información adicional disponible y el nivel de identificación requerida, es posible separar estas especies utilizando su distribución conocida, *Abies guatemalensis* cuenta con reportes de distribución en el municipio de Zimapán, *A. hidalgensis* en los municipios de Acaxochitlán y Agua Blanca, mientras que *A. religiosa* esta distribuida en Cuauhtepic de Hinojosa, Mineral del Chico, Mineral del Monte, Pachuca de Soto, Singuilucan, Tepeapulco y Tulancingo de Bravo. *Callitropsis lusitanica* no puede ser diferenciada de *C. benthamii* una especie cercanamente relacionada mediante los códigos de barras genéticos, pero puede ser identificada con características morfológicas tales como sus ramillas en brotes aplanados, que surgen de solo dos filas de hojas y sus conos ovulados que generalmente exhiben tres pares de escamas (Eckenwalder, 2009).

#### **2.4.4 Identificación de especies basada en árboles filogenéticos**

Las relaciones filogenéticas recuperadas entre los taxa corresponden con esquemas filogenéticos previos (Farjon, 2008; David S. Gernandt et al., 2005; Lin, Huang, Wu, Hsu, y Chaw, 2010; Little, 2006; Mao, Hao, Liu, Adams, y Milne, 2010; Semerikova y Semerikov, 2014). Los especímenes fueron asignados de manera no ambigua a una de las cuatro familias y uno de los ocho géneros, independientemente de la región analizada o su combinación (*matK*, *rbcL*, o *matK+rbcL*).

Mediante la aplicación del método del clado menos inclusivo es posible identificar exitosamente el 32% de las especies mediante el análisis de la región *matK* así como el 28% con la región *rbcL*. Un estudio regional para las 25 especies de coníferas de Italia logró un 54% de identificación exitosa mediante la combinación de *matK* y *rbcL*, sin registrar variación intraespecífica (Armenise et al., 2012). Dicho estudio utilizó métodos de distancias genéticas, similitud y filogenia para evaluar la identificación de los taxa y comprobar la monofilia de las especies, obteniendo 100% de identificación de especímenes con la combinación de las regiones *rbcL+trnH-psbA*, sin embargo es muy importante notar que en dicho estudio el 100% está calculado únicamente con base en especies con múltiples especímenes (e.g. *Abies alba* es la única especie de cuatro totales para el género que cumple con dicha condición y al tomar en cuenta la p-distancia mínima interespecífica y la p-distancia máxima intraespecífica es posible la identificación de dicha especie, lo anterior es tomado como el 100% de éxito de identificación para el género *Abies*); por otra parte resultado de su análisis de máxima verosimilitud este 100% está sesgado a especies con múltiples individuos por especie (debido a que es imposible aplicar el criterio del clado menos inclusivo a especies con un solo individuo), por lo que el resto de las especies con un único espécimen por especie no son tomadas en cuenta para calcular el porcentaje de éxito de identificación, si se toman en cuenta el resto de las especies el porcentaje de éxito disminuye.

Dong et al., (2015) obtuvieron 16% de identificación para especímenes de *Pinus* utilizando la región *matK* y métodos basados en árboles (NJ, UPGMA). En otro estudio la combinación de *matK + rbcL* resultó en 8% de identificación de especímenes para la subsección *Trifoliae* del género *Pinus*, uno de los grupos de mayor importancia ecológica y económica de árboles, el porcentaje de identificación aumenta a 46.9%

cuando se incorporan las regiones *trnH-psbA*, *trnD-trnY-trnE*, *ycf1*, (Hernández-León et al., 2013).

En una evaluación de árboles tropicales el éxito de identificación fue de entre 35.3% y 42.9% para *matK* y de entre 28.5% y 43.6% para *rbcL* (Huang, Ci, Conran, y Li, 2015; Kang, Deng, Zang, y Long, 2017; Tripathi et al., 2013).

Los métodos de identificación basados en árboles (e.g. el clado menos inclusivo) han presentado un menor desempeño en comparación con los métodos basados en caracteres [e.g. BRONX (Little et al., 2013)] lo cual resulta consistente con los resultados de este estudio; además de que para poder aplicar el criterio del clado menos inclusivo se requieren de al menos dos secuencias para poder comprobar la monofilia de la especie. En comparación con el método del clado menos inclusivo, el método BRONX permitió identificar de manera adicional a *Juniperus deppeana*, *Pinus greggii* y *P. montezumae*. Con el método Neighbor joining es posible obtener porcentajes de identificación similares a los obtenidos con el algoritmo BRONX, sin embargo la interpretación del árbol resultante puede ser difícil debido a que algunas especies son recuperadas como parafiléticas (e.g. *Juniperus deppeana*, *Pinus greggii*). Los especímenes morfológicamente asignados a *P. montezumae* fueron consistentemente identificados utilizando BRONX, a pesar de reportes previos de hibridación para esta especie con *P. hartwegii* y *P. pseudostrobus* para otras localidades de México (Delgado et al., 2007; Matos y Schaal, 2000).

Las tasas de diversificación bajas reportadas para complejos de especies del género *Abies* parecen ser una constante a nivel mundial y pudieran explicar por qué *A. hidalgensis*, *A. guatemalensis* y *A. religiosa* no pueden ser identificadas mediante los códigos de barras genéticos (É. Aguirre-Planter et al., 2012; Jaramillo-Correa et al.,

2008). Estas tres junto con el resto de las especies mesoamericanas comparten un ancestro común y su divergencia y especiación están determinadas por deriva genética y aislamiento durante los periodos cálidos interglaciales, con niveles históricos de flujo genético limitados, solo algunas poblaciones localizadas en los límites de los rangos de su distribución en la faja volcánica transmexicana han demostrado ser divergentes, mientras que el resto de las poblaciones forman un único agrupamiento cuando se analizan mediante marcadores del genoma del plastidio (Jaramillo-Correa et al., 2008).

Los principales centros de diversidad del género *Abies* se encuentran al oeste de América del Norte y Mesoamérica con aproximadamente 14 especies (Semerikova y Semerikov, 2014). En términos generales la clasificación de *Abies* es difícil debido a sus altos niveles de variación morfológica y su habilidad para formar híbridos en zonas de contacto secundario (Liu, 1971). Además, el estatus a nivel de especies de muchos taxa de *Abies* aun es controversial (Strandby, Christensen, y Sørensen, 2009). Por otro lado el número de especies de coníferas observadas en los subtrópicos puede ser el resultado tanto de una especiación acelerada debido a aislamiento histórico incrementado o simplemente un artefacto producido por una plasticidad fenotípica alta que ha confundido a los expertos cuando describen especies basados únicamente en variación morfológica (e.g. De Kroon et al., 2005; Prada et al., 2008). Las regiones del genoma del cloroplasto en *Abies flinckii*, *A. guatemalensis*, *A. hickelii* y *A. religiosa* presentan correspondencia con la taxonomía de las especies, así como una mayor estructura filogeográfica, posiblemente relacionada con la baja capacidad de dispersión del polen en *Abies*, la mitocondria por su parte presenta una fuerte estructura geográfica (Jaramillo-Correa et al., 2008).

La observación de que *A. guatemalensis* (*A. hidalgensis*), *A. hickelii* y *A. religiosa* compartan un bagaje citoplasmático (cpDNA, mtDNA) y nuclear común

constituye evidencia de posible hibridación y sugiere que pueden representar estados incipientes de especiación borrosa por episodios pasados de introgresión durante los diferentes periodos glaciales (E. Aguirre-Planter, Furnier, y Eguiarte, 2000; É. Aguirre-Planter et al., 2012; Jaramillo-Correa et al., 2008). En términos generales se considera como un prerequisite para la hibridación un territorio compartido por especies cercanamente relacionadas, debido a que la recombinación entre linajes distantes sugiere la presencia de flujo genético histórico así como el origen híbrido de algunas especies (Ma, Szmidt, y Wang, 2006; Semerikova y Semerikov, 2014).

En contraste con lo anterior el uso de regiones variables del genoma del cloroplasto (*atpF-atpI* y *rbcL-psaI*, *trnQ-trnR*, *rps18-rp120* y *trnL-trnF*) recuperan a *Abies guatemalensis* y *A. religiosa*, como polifiléticas por É. Aguirre-Planter et al., (2012) en clados separados con altos valores de soporte bootstrap (Semerikova y Semerikov, 2014).

*Callitropsis benthamii* y *C. lusitanica* son otras dos especies que no pueden ser separadas mediante los códigos de barras genéticos centrales en plantas; su estatus a nivel de especie aun permanece en discusión entre diferentes autores (Eckenwalder, 2009; Farjon, 2010; Little, 2006), mientras que otros tales como Eckenwalder, 2009 consideran que *C. benthamii* es una variedad de *C. lusitanica* (*Cupressus lusitanica* var. *benthamii* (Endlicher) Carrière), estas dos especies al parecer forman parte de un complejo que representan un problema taxonómico significativo y persistente, enraizado en el noroeste de México, un área que requiere de mayor estudio (Earle, 2015). Los haplotipos compartidos entre especies pueden reflejar caracteres morfológicos mínimos e insuficientes para justificar el estatus específico de algunas especies, lo cual genera un desacuerdo sobre el estatus taxonómico entre diferentes autores (Hernández-León et al., 2013).

De acuerdo con Little (2006) con base en un análisis de evidencia total, la combinación de cambios de estado de caracteres morfológicos homoplásicos permite diferenciar a *C. benthamii* del resto de los *Cupressus*. De las regiones que son variables para *C. benthamii*, *C. lusitanica* y otras especies cercanamente relacionadas es muy importante secuenciar múltiples individuos por especie a lo largo de su rango de distribución para tener una idea de la variación intra e interespecífica, así como la presencia de posibles haplotipos compartidos (Little, 2006).

Es posible que la variación en algunas regiones permita la separación de *C. benthamii* y *C. lusitanica* al incorporar múltiples individuos por especie por ejemplo la región *tRNA-Lys (trnK)* (Little, 2006). La variación que existe en la región *trnL* del genoma del cloroplasto no ayuda a distinguir entre *C. benthamii* y *C. lusitanica*, solo existe variación intra específica en un espécimen de *C. lusitanica* con un indel de 5 pares de bases. Mediante tres secuencias de la región nuclear *NEEDLY* intron 2; dos correspondientes a *Cupressus lusitanica* y una a *C. benthamii* se logra observar al menos tres sustituciones de posible utilidad para diferenciar entre estas dos especies, las secuencias tienen un aproximado de 710 pares de bases de longitud. En la región nuclear *nrITS* existe variación que pudiera ser útil para separar las dos especies, sin embargo asumiendo que las identificaciones sean correctas existe una gran cantidad de variación intra específica (Little, 2006).

Las especies cercanamente relacionadas de *Pinus* subsección *Ponderosae* (*Pinus devoniana*, *P. hartwegii*, y *P. pseudostrobus*), así como de *Pinus* subsección *Australes* (*P. leiophylla*, *P. oocarpa*, *P. patula*, y *P. teocote*; Gernandt et al., 2005) no pudieron ser identificadas debido a la ausencia de variación en *matK* y *rbcL*.

Previamente ha sido reportado que *P. hartwegii*, *P. montezumae* y *P. pseudostrobus* de la sección *Trifoliae* subsección *Ponderosae* exhiben una muy baja divergencia de secuencias y haplotipos compartidos (Delgado et al., 2007; Hernández-León et al., 2013; Matos y Schaal, 2000) y que la reciente diversificación vía especiación alopátrica seguida de contacto secundario pudo haber promovido flujo genético interespecífico en *Pinus* sección *Trifoliae*, en este grupo existe una gran variación morfológica y molecular y existen casos de haplotipos compartidos entre diferentes clados particularmente en la subsección *Ponderosae* (Hernández-León et al., 2013), posiblemente como resultado de débiles barreras intrínsecas e introgresión (Delgado et al., 2007; Duffield, 1952).

De acuerdo con Hernández-León et al., 2013 en el género *Pinus* la incongruencia entre los límites de especies y los árboles de genes pueden ser resultado de introgresión o sorteo de linajes incompleto, pero pueden acentuarse por conceptos de especie inadecuados o bajos niveles de variación molecular.

Independientemente del método utilizado para el análisis separado o combinado de las regiones *matK* y *rbcL* las especies únicas de las familias Podocarpaceae y Taxaceae en el área de estudio fueron totalmente diagnosticables. Descartando conceptos de especie inadecuados en el género *Pinus*, la principal razón de la no identificación fue la falta de variación, en este caso la secuenciación de marcadores variables del genoma del cloroplasto, con primers universales para coníferas pudiera aumentar el éxito de identificación, sin embargo en grupos con evidencia de haplotipos compartidos entre especies como consecuencia de hibridación con subsecuente introgresión o sorteo de linajes incompleto debido a reciente surgimiento, (e.g. especies cercanamente relacionadas del género *Pinus*), favorecidos por rasgos de vida tales como grandes tamaños efectivos de población, bajas tasas de mutación y largos tiempo

generacionales (Ledig, 1998; Little y Stevenson, 2007; Syring et al., 2007) , lo más recomendable es comparar y completar datos del genoma del cloroplasto con fuentes independientes de información tales como distribución geográfica, caracteres morfo-anatómicos o secuencias provenientes del genoma nuclear o mitocondrial (Hernández-León et al., 2013). En presencia de variación compartida entre taxa provocada por polimorfismo ancestral y/o introgresión no es posible la identificación de ejemplares utilizando un solo marcador de los códigos de barras genéticos (Little, 2011).

Cupressaceae y Pinaceae son las familias representadas con la mayor cantidad de especies en el estado de Hidalgo, en Cupressaceae se observó una mayor proporción de especies diagnosticables aun en especies cercanamente relacionadas (e.g. *Juniperus flaccida* y *J. monticola*; (Mao et al., 2010)) en contraste con la familia Pinaceae en la que las especies cercanamente relacionadas del género *Pinus*, presentaron una menor proporción de especies diagnosticables.

Los códigos de barras genéticos no serán capaces de proporcionar una identificación rápida y confiable en clados que han evolucionado rápidamente como resultado de radiaciones masivas con mínima divergencia de secuencias, situación que también es problemática para la identificación morfológica de este tipo de grupos. Por otro lado se debe contemplar que una divergencia genética no constituye una prueba suficiente para postular la existencia de una especie y debe ser corroborada por otro tipo de datos. Las especies cercanamente relacionadas que son difíciles de identificar por medios tradicionales también son genéticamente similares y pueden pasar por desapercibidas por el código de barras genético como única aproximación (Wiemers y Fiedler, 2007).

La lenta evolución de las coníferas es evidenciada por la abundancia de eventos de hibridación natural o doméstica, sugiriendo aislamiento reproductivo incompleto entre especies (Prunier et al., 2016). En este estudio la falta de variación interespecífica fue la causa principal para la no identificación de especímenes de especies cercanamente relacionadas de los géneros *Abies*, *Callitropsis*, *Juniperus* y *Pinus*, al interior de estos géneros se han reportado previamente haplotipos compartidos para *rbcL* (Armenise et al., 2012; Hernández-León et al., 2013).

Los especímenes que no pueden ser identificados, pueden ser asignados a uno de los siguientes grupos de especies **Grupo I:** *Cupressus lusitánica*; *C. benthamii*, **Grupo II:** *Abies hidalgensis*, *A. guatemalensis* o *A. religiosa*, **Grupo III:** *P. devoniana*, *P. hartwegii* o *P. pseudostrabus*, **Grupo IV:** *Pinus leiophylla*, *P. oocarpa*, *P. patula* o *P. teocote*, **Grupo V:** *J. flaccida* o *J. monticola*. La identificación a nivel de especies al interior de los grupos sería posible utilizando caracteres morfológicos o geográficos adicionales. La identificación basada exclusivamente los códigos de barras centrales en plantas debe aceptar una mayor proporción de error en comparación con la identificación en la que se utilizan múltiples fuentes de caracteres. El asignar especímenes a haplotipos del plastidio o a especie de manera preliminar constituye un avance metodológico y tiene un gran valor intrínseco para una amplia gama de aplicaciones prácticas de la identificación mediante los códigos de barras genéticos, particularmente cuando implican estadios de vida o fragmentos de individuos, tales como la identificación de etapas tempranas de regeneración, el cambio y la dinámica de la vegetación, la interacción planta-animal, la interacción patógeno-hospedero, la diferenciación entre germoplasma nativo e introducido, la regulación de mercado y certificación de madera, productos secundarios y germoplasma, el etiquetado de especies útiles o tóxicas, la identificación de tala ilegal, entre otros (Armenise et al.,

2012; Bergsten et al., 2012; Chase et al., 2005; Hernández-León et al., 2013; Wiemers y Fiedler, 2007).

A mayor distancia filogenética se espera un mayor éxito de identificación como resultado de una mayor variación interespecífica, menor potencial de hibridación y por lo tanto un menor flujo genético entre las especies. Para las especies en las que se ha registrado hibridación la identificación con base en el DNA del plastidio corresponderá con la especie que ha proporcionado el polen. Los genes del cloroplasto se transmitirán y homogeneizarán en especies cercanamente emparentadas siempre y cuando se encuentren geográficamente cercanas para formar híbridos, en dichos casos no será posible la asignación de especímenes aun utilizando el genoma del cloroplasto completo.

La base de datos de referencia generada en este estudio tiene suficiente profundidad de muestreo y variación de secuencias como para producir identificaciones razonablemente precisas a escala local. Estudios previos sugieren el uso de la distribución geográfica de la especie como un factor adicional de eliminación (e.g. in BOLD Systems) que puede resultar en un mayor éxito de identificación (Bergsten et al., 2012; Little y Stevenson, 2007). La disminución del área geográfica en animales ha resultado en mayor éxito de identificación utilizando la región mitocondrial CO1 pero esto no aplica necesariamente para todas las especies (Bergsten et al., 2012), particularmente aquellas cercanamente relacionadas que por simpatria presenten flujo genético interespecífico y evidencia de hibridación con subsecuente introgresión o sorteo de linajes incompleto. Los resultados obtenidos en esta investigación, sugieren que es importante considerar que en centros de diversificación de especies (e.g. México para las especies del género *Pinus*), una reducción en el área geográfica no

necesariamente resultará en un mayor éxito de identificación mediante los códigos de barras genéticos.

Mediante el uso de una biblioteca de referencia regional se pretende evitar la variación intraespecífica total de la especie, reducir el número de especies de un mismo género, incluir especies evolutivamente distantes, genéticamente diferentes y sin haplotipos compartidos en los loci de los códigos de barras genéticos, facilitando así su identificación. A nivel planetario se espera que la identificación de especies mediante los códigos de barras genéticos constituya un mayor desafío, tomando en cuenta que la divergencia interespecífica mínima disminuye conforme aumenta la escala geográfica (Bergsten et al., 2012; Little y Stevenson, 2007). La aplicación de los códigos de barras genéticos debe limitarse a instancias en las que la base de datos de referencia contenga suficiente profundidad de muestreo y variación de secuencias para producir identificaciones razonablemente exactas (Little y Stevenson, 2007). Los estimados de variación interespecíficos y la proporción de especies no monofiléticas debe incrementar conforme incrementa el muestreo intraespecífico (Bergsten et al., 2012), por lo que en algunos grupo de coníferas (e.g. *Pinus* sección *Trifoliae*) un incremento en el muestreo intraespecífico a lo largo del rango de distribución geográfico de las especies, puede resultar en más casos de haplotipos compartidos entre especies (Hernández-León et al., 2013), particularmente en áreas donde se ha documentado hibridación con subsecuente introgresión o sorteo de linajes incompleto. Es altamente probable que los porcentajes de identificación obtenidos se mantendrán para el área de estudio considerando que se incluyeron todas las coníferas conocidas para el estado de Hidalgo con múltiples individuos por especie.

La aplicación de las regiones de los códigos de barras genéticos centrales en plantas para la identificación de especies arbóreas ha demostrado ser más difícil que

para floras no leñosas (CBOL Plant Working Group et al., 2009), tal como se evidencia en estudios en *Ficus*, *Quercus*, y *Salix* (H. Q. Li, Chen, Wang, y Xiong, 2012; Percy et al., 2014; Simeone et al., 2013) además de los resultados obtenidos en este estudio.

## **2.5 Conclusiones**

Esta investigación contribuye con un listado actualizado de especies de coníferas que se distribuyen de manera natural en el estado de Hidalgo, así como a aumentar el número de secuencias (biblioteca de referencia local), hasta el momento escasas para las regiones de los códigos de barras centrales en coníferas, incorporando un contexto regional de alta diversidad y centro de diversificación del género *Pinus*. Mediante la biblioteca de referencia generada en este estudio es posible establecer programas de manejo, conservación y monitoreo a escala local para al menos la mitad de las especies en categorías de riesgo.

La identificación mediante los códigos de barras genéticos a nivel regional, incluso por parte de no especialistas y en tiempos relativamente cortos, es de suma importancia para estudios de ecología de comunidades, estimados y distribución de la biodiversidad, el establecimiento de programas efectivos de restauración, manejo, aprovechamiento y conservación, control de comercio y certificación de especies, proyectos de conservación de la biodiversidad en vista del cambio climático global, particularmente para aquellas en categorías de riesgo, sobreexplotadas y con rangos de distribución limitados o tamaños poblacionales reducidos (Armenise et al., 2012). La dependencia de la escala sobre la identificación constituye un gran incentivo para llevar a cabo esfuerzos regionales con los códigos de barras genéticos tomando en cuenta que cada caso presentará condiciones particulares (Bergsten et al., 2012).

El uso de primers específicos, regiones adicionales o fuentes de caracteres no moleculares, no debe ser un impedimento en aquellos casos en los que la identificación precisa es el objetivo final. De acuerdo a los niveles de identificación requeridos para situaciones específicas y para grupos taxonómicos particulares, se puede decidir entre usar las regiones centrales de los códigos de barras genéticos o desarrollar códigos de barras genéticos específicos a partir del análisis del genoma del plastidio completo lo cual pudiera mejorar nuestra habilidad para distinguir a nivel de especie o incluso a nivel de población entre taxa de plantas cercanamente relacionados (X. Li et al., 2015).

Existe mayor probabilidad de que las especies lejanamente relacionadas exhiban mayores tasas de éxito de identificación mediante los códigos de barras genéticos únicos o pareados y que las especies cercanamente relacionadas sin hibridación o sorteo de linajes incompleto puedan ser diagnosticables analizando la variación en el genoma completo del plastidio, sin embargo es importante tomar en cuenta que para aquellas especies cercanamente relacionadas con evidencia de hibridación o sorteo incompleto de linajes, este no permitirá su identificación debido al efecto de los haplotipos compartidos. La secuenciación del genoma completo del plastidio implica mayor tiempo y costo si se compara con la secuenciación Sanger de las regiones de los códigos de barras genéticos, además de que su análisis puede ser complejo y difícil de estandarizar.

## **2.6 Recomendaciones**

Es necesaria mayor investigación sistemática para clarificar el estatus taxonómico de *A. hidalgensis* y promover esfuerzos de conservación a corto plazo. Las bibliotecas de referencia están aún muy lejos de estar completas, por lo que para aplicar la identificación a nivel de género o especie en investigaciones sobre biodiversidad, será

necesario acelerar el registro de secuencias de referencia, así como el uso de factores adicionales para lograr una aproximación más precisa a la clasificación de los especímenes particularmente para especies con registros previos de falta de variación molecular o haplotipos interespecíficos compartidos (Tanabe y Toju, 2013).

## 2.7 Literatura citada

Abascal, F., Zardoya, R., y Telford, M. J. (2010). TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*, 38, w7–w13. <http://doi.org/10.1093/nar/gkq291>

Aguirre-Planter, E., Furnier, G. R., y Eguiarte, L. E. (2000). Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany*, 87(3), 362–371.

Aguirre-Planter, É., Jaramillo-Correa, J. P., Gómez-Acevedo, S., Khasa, D. P., Bousquet, J., y Eguiarte, L. E. (2012). Phylogeny, diversification rates and species boundaries of Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. *Molecular Phylogenetics and Evolution*, 62, 263–274. <http://doi.org/10.1016/j.ympev.2011.09.021>

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., y Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–10. [http://doi.org/10.1016/S0022-2836\(05\)80360-2](http://doi.org/10.1016/S0022-2836(05)80360-2)

Armenise, L., Simeone, M. C., Piredda, R., y Schirone, B. (2012). Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *European Journal of Forest Research*, 131, 1337–1353. <http://doi.org/10.1007/s10342-012-0602-0>

- Aubriot, X., Lowry, P. P., Cruaud, C., Couloux, A., & Haevermans, T. (2013). DNA barcoding in a biodiversity hot spot: potential value for the identification of Malagasy *Euphorbia* L. listed in CITES Appendices I and II. *Molecular Ecology Resources*, *13*, 57–65.
- Baker, D. A., Stevenson, D. W., y Little, D. P. (2012). DNA barcode identification of black cohosh herbal dietary supplements. *Journal of AOAC International*, *95*, 1023–1034. <http://doi.org/10.5740/jaoacint.11-261>
- Benjamini, Y., y Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300. <http://www.jstor.org/stable/2346101>
- Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., ... Vogler, A. P. (2012). The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology*, *61*, 851–869. <http://doi.org/10.1093/sysbio/sys037>
- Cardinale, B., Duffy, J., y Gonzalez, A. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67. <http://doi.org/10.1038/nature11148>
- CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 12794–7. <http://doi.org/10.1073/pnas.0905845106>
- Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N., y Savolainen, V. (2005). Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*, 1889–1895. <http://doi.org/10.1098/rstb.2005.1720>

Chiou, S. J., Yen, J. H., Fang, C. L., Chen, H. L., y Lin, T. Y. (2007). Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers. *Planta Medica*, 73, 1421–1426. <http://doi.org/10.1055/s-2007-990227>

Coordinators, N. R. (2017). Database resources of the national center for biotechnology information. *Nucleic acids research*, 45 (Database issue), D12.

De Kroon, H., Huber, H., Stuefer, J. F., y Van Groenendael, J. M. (2005). A modular concept of phenotypic plasticity in plants. *New phytologist*, 166(1), 73-82.

Delgado, P., Salas-Lizana, R., Vázquez-Lobo, A., Wegier, A., Anzidei, M., Álvarez-Buylla, E. R., ... Piñero, D. (2007). Introgressive hybridization in *Pinus montezumae* Lamb and *Pinus pseudostrobus* Lindl. (Pinaceae): Morphological and molecular (cpSSR) evidence. *International Journal of Plant Sciences*, 168, 861–875. <http://doi.org/10.1086/518260>

Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., ... Zhou, S. (2015). *ycf1*, the most promising plastid DNA barcode of land plants. *Scientific Reports*, 5, 1-5. <http://doi.org/10.1038/srep08348>

Doyle, J. J., y Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11–15. <http://doi.org/10.2307/4119796>

Duffield, J. W. (1952). Relationships and species hybridization in the genus *Pinus*. *Zeitschrift für Forstgenetik und Forstpflanzenzuchtung*, 1(4), 93-100.

Earle, C. J. (2015). The gymnosperm database. Recuperado de: <https://www.conifers.org/> (accesado 16 Agosto 2018).

Eckenwalder, J. (2009). *Conifers of the world: the complete reference*. Timber Press.

- FAO. (2010). *Global Forest Resources Assessment 2010: Main report*. FAO Forestry Paper (Vol. 163).
- Farjon, A. (2008). *A natural history of conifers*. Portland, OR: Timber Press.
- Farjon, A. (2010). *A handbook of the world's conifers*. Leiden: Brill.
- Farjon, A., y Page, C. (1999). *Conifers: Status Survey and Conservation Action Plan*. Washington D.C.
- Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D., ... Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *Public Library of Science ONE*, 3, e2802. <http://doi.org/10.1371/journal.pone.0002802>
- Gernandt, D. S., Geada López, G., Ortiz García, S., y Liston, A. (2005). Phylogeny and classification of *Pinus*. *Taxon*, 54, 29–42. <http://doi.org/10.2307/25065300>
- Gernandt, D. S., Hernández-León, S., Salgado-Hernández, E., y Pérez de la Rosa, J. A. (2009). Phylogenetic relationships of *Pinus* subsection *Ponderosae* inferred from rapidly evolving cpDNA regions. *Systematic Botany*, 34, 481–491. <http://doi.org/10.1600/036364409789271290>
- Gernandt, D. S., Willyard, A., Syring, J., y Liston, A. (2011). The conifers (Pinophyta). In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, Genomics and Breeding of Conifers* (1st ed., pp. 1–39). USA: Science Publishers
- Gernandt, D. S., y Pérez-de La Rosa, J. A. (2014). Biodiversidad de Pinophyta (coníferas) en México. *Revista Mexicana de Biodiversidad*, 85, S126–S133. <http://doi.org/10.7550/rmb.32195>

Goloboff, P. A., y Catalano, S. A. (2016). TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* 32, 221–238. <http://doi.org/10.1111/cla.12160>

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98. <http://doi.org/citeulike-article-id:691774>

Hanner, R. (2009). Data Standards for BARCODE Records in INSDC (BRIs). Recuperado de: <http://www.barcoding.si.edu/PDF/Guidelines%20for%20non-CO1%20selection%20FINAL.pdf> (accessed 19 December 2017)

Hebert, P. D. N., Cywinska, A., Ball, S. L., y deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270, 313–321. <http://doi.org/10.1098/rspb.2002.2218>

Hernández-León, S., Gernandt, D. S., Pérez de la Rosa, J. A., y Jardón-Barbolla, L. (2013). Phylogenetic relationships and species delimitation in *Pinus* section *Trifoliae* inferred from plastid DNA. *Public Library of Science ONE*, 8, 1–14. <http://doi.org/10.1371/journal.pone.0070501>

Hollingsworth, P. M. (2007). DNA barcoding: potential users. *Policy*, 3, 44–47.

Hollingsworth, P. M. (2011). Refining the DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 108, 19451–19452. <http://doi.org/10.1073/pnas.1116812108>

Jaramillo-Correa, J. P., Aguirre-Planter, E., Khasa, D. P., Eguiarte, L. E., Piñero, D., Furnier, G. R., y Bousquet, J. (2008). Ancestry and divergence of subtropical montane forest isolates: Molecular biogeography of the genus *Abies* (Pinaceae) in southern

México and Guatemala. *Molecular Ecology*, 17, 2476–2490.  
<http://doi.org/10.1111/j.1365-294X.2008.03762.x>

Jeanson, M. L., Labat, J. N., y Little, D. P. (2011). DNA barcoding: A new tool for palm taxonomists?. *Annals of Botany*, 108, 1445–1451.  
<http://doi.org/10.1093/aob/mcr158>

Kang, Y., Deng, Z., Zang, R., y Long, W. (2017). DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. *Scientific Reports*, 7, 12564. <http://doi.org/10.1038/s41598-017-13057-0>

Katoh, K., Asimenos, G., y Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Methods in Molecular Biology*, 537, 39–64. [http://doi.org/10.1007/978-1-59745-251-9\\_3](http://doi.org/10.1007/978-1-59745-251-9_3)

Katoh, K., Misawa, K., Kuma, K., y Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30, 3059–3066. <http://doi.org/10.1093/nar/gkf436>

Knopf, P., Schulz, C., Little, D. P., Stützel, T., y Stevenson, D. W. (2012). Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. *Cladistics*, 28, 271–299.  
<http://doi.org/10.1111/j.1096-0031.2011.00381.x>

Ledig, F. (1998). Genetic variation in *Pinus*. In D. Richardson (Ed.), *Ecology and biogeography of Pinus*. Cambridge, UK: Cambridge University Press.

Levin, R. A., Wagner, W. L., Hoch, P. C., Nepokroeff, M., Pires, J. C., Zimmer, E. A., & Sytsma, K. J. (2003). Family-level relationships of Onagraceae based on chloroplast

*rbcL* and *ndhF* data. *American Journal of Botany*, 90, 107–115.  
<http://doi.org/10.3732/ajb.90.1.107>

Li, H. Q., Chen, J. Y., Wang, S., y Xiong, S. Z. (2012). Evaluation of six candidate DNA barcoding loci in *Ficus* (Moraceae) of China. *Molecular Ecology Resources*, 12, 783–790. <http://doi.org/10.1111/j.1755-0998.2012.03147.x>

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., y Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biological Reviews of the Cambridge Philosophical Society*, 90, 157–166 <http://doi.org/10.1111/brv.12104>

Li, D. Z., Gao, L. M., Li, H. T., Wang, H., Ge, X. J., Liu, J. Q., Chen, Z. D., ... Duan, G.-W. (2011). From the Cover: Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences*, 108, 19641–19646. <http://doi.org/10.1073/pnas.1104551108>

Li, Y., Gao, L. M., Poudel, R. C., Li, D. Z., y Forrest, A. (2011). High universality of *matK* primers for barcoding gymnosperms. *Journal of Systematics and Evolution*, 49, 169–175. <http://doi.org/10.1111/j.1759-6831.2011.00128.x>

Librado, P., y Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452. <http://doi.org/10.1093/bioinformatics/btp187>

Little, D. P. (2006). Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus*). *Systematic Botany*, 31, 461–480. <http://doi.org/10.1600/036364406778388638>

Little, D. P. (2010). A unified index of sequence quality and contig overlap for DNA barcoding. *Bioinformatics* (Oxford, England), 26, 2780–1. <http://doi.org/10.1093/bioinformatics/btq507>

Little, D. P. (2011). DNA barcode sequence identification incorporating taxonomic hierarchy and within taxon variability. *Public Library of Science ONE*, 6, e20552. <http://doi.org/10.1371/journal.pone.0020552>

Little, D. P. (2012). BRONX: Barcode Recognition Obtained with Nucleotide eXposés. Recuperado de: <http://www.nybg.org/files/scientists/dlittle/BRONX2.html> (accessed 19 October 2017).

Little, D. P., Knopf, P., y Schulz, C. (2013). DNA barcode identification of Podocarpaceae-The second largest conifer family. *Public Library of Science ONE*, 8, e81008. <http://doi.org/10.1371/journal.pone.0081008>

Little, D. P., y Stevenson, D. W. (2007). A comparison of algorithms for the identification of specimens using DNA barcodes: Examples from gymnosperms. *Cladistics*, 23, 1–21. <http://doi.org/10.1111/j.1096-0031.2006.00126.x>

Liu, T. (1971). *A monograph of the genus Abies*. Department of Forestry College of Agriculture, National Taiwan University, Taipei, Taiwan.

Ma, X. F., Szmidt, A. E., y Wang, X. R. (2006). Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci. *Molecular Biology and Evolution*, 23(4), 807-816.

Maia, V. H., da Mata, C. S., Franco, L. O., Cardoso, M. A., Cardoso, S. R. S., Hemerly, A. S., y Ferreira, P. C. G. (2012). DNA barcoding Bromeliaceae: Achievements and

pitfalls. *Public Library of Science ONE*, 7, e29877.  
<http://doi.org/10.1371/journal.pone.0029877>

Mao, K., Hao, G., Liu, J., Adams, R. P., y Milne, R. I. (2010). Diversification and biogeography of *Juniperus* (Cupressaceae): Variable diversification rates and multiple intercontinental dispersals. *New Phytologist*, 188, 254–272.  
<http://doi.org/10.1111/j.1469-8137.2010.03351.x>

Matos, J. A., y Schaal, B. A. (2000). Chloroplast evolution in the *Pinus montezumae* complex: A coalescent approach to hybridization. *Evolution*, 54, 1218–1233.  
<http://doi.org/10.1111/j.0014-3820.2000.tb00556.x>

Meier, R., Zhang, G., y Ali, F. (2008). The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Systematic Biology*, 57, 809–813. <http://doi.org/10.1080/10635150802406343>

Meyer, C. P., y Paulay, G. (2005). DNA barcoding: Error rates based on comprehensive sampling. *Public Library of Science Biology*, 3, 1–10.  
<http://doi.org/10.1371/journal.pbio.0030422>

Nixon, K. C. (1999). The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics*, 15, 407–414. <http://doi.org/10.1006/clad.1999.0121>

Percy, D. M., Argus, G. W., Cronk, Q. C., Fazekas, A. J., Kesanakurti, P. R., Burgess, K. S., Husband, B. C., ... Graham, S. W. (2014). Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this result from a trans-specific selective sweep? *Molecular Ecology*, 23, 4737–4756. <http://doi.org/10.1111/mec.12837>

Prada, C., Schizas, N. V., y Yoshioka, P. M. (2008). Phenotypic plasticity or speciation? A case from a clonal marine organism. *BMC Evolutionary Biology*, 8(1), 47.

Prunier, J., Verta, J. P., y Mackay, J. J. (2016). Conifer genomics and adaptation: At the crossroads of genetic diversity and genome function. *New Phytologist*. <http://doi.org/10.1111/nph.13565>

Rice, P., Longden, I., y Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends in Genetics: TIG*, 16, 276–277. [http://doi.org/10.1016/S0168-9525\(00\)02024-2](http://doi.org/10.1016/S0168-9525(00)02024-2)

Salinas, N. R., y Little, D. P. (2014). 2matrix: A utility for indel coding and phylogenetic matrix concatenation. *Applications in Plant Sciences*, 2, 1300083. <http://doi.org/10.3732/apps.1300083>

Semerikova, S. A., y Semerikov, V. L. (2014). Molecular phylogenetic analysis of the genus *Abies* (Pinaceae) based on the nucleotide sequence of chloroplast DNA. *Russian Journal of Genetics*, 50, 7–19. <http://doi.org/10.1134/S1022795414010104>

Simeone, M. C., Piredda, R., Papini, A., Vessella, F., y Schirone, B. (2013). Application of plastid and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (*Quercus*, Fagaceae): Problems, prospects and phylogenetic implications. *Botanical Journal of the Linnean Society*, 172, 478–499. <http://doi.org/10.1111/boj.12059>

Simmons, M. P., y Ochoterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, 49, 369–381. <http://doi.org/10.1080/10635159950173889>

Spearman, C. (1904). Spearman's rank correlation coefficient. *The American Journal of Psychology*, 15, 72–101. <http://doi.org/10.1136/bmj.g7327>

Steele, P. R., y Pires, J. C. (2011). Biodiversity assessment: State-of-the-art techniques in phylogenomics and species identification. *American Journal of Botany*, 98, 415–425. <http://doi.org/10.3732/ajb.1000296>

Strandby, U., Christensen, K. I., y Sørensen, M. (2009). A morphometric study of the *Abies religiosa-hickelii-guatemalensis* complex (Pinaceae) in Guatemala and Mexico. *Plant Systematics and Evolution*, 280, 59–76. <http://doi.org/10.1007/s00606-009-0164-x>

Syring, J., Farrell, K., Businský, R., Cronn, R., y Liston, A. (2007). Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. *Systematic Biology*, 56, 163–181. <http://doi.org/10.1080/10635150701258787>

Tanabe, A. S., y Toju, H. (2013). Two new computational methods for universal DNA barcoding: A benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land plants. *Public Library of Science ONE*, 8, e76910. <http://doi.org/10.1371/journal.pone.0076910>

Tripathi, A. M., Tyagi, A., Kumar, A., Singh, A., Singh, S., Chaudhary, L. B., y Roy, S. (2013). The internal transcribed spacer (ITS) region and *trnH-psbA* are suitable candidate loci for DNA barcoding of tropical tree species of India. *Public Library of Science ONE*, 8, e57934. <http://doi.org/10.1371/journal.pone.0057934>

Wang, X. R., Tsumura, Y., Yoshimaru, H., Nagasaka, K., y Szmidt, A. E. (1999). Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL*, *matK*, *rpl20-rps18* spacer, and *trnV* intron sequences. *American Journal of Botany*, 86, 1742–1753. <http://doi.org/10.2307/2656672>

Welch B. L. (1947). The generalization of student's problems when several different population variances are involved. *Biometrika*, 34, 28–35.

Wiemers, M., y Fiedler, K. (2007). Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology*, 4, 8–23. <http://doi.org/10.1186/1742-9994-4-8>

Yan, H. F., Hao, G., Hu, C. M., y Ge, X. J. (2011). DNA barcoding in closely related species: A case study of *Primula* L. sect. *Proliferae* Pax (Primulaceae) in China. *Journal of Systematics and Evolution*, 49, 225–236. <http://doi.org/10.1111/j.1759-6831.2011.00115.x>

Yang, J. B., Wang, Y. P., Möller, M., Gao, L. M., y Wu, D. (2012). Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). *Molecular Ecology Resources*, 12, 267–275. <http://doi.org/10.1111/j.1755-0998.2011.03095.x>

## Capítulo III

### Clave de identificación taxonómica con caracteres anatómicos foliares para especies de *Pinus* L. en Hidalgo

#### Resumen

En el estado de Hidalgo se distribuyen naturalmente 12 especies de pinos, de las cuales, *Pinus pinceana* y *P. greggii* se encuentran en categorías de riesgo. Dentro de este contexto es necesario generar información para la identificación taxonómica de *Pinus* utilizando únicamente caracteres anatómicos de acículas en individuos jóvenes que carecen de estróbilos femeninos. El objetivo de este estudio fue establecer una clave dicotómica para la identificación de pinos del estado de Hidalgo, con base en la anatomía foliar. Se recolectaron ejemplares en diversas localidades, se obtuvieron cortes transversales de las acículas, a partir de los cuales se codificó una matriz de caracteres, la cual fue sometida a un análisis de similitud UPGMA, se elaboró una clave de identificación taxonómica, con la cual se logran diferenciar 10 de las 12 especies, incluyendo *Pinus greggii* una especie bajo categoría de riesgo. *Pinus cembroides* se diferencia de *P. pinceana*, únicamente cuando la forma de la acícula en corte transversal es semicircular. Los canales resiníferos externos se observaron únicamente en pinos del subgénero *Strobus*, los septales solamente en *P. oocarpa*. El número de canales resiníferos es utilizado para identificar de manera complementaria debido a su alta variación. El análisis UPGMA resultó en dos agrupaciones principales, en la primera se recuperaron dos subgrupos correspondientes con las especies del subgénero *Strobus* y del subgénero *Pinus*, en la segunda se recuperaron especies del subgénero *Pinus*. Se concluye que la clave de identificación elaborada es robusta para la identificación de

especies de pinos a escala local utilizando exclusivamente caracteres anatómicos foliares.

**Palabras clave:** Acículas, anatomía foliar, canales resiníferos, hipodermis, *Pinus greggii* Engelm. ex Parl., *Pinus pinceana* Gordon.

### **Abstract**

Twelve pine species occur naturally in the state of Hidalgo, of which, *Pinus pinceana* and *P. greggii* are listed as threatened. In this context it is necessary to generate information for *Pinus* species identification especially for immature individuals without seed cones. In this study we establish an identification key based on leaf anatomy for the identification of pine species occurring in the state of Hidalgo. Specimens were collected from different localities, permanent slides were prepared from transverse leaf sections from which a character matrix was codified from leaf anatomy features and analyzed with an UPGMA. The matrix was used to construct an identification key. The key allows for the identification of 10 out of the 12 species, including *Pinus greggii* a species under risk. *Pinus cembroides* can only be differentiated from *P. pinceana* when it exhibits a semicircular shape in transverse section. External resin canals were observed only in *Pinus* subgenus *Strobus* and septal resin canals were only recorded for *P. oocarpa*. The number of resin canals may only be used for identification as complementary because of its high variation. UPGMA analysis generated two principal clusters, the first group included two interior clusters corresponding with subgenus *Strobus* species and subgenus *Pinus* species, the second principal cluster included only species from subgenus *Pinus*. This study contributes with a robust identification key for pine species at a local scale using exclusively secondary leaf anatomy characters.

**Keywords:** Needles, leaf anatomy, resin canals, hypodermis, *Pinus greggii* Engelm. ex Parl., *Pinus pinceana* Gordon.

### 3.1 Introducción

Los pinos son importantes ecológicamente en los bosques tropicales cercanos a nivel del mar, zonas desérticas y bosques de las tierras altas de México (Farjon *et al.*, 1997). México es considerado como el centro secundario de diversificación del género *Pinus*, con 50 (41.3%) de las 121 especies reportadas en el mundo (Gernandt y Pérez-De La Rosa, 2014; Pérez-De La Rosa y Gernandt, 2017). En el estado de Hidalgo se distribuyen de manera natural 12 especies: *Pinus ayacahuite* Ehrenb. ex Schltdl., *P. cembroides* Zucc., *P. devoniana* Lindl., *P. greggii* Engelm. ex Parl., *P. hartwegii* Lindl., *P. leiophylla* Schiede. ex Schltdl. & Cham., *P. montezumae* Lamb., *P. oocarpa* Schiede., *P. patula* Schiede. ex Schltdl. & Cham., *P. pinceana* Gordon., *P. pseudostrobus* Lindl. y *Pinus teocote* Schiede. ex Schltdl. & Cham. De estas especies, *Pinus greggii*, *P. patula*, *P. pinceana* y *P. teocote* son endémicas de México. *Pinus pinceana* se encuentra en el listado de especies de la Norma Oficial Mexicana 059-SEMARNAT-2010, en la categoría de riesgo de extinción (P) (SEMARNAT, 2010) y *Pinus greggii* en la categoría de Vulnerable en la lista roja de especies de la IUCN (Farjon, 2013).

Los pinos presentan cuatro tipos de hojas durante sus etapas de desarrollo: cotiledonares, primarias, catafilos y secundarias. Estas últimas, también son conocidas como acículas y comienzan a aparecer hacia el final de la primera estación de crecimiento o en algunas especies hasta la segunda estación de crecimiento. Son siempre verdes y metabólicamente activas, se desarrollan en brotes enanos axilares a los

catafilos, las acículas están agrupadas de 1-8 formando fascículos (Farjon y Styles, 1997). El número de hojas por fascículo varía según la especie debido a las condiciones del medio en el que se desarrollan siendo tres y cinco las cifras más comunes (Martínez, 1945).

La anatomía de las acículas, comprende tres regiones que están definidos por su función: la dermis, el mesófilo, y el tejido vascular (Farjon y Styles, 1997). La región de la dermis está constituida por la cutícula, epidermis e hipodermis; la región del tejido del mesófilo se encuentra estructurado por largas células de pared delgada, que llenan el espacio entre la hipodermis y la endodermis y en él se desarrollan los canales resiníferos que se sitúan en diferentes posiciones; la región del tejido vascular está conformada por la endodermis (capa de células ovals que rodean al tejido vascular), y en él se encuentran los haces vasculares (xilema y floema) (Farjon y Styles, 1997).

Los caracteres anatómicos foliares son fuentes importantes de información para la identificación de especies mediante claves taxonómicas. La clave dicotómica es un método para identificar especies, mediante breves descripciones, seleccionado uno de dos enunciados alternos. Las claves de identificación son de dos tipos: regional o por grupo morfológico (Farjon et al., 1997).

En relación a la identificación taxonómica de pinos es importante generar información para identificar especies con base en la anatomía de acículas, particularmente en individuos jóvenes que carecen de estróbilos femeninos. Por esta razón el objetivo de esta investigación fue establecer una clave dicotómica para la identificación de los pinos con distribución natural en el estado de Hidalgo, con base en la anatomía de las acículas.

### **3.2 Materiales y métodos**

El estado de Hidalgo cuenta con aproximadamente 20,905 km<sup>2</sup>, lo que representa el 1.1% del territorio nacional, y se encuentra ubicado dentro de la zona de transición mexicana incluyendo las provincias biogeográficas de la Sierra Madre Oriental, Altiplano Mexicano, Golfo de México y la Faja Volcánica Transmexicana (Rzedowski, 1978; CONABIO, 1997; Morrone, 2005; INEGI, 2016). Gernandt y Pérez-de la Rosa (2014) proponen que la Sierra Madre Oriental se encuentra entre las regiones con mayor diversidad de coníferas.

### **3.2.1 Trabajo de campo**

La colecta de ejemplares de *Pinus* se realizó en diversas localidades del estado de Hidalgo (Cuadro 1).

### **3.2.2 Cortes foliares y tinción**

Las acículas se fijaron en FAA (Formaldehído: Ácido acético glacial: Alcohol etílico al 50%, en proporción 5: 5: 90) por tiempo indefinido. Se obtuvieron cortes transversales de forma manual con una navaja de un solo filo seleccionando los mejores para tinción histológica mediante el procedimiento utilizado por Hernández-León (2011); Citrisolv 2 (15 min); EtOH 100% (5 min); EtOH 96% (5 min); EtOH 70% (5 min), EtOH 50% (5 min); ddH<sub>2</sub>O (5 min); Alcian blue (30 min); ddH<sub>2</sub>O (5 min); EtOH 50% (3 min); Safranina (2 hr); 50% EtOH (3 min); 70% EtOH (3 min); 96% EtOH (3 min); 100% EtOH (3 min); Citrisolv (15 min). Una vez teñidos los cortes, se montaron en preparaciones permanentes con la resina sintética Entellan (Merk Co., Alemania).

**Cuadro 3.1.** Datos detallados de recolecta de especies de *Pinus* del estado de Hidalgo.

Especie	No. de colecta	Localidad de colecta	Coordenadas	
			Latitud	Longitud
<i>Pinus ayacahuite</i>	SHL089	Acaxochitlán	20° 13.601'N	98° 13.351'W
	SHL121	Zimapán	20° 51.991'N	99° 12.916'W
<i>P. cembroides</i>	SHL114	Jacala	20° 47.377'N	99° 17.139'W
	SHL036	Zimapán	20° 51.092'N	99° 15.784'W
<i>P. devoniana</i>	SHL092	Huasca de Ocampo	20° 12.259'N	98° 36.995'W
	SHL115	Jacala	20° 48.462'N	99° 15.591'W
<i>P. greggii</i>	SHL137	Tlanchinol	20° 59.037'N	98° 38.478'W
	SHL095	San Bartolo Tutotepec	20° 23.865'N	98° 14.866'W
	SHL143	Molango	20° 44.579'N	98° 42.520'W
<i>P. hartwegii</i>	SHL149	Singuilucan	19° 58.245'N	98° 28.314'W
	SHL077	Mineral del Chico	20° 10.264' N	98° 43.873' W
<i>P. leiophylla</i>	SHL152	Singuilucan	19° 58.323'N	98° 26.960'W
	SHL022	Cuautepec de Hinojosa	19° 57.007'N	98° 15.6263'W
<i>P. montezumae</i>	SHL150	Singuilucan	19° 58.675'N	98° 26.994'W
	SHL155	Singuilucan	19° 57.348'N	98° 24.768'W
	EUAC002	Zimapán	20° 50.246'N	99° 15.314'W
<i>P. oocarpa</i>	SHL138	Calnali	20° 52.934'N	98° 35.247'W
	SHL145	Xochicoatlán	20° 44.321'N	98° 42.443'W
	SHL146	Xochicoatlán	20° 44.302'N	98° 42.461'W
	SHL144	Molango	20° 44.581'N	98° 42.513'W
<i>P. patula</i>	SHL030	Huasca de Ocampo	20° 09.064'N	98° 33.637'W
	SHL100	San Bartolo Tutotepec	20° 23.974'N	98° 14.955'W
<i>P. pinceana</i>	APO001	El Cardonal	20° 38.308'N	98° 59.436'W

	AAE002	El Cardonal	20° 38.431'N	98° 59.332'W
	No.003	El Cardonal	20° 38.431'N	98° 59.332'W
<i>P. pseudostrabus</i>	SHL123	Zimapán	20° 51.989'N	99° 13.742'W
	SHL130	Metepéc	20° 28.234'N	98° 18.835'W
	SHL154	Singuilucan	19° 58.329'N	98° 26.312'W
<i>P. teocote</i>	SHL131	Metepéc	20° 16.906'N	98° 18.939'W
	SHL021	Cuautepec de Hinojosa	19° 57.026'N	98° 15.660'W

---

### 3.2.3 Codificación de caracteres

Mediante una revisión de la literatura y la observación de las preparaciones permanentes se establecieron nueve caracteres foliares y sus correspondientes estados de carácter como se muestra en el cuadro 2 (Coulter y Rose, 1886; Shaw, 1914; Martínez, 1945; Lanyon, 1966; Mirov, 1967; Farjon y Styles, 1997; Ortiz, 1999). Los caracteres y estados de carácter fueron descritos con observación de las preparaciones permanentes a diferentes aumentos con ayuda de un microscopio óptico Axioscope 40 (Carl Zeiss). Los cortes foliares fueron fotografiados con un sistema de cámara digital acoplado al microscopio, y las imágenes fueron editadas en Adobe®Photoshop® CS6 Extended, versión 13.0.1 x32 (2017) para destacar los rasgos anatómicos foliares relevantes. Se construyó una matriz de datos morfoanatómicos a partir de las observaciones.

**Cuadro 3.2.** Caracteres y estados de carácter utilizados para la descripción anatómica foliar de las especies de *Pinus* del estado de Hidalgo.

<b>Caracter</b>	<b>Estados de caracter</b>
1. Forma del corte transversal ( <b>FCT</b> )	(0) Semicircular, (1) Transversalmente triangular, (2) Ampliamente triangular.
2. Posición de estomas en las acículas ( <b>PAE</b> )	(0) Epistomática, (1) Amfistomática.
3. Espesor de las paredes de las células de la hipodermis ( <b>EPCH</b> )	(0) Uniformemente delgadas, (1) uniformemente gruesas (2) delgadas en los primeros estratos y engrosadas en el resto, (3) aumentando gradualmente en grosor hacia el interior.
4. Intrusiones hipodermales ( <b>IH</b> )	(0) Ausentes, (1) Presentes.
5. Número de canales resiníferos ( <b>NCR</b> )	(0) Cero, (1) Uno, (2) Dos, (3) Tres, (4) Cuatro, (5) Cinco, (6) Seis, (7) Siete, (8) Ocho.
6. Posición de los canales resiníferos ( <b>PCR</b> )	(0) Externa, (1) Interna, (2) Media, (3) Septal.
7. Paredes de las células de la endodermis ( <b>PCE</b> )	(0) Uniformemente engrosadas, (1) Externamente engrosadas.
8. Haces vasculares ( <b>HV</b> )	(0) Separados, (1) Fusionados únicamente en la región del xilema, (2) Conados o fusionados.
9. Esclerénquima en la región vascular	(0) Ausente, (1) Presente.

### **3.2.3 Análisis de datos**

Los datos anatómicos foliares de las 12 especies fueron analizados con el método de Agrupamiento por Pares no Ponderados con Media Aritmética (UPGMA) (Sokal y Michener, 1958); mediante una matriz de presencia/ausencia (Cuadro 3). El análisis utilizó el coeficiente de similitud de Jaccard mediante la herramienta en línea DendroUPGMA (García-Vallvé., 1999).

## **3.3 Resultados y discusión**

### **3.3.1 Trabajo de campo**

Se recolectaron treinta ejemplares durante el trabajo de campo, los cuales fueron procesados y depositados en la colección del Herbario del Área Académica de Ciencias Agrícolas y Forestales del Instituto de Ciencias Agropecuarias de la Universidad Autónoma del Estado de Hidalgo campus Tulancingo y en el Herbario Nacional de México (MEXU). Los ejemplares corresponden con 1 - 3 individuos para las 12 especies de pinos en Hidalgo (Cuadro 1).

#### *Cortes foliares y tinción*

Las treinta preparaciones permanentes de cortes transversales de acículas fueron depositadas en la colección del Herbario del Área Académica de Ciencias Agrícolas y Forestales del Instituto de Ciencias Agropecuarias de la Universidad Autónoma del Estado de Hidalgo y se encuentran disponibles para consultas posteriores.

### 3.3.2 Codificación de caracteres

La matriz de datos de los caracteres y estados de carácter se presenta en el Cuadro 3. La forma de las células de la epidermis presentó variación intraespecífica e interespecífica lo cual es consistente con reportes previos (Farjon *et al.*, 1997).

Ocho de las especies estudiadas presentaron la vaina basal del fascículo permanente. Las especies que presentaron la vaina basal de fascículo decidua fueron *Pinus ayacahuite*, *P. cembroides*, *P. leiophylla* y *P. pinceana*, con la excepción de *P. leiophylla*, las otras especies pertenecen a los pinos blandos del subgénero *Strobus*. *Pinus leiophylla* es una de las pocas especies de pinos duros del subgénero *Pinus* con la vaina basal de los fascículos decidua (Gernandt *et al.*, 2005). El estado de carácter de la vaina basal del fascículo es de utilidad para la asignación de ejemplares a grupos, pero en combinación con otras características puede ser efectiva para la identificación de algunas especies (Farjon *et al.*, 1997).

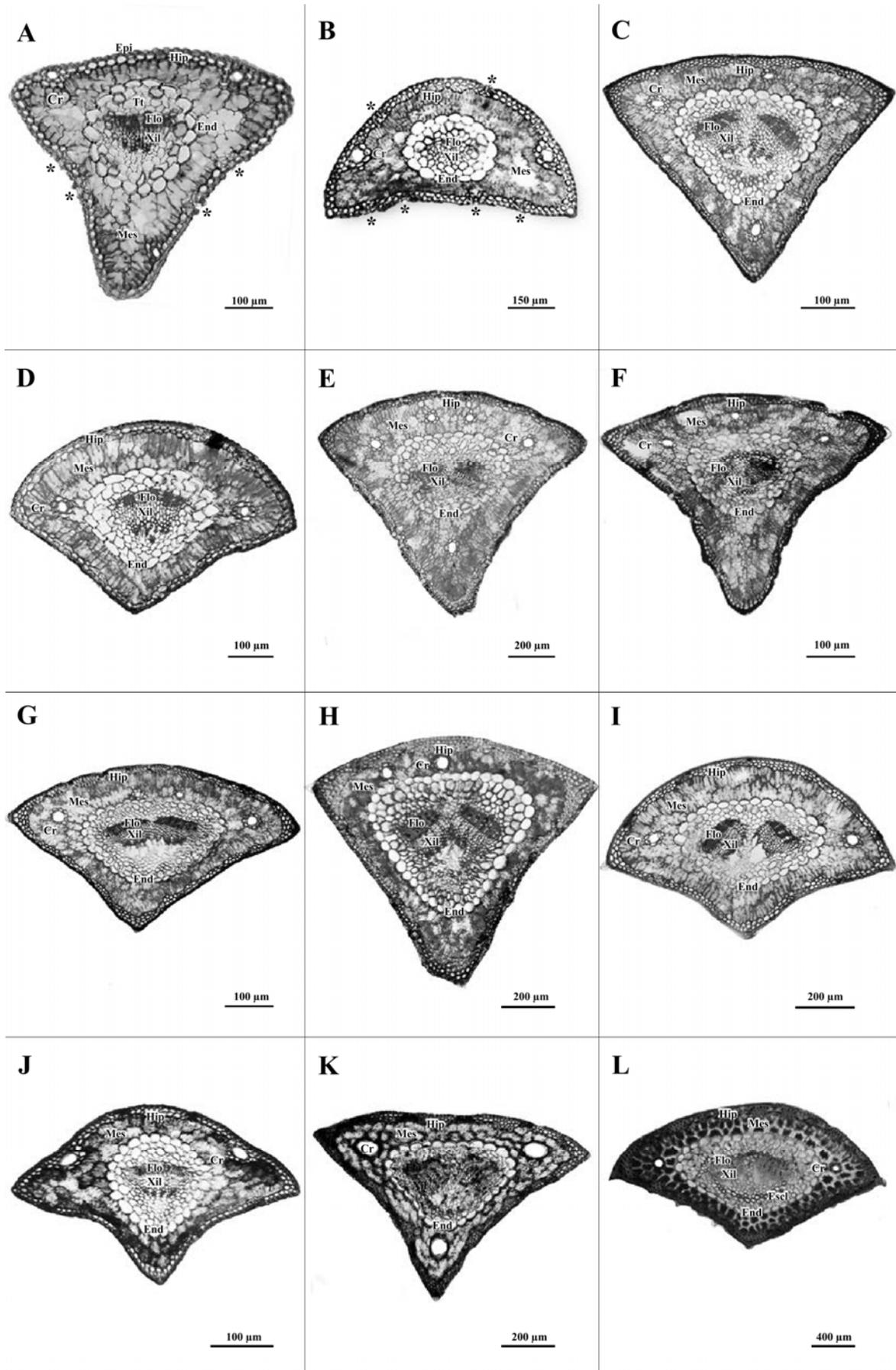
Las especies presentaron dos, tres y hasta cinco acículas por fascículo. El número de acículas por fascículo ha sido ampliamente utilizado para la identificación de especies de pinos generalmente en combinación con otros rasgos (Farjon y Styles, 1997; Farjon *et al.*, 1997; Eckenwalder, 2009). Las acículas de un fascículo forman un cilindro, por lo que el número de acículas por fascículo determina su forma transversal, correspondiendo a una sección del cilindro por ejemplo en fascículos con una acícula la forma transversal será circular, en fascículos con dos acículas semicircular (Figura 3.1B), en fascículos de tres será transversalmente triangular (Figura 3.1D, G, I, J, L) y en fascículos con cuatro o más presentará un forma ampliamente triangular (Figura 3.1A, C, E, F, H, K; Martínez, 1953; Farjon y Styles, 1997).

**Cuadro 3.3.** Matriz de datos morfoanatómicos de acículas para las especies de *Pinus* naturalmente distribuidas en el estado de Hidalgo.

Especie	No. de colecta	Clave dendrograma	Caracter								
			FCT (1)	PAE (2)	EPCH (3)	IH (4)	NCR (5)	PCR (6)	PCE (7)	ERV (9)	
<b>Subgénero <i>Strobus</i></b>											
<i>Pinus ayacahuite</i>	SHL089	Pa1	1	0	0	0	2	0	0	0	
	SHL121	Pa2	1	0	0	0	6	0	0	0	
<i>P. cembroides</i>	SHL114	Pc1	0	1	1	0	2	0	0	0	
	SHL036	Pc2	1	1	1	0	2	0	0	0	
<i>P. pinceana</i>	APO001	Ppi1	1	1	2	0	2	0	0	0	
	AAE002	Ppi2	1	1	2	0	2	0	0	0	
	No.003	Ppi3	1	1	1	0	2	0	0	0	
<b>Subgénero <i>Pinus</i></b>											
<i>P. devoniana</i>	SHL092	Pd1	2	1	3	1	4	2,0	1	0	
<i>P. greggii</i>	SHL115	Pg1	1	1	0	0	2	2	0	0	
	SHL137	Pg2	1	1	0	0	2	2	0	0	
	SHL095	Pg3	1	1	0	0	2	2	0	0	
	SHL143	Pg4	1	1	0	0	2	2	0	0	
<i>P. hartwegii</i>	SHL149	Ph1	2	1	2	0	5	2	1	1	
	SHL077	Ph2	2	1	2	0	4	2	0	1	
<i>P. leiophylla</i>	SHL152	Pl1	2	1	0	0	2	2	1	1	
	SHL022	Pl2	2	1	0	0	3	2	1	1	
<i>P. montezumae</i>	SHL150	Pm1	2	1	0	0	4	2,1	1	1	

	SHL155	Pm2	2	1	0	0	4	2,1	1	1
	EUAC002	Pm3	2	1	0	0	4	2,1	1	1
<i>P. oocarpa</i>	SHL138	Po1	2	1	2	0	6	3	0	1
	SHL145	Po2	2	1	2	0	1	3	0	1
	SHL146	Po3	2	1	2	0	3	3	1	1
<i>P. patula</i>	SHL144	Ppa1	1	1	0	0	5	2,1	1	1
	SHL030	Ppa2	1	1	0	0	2	2,1	1	1
	SHL100	Ppa3	1	1	0	0	2	2,1	0	1
<i>P. pseudostrobus</i>	SHL123	Pps1	2	1	3	0	3	2	1	1
	SHL130	Pps2	2	1	3	1	3	2	1	1
<i>P. teocote</i>	SHL154	Pt1	1	1	2	0	2	2	1	0
	SHL131	Pt2	1	1	2	0	4	2	1	0
	SHL021	Pt3	1	1	2	0	4	2	0	0

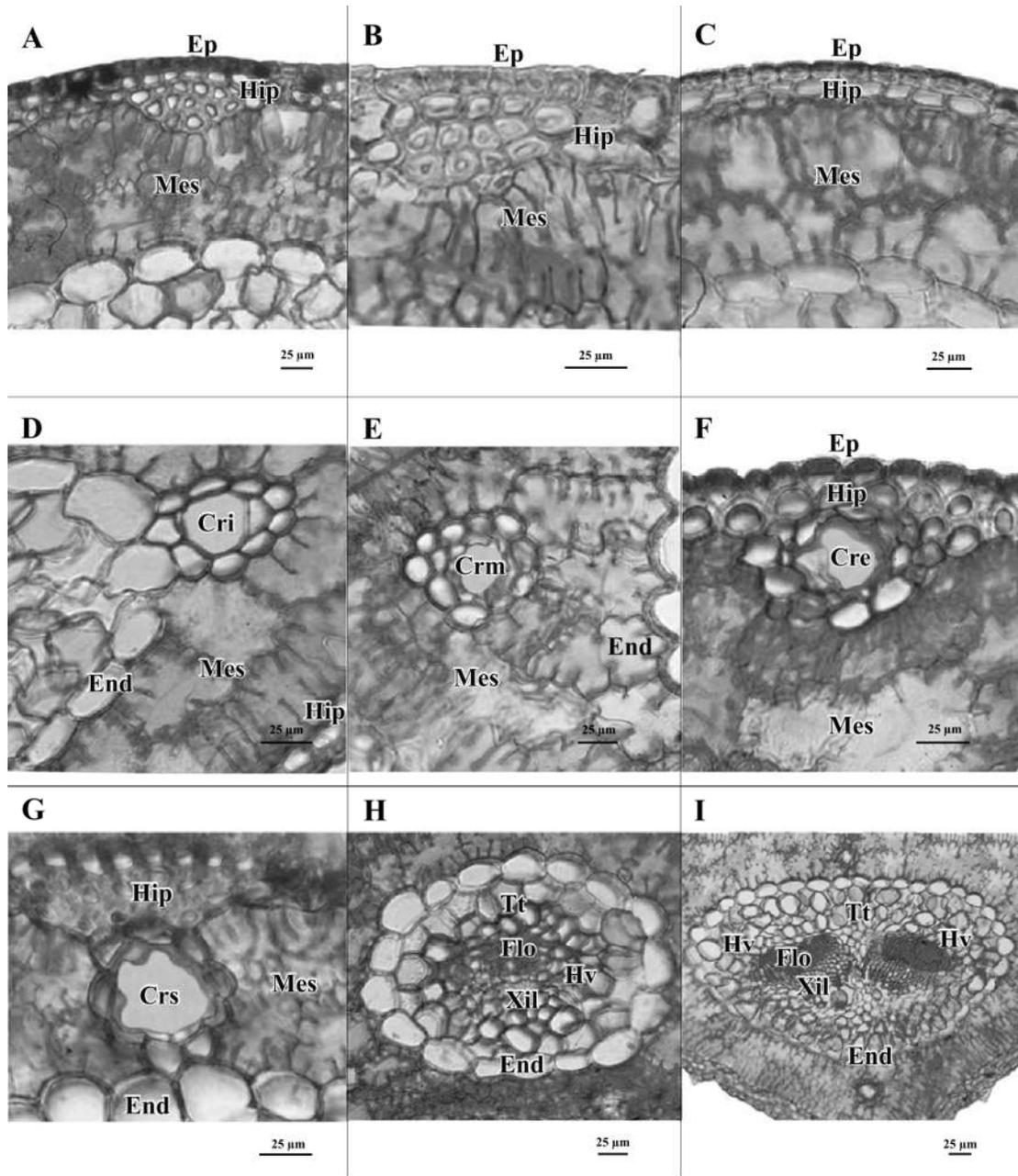
---



**Figura 3.1.** Secciones transversales de acículas de especies de *Pinus* del estado de Hidalgo. **A** *Pinus ayacahuite*. **B** *P. cembroides*. **C** *P. devoniana*. **D** *P. greggii*. **E** *P. hartwegii*. **F** *P. leiophylla*. **G** *P. montezumae*. **H** *P. oocarpa*. **I** *P. patula*. **J** *P. pinceana*. **K** *P. pseudostrobus*. **L** *P. teocote*. \* indica la posición de los estomas. Abreviaciones: **Cr**, Canal resinífero; **End**, Endodermis; **Epi**, Epidermis; **Escl**, Esclerenquima; **Flo**, Floema; **Hip**, Hipodermis; **Mes**, Mesófilo; **Tt**, Tejido de transfusión; **Xil**, Xilema.

*Pinus cembroides* fue la única especie con dos formas de la hoja en corte transversal. Por su forma semicircular puede diferenciarse del resto de las especies (Figura 3.1B); sin embargo, cuando presenta forma transversalmente triangular resulta imposible de separar de *P. pinceana* (Figura 3.1J). *Pinus ayacahuite* fue la única especie en presentar la disposición de las estomas epistomática (Figura 3.1A), mientras que el resto de las especies presentó una disposición anfistomática.

Para el carácter espesor de las paredes de las células de la hipodermis se registró con mayor frecuencia el estado de carácter "espesor uniformemente delgado de las paredes de las células de la hipodermis" (e.g., *P. ayacahuite*, *P. greggii* y *P. leiophylla*; Figura 3.2C), mientras que el "grosor de las paredes de las células de la hipodermis aumentando gradualmente hacia el interior" fue observado con menor frecuencia (e.g., *P. devoniana* y *P. pseudostrobus*; Figura 3.2A). *Pinus patula* y *P. pinceana* exhibieron variación intraespecífica para este carácter (Cuadro 3). Farjon y Styles (1997) proponen que las hojas largas, con estratos de células de paredes delgadas en la hipodermis están restringidas principalmente a climas templados o subtropicales mésicos (e.g., *P. ayacahuite*, *P. greggii*, *P. patula*).



**Figura 3.2.** Secciones transversales de acículas de especies de *Pinus* del estado de Hidalgo. **A** *Pinus devoniana*. **B, E, I** *P. teocote*. **C-D** *P. greggii*. **F, H** *P. cembroides*. **G** *P. oocarpa*. Abreviaciones de los estados de caracteres morfoanatómicos: **Cre**, Canal resinífero externo; **Cri**, Canal resinífero interno; **Crn**, Canal resinífero medio; **Crs**, Canal resinífero septal; **End**, Endodermis; **Epi**, Epidermis; **Flo**, Floema; **Hip**, Hipodermis; **Hv**, Haz vascular; **Mes**, Mesófilo; **Tt**, Tejido de transfusión; **Xil**, Xilema.

La presencia de diferentes grados de intrusiones hipodermales en el mesófilo ha sido previamente observada (Martínez, 1945). De las especies en esta investigación *Pinus devoniana* y *P. pseudostrobus* presentaron intrusiones hipodermales (Cuadro 3.3, Figuras 3.1C, K y 3.2A), mientras que en el resto las especies las intrusiones se encuentran ausentes (e.g., *P. greggii*; Figura 3.2C). Se ha observado que existen concentraciones de tejido hipodermal en áreas marginales mientras que en el resto de la hoja la hipodermis persiste en dos estratos celulares (Farjon y Styles, 1997).

En algunas especies como *Pinus oocarpa* el número de canales resiníferos fue altamente variable, presentando de 1-6, mientras que *P. pinceana* por ejemplo presentó 2 canales resiníferos de manera constante (Figura 3.1J). En este último caso el carácter puede ser de utilidad para la identificación de la especie a nivel de estado (Cuadro 3) (Perry, 1991; Farjon y Styles 1997; Cole *et al.*, 2008). Estudios previos reportan un rango de 1-12 canales resiníferos en las hojas de especies de pinos en América Latina, y para el caso del estado de Hidalgo se encontró un rango de 1-6 canales resiníferos por acícula (Farjon y Styles, 1997).

*Pinus oocarpa* y *P. ayacahuite* presentaron el mayor número de canales resiníferos (6), mientras que en cortes de acículas de *P. oocarpa* se observó el menor número de canales resiníferos que fue de uno. De acuerdo con Helmers (1943) el número de canales resiníferos tiene una estrecha relación con el medio ambiente. Eventualmente algunas acículas pueden no presentar canales resiníferos, pero esta condición no está asociada con ninguna especie en particular (*P. donnell-smithii*, puede ser la excepción de ser reconocido a nivel específico, ya que muchas acículas se encuentran sin canales de resina (Perry, 1991). Algunos autores consideran que el número de canales resiníferos por ser altamente variable no puede tomarse como elemento sistemático (Martínez, 1945).

La posición de los canales resiníferos constituye un carácter comúnmente utilizado para la clasificación e identificación de especies desde Engelmann (1880) y Coulter y Rose (1886) debido a que la mayoría de las especies se caracteriza por presentar una sola posición, aunque existen combinaciones como las reportadas en este estudio para *Pinus devoniana* y *P. montezumae*; (Figura 3.1C, G). Pero nunca se han encontrado combinaciones de más de dos posiciones (Farjon y Styles, 1997). La posición y el número de los canales resiníferos usados en combinación son muy útiles para la identificación de muchas especies (Perry, 1991). En especies con número variable de canales resiníferos se consideran primarios los de mayor tamaño y posición fija y canales resiníferos subsidiarios a los de menor tamaño (Farjon y Styles, 1997). La posición de los canales al ser menos variable que su número puede tener mayor utilidad para fines de identificación (Martínez, 1945). La posición media de los canales resiníferos fue predominante para las especies de pinos del estado de Hidalgo (e.g., *Pinus leiophylla*, *P. montezumae*, *P. pseudostrobus*; (Martínez, 1945; Farjon y Styles, 1997) (Figura 3.1F, G, K), mientras que la posición septal únicamente se observó en *P. oocarpa*, por lo que este estado de caracteres diagnóstico para esta especie, al menos dentro del estado de Hidalgo (Figuras 3.1H, 2G).

El engrosamiento externo de la pared de las células de la endodermis (Figura 3.2I) ha sido utilizado como carácter diagnóstico (Farjon y Styles, 1997). Sin embargo, en las especies con múltiples individuos como las incluidas en este estudio existe polimorfismo intraespecífico e interespecífico, por lo que este carácter no se considera adecuado para la identificación taxonómica.

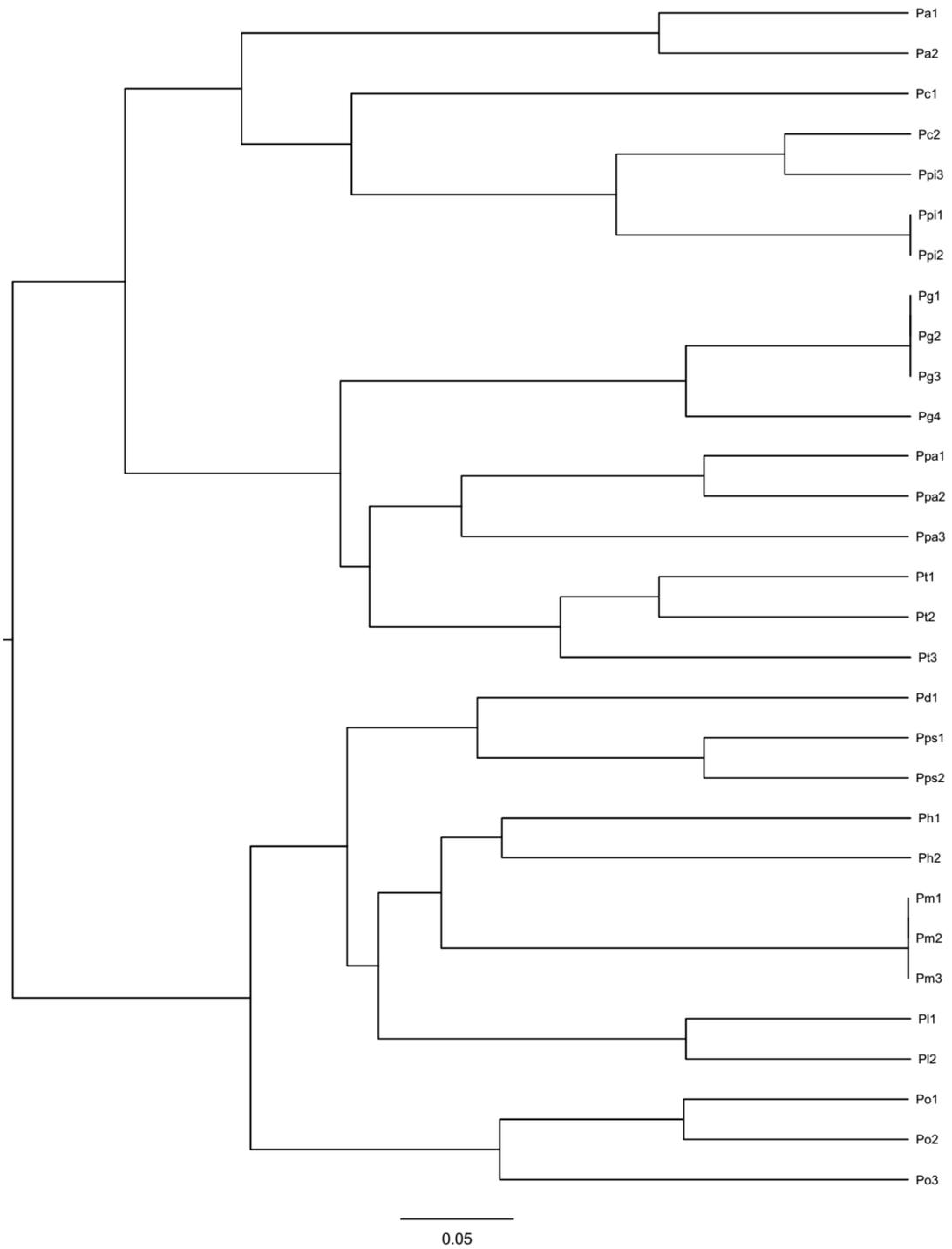
La presencia de uno (Figura 3.2H) o dos (Figura 3.2I) haces vasculares sustenta la propuesta de Koehne (1893) de dividir en dos secciones al género *Pinus*, sección

Haploxilon y Diploxilon, que corresponden con los subgéneros *Strobus* y *Pinus* (Gernandt *et al.*, 2005), lo cual fue consistente con los resultados de esta investigación.

El esclerénquima que en la región vascular rodea y en algunas ocasiones dividen los haces vasculares (Farjon y Styles, 1997), fue observado en algunas especies tales como *Pinus oocarpa* y *P. teocote* (Figura 3.1L). Sin embargo, no se considera un carácter diagnóstico debido a su variación.

### **3.3.3 Análisis de agrupamiento**

El análisis de similitud de los caracteres morfoanatómicos de las acículas resultó en dos grupos principales (Figura 3.3). En el grupo uno se asocian las especies de *Pinus ayacahuite*, *P. cembroides* y *P. pinceana* del subgénero *Strobus* más cercanamente relacionas que con otras especies y por otro lado el *P. greggii*, *P. patula* y *P. teocote* del subgénero *Pinus*. El segundo grupo principal está formado por especies del subgénero *Pinus* tales como *P. devoniana*, *P. hartwegii*, *P. leiophylla*, *P. montezumae*, *P. oocarpa* y *P. pseudostrobus*; *P. oocarpa* se separada del resto de las especies al interior del grupo. En las terminales se forman grupos por especie excepto para una muestra de *Pinus cembroides* que se agrupa con *P. pinceana*. De acuerdo con Eckenwalder (2009), la convergencia en las características de las hojas entre especies no relacionadas está provocada por la fuerte asociación entre los rasgos foliares con las condiciones medioambientales, de tal manera que los rasgos de la hojas pueden reflejar los mecanismos de adaptación de las plantas a su medioambiente (Tian *et al.*, 2015).



**Figura 3.1.** Dendrograma obtenido por análisis de agrupamiento UPGMA, con el uso del coeficiente de similitud de Jaccard en 12 especies de *Pinus* distribuidas en el estado de Hidalgo. Coeficiente de correlación cofenético = 0.847018494508054. El número en cada terminal corresponde con el número del ejemplar por especie.

### 3.3.4 Clave de identificación

Mediante la clave dicotómica producto de esta investigación es posible diferenciar 10 de las 12 especies de pinos del estado de Hidalgo. *P. cembroides* y *P. pinceana* únicamente pueden separarse cuando la forma del corte transversal es semicircular en la primera especie.

1.- Canales resiníferos externos.....	2
2.- Posición de estomas epistomática.....	<i>P. ayacahuite</i>
2'.- Posición anfistomática de estomas.....	3
3.- Forma de la hoja transversalmente triangular.....	<i>P. cembroides</i> y <i>P. pinceana</i>
3'.- Forma de la hoja semicircular.....	<i>P. cembroides</i>
1'.- Canales resiníferos medios, internos o septales.....	4
4.- Canales resiníferos septales.....	<i>P. oocarpa</i>
4'.- Canales resiníferos medios o medios e internos.....	5
5.- Intrusiones hipodermales presentes.....	6
6.- Canal resiníferos medios.....	<i>P. pseudostrobus</i>
6'.- Canales resiníferos medios e internos.....	<i>P. devoniana</i>
5'.- Intrusiones hipodermales ausentes.....	7
7.- Hipodermis delgada en el primer estrato y engrosada en el resto.....	8
8.- Forma de la hoja ampliamente triangular.....	<i>P. hartwegii</i>
8'.- Forma de la hoja transversalmente triangular.....	<i>P. teocote</i>
7'.- Hipodermis uniformemente delgada.....	9
9.- Forma de la hoja ampliamente triangular.....	10
10.- Canales resiníferos medios.....	<i>P. leiophylla</i>
10'.- Canales resiníferos medios e internos.....	<i>P. montezumae</i>
9'.- Forma de la hoja transversalmente triangular.....	11
11.- Canales resiníferos medios.....	<i>P. greggii</i>
11'.- Canales resiníferos medios e internos.....	<i>P. patula</i>

### 3.4 Conclusión

Este trabajo contribuye a la generación de datos anatómicos foliares para especies del género *Pinus*. La generación de herramientas de identificación taxonómica a partir de caracteres anatómicos foliares es de suma importancia para la identificación de especies

cuando no se cuenta con otra fuente de identificación, como puede ser de plantas producidas en vivero, individuos jóvenes, plantaciones comerciales, reforestaciones con fines de restauración en etapa de brinzal y árboles sin presencia de estróbilos femeninos. Los resultados obtenidos permitieron diferenciar a las especies y generar una clave de identificación robusta. Por lo que la aproximación metodológica usada en este estudio es una alternativa para la identificación de especies de pinos distribuidos en otros estados o regiones del país.

### 3.5 Literatura citada

Cole, K. L., J. Fisher, S. T. Arundel, J. Cannella y S. Swift. 2008. Geographical and climatic limits of needle types of one and two needled pinyon pines. *Journal of Biogeography*. 35(2): 257.

CONABIO. 1997. Provincias biogeográficas de México. Escala 1:4 000 000. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México, D. F.

Coulter, J. M., y J. N. Rose. 1886. Synopsis of North American pines, based upon leaf-anatomy. I. *Botanical Gazette*. 11(10): 256–269. <https://doi.org/10.1111/j.13652699.2007.01786.x>

Eckenwalder, J. E. 2009. *Conifers of the world: the complete reference*. TimberPress. <https://doi.org/citeulike-article-id:13426181>

Engelmann, G. 1880. Revision of the genus *Pinus*, and description of *Pinus elliottii*. *Transactions of the St. Louis Academy of Science*. 4: 161–190.

Farjon, A. 2013. *Pinus greggii*. The IUCN Red List of Threatened Species 2013:e.T32630A2822678. Recuperado a partir de <http://www.iucnredlist.org/details/32630/0> (2 de octubre de 2017).

Farjon, A., J. Pérez de la Rosa, y B. T. Styles. 1997. Guía de campo de los pinos de México y América Central. The Royal Botanic Gardens, Kew. Richmond, Londres. Inglaterra. 151 p.

Farjon, A. y B. T. Styles. 1997. Monograph 75. *Pinus* (Pinaceae). Flora Neotropica. New York, USA. 291 p.

García-Vallve, S., J. Palau y A. Romeu. 1999. Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. *Molecular Biology and Evolution*. 16(9):1125-1134.

Gernandt, D. S., G. Geada López, S. Ortiz García, y A. Liston. 2005. Phylogeny and classification of *Pinus*. *Taxon*. 54(1): 29–42. <https://doi.org/10.2307/25065300>.

Gernandt, D. S. y J. A. Pérez-de la Rosa. 2014. Biodiversidad de Pinophyta (coníferas) en México. *Revista Mexicana de Biodiversidad*, 85(SUPPL.). <https://doi.org/10.7550/rmb.32195>

Ghimire, B., C. Lee, J. Yang, y K. Heo. 2014. Comparative leaf anatomy of native and cultivated *Pinus* (Pinaceae) in Korea: implication for the subgeneric classification. *Plant Systematics and Evolution*. 301(2): 531–540. <https://doi.org/10.1007/s00606-014-1090-0>

Helmers, A. E. 1943. The ecological anatomy of ponderosa pine needles. *The American Midland Naturalist*. 29: 55–71.

Hernández-León, S. 2011. Filogenia molecular y evolución de caracteres foliares de *Pinus* sección *Trifoliae*. Tesis de maestría. Universidad Nacional Autónoma de México. 70 p.

Instituto Nacional de Estadística y Geografía (Inegi). 2016. Anuario estadístico y geográfico de Hidalgo 2016. Instituto Nacional de Estadística y Geografía. Recuperado a partir de

[http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/anuarios\\_2016/702825083748.pdf](http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/anuarios_2016/702825083748.pdf) (25 de septiembre de 2017).

Koehne, B. A. E. 1893. *Deutsche Dendrologie*. Verlag von Ferdinand Enke. Stuttgart, Baden-Württemberg, Deutschland. 602 p.

Lanyon, J. 1966. A card key to *Pinus* based on needle anatomy. Ministry of Conservation, NSW, Australia. 27: 5–27.

Martínez, M. 1945. Las pináceas mexicanas. *Anales del Instituto de Biología Universidad Nacional de México*, 1(16): 1–35.

Mirov, N. T. 1967. *The genus Pinus*. New York: The Ronald Press Company. New York, USA. 602 p.

Morrone, J. J. 2005. Hacia una síntesis biogeográfica de México. *Revista Mexicana de Biodiversidad*, 76(2): 207–252. <https://doi.org/10.1111/j.1475-4932.1986.tb00883.x>

Ortiz G., S. 1999. Evolución y filogenia en pinos y sus hongos endófitos: aspectos sistemáticos de la coespeciación. Tesis de doctorado. Instituto de Ecología, Universidad Nacional Autónoma de México. 120 p.

Perry Jr, J. P. 1991. *The pines of Mexico and Central America*. Timber Press. Portland, Oregon, USA. 231 p.

Pérez-De La Rosa, J. A., & Gernandt, D. S. (2017). *Pinus vallartensis* (Pinaceae), a new species from western Jalisco, Mexico. *Phytotaxa*. 233–242. <https://doi.org/ISSN 1179-3155>

Rzedowski, J. 1978. *Vegetación de México*. Limusa. México, D.F., México. 432 p.

Secretaría de Medio Ambiente y Recursos Naturales (Semarnat). 2010. Norma Oficial Mexicana NOM-059-SEMARNAT-2010. Secretaría de Medio Ambiente y Recursos Naturales, México, , D.F., México. 77 p.

Shaw, G. 1914. The genus *Pinus*. Arnold Arboretum Publications 5. Forage Village, Mass. The Murray Printing Co.

Sokal, R. R. y C. D. Michener. 1958. A statistical method for evaluating systematic relationships. The university of Kansas science bulletin. Vol XXVIII, Pt. II. No. 22. 1409-1438.

Tian, M., G.Yu, N. He and J. Hou. 2015. Leaf morphological and anatomical traits from tropical to temperate coniferous forests: Mechanisms and influencing factors. Scientific Reports, 6, 19703. <https://doi.org/10.1038/srep19703> (2 de octubre de 2017).

## Capítulo IV

### Conclusiones generales y perspectivas

En este estudio se determinó la presencia de 25 especies de coníferas en el estado de Hidalgo, los resultados contribuyen a aumentar el número de secuencias para los códigos de barras centrales (*matK* y *rbcL*), las cuales hasta el momento son escasas para coníferas, también ayuda a predecir el desempeño de los códigos de barras en grupos con historias de vida similares. La región *rbcL* presentó la mayor calidad en las secuencias y mayor complejidad lingüística, mientras que *matK* presentó un mayor porcentaje del gap de códigos de barras así como mayor variación. El método basado en caracteres BRONX es el mejor para identificación.

Utilizando la identificación mediante los códigos de barras genéticos será posible establecer planes de manejo, conservación y monitoreo efectivos a nivel local para al menos la mitad de especies en categorías de riesgo. Este estudio contribuye a predecir el desempeño de los códigos de barras centrales en plantas a nivel local en centros de diversidad de especies. Utilizando solo una región *matK* se obtienen los mismos resultados permitiendo menor inversión de tiempo y dinero para identificación de este grupo. La clave de identificación es robusta para la identificación de especies de pinos a escala local, utilizando exclusivamente caracteres anatómicos foliares. Se describieron anatómicamente las acículas de las 12 especies de pinos de Hidalgo. Un análisis cuantitativo de la anatomía foliar puede contribuir para la identificación del 100% de las especies.