



UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE FITOTECNIA

DOCTORADO EN CIENCIAS EN HORTICULTURA

FILOGENIA Y BIOGEOGRAFÍA DE *Sechium* P. BROWN

TESIS

Que como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS EN HORTICULTURA

PRESENTA:

LUIS ÁNGEL BARRERA GUZMÁN

BAJO LA SUPERVISIÓN DE:

DR. JUAN PORFIRIO LEGARIA SOLANO



APROBADA



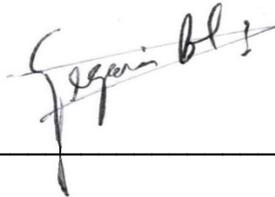
Chapingo, Estado de México, diciembre de 2021.

FILOGENIA Y BIOGEOGRAFÍA DE *Sechium* P. BROWN

Tesis realizada por el C. Luis Ángel Barrera Guzmán bajo la supervisión del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito para el obtener el grado de:

DOCTOR EN CIENCIAS EN HORTICULTURA

DIRECTOR: _____



DR. JUAN PORFIRIO LEGARIA SOLANO

ASESOR: _____



DR. JAIME SAHAGÚN CASTELLANOS

ASESOR: _____



DR. JORGE CADENA IÑIGUEZ

LECTOR EXTERNO: _____



DRA. MA. DE LOURDES C. ARÉVALO GALARZA

CONTENIDO

	Página
ÍNDICE DE CUADROS.....	vii
ÍNDICE DE FIGURAS	ix
DEDICATORIA.....	xv
AGRADECIMIENTOS	xvi
DATOS BIOGRÁFICOS	xvii
RESUMEN GENERAL	xviii
GENERAL ABSTRACT.....	xix
CAPÍTULO I. INTRODUCCIÓN GENERAL.....	16
1.1. Objetivos	18
1.2 Hipótesis	18
1.2 LITERATURA CITADA	21
CAPÍTULO II. FILOGENÉTICA DEL GÉNERO <i>Sechium</i> P. Brown: UNA REVISIÓN	25
RESUMEN	25
ABSTRACT	26
2.1 INTRODUCCIÓN.....	27
2.2. ORIGEN Y ECOGEOGRAFÍA	29
2.3. ANTECEDENTES DE LA CLASIFICACIÓN TAXONÓMICA.....	47
2.4. INFERENCIAS FILOGENÉTICAS EN <i>SECHIUM</i>	51
2.5. CONCLUSIONES	59
2.6. LITERATURA CITADA	60
CAPÍTULO III. RELACIONES FILOGENÉTICAS ENTRE ESPECIES MEXICANAS DEL GÉNERO <i>Sechium</i> (CUCURBITACEA)	72
RESUMEN	72
ABSTRACT	73

3.1. INTRODUCCIÓN.....	74
3.2. MATERIALES Y MÉTODOS.....	77
3.2.1. Material genético	77
3.2.2. Extracción y calidad de ADN.....	79
3.2.3. Condiciones de PCR y amplificación de fragmentos	79
3.2.4. Electroforesis y purificación de fragmentos	80
3.2.5. Cuantificación y secuenciación	81
3.2.6. Alineamiento y análisis estadístico.....	81
3.3. RESULTADOS Y DISCUSIÓN	83
3.3.1. Análisis filogenético <i>CoxIII</i>	84
3.3.2. Análisis filogenético <i>rbcl</i>	85
3.3.3. Análisis filogenético de la región ITS	86
3.3.4. Relaciones filogenéticas de especies mexicanas de <i>Sechium</i>	89
3.4. CONCLUSIONES	92
3.5. LITERATURA CITADA	92
CAPÍTULO IV. MODELOS DE DISTRIBUCIÓN POTENCIAL DE <i>Sechium</i> <i>tacaco</i> (PITTIER) C. JEFFREY EN COSTA RICA.....	97
RESUMEN	97
ABSTRACT	98
4.1. INTRODUCCIÓN.....	99
4.2. MATERIALES Y MÉTODOS.....	101
4.2.1. Datos de ocurrencia e información ambiental	101
4.2.2. Modelos de distribución	103
4.3. RESULTADOS Y DISCUSIÓN	104
4.4. CONCLUSIONES	110
4.5. LITERATURA CITADA	110
CAPÍTULO V. MODELOS DE DISTRIBUCIÓN POTENCIAL DE <i>Sechium edule</i> (CUCURBITACEAE) EN MÉXICO	115

RESUMEN.....	115
ABSTRACT	116
5.1. INTRODUCCIÓN.....	117
5.2. MATERIALES Y MÉTODOS.....	120
5.2.1. Datos de ocurrencia e información ambiental	120
5.2.2. Modelos de distribución.	121
5.3. RESULTADOS Y DISCUSIÓN	122
5.3.1. Información edafoclimática	122
5.3.2. Modelos de distribución	124
5.4. CONCLUSIONES	131
5.5. LITERATURA CITADA	131
CAPÍTULO VI. DESCRIPTORES EDAFOCLIMÁTICOS Y MODELOS DE DISTRIBUCIÓN DEL CLADO CENTROAMERICANO DE <i>Sechium</i> P. BROWN	139
RESUMEN.....	139
ABSTRACT	140
6.1. INTRODUCCIÓN.....	141
6.2. MATERIALES Y MÉTODOS.....	142
6.2.1. Puntos de ocurrencia y variables bioclimáticas	142
6.2.2. Análisis descriptivo	142
6.2.3. Análisis de agrupamiento.....	143
6.2.4. Modelos de distribución	143
6.3. RESULTADOS Y DISCUSIÓN	144
6.3.1. Tipos de clima y suelo	144
6.3.2. Análisis ACP y clúster k-mediooides.....	148
6.3.3. Modelos de distribución	150
6.4. CONCLUSIONES	160
6.5. LITERATURA CITADA	160

Apéndice 1. Estadísticos descriptivos para las variables bioclimáticas de las especies centroamericanas de <i>Sechium</i>	163
CAPÍTULO VII. DESCRIPTORES EDAFOCLIMÁTICOS Y MODELOS DE DISTRIBUCIÓN DEL CLADO MEXICANO DE <i>Sechium</i> P. BROWN.....	165
RESUMEN.....	165
ABSTRACT	166
7.1. INTRODUCCIÓN.....	167
7.2. MATERIALES Y MÉTODOS.....	169
7.2.1. Puntos de ocurrencia y variables bioclimáticas	169
7.2.2. Análisis descriptivo	169
7.2.3. Análisis de agrupamiento.....	169
7.2.4. Modelo Maxent	170
7.3. RESULTADOS Y DISCUSIÓN	170
7.3.1. Tipos de clima y suelo	170
7.3.2. Análisis ACP y clúster k-mediooides.....	172
7.3.3. Análisis de agrupamiento para el género <i>Sechium</i>	176
7.3.4. Modelos de distribución	177
7.4. CONCLUSIONES	180
7.5. LITERATURA CITADA	180
Apéndice 2. Estadísticos descriptivos para las variables bioclimáticas de las especies mexicanas de <i>Sechium</i>	183
CAPÍTULO VIII. ANÁLISIS DESCRIPTIVO Y MULTIVARIADO DE LA VARIACIÓN INFRAESPECÍFICA DE <i>S. edule</i> EN UN BANCO DE GERMOPLASMA <i>ex situ</i>.....	185
RESUMEN.....	185
ABSTRACT	186
8.1. INTRODUCCIÓN.....	187
8.2. MATERIALES Y MÉTODOS.....	188
8.2.1. Datos de pasaporte.....	188

8.2.2. Análisis multivariado	188
8.3. RESULTADOS Y DISCUSIÓN	189
8.3.1. Análisis descriptivo	189
8.3.2. Análisis de correspondencias múltiples	194
8.4. CONCLUSIONES	201
8.5. LITERATURA CITADA	201
CONCLUSIONES GENERALES.....	206
Apéndice 3. Variación infraespecífica de <i>S. edule</i>	207

ÍNDICE DE CUADROS

	Página
Cuadro 1. Clasificación dinámica de las especies de <i>Sechium</i>	49
Cuadro 2. Datos de pasaporte de los materiales genéticos empleados.	77
Cuadro 3. Características de los pares de iniciadores y temperatura de metilación (TM).....	80
Cuadro 4. Origen de los materiales genéticos y números de accesión del GenBank.	82
Cuadro 5. Resumen general de los parámetros genéticos para cada <i>locus</i>	84
Cuadro 6. Lista de variables utilizadas para el desarrollo de modelos de distribución de <i>S. tacaco</i> en Costa Rica.....	102
Cuadro 7. Parámetros estadísticos para los modelos de distribución de <i>S. tacaco</i> en Costa Rica.	105
Cuadro 8. Contribución (%) de las variables ambientales a los modelos de distribución de <i>S. tacaco</i> en Costa Rica.....	106
Cuadro 9. Evaluación y rendimiento de los modelos empleando los estadísticos AUC y TSS.	125

Cuadro 10. Parámetros estadísticos para los nueve modelos de distribución y ensamble para <i>S. panamense</i>	150
Cuadro 11. Prueba Jacknife para evaluar la contribución de las variables a los modelos en <i>S. panamense</i>	150
Cuadro 12. Parámetros estadísticos para los nueve modelos de distribución y ensamble para <i>S. pittieri</i>	152
Cuadro 13. Prueba Jacknife para evaluar la contribución de las variables a los modelos de distribución de <i>S. pittieri</i>	152
Cuadro 14. Parámetros estadísticos para los nueve modelos de distribución y ensamble para <i>S. talamancense</i>	154
Cuadro 15. Prueba Jacknife para evaluar la contribución de las variables a los modelos de distribución de <i>S. talamancense</i>	154
Cuadro 16. Parámetros estadísticos para los nueve modelos de distribución y ensamble para <i>S. venosum</i>	156
Cuadro 17. Prueba Jacknife para evaluar la contribución de las variables a los modelos de distribución de <i>S. venosum</i>	156
Cuadro 18. Parámetros estadísticos para los nueve modelos de distribución y ensamble para <i>S. villosum</i>	158
Cuadro 19. Prueba Jacknife para evaluar la contribución de las variables a los modelos de distribución de <i>S. villosum</i>	158
Cuadro 20. Estadísticos para el modelo Maxent de las especies mexicanas de <i>Sechium</i>	177
Cuadro 21. Estadísticos de tendencia central y dispersión para la variable altitud.	193
Cuadro 22. Eigenvalores y varianza explicada para los primeros 10 componentes del ACM.....	194
Cuadro 23. Vectores propios para las variables empleados en el ACM.	195

ÍNDICE DE FIGURAS

Página

Figura 1. Distribución geográfica de las especies de <i>Sechium</i> en Mesoamérica. Fuente: elaboración propia con datos de pasaporte de la Global Diversity Information Facility (GBIF) (https://www.gbif.org/).....	31
Figura 2. Estructuras morfológicas de <i>S. edule</i> : a) hoja (PH joevaccaro), b) zarcillos (PH Bodo), c) huerto comercial de chayote (PH Thalina García), d) inflorescencias (PH ming110), e) - h) variación morfológica de frutos (PH elizabeth1, PH César Lezama, PH Horacio Sirolli, PH katabenn).	33
Figura 3. Estructuras morfológicas de <i>S. compositum</i> : a) colecta de herbario (creada por Ingrid P. Lin) y b) variación morfológica de sus frutos [Fuente: Avendaño-Arrazate <i>et al.</i> (2017)].	37
Figura 4. Estructuras morfológicas de <i>S. chinantlense</i> : a) inflorescencia (Fuente: PH Javier Castrejón Lira) y b) variación morfológica de sus frutos [Fuente: Avendaño-Arrazate <i>et al.</i> (2017)].	38
Figura 5. Espécimen conservado de <i>S. hintonii</i> . Fuente: G.B. Hinton, The New York Botanical Garden.....	39
Figura 6. Estructuras morfológicas de <i>S. mexicanum</i> : a) hojas, zarcillos, frutos y b) inflorescencias. Fotografías de Rafael Lira/CONABIO.....	40
Figura 7. Hojas, zarcillos, inflorescencias y fruto de <i>S. panamense</i> . PH Thomas Bernad Croat.	41
Figura 8. Estructuras morfológicas de <i>S. tacaco</i> : a) – b) plantas y frutos (PH Barry Hammel y PH Arthur de Jesús); c) – d) estructuras florales (PH Jim Conrad) y e) – f) diversidad de frutos (PH Barry Hammel).....	42
Figura 9. Espécimen conservado de <i>S. talamancense</i> . CUCU F 1841927, Rapid Reference Collection (RRC) Field Museum of Natural History - Keller Action Science Center.	43
Figura 10. Espécimen conservado de <i>S. venosum</i> . Descrita por Gómez Pignataro, 20483, Costa Rica. Missouri Botanical Garden.....	44
Figura 11. Espécimen conservado de <i>S. villosum</i> . The New York Botanical Garden.	45

Figura 12. Espécimen conservado de <i>S. pittieri</i> . Barcode 02728781. PH Covenyor Belt. US National Herbarium, Department of Botany, Smithsonian Institution.	46
Figura 13. Relaciones filogenéticas de la tribu Sicyoeae. La estimación del tiempo de divergencia para todos los taxones del <i>TimeTree of Life</i> (TTOL) está precalculada y almacenada en la base de datos de <i>TimeTree</i> . La línea de tiempo evolutiva que se muestra se construyó ubicando primero las especies consultadas o taxones superiores en el TTOL y luego cruzando el árbol hacia la raíz, obteniendo así la estimación del tiempo de divergencia en cada nodo del árbol. <i>TimeTree</i> se tomó de Kumar <i>et al.</i> (2017) en MEGA 7.0 (Kumar <i>et al.</i> , 2016).	50
Figura 14. Cronología evolutiva de <i>Sechium</i> con tiempos de divergencia en millones de años. La estimación del tiempo de divergencia para todos los taxones del <i>TimeTree of Life</i> (TTOL) está precalculada y almacenada en la base de datos de <i>TimeTree</i> . La línea de tiempo evolutiva que se muestra se construyó ubicando primero las especies consultadas o taxones superiores en el TTOL y luego cruzando el árbol hacia la raíz, obteniendo así la estimación del tiempo de divergencia en cada nodo del árbol. <i>TimeTree</i> se tomó de Kumar <i>et al.</i> (2017) en MEGA 7.0 (Kumar <i>et al.</i> , 2016). Eras: CZ = Cenozoico, Z = Neo-Proterozoico, Y = Meso-Proterozoico, X = Paleo-Proterozoico, W = Neoarqueano, V = Mesoarqueano, U = Paleoarqueano, EA = Eoarqueano; Eones: Pr = Proterozoico, A = Arcaico, HDE = Hadeon.	55
Figura 15. Cronología evolutiva de las especies mexicanas de <i>Sechium</i> con tiempos de divergencia en millones de años. Edades: SER = Serravallian, PIA = Piacenzian, GEL = Gelasian; Épocas: PLE: Pleistoceno. Tomado de Sebastian <i>et al.</i> (2012) y elaborado con la herramienta <i>TimeTree</i> por Kumar <i>et al.</i> (2017) en MEGA 7.0 (Kumar <i>et al.</i> , 2016).	56
Figura 16. Cronología evolutiva de las especies centroamericanas de <i>Sechium</i> (<i>Frantzia</i>) con tiempos de divergencia en millones de años. Tomado de Sebastian <i>et al.</i> (2012) y elaborado con la herramienta <i>TimeTree</i> por Kumar <i>et al.</i> (2017) en MEGA 7.0 (Kumar <i>et al.</i> , 2016).	58
Figura 17. Árbol filogenético resultante de secuencias <i>CoxIII</i> . El árbol de consenso de arranque inferido de 1000 réplicas se utiliza para representar la historia evolutiva de los taxones analizados. El porcentaje de árboles replicados, en los que los taxones asociados se agrupan en la prueba de arranque, se	

muestra junto a las ramas, mostrando los 50 árboles más parsimoniosos (longitud = 1502); el índice de consistencia es 0.7 con un índice de retención de 0.6.85

Figura 18. Árbol filogenético resultante de secuencias *rbcL*. El árbol de consenso de arranque inferido de 1000 réplicas se utiliza para representar la historia evolutiva de los taxones analizados. El porcentaje de árboles replicados, en los que los taxones asociados se agrupan en la prueba de arranque, se muestra junto a las ramas. El análisis de la parsimonia produjo 194 árboles más parsimoniosos con una longitud de 1602; el índice de consistencia es de 0.75 con un índice de retención de 0.78.87

Figura 19. Árbol filogenético IB resultante de secuencias ITS1-5.8S-ITS2 obtenidas de 100.000 generaciones. Las probabilidades posteriores se indican debajo de los nodos, y los valores máximos de soporte de arranque de parsimonia aparecen encima de los nodos. El análisis de parsimonia produjo 83 árboles más parsimoniosos con una longitud de 518; el índice de consistencia es de 0.65 con un índice de retención de 0.79.88

Figura 20. Modelos de distribución para *S. tacaco* en Costa Rica. **A.** GLM. **B.** GAM. **C.** GBM. **D.** CTA. 107

Figura 21. Modelos de distribución de *S. tacaco* en Costa Rica. Técnicas basadas en aprendizaje automático, **A.** RF, **B.** Maxent, **C.** SVM y **D.** Ensamble de los siete modelos. 108

Figura 22. Proporción de puntos de ocurrencia para las subprovincias fisiográficas de México. 123

Figura 23. Proporción de A) tipos de climas y B) tipos de suelo en las subprovincias fisiográficas de México respecto a *S. edule*. 124

Figura 24. Modelos de distribución para *S. edule* en México. A) Bioclim, B) Dominio, C) Mahalanobis y D) MLG. 126

Figura 25. Modelos de distribución para *S. edule* en México. A) Maxent, B) MVS, C) Bosque aleatorio y D) Ensamble de coincidencia para los siete modelos. . 128

Figura 26. Correlograma de las variables cuantitativas para las especies centroamericanas de *Sechium*. 144

Figura 27. Proporción de **A)** tipos de suelo y **B)** tipos de clima en relación con los puntos de ocurrencia para cada especie. 145

Figura 28. **A)** Agrupación k-medioídes, **B)** principales variables bioclimáticas que contribuyen en la distribución de las especies centroamericanas de *Sechium*, **C)**

proporción de los puntos de ocurrencia de las especies en cada clúster y D) consistencia del agrupamiento mediante el coeficiente de la Silueta.	149
Figura 29. Ensamble de los nueve modelos de distribución de especies para <i>S. panamense</i> en Panamá.	151
Figura 30. Ensamble de los nueve modelos de distribución de especies para <i>S. pittieri</i> en Costa Rica-Panamá.	153
Figura 31. Ensamble de los nueve modelos de distribución de especies para <i>S. talamancense</i> en Costa Rica.	155
Figura 32. Ensamble de los nueve modelos de distribución de especies para <i>S. venosum</i> en Panamá-Costa Rica.	157
Figura 33. Ensamble de los nueve modelos de distribución de especies de <i>S. villosum</i> en Costa Rica.	159
Figura 34. Proporción de A) tipos de suelo y B) tipos de clima en relación con los puntos de ocurrencia para cada especie.	172
Figura 35. A) Agrupación k-medioides, B) principales variables bioclimáticas que contribuyen en la distribución de las especies mexicanas de <i>Sechium</i> , C) proporción de los puntos de ocurrencia de las especies en cada clúster y D) consistencia del agrupamiento mediante el coeficiente de la Silueta.	174
Figura 36. Histogramas de frecuencias, cajas de bigotes, gráficos de densidad y correlaciones para las variables bio1, bio4, bio12, bio15 y elevación de las especies mexicanas de <i>Sechium</i>	175
Figura 37. Dendrogramas calculados a partir de las medianas a partir de las 19 variables bioclimáticas y la variable elevación para las especies de <i>Sechium</i> . A) k = 2, índice de consistencia promedio de 0.48 y B) k = 3 con índice de consistencia de 0.55.	176
Figura 38. Modelo de distribución Maxent de <i>S. chinantlense</i> en México.	177
Figura 39. Modelo de distribución Maxent de <i>S. compositum</i> para México-Guatemala.	178
Figura 40. Modelo de distribución Maxent de <i>S. hintonii</i> para México.	179
Figura 41. Modelo de distribución Maxent de <i>S. mexicanum</i> para México.	179
Figura 42. Colectas de <i>S. edule</i> en función a sus A) complejos varietales y B) lugar de procedencia.	189

Figura 43. Proporciones y distribución de <i>Sechium</i> en función al lugar de procedencia.	190
Figura 44. Establecimiento de la variación infraespecífica de <i>S. edule</i> y las especies <i>S. compositum</i> y <i>S. chinantlense</i>	191
Figura 45. Propietario y manejo de los complejos varietales de <i>S. edule</i> y las especies <i>S. chinantlense</i> y <i>S. compositum</i>	191
Figura 46. Destino final de la variación infraespecífica de <i>S. edule</i>	192
Figura 47. Cajas de bigotes para las colecciones de <i>Sechium</i> referente a la altitud.	193
Figura 48. ACM para las variables A) Condición y B) Propietario.	196
Figura 49. ACM para las variables A) Color del fruto y B) Textura del fruto. ..	196
Figura 50. ACM para las variables A) Tamaño y B) Sabor de frutos.	197
Figura 51. ACM para la variable forma de fruto.	198
Figura 52. <i>Sechium edule</i> var. <i>nigrum xalapensis</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	207
Figura 53. <i>Sechium edule</i> var. <i>nigrum spinosum</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	208
Figura 54. <i>Sechium edule</i> var. <i>nigrum levis</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	209
Figura 55. <i>Sechium edule</i> var. <i>nigrum minor</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	210
Figura 56. <i>Sechium edule</i> var. <i>nigrum maxima</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	211
Figura 57. <i>Sechium edule</i> var. <i>nigrum conus</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	212
Figura 58. <i>Sechium edule</i> var. <i>virens levis</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	213
Figura 59. <i>Sechium edule</i> var. <i>albus dulcis</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	214
Figura 60. <i>Sechium edule</i> var. <i>albus minor</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	215

Figura 61. <i>Sechium edule</i> var. <i>albus levis</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	216
Figura 62. <i>Sechium edule</i> var. <i>albus spinosum</i> . Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).	217

DEDICATORIA

En memoria de mi madre María Elena Barrera-Guzmán †.

A Lidia Barrera-Guzmán por su gran apoyo incondicional, por enseñarme e inculcarme las bases y principios para mi formación personal y académica. Gracias por todo.

A mis primos Diego y José Alfredo por ser una fuente de inspiración para hacerme de mis logros.

A mis abuelos, Esther † y Máximo † que han sido como mis padres y me enseñaron y aconsejaron en toda etapa de mi vida.

A mis hermanas Lucero y Luz del Carmen por siempre motivarme y estar ahí en los momentos más difíciles.

A mis tíos Juan, Margarita y Trinidad que también han sido como mis padres.

A mi familia en general por su gran apoyo incondicional.

AGRADECIMIENTOS

A CONACyT (Consejo Nacional de Ciencia y Tecnología) por brindarme el financiamiento necesario para cursar mis estudios de doctorado.

A la Universidad Autónoma Chapingo por otorgarme todos los medios necesarios para concluir mis estudios de posgrado y disfrutar nuevamente la estadía dentro de ella.

Al Dr. Juan Porfirio Legaria Solano por su asesoría y por brindarme todos los recursos y conocimientos necesarios para la finalización del presente trabajo.

Al Dr. Jorge Cadena Iñiguez por brindarme sus conocimientos, compromiso y asesoría, mostrando así una buena actitud de colaboración para este trabajo.

Al Dr. Jaime Sahagún Castellanos por su apoyo en la culminación de este trabajo.

A GISeM (Grupo Interdisciplinario de Investigación en *Sechium edule* en México) por otorgarme las facilidades y los recursos con los cuales se ha llevado a cabo parte de este proyecto de investigación.

DATOS BIOGRÁFICOS

Datos personales

Nombre: Luis Ángel Barrera Guzmán

Fecha de nacimiento: 13 de octubre de 1990

Lugar de nacimiento: Tlapacoyan, Veracruz.

CURP: BAGL901013HVZRZS09

Profesión: Maestro en Ciencias en Biotecnología Agrícola

Cédula profesional: 11511785



Desarrollo académico

Licenciatura 2009 – 2014 Universidad Autónoma Chapingo

Maestría 2016 – 2017 Universidad Autónoma Chapingo

Doctorado 2018 – 2021 Universidad Autónoma Chapingo

RESUMEN GENERAL

FILOGENIA Y BIOGEOGRAFÍA DE *Sechium* P. Brown¹

Sechium P. Brown (Cucurbitaceae) es un género de plantas originado en las regiones montañosas de Mesoamérica. *Sechium edule* y *Sechium tacaco* son las especies más aprovechadas. La especialización adaptativa de *S. edule* dio origen a la formación de complejos varietales que difieren en características morfológicas, principalmente del fruto, generando grupos varietales. Los genotipos agrupados en *var. nigrum* se consideran los más cercanos a su pariente silvestre, mientras los complejos del grupo *albus* son considerados lo más evolucionados y distantes de su ancestro. Los análisis de ADN extranuclear mostraron alta afinidad filogenética entre *S. chinantlense*, *S. compositum* y *S. edule*, sugiriendo que estas especies pudieron haberse originado a través de hibridaciones. Los modelos de distribución basados en aprendizaje automático arrojaron buenos resultados en términos de Área bajo la curva (AUC) y el estadístico de habilidad verdadera (TSS). Las variables relacionadas con la precipitación son las limitantes principales de la distribución de las especies centroamericanas; mientras que las variables relacionadas a la temperatura son las limitantes para las especies mexicanas. Los climas predominantes para especies centroamericanas fueron Af (29.2 %) y Cfb (27.3 %); los tipos de suelo predominantes fueron andosol (38.3 %) y cambisol (30.6 %). Los climas predominantes para especies mexicanas fueron Cwb (29.3 %) y Aw (16.1 %); los suelos predominantes fueron leptosol (29.3 %) y phaeozem (15.4 %). El grupo *var. virens levis* es el más representado en el Banco de Germoplasma de *Sechium edule* (BANGESE) y sus características morfológicas de este complejo varietal lo hace ideal para el mercado nacional e internacional. Respecto a la información del BANGESE, la mujer rural juega un papel determinante en la conservación, aislamiento reproductivo, preferencia de variedades y nuevos morfotipos de *S. edule*.

Palabras clave: *Sechium* P. Brown, complejos varietales, afinidad filogenética, modelos de distribución, aprendizaje automático, diversidad genética.

¹ Tesis de Doctorado en Ciencias en Horticultura, Universidad Autónoma Chapingo
Autor: Luis Ángel Barrera Guzmán
Director de tesis: Dr. Juan Porfirio Legaria Solano

GENERAL ABSTRACT

PHYLOGENY AND BIOGEOGRAPHY OF *Sechium* P. BROWN²

Sechium P. Brown (Cucurbitaceae) is a plant genus that originated in the mountainous regions of Mesoamerica. *Sechium edule* and *Sechium tacaco* are the most widely used species. The adaptive specialization of *S. edule* gave rise to the formation of varietal complexes that differ in morphological characteristics, mainly of the fruit, generating varietal groups. Genotypes grouped in *var. nigrum* are considered the closest to their wild relative, while the complexes of the *albus group* are considered the most evolved and distant from their ancestor. Extranuclear DNA analysis showed high phylogenetic affinity between *S. chinantlense*, *S. compositum* and *S. edule*, suggesting that these species may have originated through hybridizations. The distribution models based on machine learning showed good results in terms of Area under the curve (AUC) and the True Skill Statistic (TSS). Variables related to precipitation are the main limitations of the distribution of Central American species; while the variables related to temperature are the limiting factors for Mexican species. The predominant climates for Central American species were Af (29.2%) and Cfb (27.3%); the predominant soil types were andosol (38.3%) and cambisol (30.6%). The predominant climates for Mexican species were Cwb (29.3%) and Aw (16.1%); the predominant soils were leptosol (29.3%) and phaeozem (15.4%). The group *var. virens levis* is the most represented in the *Sechium edule* Germplasm Bank (BANGESE) and its morphological characteristics of this varietal complex make it ideal for the national and international market. Regarding the BANGESE information, rural women have a determining role in the conservation, reproductive isolation, variety preference and new morphotypes of *S. edule*.

Key words: *Sechium* P. Brown, varietal complexes, phylogenetic affinity, distribution models, machine learning, genetic diversity.

² PhD Thesis of Science in Horticulture, Universidad Autónoma Chapingo
Author: Luis Ángel Barrera Guzmán
Advisor: Dr. Juan Porfirio Legaria Solano

CAPÍTULO I. INTRODUCCIÓN GENERAL

Jeffrey (1966) reportó por primera vez que la familia Cucurbitaceae agrupaba alrededor de 700 especies distribuidas en 90 géneros. Actualmente se tienen aceptadas 1000 especies incluidas en 134 géneros (The Plant List, 2010), de los cuales destacan por su importancia económica los géneros *Cucurbita*, *Cucumis*, *Citrullus*, *Luffa*, *Lagenaria*, *Momordica*, *Trichosanthes*, *Benincasa* y *Coccinia*. De estos géneros, 10 especies son de importancia económica, ya que son cultivadas en distintas partes del mundo y se les considera como cultivos principales [*Cucurbita argyrosperma* C. Huber, *Cucurbita maxima* Duchesne, *Cucurbita moshata* Duchesne, *Cucurbita pepo* L., *Citrullus lanatus* (Thunb.) Matsum. & Nakai, *Benincasa hispida* (Thunb.) Cogn., *Cucumis melo* L., *Cucumis sativus* L., *Lagenaria siceraria* (Molina) Standl. y *Momordica charantia* L.]; 23 especies más se consideran cultivos de importancia local y cultivadas, pero no todas están completamente domesticadas, incluyéndose a *S. edule*, que particularmente es cultivada en varias partes del mundo como especie introducida (Chomicki *et al.*, 2020).

El género *Sechium* P. Brown incluye las siguientes especies: *S. chinantlense* Lira & F. Chiang, *S. compositum* (Donn. Sm) C. Jeffrey, *S. edule* (Jacq.) Swartz, *S. hintonii* (P.G. Wilson) C. Jeffrey, *S. mexicanum* Lira & Nee, *S. panamensis* (Wunderlin) Lira & F. Chiang, *S. pittieri* (Cogn.) C. Jeffrey, *S. tacaco* (Pittier) C. Jeffrey, *S. talamancensis* (Wunderlin) C. Jeffrey, *S. venosum* (L.D. Gómez) Lira & Chiang y *S. villosum* (Wunderlin) C. Jeffrey. Estas especies se originaron, domesticaron y diversificaron en Mesoamérica (Newstrom, 1986, 1990, 1991; Lira y Chiang, 1992; Lira, 1995; Lira *et al.*, 1997; Lira y Nee, 1999; Cross *et al.*, 2006).

La importancia del género radica en la utilización de sus estructuras morfológicas (frutos, raíces, hojas) para fines alimenticios y medicinales (Newstrom, 1990; Lira, 1995, 1996). Las cucurbitacinas, uno de sus componentes químicos, son terpenos tetracíclicos presentes en las cucurbitáceas que confieren un sabor

amargo a sus estructuras morfológicas y muestran actividad antitumoral contra varios tipos de cáncer (Zhou *et al.*, 2019). Los frutos de *S. edule* han sido estudiados con mayor precisión debido a sus compuestos que poseen propiedades antineoplásicas (Monroy-Vázquez *et al.*, 2009; Cadena-Iñiguez *et al.*, 2013b; Aguiñiga-Sánchez *et al.*, 2015, 2017; Salazar-Aguilar *et al.*, 2017).

Mediante estudios morfológicos y moleculares se concluyó que *S. edule* se originó en el sureste de México (Cross *et al.*, 2006). El resto de las especies de *Sechium* son endémicas de Costa Rica y Panamá (Monge y Loría, 2017). Respecto a *S. edule*, se han formado complejos varietales que difieren en características morfológicas principalmente de frutos, las cuales han sido producto de la especialización adaptativa y de la riqueza cultural y ecogeográfica de México (Cadena-Iñiguez *et al.*, 2008; Cadena-Iñiguez y Arévalo-Galarza, 2011).

Las especies centroamericanas han sido poco estudiadas, sin embargo, se sabe que prosperaran en ambientes adversos, lo que podría suponer que su estructura genética les confiere cierta tolerancia/resistencia. En un futuro, se espera que las especies centroamericanas también tengan utilidad para aportar mejoras a las especies cultivadas. Los sistemas de información geográfica junto con los modelos de distribución de especies permiten aportar información nueva sobre las características edafoclimáticas donde éstas se desarrollan. Son importantes ya que permiten formular estrategias para la búsqueda de germoplasma e identificar áreas prioritarias (Phillips *et al.*, 2006, 2009; Hijmans *et al.*, 2017; Manzoor *et al.*, 2018; Zhang *et al.*, 2019). Sin embargo, existen algunas dudas respecto a la filogenia del género *Sechium* y los complejos varietales de la especie *S. edule*. Considerando la facilidad de contar con un banco de germoplasma de *S. edule*, se partió para realizar diferentes acercamientos a este tema filogenético con el fin de contribuir al conocimiento, distinción taxonómica y acciones de conservación. Con base en lo anterior, se plantearon diferentes objetivos desarrollados como capítulos.

1.1. Objetivos

1.- Dilucidar las relaciones filogenéticas del género *Sechium* P. Br., basadas en estudios moleculares y morfológicos para contribuir al conocimiento taxonómico sobre estas especies, en aras de formular estrategias de conservación y mejoramiento genético.

2.- Analizar filogenéticamente especies del género *Sechium* spp., y los complejos varietales de *S. edule* mediante secuenciación de ADN extranuclear y ribosomal para establecer relaciones de parentesco que ayuden a formular las bases de un programa de mejoramiento genético de largo plazo y usos múltiples

3.- Determinar la idoneidad geográfica de las especies de *Sechium* mediante modelos de distribución, con el fin de analizar las condiciones edafoclimáticas bajo las cuales se desarrollan, esto para contribuir a su estudio e incentivar estrategias de búsqueda de colectas y medidas de conservación en el manejo de los recursos fitogenéticos de *Sechium*.

4. Analizar mediante técnicas descriptivas y multivariadas las accesiones del banco de germoplasma de *S. edule* con la finalidad de encontrar asociaciones y patrones de similitud que permitan gestionar eficientemente los recursos fitogenéticos.

1.2 Hipótesis

1. La revisión y análisis en conjunto de estudios moleculares y morfológicos permitirá sentar las bases sobre el origen, evolución, domesticación y diversificación de las especies de *Sechium*, particularmente las del clado mexicano.

2. Los *loci* de ADN extranuclear *coxIII*, *rbcL* y las regiones ITS proporcionarán los árboles filogenéticos que permitan explicar las relaciones de parentesco entre las especies de *Sechium*, mismas que contribuirán a la taxonomía y el establecimiento de bases en la mejora genética de los chayotes

3. Los mapas de distribución potencial ayudarán a determinar cuáles son las áreas idóneas para establecer y expandir las poblaciones de *S. edule*, además de proporcionar información sobre la ubicación de posibles poblaciones silvestres que pueden brindar nuevas pistas sobre la evolución de esta especie.

4. Los análisis descriptivos y multivariantes sobre las accesiones del banco de germoplasma permitirán tener un consenso sobre sus características, así como la relación e importancia que tienen con los asentamientos humanos.

El capítulo II se analizó y detalló los estudios moleculares y morfológicos referentes a *Sechium*. El reloj molecular indica que las especies *S. chinantlense*, *S. compositum* y *S. edule* divergieron hace aproximadamente 1.87 millones de años durante el pleistoceno. Las diferencias en la clasificación taxonómica de las especies de *Sechium* son producto de la amplia variabilidad morfológica presente entre las muestras estudiadas para cada especie. Esta variación se debe a los procesos de selección natural y artificial, hibridaciones espontáneas, ubicación geográfica de las muestras recolectadas y flujo genético entre especies; además, de que la existencia de rasgos plesiomórficos y apomórficos también provoca cambios en la clasificación taxonómica.

En el capítulo III se analizó a nivel de ADN extranuclear especies mexicanas de *Sechium* y los complejos varietales de *S. edule*. Los complejos varietales del grupo *nigrum* son los más antiguos y cercanos a un ancestro silvestre; sin embargo, debido a las múltiples variantes fenotípicas que existen en el estado de Veracruz, México, es necesario evaluar más poblaciones e individuos que ayuden a comprender la historia evolutiva de las poblaciones silvestres de *S. edule*. Los complejos varietales del grupo *albus* son los más evolucionados. Las especies *S. chinantlense*, *S. compositum*, *S. hintonii* y *S. edule* formaron un grupo dentro del clado mexicano de *Sechium*. La hibridación entre *S. edule*, *S. compositum* y *S. chinantlense* originó nuevas variantes fenotípicas. Los

complejos varietales de *S. edule* se formaron mediante especialización adaptativa, aislamiento reproductivo en vías de su domesticación.

En los capítulos IV y V se estudió la idoneidad climática para *S. tacaco* y *S. edule*, respectivamente. Las variables relacionadas a la temperatura son las principales responsables de la distribución de *S. tacaco* en Costa Rica; mientras para *S. edule*, aquellas variables relacionadas a la precipitación son la que limitan su distribución. Las zonas con alta idoneidad para *S. edule* se registraron en los estados de Veracruz, Chiapas, Guerrero, Oaxaca, Hidalgo y Puebla, aunque en el occidente de México y hacia la península de Yucatán también se encontró idoneidad moderada.

En los capítulos VI y VII se analizaron las especies de *Sechium* en los clados mexicano y centroamericano. La esencia de estos capítulos es muy similar a la analizada en los capítulos IV y V, sin embargo, se extrajo información inédita sobre las condiciones climáticas y edáficas bajo las cuales estas especies prosperan.

El capítulo VIII dilucidó la información de los datos de pasaporte de un banco de germoplasma de los complejos varietales de *S. edule*. El complejo varietal *virens levis* es el más representado en el BANGESe y sus características morfológicas las hacen ideales para el mercado nacional e internacional. La mujer es el actor principal en el manejo, conservación y flujo de la diversidad genética de los recursos fitogenéticos de *S. edule*. El estudio de las accesiones en los bancos de germoplasma es de vital importancia para tener un consenso sobre los materiales y tomar decisiones encaminadas a su conservación y mejoramiento genético, estas mismas accesiones deben ir acompañadas de una caracterización morfológica, química y molecular.

1.2 LITERATURA CITADA

- Aguiñiga-Sánchez, I., Cadena-Iñiguez, J., Santiago-Osorio, E., Gómez-García, G., Mendoza-Núñez, V. M., Rosado-Pérez, J., Ruíz-Ramos, M., Cisneros-Solano, V. M., Ledesma-Martínez, E., Delgado-Bordonave, A. de J., & Soto-Hernández, R. M. (2017). Chemical analyses and in vitro and in vivo toxicity of fruit methanol extract of *Sechium edule* var. *nigrum spinosum*. *Pharmaceutical Biology*, 55(1), 1638-1645. <https://doi.org/10.1080/13880209.2017.1316746>
- Aguiñiga-Sánchez, I., Soto-Hernández, M., Cadena-Iñiguez, J., Ruíz-Posadas, L. del M., Cadena-Zamudio, J. D., González-Ugarte, A. K., Steider, B. W., & Santiago-Osorio, E. (2015). Fruit extract from a *Sechium edule* hybrid induce apoptosis in leukaemic cell lines but not in normal cells. *Nutrition and Cancer*, 67(2), 250-257. <https://doi.org/10.1080/01635581.2015.989370>
- Cadena-Iñiguez, J., & Arévalo-Galarza, M. L. C. (2011). *Las variedades de Chayote (Sechium edule (Jacq.) Sw.) y su comercio mundial* (1 st ed.). Biblioteca Básica de Agricultura. Ed. Colegio de Postgraduados. 98 p.
- Cadena-Iñiguez, J., Soto-Hernández, M., Torres-Salas, A., Aguiñiga-Sánchez, I., Ruíz-Posadas, L., Rivera-Martínez, A. R., Avendaño-Arrazate, C. H., & Santiago-Osorio, E. (2013). The antiproliferative effect of chayote varieties (*Sechium edule* (Jacq.) Sw.) on tumour cell lines. *Journal of Medicinal Plants Research*, 7(8), 455-460. <https://doi.org/10.5897/JMPR12.866>
- Chomicki, G., Schaefer, H., & Renner, S. S. (2020). Origin and domestication of Cucurbitaceae crops: Insights from phylogenies, genomics and archaeology. *New Phytologist*, 226(5), 1240-1255. <https://doi.org/10.1111/nph.16015>
- Cross, H., Lira, S. R., & Motley, T. J. (2006). Origin and diversification of chayote. En T. J. Motley, N. Zerega, & H. Cross (Eds.), *Darwin's Harvest: New*

Approaches to the Origins, Evolution, and Conservation of Crops (1°, pp. 171-194). Columbia University Press.

Hijmans, R. J., Phillips, S. J., Leatwick, J., & Elith, J. (2017). *Dismo: Species Distribution Modeling. R package version 1.1-4*. <https://CRAN.R-project.org/package=dismo>

Jeffrey, C. (1966). On the Classification of the Cucurbitaceae. *Kew Bulletin*, 20(3), 417-426. JSTOR. <https://doi.org/10.2307/4108235>

Lira, R. (1996). Chayote. *Sechium edule* (Jacq.) Sw., promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute., Rome, Italy.

Lira, Rafael, & Chiang, F. (1992). Two New Combinations in *Sechium* (Cucurbitaceae) from Central America, and a New Species from Oaxaca, Mexico. *Novon*, 2(3), 227-231. JSTOR. <https://doi.org/10.2307/3391556>

Lira, Rafael, & Nee, M. (1999). A new species of *Sechium* sect. *Frantzia* (Cucurbitaceae, *Sicyeae*, *Sicyinae*) from Mexico. *Brittonia*, 51(2), 204-209. <https://doi.org/10.2307/2666628>

Lira, Rafael, Villaseñor, J. L., & Dávila, P. D. (1997). A Cladistic Analysis of the Subtribe *Sicyinae* (Cucurbitaceae). *Systematic Botany*, 22(3), 415-425. JSTOR. <https://doi.org/10.2307/2419817>

Lira, S. R. (1995). Estudios taxonómicos en el género *Sechium* P. Br. Cucurbitaceae [Tesis Doctoral]. Universidad Nacional Autónoma de México. 267 pp.

- Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size—finer may not always be better. *Scientific Reports*, 8(1), 7168. <https://doi.org/10.1038/s41598-018-25437-1>
- Newstrom, L. (1986). Studies in the origin and evolution of chayote, *Sechium edule* (Jacq.) Sw. (Cucurbitaceae) [Tesis Doctoral]. University of California.
- Newstrom, L. (1990). Origin and evolution of chayote, *Sechium edule*. In D. M. Bates, R. W. Robinson, & C. Jeffrey (Eds.), *Biology and Utilization of the Cucurbitaceae* (1º, pp. 141-149). Cornell University Press.
- Newstrom, L. (1991). Evidence for the origin of chayote, *Sechium edule* (Cucurbitaceae). *Economic Botany*, 45(3), 410-428. <https://doi.org/10.1007/BF02887082>
- Phillips, S.J. (2009). *A Brief Tutorial on Maxent*. Network of Conservation Educators and Practitioners, Center for Biodiversity and Conservation, American Museum of Natural History. Lessons in Conservation. pp. 108-135.
- Phillips, Steven J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Rosado-Pérez, J., Aguiñiga-Sánchez, I., Santiago-Osorio, E., & Mendoza-Núñez, V. M. (2019). Effect of *Sechium edule* var. *nigrum spinosum* (Chayote) on oxidative stress and pro-Inflammatory markers in older adults with metabolic syndrome: an exploratory study. *Antioxidants (Basel, Switzerland)*, 8(5), 146. <https://doi.org/10.3390/antiox8050146>
- Salazar-Aguilar, S., Ruiz-Posadas, L. D. M., Cadena-Iñiguez, J., Soto-Hernández, M., Santiago-Osorio, E., Aguiñiga-Sánchez, I., Rivera-

Martínez, A. R., & Aguirre-Medina, J. F. (2017). *Sechium edule* (Jacq.) Swartz, a New Cultivar with Antiproliferative Potential in a Human Cervical Cancer HeLa Cell Line. *Nutrients*, 9(8), E798. <https://doi.org/10.3390/nu9080798>

The Plant List. (2010). *Version 1*. <http://www.theplantlist.org/>

Zhang, J., Jiang, F., Li, G., Qin, W., Li, S., Gao, H., Cai, Z., Lin, G., & Zhang, T. (2019). Maxent modeling for predicting the spatial distribution of three raptors in the Sanjiangyuan National Park, China. *Ecology and Evolution*, 9(11), 6643-6654. <https://doi.org/10.1002/ece3.5243>

Zhou, J., Liu, M., Chen, Y., Xu, S., Guo, Y., & Zhao, L. (2019). Cucurbitacin B suppresses proliferation of pancreatic cancer cells by ceRNA: Effect of miR-146b-5p and lncRNA-AFAP1-AS1. *Journal of Cellular Physiology*, 234(4), 4655-4667. <https://doi.org/10.1002/jcp.27264>

CAPÍTULO II. FILOGENÉTICA DEL GÉNERO *Sechium* P. Brown: UNA REVISIÓN

PHYLOGENETICS OF THE GENUS *Sechium* P. Brown: A REVIEW

RESUMEN

El género *Sechium* P. Brown se integra de 11 especies, las cuales se originaron, domesticaron y diversificaron en las regiones montañosas de Mesoamérica. Estas especies conforman dos grandes grupos: los clados mexicano y centroamericano. Los estudios moleculares y morfológicos han demostrado que las especies del clado mexicano se formaron mediante hibridaciones interespecíficas y flujo génico, excepto *Sechium mexicanum*, que está fuertemente relacionada con el género *Sicyos*. El objetivo de esta revisión fue analizar las relaciones filogenéticas de *Sechium* basadas en estudios moleculares y morfológicos, lo cual contribuirá a mejorar el conocimiento taxonómico sobre estas especies, en aras de formular estrategias de conservación y mejoramiento genético. El clado centroamericano está bien soportado con estudio moleculares, más no así con estudios morfológicos; las especies de este clado están aisladas geográficamente y son endémicas. *Sechium edule* y *S. tacaco* son las especies más aprovechadas en los sectores agrícolas e industrial; ambas tienen diversidad genética y fenotípica que les permite adaptarse a una extensa gama de nichos ecológicos. Finalmente, las especies del clado centroamericano prosperan en ambientes adversos con bajas temperaturas en bosques mesófilos con alta humedad relativa, con características que pueden otorgar resistencia a heladas y agentes fitopatógenos a las especies cultivadas de este género.

Palabras clave: flujo genético, hibridaciones, diversidad genética, nichos ecológicos, conservación, mejoramiento genético.

ABSTRACT

The *Sechium* P. Br. genus composed of 11 species, which originated from the mountainous regions of Mesoamerica, have been domesticated and diversified. These species are clustered in two large groups: the Mexican clade and the Central American clade. Morphological and molecular studies have shown that species of the Mexican clade are formed through interspecific hybridizations and genetic flow, with the exception of *Sechium mexicanum*, which is strongly linked to the genus *Sicyos*. The objective of this review was to analyze the phylogenetics of *Sechium* based on morphological and molecular studies, which contributed to taxonomic knowledge and utilization, thereby favoring its conservation and improvement. The Central American clade is well supported with molecular data, but not so with morphological data. The species in this clade were geographically isolated and endemic. *Sechium edule* and *Sechium tacaco* are exploited species in the agricultural and industrial sectors, and both have an extensive genetic and phenotypic diversity that has allowed them to diversify and expand into different ecological niches. Finally, the Central American species of *Sechium* thrive in adverse environments of temperatures of mesophyll forest and high relative humidity, with characteristics that can give resistance to frosts and phytopathogenic agents, as well as cultivated species of this genus.

Keywords: genetic flow, hybridisations, genetic diversity, ecological niches, conservation, breeding.

2.1 INTRODUCCIÓN

Jeffrey (1966) reportó por primera vez que la familia Cucurbitaceae agrupaba alrededor de 700 especies distribuidas en 90 géneros. Actualmente se tienen aceptadas 1000 especies incluidas en 134 géneros (The Plant List, 2010), de los cuales destacan por su importancia económica los géneros *Cucurbita*, *Cucumis*, *Citrullus*, *Luffa*, *Lagenaria*, *Momordica*, *Trichosanthes*, *Benincasa* y *Coccinia*. De estos géneros, 10 especies son de importancia económica, ya que son cultivadas en distintas partes del mundo y se les considera como cultivos principales [*Cucurbita argyrosperma* C. Huber, *Cucurbita maxima* Duchesne, *Cucurbita moshata* Duchesne, *Cucurbita pepo* L., *Citrullus lanatus* (Thunb.) Matsum. & Nakai, *Benincasa hispida* (Thunb.) Cogn., *Cucumis melo* L., *Cucumis sativus* L., *Lagenaria siceraria* (Molina) Standl. y *Momordica charantia* L.]; 23 especies más se consideran como cultivos de importancia local y cultivadas, pero no todas están completamente domesticadas, incluyéndose a *S. edule*, que particularmente es cultivada en varias partes del mundo como especie introducida (Chomicki *et al.*, 2019). Se estima que la familia Cucurbitaceae se originó en Asia hace aproximadamente 60 millones de años y paulatinamente fue dispersándose al resto de los continentes (Schaefer *et al.*, 2009).

El género *Sechium* P. Brown incluye las especies siguientes: *S. chinantlense* Lira & F. Chiang, *S. compositum* (Donn. Sm) C. Jeffrey, *S. edule* (Jacq.) Swartz, *S. hintonii* (P.G. Wilson) C. Jeffrey, *S. mexicanum* Lira & Nee, *S. panamensis* (Wunderlin) Lira & F. Chiang, *S. pittieri* (Cogn.) C. Jeffrey, *S. tacaco* (Pittier) C. Jeffrey, *S. talamancensis* (Wunderlin) C. Jeffrey, *S. venosum* (L.D. Gómez) Lira & Chiang y *S. villosum* (Wunderlin) C. Jeffrey. Estas especies se originaron, domesticaron y diversificaron en Mesoamérica (Newstrom, 1986, 1990, 1991; Lira y Chiang, 1992; Lira, 1995; Lira *et al.*, 1997a; Lira y Nee, 1999; Cross *et al.*, 2006).

La importancia del género radica en la utilización de sus estructuras morfológicas (frutos, raíces, hojas) para fines alimenticios y medicinales (Newstrom, 1990; Lira, 1995, 1996). Las cucurbitacinas, uno de sus componentes químicos, son

terpenos tetracíclicos presentes en Cucurbitaceae que confieren un sabor amargo a sus estructuras morfológicas y muestran actividad antitumoral contra varios tipos de cáncer (Zhou *et al.*, 2019). Los frutos de *S. edule* han sido estudiados con mayor precisión debido a que sus compuestos alcohólicos poseen propiedades antineoplásicas (Monroy- Vázquez *et al.*, 2009; Cadena-Iñiguez *et al.*, 2013b; Aguiñiga-Sánchez *et al.*, 2015, 2017; Salazar-Aguilar *et al.*, 2017).

Las especies *S. edule* y *S. tacaco* son las más explotadas de su género en el sector agrícola. Particularmente, las poblaciones de *S. edule* muestran una gran variabilidad genotípica y fenotípica, las cuales son producto de los efectos de la domesticación y de la interacción con el ambiente. De estas relaciones se han formado complejos varietales y se han agrupado por sus características bioquímicas y morfológicas, las cuales son estables y hereditarias. Algunas de las características están vinculadas a sus frutos: color, forma, textura, sabor, tamaño, composición bromatológica (Cadena-Iñiguez *et al.*, 2008, 2011; Cadena-Iñiguez y Arévalo-Galarza, 2011), diversidad genética (Machida-Hirano *et al.*, 2015) y estructura citogenética (Palacios, 1987; Mercado *et al.*, 1993; De Donato y Cequea, 1994; Olvera-Vázquez *et al.*, 2019).

Las nueve especies restantes de *Sechium* han sido brevemente descritas en términos morfológicos y ecogeográficos. Las especies del clado centroamericano (*S. talamancensis*, *S. panamensis*, *S. pittieri*, *S. venosum*, *S. villosum* y *S. tacaco*) tienen características que pueden ser de interés agronómico y ser transferidas a las especies de chayote cultivado (Lira, 1995). El reconocimiento de los ancestros silvestres de las especies cultivadas es de importancia para comprender y realizar mejoramiento genético. Sin embargo, las divergencias morfológicas entre especies cultivadas y silvestres pueden provocar problemas taxonómicos debido a la existencia de estados pleisomórficos y apomórficos. Las secuencias de ADN procedentes de plastidios y ADN ribosomal tienen bajas tasas de mutación y están libres de selección (Patwardhan *et al.*, 2014). Su estudio ha sensibilizado a los científicos sobre los eventos de especiación, domesticación y diversificación

de las especies (Sanjur *et al.*, 2002; Smýkal *et al.*, 2015; Chen *et al.*, 2016; Zhang *et al.*, 2017).

Comprender la filogenia de las especies es importante para la mejora de cultivos. En general, las especies silvestres tienen una amplia variabilidad genética, lo que permite que se adapten y sobrevivan frente a condiciones ambientales adversas (Doebley *et al.*, 1987; EL-Bakatoushi y Ahmed, 2018; Gao *et al.*, 2018). Los genes de interés pueden transmitirse a poblaciones cultivadas utilizando diferentes métodos o técnicas de ingeniería genética. Sin embargo, el éxito de la incorporación de nuevos genes depende de la proximidad filogenética entre especies, asegurando así que la progenie sea fértil y estable a lo largo del tiempo (Patwardhan *et al.*, 2014).

2.2. ORIGEN Y ECOGEOGRAFÍA

Mediante el análisis de los genes cloroplásticos que codifican para la ribulosa bifosfato carboxilasa de cadena larga (*rbcl*), maturasa-K (*matK*), *trnL* y los espaciadores intergénicos *trnL-F* y *rpl20-rps12* en 240 especies de 114 géneros de la familia Cucurbitaceae, y en conjunción con un análisis biogeográfico de parsimonia para determinar la vicarianza y los eventos de dispersión, se encontró que Asia es el centro de origen más probable de los linajes de las cucurbitáceas, los cuales fueron distribuyéndose al resto de los continentes por medio de dispersiones transoceánicas (Schaefer *et al.*, 2009).

La evidencia lingüística, el endemismo de la mayoría de las especies de *Sechium*, apuntan a que todas ellas se originaron en Mesoamérica. Por otra parte, en esta región geográfica se encuentra la mayoría de la diversidad morfológica de estas especies, y es aquí donde únicamente se han encontrado poblaciones silvestres de las mismas (Newstrom, 1990, 1991; Lira y Chiang, 1992; Lira, 1995; Lira *et al.*, 1997b; Lira y Nee, 1999; Cross *et al.*, 2006; Cadena-Iñiguez *et al.*, 2008, 2011, Avendaño-Arrazate *et al.*, 2012, 2017). En México y Guatemala destacan *S. edule*, *S. mexicanum*, *S. compositum*, *S. chinantlense* and *S. hintonii*; en Centroamérica se encuentran *S. panamensis*, *S. talamancensis*, *S. venosum*, *S.*

villosum, *S. pittieri* y *S. tacaco* (Figura 1). La mayoría de las especies del clado centroamericano tienen altos grados de endemismo, donde los factores altitudinales influyen en las características de los nichos ecológicos donde éstas se desarrollan (Wunderlin, 1976; Lira, 1995; Lira *et al.*, 1999; Lira y Nee, 1999). Los frutos, las raíces y las hojas de *Sechium* son ampliamente usados en la gastronomía (Newstrom, 1990; Lira, 1995) debido a que son ricas en agua, vitaminas, azúcares, minerales, fibra y carbohidratos (Aung *et al.*, 1990; Lira, 1996; Islam *et al.*, 2018). Los frutos de *Sechium* también se emplean en la industria para la elaboración de harinas en dietas balanceadas de porcinos (Lalthansanga y Samanta, 2015) y como fuente de endotelina, que es un poderoso vasoconstrictor y mitógeno que se emplea como agente blanqueador de piel (Okamoto *et al.*, 2008).

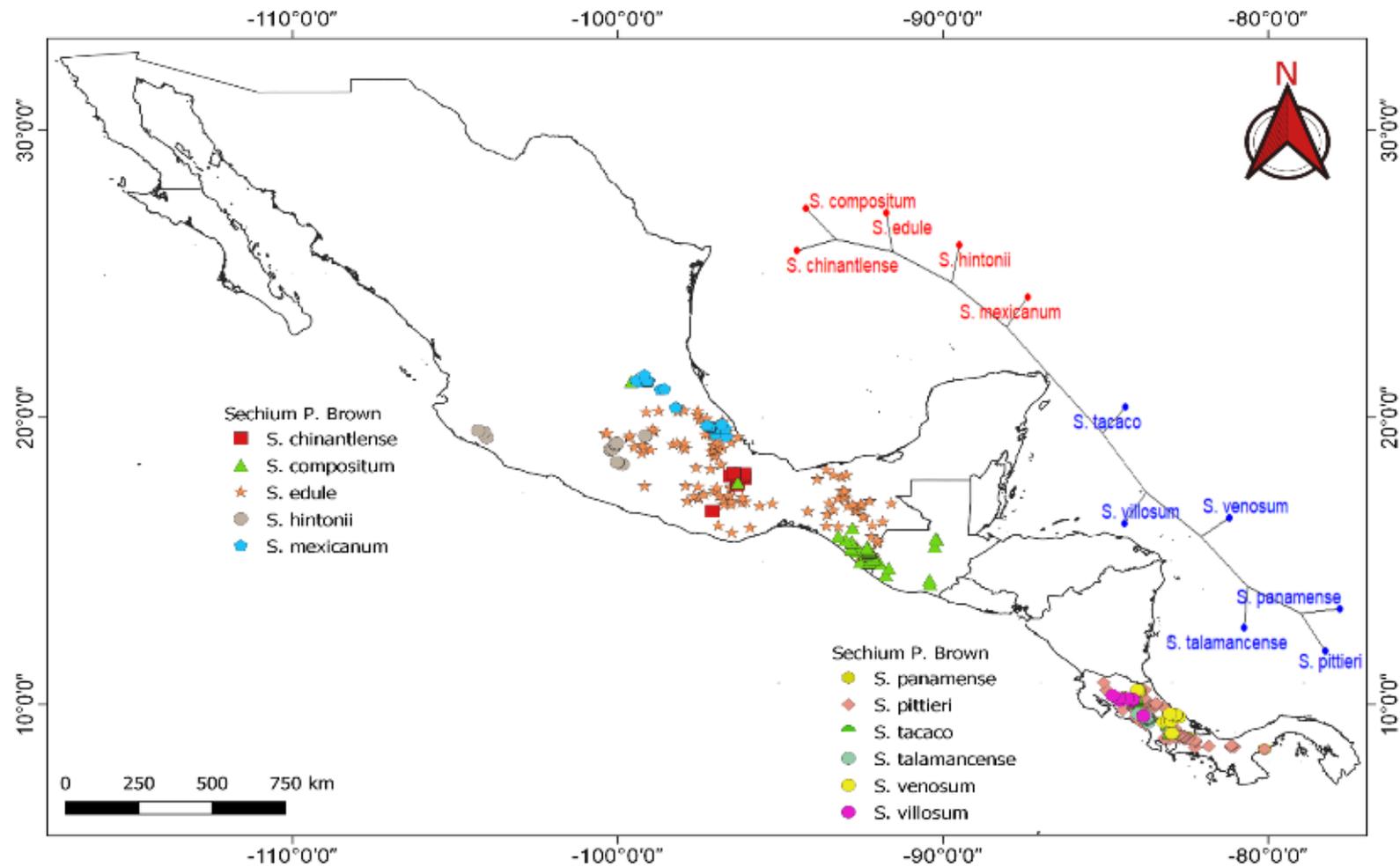


Figura 1. Distribución geográfica de las especies de *Sechium* en Mesoamérica. Fuente: elaboración propia con datos de pasaporte de la Global Diversity Information Facility (GBIF) (<https://www.gbif.org/>)

Sechium edule. - Es la especie representativa de su género debido a su importancia agrícola e industrial. Conocido coloquialmente como “chayote”, palabra proveniente del Náhuatl “*huizt ayotl*” (calabaza con espinas). La amplia riqueza morfológica (Figura 2) y genética de *S. edule* presentes en el sureste de México y Guatemala, además de la presencia de *S. compositum*, que es una especie silvestre altamente relacionada, indican que el chayote fue domesticado en la región de Chiapas-Oaxaca, México (Newstrom, 1991).

Cadena-Iñiguez *et al.* (2011) estudiaron ocho poblaciones de *S. edule* (siete cultivadas y una silvestre) de Veracruz, México. Los estudios bioquímicos han demostrado que la accesión silvestre sintetiza una mayor cantidad de clorofila y cucurbitacinas con respecto a las poblaciones cultivadas. El análisis de componentes principales (ACP) separó las poblaciones de chayote por color, textura y contenido de metabolitos secundarios. Utilizando las mismas poblaciones de *S. edule*, Cadena-Iñiguez *et al.* (2008) las evaluaron mediante 27 variables morfológicas de hojas, flores y frutos. El primer componente del ACP se inclinó para las variables de forma de hojas, tricomas, color del fruto, tamaño y color de la flor, color del pecíolo y estructura de las nervaduras; mientras que, para el segundo componente, se obtuvo una mayor variación para las características del fruto (largo y ancho). El análisis cladístico basado en la morfología y la anatomía mostró que la población silvestre es el ancestro más remoto, pero el más probable de los chayotes cultivados. Yuan *et al.* (2016) compararon los genomas de pepino, melón y sandía, encontrando un grupo de genes que regulan la síntesis de cucurbitacinas, los cuales están asociados a la domesticación y selección; por lo tanto, algunas poblaciones silvestres de cucurbitáceas pueden reducir significativamente el amargor de sus frutos.

Castro *et al.* (2015) evaluaron la diversidad morfológica y bioquímica de siete accesiones de *S. edule* cultivadas en España. La longitud del fruto fue de 9.2-13.1 cm; Cadena *et al.* (2008) reportaron un rango superior de 3-15 cm, en el que otras variables como el peso del fruto y la morfometría de la semilla también se vieron afectadas por la longitud del fruto. Sin embargo, Castro *et al.* (2015)

encontraron una gran diversidad en el color del fruto, desde el verde claro al verde oscuro. Por otro lado, Castro *et al.* (2015) obtuvieron un promedio de 2.27 °Brix, que fue relativamente bajo en comparación con Cadena *et al.* (2011), quienes obtuvieron valores superiores a 10 °Brix, lo cual se debe a la presencia de chayotes con frutos amarillos que tienden a tener un sabor ligeramente dulce.

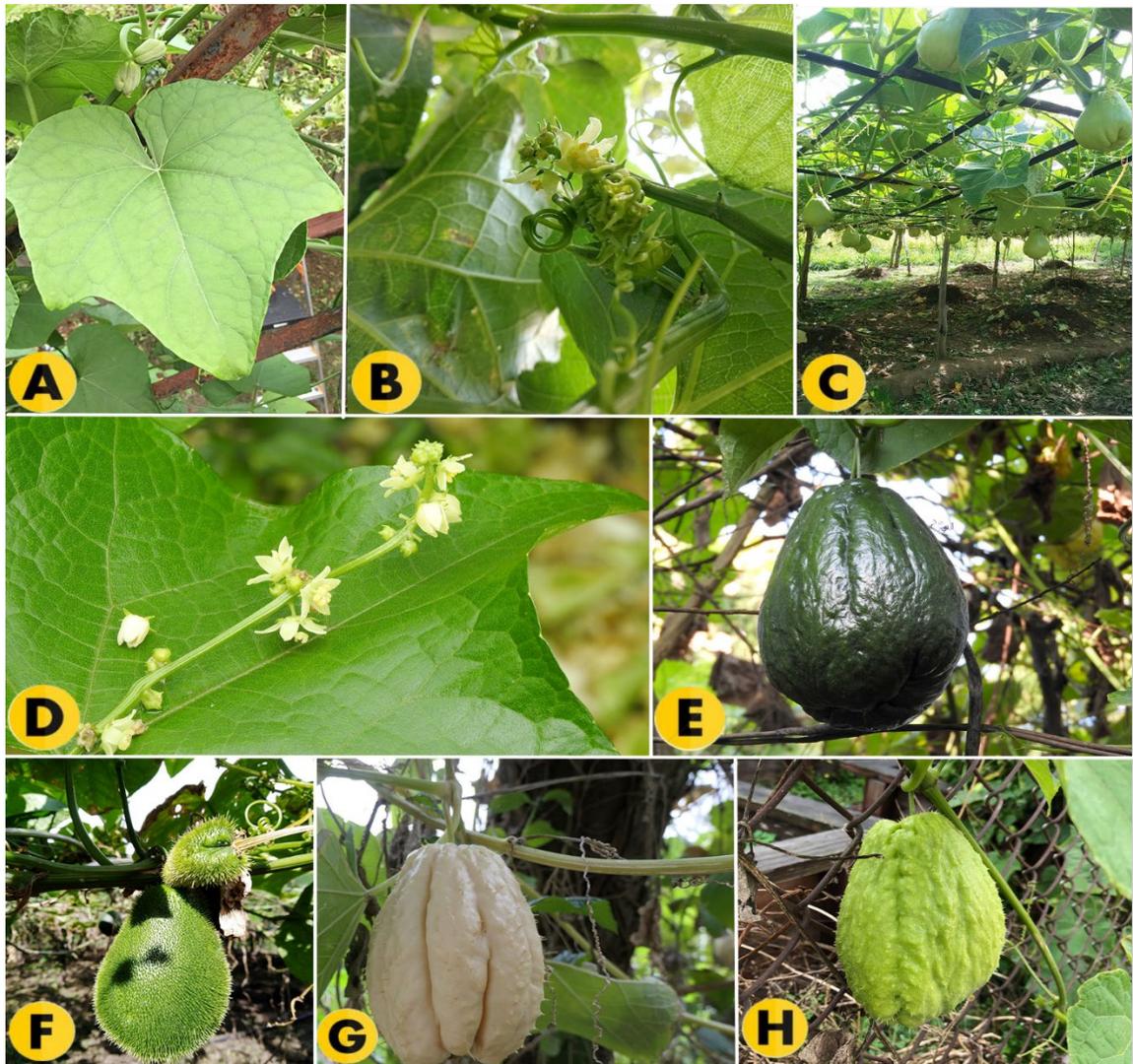


Figura 2. Estructuras morfológicas de *S. edule*: a) hoja, b) zarcillos, c) huerto comercial de chayote, d) inflorescencias, e) - h) variación morfológica de frutos (PH). Fotografías de joevaccaro, Bodo, Thalina García, ming110, elizabeth1, César Lezama, Horacio Sirolli y katabenn.

El chayote es un cultivo introducido en la India, pero con altos índices de diversidad morfológica. Jain *et al.* (2017) evaluaron 36 poblaciones de *S. edule* con 18 marcadores morfológicos relacionados a los frutos y 97 bandas polimórficas obtenidas de marcadores moleculares de Amplificación Directa con ADN de Microsatélites (DAMD). Los resultados del ACP son similares a los obtenidos por Cadena-Iñiguez *et al.* (2008), dado que los componentes principales dan mayor variación a las características del fruto, pero no así para la morfología en flores y hojas, donde Cadena-Iñiguez *et al.* (2008) encontraron una mayor variación. Verma *et al.* (2017) también encontraron diversidad genética moderada en 74 variedades de chayote en la región North Eastern Hill de la India con Inter-Secuencias Simples Repetidas (ISSR) y Polimorfismo de ADN Amplificados al Azar (RAPD). Cabe mencionar que las variedades de color verde oscuro obtuvieron los contenidos más altos de fenoles y vitamina C, y se recomiendan para selección e hibridación a futuro.

Abdelnour y Rocha (2008) evaluaron la diversidad bioquímica con isoenzimas en 42 poblaciones de *S. edule* nativas de Costa Rica y encontraron que 57 % de los 14 *loci* eran polimórficos; 24 poblaciones eran heterocigotas para un solo *locus*, y las otras tenían polimorfismos en dos o tres *loci*. Avendaño-Arrazate *et al.* (2012) reportaron un polimorfismo del 59.8 % en 27 poblaciones cultivadas de *S. edule* y tres poblaciones silvestres con isoenzimas; se encontró que las poblaciones de *S. edule* con frutos y espinas de color verde oscuro tenían los valores de heterocigocidad más altos observados y esperados; aun así, la heterocigocidad media observada fue baja (0.05). La mayor variación genética se encontró mediante análisis isoenzimático en materiales con frutos de color verde oscuro. Actualmente las isoenzimas no son recomendadas para estudios de diversidad genética, ya que su expresión está regulada por el ambiente y los resultados pueden distorsionarse.

La naturaleza recalcitrante de las semillas de *S. edule* impide su conservación a largo plazo, dificultando la obtención de registros arqueológicos que permitan profundizar en temas relacionados con su origen. Sin embargo, mediante

técnicas etnohistóricas y lingüísticas (Guevara *et al.*, 2014), se ha determinado que los aztecas usaban y cultivaban el chayote (Newstrom, 1991; Lira, 1996). Por ser la especie de *Sechium* con mayor grado de domesticación en una región de alta diversidad ecogeográfica y cultural como México, dentro de *S. edule* se han formado complejos varietales que difieren principalmente en las características fisiológicas, químicas y anatómicas relacionadas con los frutos (Cadena-Iñiguez y Arévalo-Galarza, 2011). El mayor grado de heterocigosidad observado se ha encontrado en materiales de color verde intenso y con espinas, características prominentes de los materiales silvestres (Avendaño-Arrazate *et al.*, 2012).

Los frutos de las poblaciones silvestres de *S. edule* tienden a tener sabor amargo y altas concentraciones de cucurbitacinas, y son ricos en flavonoides y compuestos fenólicos que sirven para capturar radicales libres (Cadena-Iñiguez *et al.*, 2011; Riviello-Flores *et al.*, 2018). Estos compuestos químicos tienen propiedades antineoplásicas e inductoras de apoptosis en las células cancerosas y pueden inhibir el proceso metastásico del cáncer de pulmón (Shukla *et al.*, 2016). Los extractos alcohólicos de los complejos varietales de chayote y sus genotipos mejorados H-387 (*S. edule amarus silvestrys* x *S. edule* var. *virens levis* 290) y Perla negra (*S. edule* var. *nigrum minor* x *S. edule* var. *amarus silvestrys*) (Cadena-Iñiguez *et al.*, 2013a) han mostrado resultados positivos en líneas celulares cancerosas como L929 (fibrosarcoma pulmonar de ratón), HeLa (carcinoma cervical humano) y P-388 (macrófagos leucémicos de ratón), con la peculiaridad de tener efectos negativos mínimos en células normales (Monroy-Vázquez *et al.*, 2009; Cadena-Iñiguez *et al.*, 2013b; Aguiñiga-Sánchez *et al.*, 2015, 2017; Salazar-Aguilar *et al.*, 2017).

Los extractos alcohólicos de frutos de *S. edule* disminuyen los niveles de agentes relacionados con enfermedades coronarias y aumentan significativamente los niveles de colesterol de alta densidad (Neeraja *et al.*, 2015). El polvo de frutos secos de *S. edule* var. *nigrum spinosum* tiene efectos significativos en la reducción del estrés oxidativo, lo que ayuda a combatir los problemas del síndrome metabólico en adultos mayores (Rosado-Pérez *et al.*, 2019). Los

extractos alcohólicos de *S. edule* provenientes de raíces y frutos tienen efectos reductores sobre la presión arterial (Lombardo-Earl *et al.*, 2014; Fauziah *et al.*, 2019); también actúan como reguladores de niveles de glucosa en sangre (Maity *et al.*, 2013; Martiar, 2017), protectores de la mucosa gastrointestinal (Sateesh *et al.*, 2012) y tienen actividad antibacteriana frente a *Escherichia coli* gram negativa ATCC 8739, *Salmonella typhimurium* ATCC 3224 y *Shigella flexneri* ATCC 12022 (Kalpana *et al.*, 2013); solo por citar algunos ejemplos.

Respecto al mejoramiento genético para aumentar el contenido de cucurbitacinas tanto en hojas y frutos, se han descubierto que los genes *bi-1* y *Bt-1* son los responsables de sintetizar dichos metabolitos secundarios. A partir de un conjunto de 2416 pares de iniciadores de Secuencias Simples Repetidas (SSR) desarrollados a partir de la secuenciación del genoma del pepino (*Cucumis sativus* L.) (Ren *et al.*, 2009), Zhang *et al.* (2013) utilizaron el análisis de *Locus* de Rasgos Cuantitativos (QTL) para estudiar 149 Líneas Endogámicas Recombinantes (RIL's) producidas por el cruce de las líneas puras 9110Gt (P1) y 9930P (P2) de pepino con el fin de detectar marcadores moleculares vinculados al sabor amargo en frutos y hojas. De esta forma, se detectó que los genes *bi* (*bi-1*) y *Bt* (*Bt-1*) están ligados y confieren sabor amargo a las estructuras mencionadas. Estos genes se encontraron en los cromosomas 5 y 6 del pepino, y podrían tener aplicación para ser detectados en otros géneros de la familia como *Sechium*. Los métodos clásicos de mejoramiento basados en la selección e hibridación entre las especies *S. compositum*, *S. chinantlense* y los complejos varietales de *S. edule*, han producido variedades e híbridos mejorados para consumo humano y con potencial farmacéutico por su alto contenido de cucurbitacinas (Cadena- Iñiguez *et al.*, 2013a; Avendaño-Arrazate *et al.*, 2014).

***Sechium compositum*.** - Su distribución geográfica es desde Chiapas, México hasta Guatemala. Puede desarrollarse hasta una altitud de 2100 m, que le permite crecer en ambientes húmedos, como bosques tropical y mesófilo. Newstrom (1990, 1991) la describió como la especie silvestre más cercana a *S. edule*, a pesar de tener diferente número de cromosomas [(*S. compositum* n =

14) (Lira *et al.*, 2009) (*S. edule* n = 12, 13 y 14) (Palacios, 1987; Mercado *et al.*, 1993; De Donato y Cequea, 1994)]. *Sechium compositum* tiene la singularidad de no tener fisura apical, característica que puede ser considerada como objeto de selección (Lira, 1995). La Figura 3 muestra la morfología de hojas y frutos de *S. compositum*.

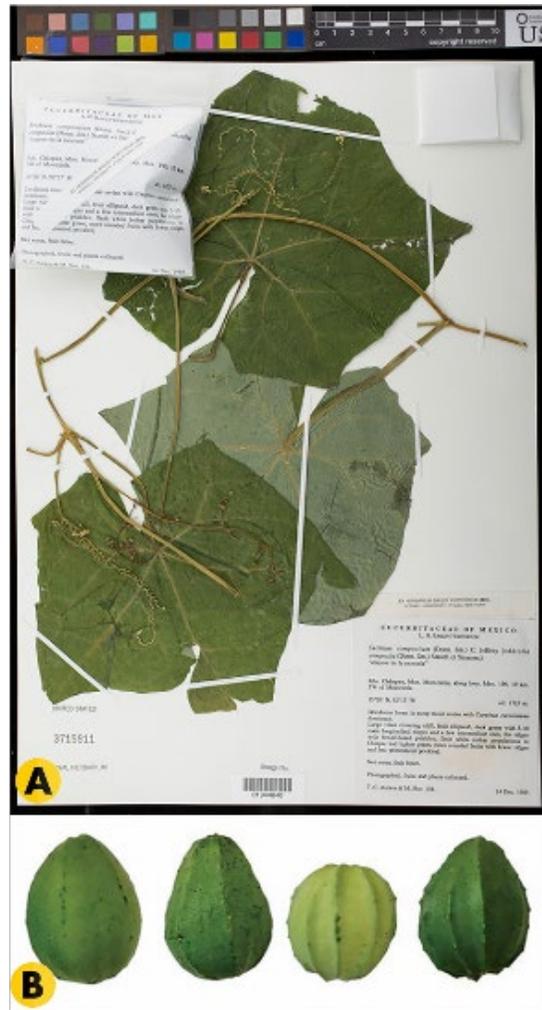


Figura 3. Estructuras morfológicas de *S. compositum*: a) colecta de herbario (creada por Ingrid P. Lin) y b) variación morfológica de sus frutos [Fuente: Avendaño-Arrazate *et al.* (2017)].

***Sechium chinantlense*.** – Esta especie se limita al norte de Oaxaca, México, en una altitud inferior a 800 m. Se desarrolla en bosques mesófilos y caducifolios. *Sechium chinantlense* está relacionada con *S. edule* y *S. compositum*, por lo que su conservación es necesaria para comprender los procesos evolutivos de las especies cultivadas (Lira, 1995; Lira *et al.*, 1999). Estudios taxonómicos de palinología de Lira *et al.* (1994) mostraron que las características físicas cualitativas de los granos de polen (granos espinosos) de *S. chinantlense* pertenecen a la subtribu *Siciynae* y comparten ciertas características con el resto de las especies de *Sechium*. La Figura 4 muestra la morfología de flores y frutos de *S. chinantlense*.



Figura 4. Estructuras morfológicas de *S. chinantlense*: a) inflorescencia (Fuente: PH Javier Castrejón Lira) y b) variación morfológica de sus frutos [Fuente: Avendaño-Arrazate *et al.* (2017)].

Sechium hintonii. - Tiene una distribución restringida en Temascaltepec, Estado de México y en el sureste del estado de Guerrero, México. Prospera en un rango de altitud limitado, fluctuando entre 1300-1500 m en bosques tropicales caducifolios y bosques de robles (Lira, 1995). Recientemente, esta especie se ha recolectado del bosque mesófilo de Veracruz a una altitud de 1340 msnm en el banco de germoplasma del Grupo Interdisciplinario de Investigación en *Sechium edule* en México (GISeM). Pocos especímenes han sido conservados en herbarios y bancos de germoplasma (Figura 5).

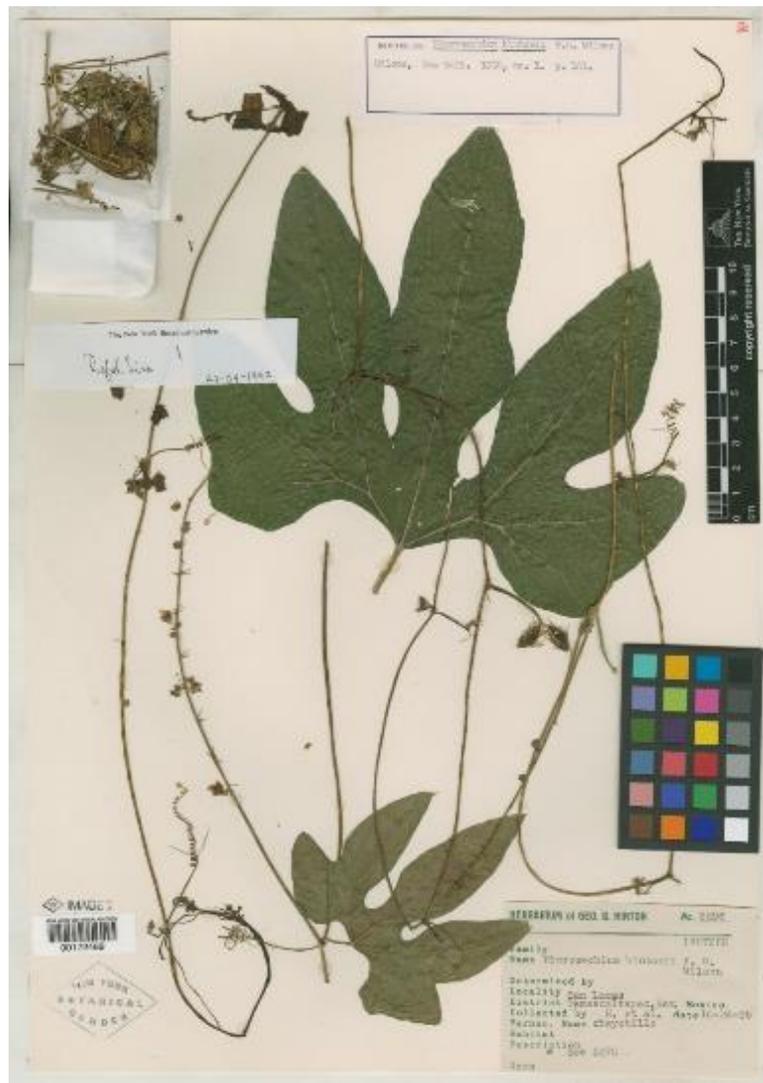


Figura 5. Especimen conservado de *S. hintonii*. Fuente: G.B. Hinton, The New York Botanical Garden.

***Sechium mexicanum*.** - Esta especie es endémica a México y se encuentra en los estados de Hidalgo, Puebla, Querétaro y Veracruz (Figura 6). Se desarrolla en altitudes de 900-2300 msnm en bosques mesófilo, de robles y pinos (Lira y Nee, 1999). Los estudios moleculares sugieren que esta especie debería reubicarse en otro género, y *Sicyos* parece ser el más cercano (Cross *et al.*, 2006; Sebastian *et al.*, 2012).



Figura 6. Estructuras morfológicas de *S. mexicanum*: a) hojas, zarcillos, frutos y b) inflorescencias. Fotografías de Rafael Lira/CONABIO.

***Sechium panamense*.** - Esta especie es endémica a Panamá, donde se desarrolla en altitudes que fluctúan los 1500-3000 msnm (Figura 7). Su descripción botánica se basa únicamente en cuatro muestras recolectadas, por lo que es necesario intensificar su búsqueda para encontrar variación y posibles

hibridaciones que contribuyan a su estudio. Su adaptación a climas fríos la hace ideal para incorporarla en programas de mejoramiento genético, lo que ayudaría a inducir su búsqueda e investigación (Lira, 1995).



Figura 7. Hojas, zarcillos, inflorescencias y fruto de *S. panamense*. PH Thomas Bernad Croat.

Sechium tacaco. - Esta especie está totalmente restringida a toda la cordillera de Talamanca, Costa Rica, en altitudes de 1000-1700 m, y se la conoce localmente como “tacaco” (Figura 8) (Wunderlin, 1976). Existen poblaciones cultivadas y semicultivadas en San José, Costa Rica, donde sus frutos son empleados como vegetales para elaborar platillos regionales (Lira, 1995). Morales (1994) estudió la morfología vegetativa y reproductiva de *S. tacaco* y encontró similitudes con *S. edule* en la estructura anatómica de órganos vegetales, tricomas multicelulares, estructura de nectarios florales y presencia de estomas anomocíticos en órganos vegetativos y reproductivos.

Lira (1995) enfatizó la baja diversidad morfológica de los frutos de *S. tacaco*, excepto por la presencia/ausencia de espinas y la cantidad de fibra. Sin embargo, en un estudio reciente, Monge y Loría (2017) describieron morfológicamente cinco variantes de *S. tacaco* en localidades geográficamente distantes, aunque en altitudes promedio similares de 1100 msnm, encontrando diversidad y diferencias significativas en las características del fruto (peso, largo y ancho). Los hallazgos más importantes de Monge y Loría (2017) se encontraron en el reporte de frutos con 6-7 suturas longitudinales completas y 2-5 suturas longitudinales incompletas, que pueden brindar nueva información para dilucidar procesos evolutivos a nivel de especie y del género *Sechium*. En la Figura 8 se muestran algunas características morfológicas de *S. tacaco*.

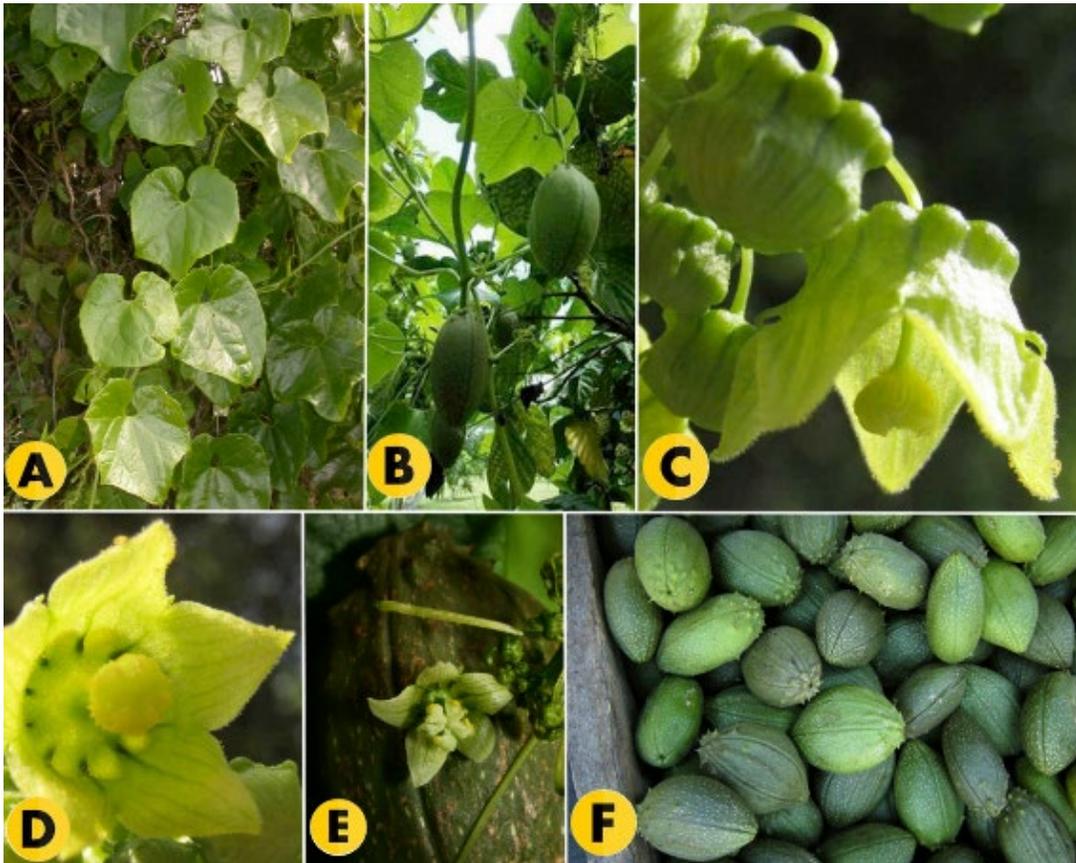


Figura 8. Estructuras morfológicas de *S. tacaco*: a) – b) plantas y frutos; c) – d) estructuras florales y e) – f) diversidad de frutos. Fotografías de Barry Hammel, Arthur de Jesús, Jim Conrad

Sechium talamancense. - También conocido como “chayotillo” y “tacaquillo”, *S. talamancense* es endémico al igual que *S. tacaco*, a la cordillera de Talamanca, Costa Rica. Se encuentra en el bosque de niebla (mesófilo) a altitudes de 2400-3200 msnm. Debido a las bajas temperaturas donde prospera, podría ser de utilidad para la transferencia de genes y mejorar materiales cultivados para resistencia a heladas (Lira, 1995). Pocos especímenes han sido conservados en herbarios (Figura 9).



Figura 9. Especimen conservado de *S. talamancense*. CUCU F 1841927, Rapid Reference Collection (RRC) | Field Museum of Natural History - Keller Action Science Center.

Sechium venosum. - Esta especie es endémica a la costa caribeña de Costa Rica y es muy similar a *S. hintonii* en cuanto a sus inflorescencias pendulares. Se adapta a condiciones con alto grado de humedad (Lira, 1995), sugiriendo que su estructura genética es resistente a fitopatógenos, que son los principales problemas en especies cultivadas de *Sechium* (Olguín-Hernández *et al.*, 2013) y uno de los enfoques más significativos en el mejoramiento genético de especies con importancia agrícola (Newstrom, 1990; Moncayo-Pérez *et al.*, 2020). Al igual que el resto de las especies centroamericanas, se encuentra poco representada en herbarios (Figura 10).



Figura 10. Espécimen conservado de *S. venosum*. Descrita por Gómez Pignataro, 20483, Costa Rica. Missouri Botanical Garden.

***Sechium villosum*.** - Esta especie es endémica a Costa Rica y prospera en ambientes perturbados de bosque tropical o bosque mesófilo, en altitudes de 1500-2000 msnm. Similar a *S. venosum*, se desarrolla en ambientes húmedos, lo que podría ser una fuente de genes de resistencia a enfermedades fúngicas (Lira, 1995). La Figura 11 muestra la morfología de hojas y frutos de *S. villosum*.



Figura 11. Espécimen conservado de *S. villosum*. The New York Botanical Garden.

Sechium pittieri. - Esta especie se distribuye desde Nicaragua hasta Costa Rica, donde puede consumirse como verdura. Su plasticidad fenotípica y genética le permite adaptarse a diversos ambientes con un amplio rango altitudinal que cubre hasta los 2500 msnm, lo que puede ser sinónimo de que *S. pittieri* tenga un amplio rango de diversidad genética, pudiendo también ser utilizado como *S. talamancensis* y *S. panamensis* para mejorar las especies cultivadas (Lira, 1995). La Figura 12 muestra algunas características morfológicas de *S. pittieri*.



Figura 12. Espécimen conservado de *S. pittieri*. Barcode 02728781. PH Covenyor Belt. US National Herbarium, Department of Botany, Smithsonian Institution.

2.3. ANTECEDENTES DE LA CLASIFICACIÓN TAXONÓMICA

Jeffrey (1978) describió siete especies de *Sechium* y las agrupó en dos secciones (*Sechium* y *Frantzia*) con base en las características de sus nectarios florales (Cuadro 1). Las anteras de la sección *Sechium* son generalmente libres, lo que les confiere un estado pleisomórfico, mientras que en *Frantzia*, las anteras están totalmente fusionadas, lo que indica un carácter apomórfico; esta es posiblemente la razón para describir estas dos secciones de *Sechium*. Newstrom (1986, 1991) solo consideró tres especies de *Sechium* (*S. compositum*, *S. edule* y *S. hintonii*) y al igual que Jeffrey (1978), no incluyó a *S. chinantlense* dentro de este grupo (Cuadro 1).

Newstrom decidió agrupar el resto de las especies centroamericanas en los géneros *Polakowskia* y *Frantzia*; estos dos géneros actualmente se encuentran como no aceptados o sin resolver (The Plant List, 2010). En los estudios de Lira y Chiang (1992), Lira (1996), Lira *et al.* (1997a), Lira *et al.* (1997b), Lira *et al.* (1999), Lira y Nee (1999), el género *Sechium* se dividió en las secciones *Sechium* y *Frantzia*. En la sección *Sechium* se incluyó la especie cultivada *S. edule*, así como las especies silvestres más cercanas (*S. compositum*, *S. chinantlense* y *S. hintonii*); también se agregaron las especies centroamericanas semidomesticadas *S. tacaco* y *S. talamancensis*, especies filogenéticamente cercanas. El resto de las especies centroamericanas y *S. mexicanum* se agruparon en la sección *Frantzia* (Cuadro 1).

Lira *et al.* (1997b) estudiaron diez especies de *Sechium* y 11 de la subtribu *Sicyynae* usando taxonomía numérica de 62 caracteres cuantitativos y cualitativos, correspondientes a la palinología y morfología general de las plantas. En el análisis de conglomerados, las diez especies de *Sechium* están bien diferenciadas de las demás; sin embargo, dentro del grupo *Sechium*, no se observó congruencia con respecto a la clasificación de Jeffrey (1978). *Sechium pittieri* y *S. villosum* (*Sechium* sección *Frantzia*) están separados por una considerable distancia euclidiana. En el análisis de componentes principales, las características de las flores y frutos fueron las de mayor variación, y el patrón de

agrupamiento no apoya la clasificación elaborada por Newstrom (1986) (*Sechium sensu stricto* sin secciones, *Polakowskia sensu lato* y *Frantzia* estricta); así, este análisis logró separar las especies de *Sechium* en dos grupos geográficos (México y Centroamérica).

Las poblaciones silvestres de *S. edule* residen principalmente en los estados de Veracruz y Oaxaca, pero con potencial de idoneidad en Querétaro, Hidalgo, Puebla y Guerrero en México. Newstrom (1991) afirmó que las poblaciones silvestres de chayotes silvestres tipo I de Oaxaca eran altamente homogéneas en sus frutos, independientemente de su forma, color verde oscuro y alta densidad de espinas; por otro lado, nombró a las poblaciones de Veracruz como chayotes silvestres tipo II, esto por tener rangos fenotípicos en las características de los frutos antes mencionados. En ambas clasificaciones, los frutos son amargos y fibrosos. Sin embargo, Lira *et al.* (1999) decidieron agruparlos como *S. edule* spp. *edule* para chayotes cultivados y *S. edule* spp. *sylvestris* (chayotes silvestres tipos I y II), los cuales se catalogaron como los antepasados más probables del chayote.

En las tres primeras clasificaciones se utilizaron marcadores morfológicos de *Sechium* (Cuadro 1). Como se mencionó anteriormente, las estructuras morfológicas sufren cambios a través de la selección natural y artificial, lo que dificulta la identificación de las especies a través del estado de rasgos plesiomórficos y apomórficos (Kück y Wägele, 2016). Los cruces e hibridaciones entre especies de *Sechium* también pueden dar lugar a problemas taxonómicos. Sin embargo, muchas de las muestras de la subespecie silvestre tienen una alta probabilidad de ser producto del cruzamiento con otras especies del género, lo que ha sido ampliamente probado en los estudios de Secuencias de Transcripción Interna (ITS) y Secuencias de Transcripción Externa (ETS) de Cross *et al.* (2006), donde las muestras silvestres de *S. edule* se agrupan con *S. chinantlense*, *S. compositum*, y particularmente, con *S. hintonii*. *Sechium* también tiene afinidad con los géneros *Sicyos*, *Microsechium*, *Sicyosperma*, *Parasicyos* y *Sechiopsis* (Figura 13).

Cuadro 1. Clasificación dinámica de las especies de *Sechium*.

Autores	Clasificación
Jeffrey (1978)	<p>Sechium section Sechium <i>S. compositum</i> <i>S. edule</i> <i>S. hintonii</i> <i>S. talamancense</i></p> <p>Sechium section Frantzia <i>S. pittieri</i> <i>S. talamancense</i></p>
Newstrom (1986)	<p>Sechium sensu stricto sin secciones <i>S. compositum</i> <i>S. edule</i> <i>S. hintonii</i></p> <p>Polakowskia sensu lato <i>S. tacaco</i> <i>S. talamancense</i></p> <p>Frantzia sensu stricto <i>F. pittieri</i> <i>F. panamense</i> <i>F. venosum</i> <i>F. villosum</i></p>
<p>Lira (1996) Lira and Nee (1999) Lira y Chiang (1992) Lira et al. (1999)</p>	<p>Sechium section Sechium <i>S. compositum</i> <i>S. chinantlense</i> <i>S. edule</i> <i>S. hintonii</i> <i>S. tacaco</i> <i>S. talamancense</i></p> <p>Sechium sección Frantzia <i>S. mexicanum</i> <i>S. panamense</i> <i>S. pittieri</i> <i>S. venosum</i> <i>S. villosum</i></p>
Cross et al. (2006)	<p>Sechium sensu stricto <i>S. compositum</i> <i>S. chinantlense</i> <i>S. edule</i> <i>S. hintonii</i></p>

Fuente: tomado de Lira et al. (2009)

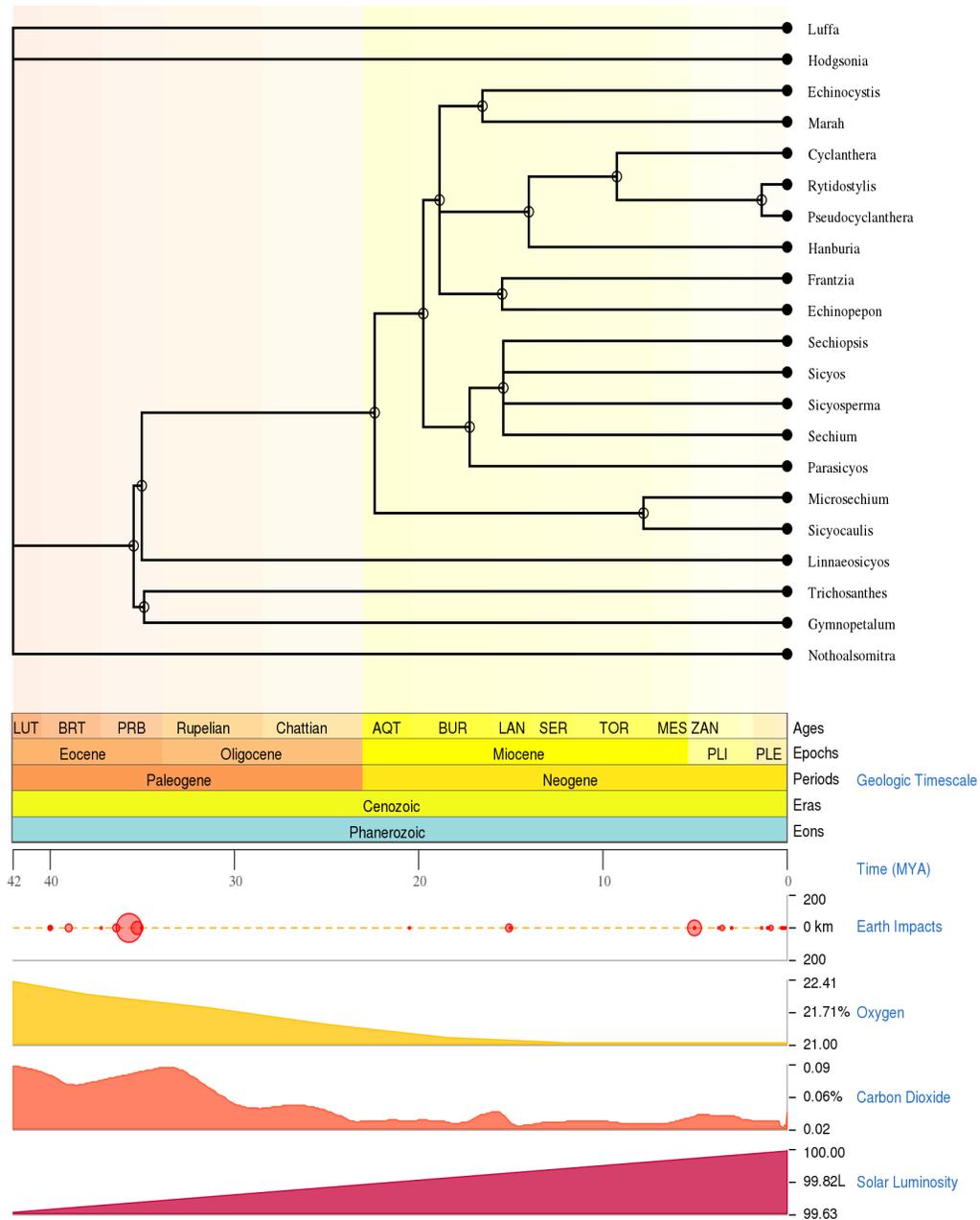


Figura 13. Relaciones filogenéticas de la tribu Sicyoeae. La estimación del tiempo de divergencia para todos los taxones del *TimeTree of Life* (TTOL) está precalculada y almacenada en la base de datos de *TimeTree*. La línea de tiempo evolutiva que se muestra se construyó ubicando primero las especies consultadas o taxones superiores en el TTOL y luego cruzando el árbol hacia la raíz, obteniendo así la estimación del tiempo de divergencia en cada nodo del árbol. *TimeTree* se tomó de Kumar *et al.* (2017) en MEGA 7.0 (Kumar *et al.*, 2016).

2.4. INFERENCIAS FILOGENÉTICAS EN *SECHIUM*

El origen filogenético de *Sechium* tiene varias vertientes dependiendo del tipo y análisis de los datos. Las conclusiones de Lira *et al.* (1997a) fueron objeto de discusión a través de los estudios moleculares de Cross *et al.* (2006), al analizar las secuencias ITS y ETS en 42 especies de la subtribu *Siciynae* [*Rytidostylis carthaginesis* (Jacq.) Kuntze, *Microsechium helleri* (Peyr.) Cogn., *Parasicyos dieterleae* Lira & R. Torres, tres especies de *Sechiopsis*, 27 poblaciones de *Sechium* representadas por *S. chinantlense*, *S. compositum*, las dos subespecies de *S. edule*, *S. hintonii*, *S. mexicanum*, *S. pittieri*, *S. tacaco*, *S. talamancense*, *S. villosum* y nueve especies de *Sicyos*].

Los métodos heurísticos y de máxima parsimonia no demuestran que las especies mexicanas de *Sechium* sean la misma especie (*S. chinantlense*, *S. compositum*, *S. edule*, *S. hintonii*); por otro lado, *S. mexicanum* se agrupó con *Sicyos*, género que a su vez está estrechamente relacionado con *Sechium*. Estos datos de secuencias ITS y ETS no brindan suficiente información sobre la filogenia de las especies mexicanas de *Sechium*, con la excepción de las especies centroamericanas que están bien soportadas.

El árbol consenso de Cross *et al.* (2006) contiene muchas politomías, lo que dificulta la resolución de las relaciones filogenéticas de *Sechium* debido a la probable evolución rápida de estas especies. Sin embargo, la prueba del reloj molecular calculada utilizando un enfoque de estimación de verosimilitud penalizada y datos fósiles de *Sicyos*, sugiere que las especies de *Sechium* divergieron en el Pleistoceno, y las fluctuaciones climáticas llevaron a las especies a aislarse y diferenciarse genéticamente entre períodos interglaciares.

Cross *et al.* (2006) obtuvieron 453 marcadores de Polimorfismos en la Longitud de Fragmentos Amplificados (AFLP) para 127 individuos de *S. edule* spp. *edule*, 21 de *S. edule* spp. *sylvestris*, 20 de *S. chinantlense* y 10 de *S. compositum*. El filograma obtenido por el método del vecino más cercano sugiere que *S. edule* spp. *sylvestris* no forma un grupo monofilético, ya que la geografía de estas

accesiones se asocia a los estados de Oaxaca y Veracruz, así como las respectivas formas cultivadas de chayote; por ejemplo, algunas bandas AFLP eran exclusivas de *S. chinantlense* y poblaciones de *S. edule* spp. *sylvestris* de Oaxaca, abriendo así la posibilidad de flujo genético entre estas especies. El análisis de la varianza molecular sugiere que solo el 7 % de la variación ocurrió entre especies, el 57 % entre poblaciones y el 36 % dentro de los individuos. El elevado número de poblaciones muestreadas puede ser responsable de la mayor variación entre poblaciones.

El número de cromosomas de *S. edule* ha sido estudiado y reportado por varios autores. Los resultados de cada autor difieren considerablemente. Palacios *et al.* (1987) informaron $n = 12$ para *S. edule* de Veracruz; Mercado *et al.* (1993) $n = 13$ para *S. edule* de Oaxaca; De Donato y Cequea (1994) reportaron $n = 14$ en seis variedades de *S. edule*. Estas diferencias pueden deberse a un proceso llamado displodidía, donde existen mutaciones meióticas que provocan translocaciones cromosómicas sin pérdida o ganancia de genoma, aunque pueden cambiar los números de cromosomas y dar lugar a eventos de especiación *a posteriori* (De Storme y Mason, 2014).

Aunque estas tres especies tienen cargas cromosómicas diferentes, parece que es posible la hibridación interespecífica entre ellas. El estudio de AFLP puede explicar parcialmente el cuestionamiento de diferentes números de cromosomas en la subespecie de *S. edule*, como propone Lira (1999). La displodidía puede desempeñar un papel importante en los cruces interespecíficos, el número de cromosomas y la estabilidad de la progenie.

Sebastian *et al.* (2012) estudiaron las secuencias de los genes *rbcL* y *trnL*, los espaciadores intergénicos *trnS-trnG*, *rpl20-rps12*, *psbA-trnH* y la región ITS del ADN ribosómico en 112 accesiones representativas de 87 especies de *Sicyoeae*, incluidas las 11 especies de *Sechium*. El objetivo de este trabajo fue estudiar la filogenia y biogeografía de las especies de *Sicyos*. El filograma de máxima verosimilitud usando el modelo GTR + Γ , como lo muestran Sebastian *et al.*

(2012), logró separar las especies de *Sechium* en dos grandes clados; sin embargo, al igual que Cross *et al.* (2006), *S. mexicanum* no se agrupó en ninguno de estos dos clados, y está más relacionado con las especies de *Sicyos*. El estudio molecular de Sebastian *et al.* (2012) no concuerda con la clasificación taxonómica de Lira (1996), Lira y Nee (1999), Lira y Chiang (1992) y Lira *et al.* (1999) (Cuadro 1), ya que *S. tacaco* y *S. talamancense* no pertenecen a la sección *Sechium* del grupo *Sechium*; además, *S. mexicanum* no debe agruparse con especies centroamericanas (*Sechium* sección *Frantzia*).

El filograma de Sebastian *et al.* (2012) no presentó politomías en el grupo *Sechium* sensu stricto (*S. compositum*, *S. chinantlense*, *S. edule* y *S. hintonii*) y está bien soportado; en este caso, *S. chinantlense* y *S. edule* representan el evento de divergencia más reciente con un millón de años, seguidos por *S. compositum* y *S. hintonii* con 1.9 y 4.9 millones de años, respectivamente. Sin embargo, todavía se requieren pruebas más sólidas para asegurar que *S. compositum* o *S. hintonii* fueron los ancestros directos o participantes en la formación de poblaciones silvestres de *S. edule*.

Las poblaciones de *S. chinantlense* adaptadas a condiciones tropicales se trasladaron a condiciones de mayor altitud, como bosques de pinos y robles, que podrían haberse originado en poblaciones aisladas en regiones montañosas de los estados de Veracruz e Hidalgo, México, y estas barreras geográficas ayudaron a la diferenciación genética entre poblaciones de *S. edule* (Cross *et al.*, 2006).

Respecto al clado centroamericano, las especies divergentes más recientes son *S. pittieri* y *S. villosum*, con aproximadamente 500 mil años. La especie más antigua dentro del clado es *S. venosum*, con una historia de más de 7.4 millones de años, seguida de *S. talamancensis* y *S. panamensis* (Sebastian *et al.*, 2012). Estas tres especies mencionadas se adaptan a las bajas temperaturas en las zonas montañosas de Costa Rica y Panamá.

Aunque el clado centroamericano apenas ha sido estudiado, sus especies tienen características importantes que pueden contribuir al mejoramiento genético de las especies cultivadas (Lira, 1995), ya sea mediante cruces interespecíficos o mediante métodos sofisticados de ingeniería genética. Los datos moleculares pueden funcionar como una herramienta en la clasificación taxonómica de especies. En este contexto, *S. mexicanum* tiene problemas para agruparse en los dos clados de *Sechium*, por lo que tiene más afinidad con especies del género *Sicyos* (Cross *et al.*, 2006, Sebastian *et al.*, 2012). Se necesitan más estudios y revaloraciones para asignar y clasificar correctamente a *S. mexicanum*. Sin embargo, para la fecha de la divergencia del género *Sechium*, se ha tomado como referencia la especie más antigua del género; en este caso, *S. mexicanum*, que comenzó a divergir hace aproximadamente 15 millones de años en la era Cenozoica (Sebastian *et al.*, 2012) (Figura 14).

En el clado mexicano, *S. edule*, *S. compositum* y *S. chinantlense* están filogenéticamente cercanas, lo que indica períodos recientes de divergencia durante el Pleistoceno, caracterizados por períodos glaciares. Sin embargo, también hubo períodos interglaciares, en los que los rayos del sol proporcionaron un clima más cálido (Figura 15).

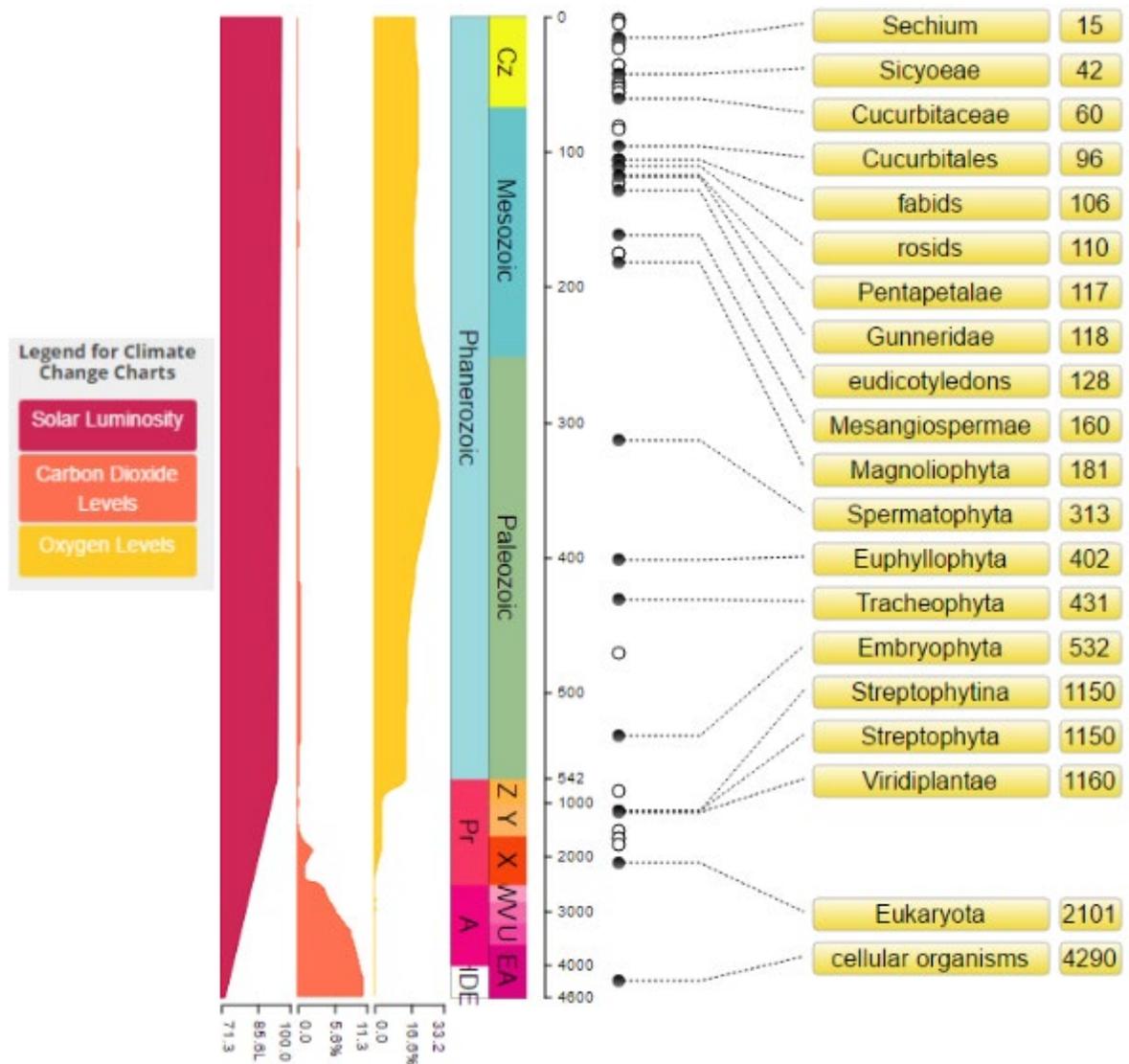


Figura 14. Cronología evolutiva de *Sechium* con tiempos de divergencia en millones de años. La estimación del tiempo de divergencia para todos los taxones del *TimeTree of Life* (TTOL) está precalculada y almacenada en la base de datos de *TimeTree*. La línea de tiempo evolutiva que se muestra se construyó ubicando primero las especies consultadas o taxones superiores en el TTOL y luego cruzando el árbol hacia la raíz, obteniendo así la estimación del tiempo de divergencia en cada nodo del árbol. *TimeTree* se tomó de Kumar *et al.* (2017) en MEGA 7.0 (Kumar *et al.*, 2016). Eras: CZ = Cenozoico, Z = Neo-Proterozoico, Y = Meso-Proterozoico, X = Paleo-Proterozoico, W = Neoarqueano, V = Mesoarqueano, U = Paleoarqueano, EA = Eoarqueano; Eones: Pr = Proterozoico, A = Arcaico, HDE = Hadeon.

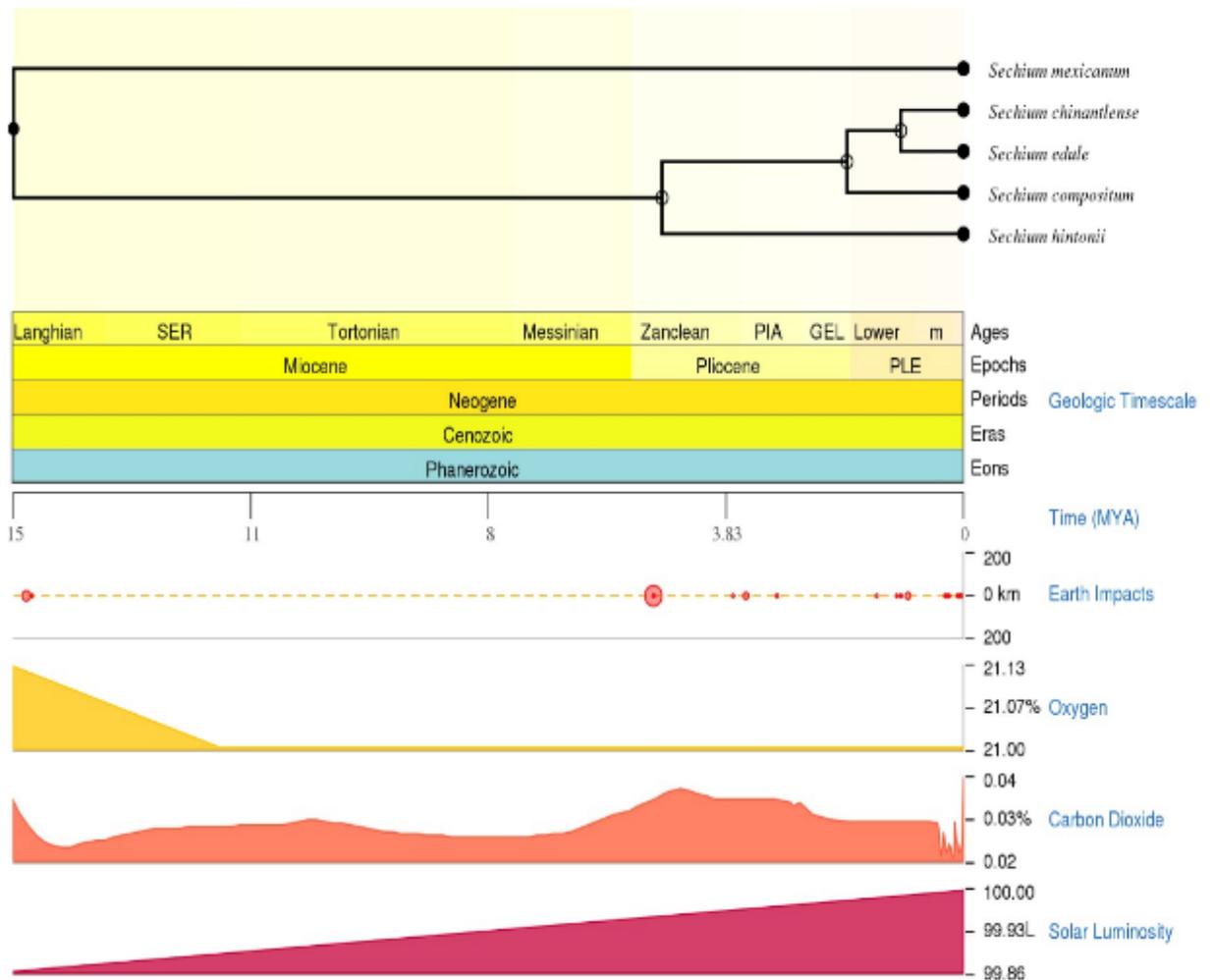


Figura 15. Cronología evolutiva de las especies mexicanas de *Sechium* con tiempos de divergencia en millones de años. Edades: SER = Serravallian, PIA = Piacenzian, GEL = Gelasian; Épocas: PLE: Pleistoceno. Tomado de Sebastian *et al.* (2012) y elaborado con la herramienta *TimeTree* por Kumar *et al.* (2017) en MEGA 7.0 (Kumar *et al.*, 2016).

Estas fluctuaciones térmicas pudieron haber provocado que las especies se contrajeran y expandieran. Durante el período de expansión, existió flujo de genes, introgresión e hibridación entre especies, lo que llevó a una gran diversificación (Ikeda *et al.*, 2012); durante el período de contracción, las especies se aislaron y se condujeron a una especiación peripátrica o diversificación *in situ*, que da origen a especies endémicas (Nevado *et al.*, 2018).

En este caso, *S. chinantlense* y *S. compositum* parecen haber experimentado una hibridación con poblaciones de *S. edule* y luego se aislaron geográficamente, ya que generalmente se encuentran en zonas restringidas. Las colecciones de chayote mexicano se han realizado en regiones montañosas (GISEM, 2015), las cuales destacan por tener una riqueza fisiográfica y climática que permite la diversificación de especies y provocar divergencia a través de la especialización adaptativa, como por ejemplo la emergencia de complejos varietales de chayote (Cadena-Iñiguez *et al.*, 2007, 2008, 2011). En el clado centroamericano, *S. villosum* y *S. pittieri* representan el evento de divergencia más reciente y, en contraste, *S. venosum* se considera la especie más antigua (Figura 16).

Los estudios taxonómicos de Jeffrey, Lira y Newstrom, citados en este estudio, se basaron principalmente en las características morfológicas de flores, frutos, hojas y granos de polen. La morfología vegetal tiende a presentar una cierta plasticidad fenotípica en función de la estructura de interés. Generalmente, las diferencias entre plantas cultivadas y silvestres radican en sus frutos, los cuales son seleccionados por grupos humanos para satisfacer sus necesidades, donde se pueden fijar las mismas características a lo largo de las generaciones. A pesar de la diferente carga cromosómica entre especies de *Sechium*, los cruces intra e interespecíficos ha sido naturalmente posibles, lo que ha generado diferentes cariotipos en las poblaciones silvestres y cultivadas de *S. edule* (Avendaño-Arrazate *et al.*, 2014).

La disploidía puede desempeñar un papel importante en el control del número de cromosomas, lo que permite que la progenie sea estable y fértil. El mejoramiento genético, la fijación de rasgos, las selecciones natural y artificial, así como el tipo de muestras de diferentes zonas geográficas, provocan cambios en la clasificación taxonómica de *Sechium*. A nivel morfológico, es muy complicado discernir el origen de las especies o determinar algún patrón morfológico predominante.

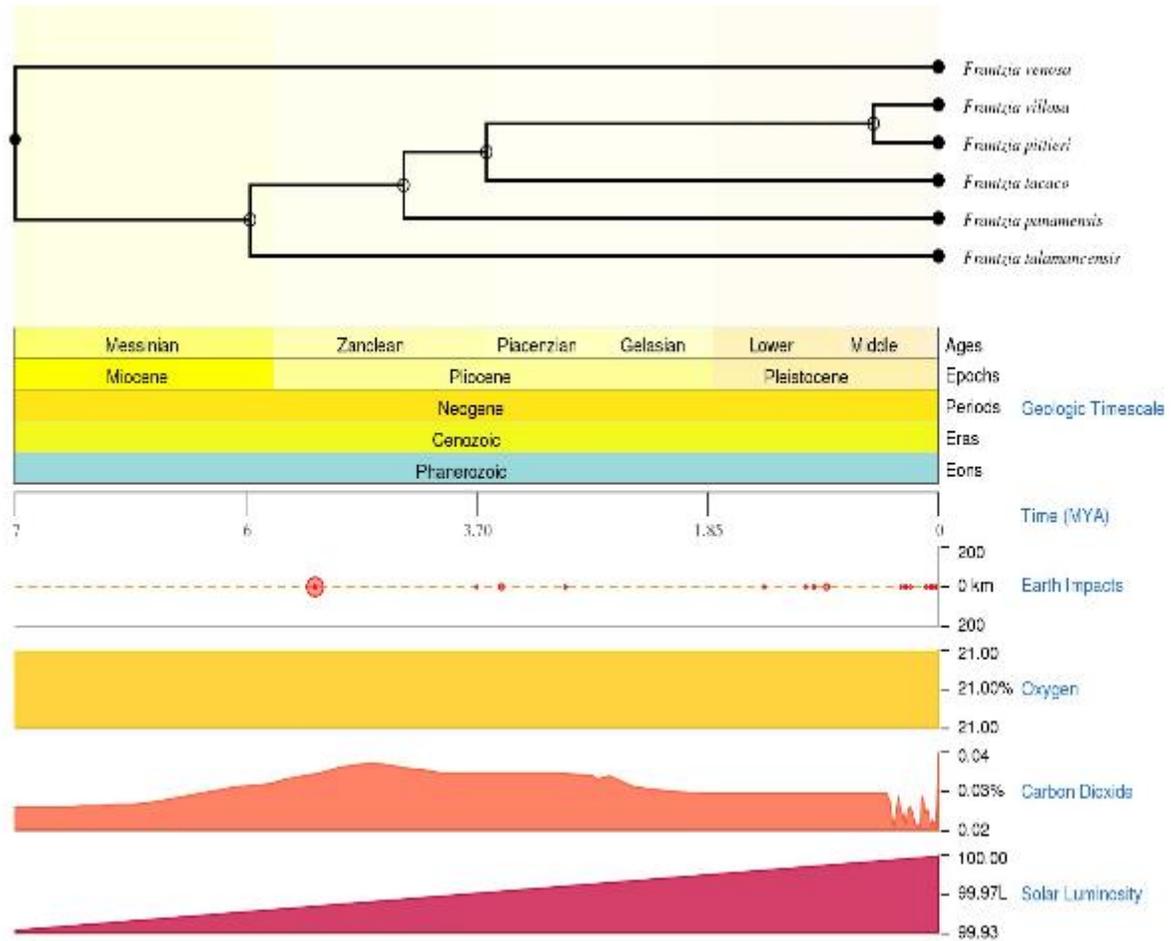


Figura 16. Cronología evolutiva de las especies centroamericanas de *Sechium* (*Frantzia*) con tiempos de divergencia en millones de años. Tomado de Sebastian *et al.* (2012) y elaborado con la herramienta *TimeTree* por Kumar *et al.* (2017) en MEGA 7.0 (Kumar *et al.*, 2016).

En el estudio de Cross *et al.* (2006), se asume que las poblaciones silvestres de *S. edule* se formaron por hibridación con las especies *S. compositum*, *S. chinantlense* y *S. hintonii*. Las tres especies son morfológicamente diferentes, lo que implicaría que las poblaciones silvestres también serían producto de la hibridación; otra explicación es que las poblaciones de estas tres especies fueron sometidas a selección artificial y evolucionaron hasta que sus frutos tuvieron morfologías similares. Actualmente, los complejos varietales de *S. edule*

descritos por Cadena-Iñiguez *et al.* (2007, 2008, 2011) puede proporcionar respuestas sobre el origen de *S. edule*. Los complejos varietales con características de frutos color verde oscuro y presencia de espinas son los más cercanos a los parientes silvestres; le siguen poblaciones con frutos color verde claro y textura lisa o con pocas espinas; mientras que los complejos varietales con frutos amarillos son considerados los más alejados de sus parientes silvestres por especialización adaptativa al ambiente, más que por domesticación. Sin embargo, algunas poblaciones del grupo *albus* tienen características ancestrales, como la presencia de espinas y frutos pequeños. El estudio de estos complejos varietales a nivel de ADN extranuclear puede proporcionar pistas para aclarar el origen filogenético y la diversificación de *S. edule*. Los estudios moleculares aún no han aclarado completamente el origen y la formación de las especies mexicanas de *Sechium*; sin embargo, *S. edule* y *S. chinantlense* constituyen el evento de divergencia más reciente, aun así, se necesita más información para determinar si *S. compositum* y *S. hintonii* participaron en la formación de poblaciones silvestres y cultivadas de *S. edule*. Los eventos de especiación, flujo genético y aislamiento geográfico de las especies mexicanas de *Sechium* probablemente ocurrieron en los períodos interglaciares del Pleistoceno. Quizás una exploración completa del genoma de estas especies pueda ser una herramienta útil. El clado centroamericano está bien respaldado con datos moleculares, pero no con datos morfológicos.

2.5. CONCLUSIONES

Las diferencias en la clasificación taxonómica de las especies de *Sechium* son producto de la amplia variabilidad morfológica presentes entre las muestras estudiadas para cada especie. Esta variación se debe a los procesos de selección natural y artificial, hibridaciones espontáneas, ubicación geográfica de las muestras recolectadas y flujo genético entre especies; además, la existencia de rasgos plesiomórficos y apomórficos también provoca cambios en la clasificación taxonómica.

Se requiere una revaloración para ubicar correctamente *S. mexicanum*, ya que en esta revisión se encontró mayor afinidad con el género *Sicyos*. Los nuevos estudios para *Sechium* requieren un enfoque integral de bases de datos para determinar con mayor precisión las relaciones filogenéticas dentro del género. Por otro lado, las especies centroamericanas de *Sechium* tienen características importantes que pueden introducirse en especies cultivadas a través del mejoramiento genético clásico y asistido por marcadores moleculares.

2.6. LITERATURA CITADA

Abdelnour, E. A., & Rocha, O. (2007). Genetic Characterization of chayote, *Sechium edule* (Jacq) Swartz, in Costa Rica by using isozyme markers. *Genetic Resources and Crop Evolution*, 55(1), 163-170.

Aguñiga-Sánchez, I., Cadena-Íñiguez, J., Santiago-Osorio, E., Gómez-García, G., Mendoza-Núñez, V. M., Rosado-Pérez, J., Ruíz-Ramos, M., Cisneros-Solano, V. M., Ledesma-Martínez, E., Delgado-Bordonave, A. de J., & Soto-Hernández, R. M. (2017). Chemical analyses and *in vitro* and *in vivo* toxicity of fruit methanol extract of *Sechium edule* var. *nigrum spinosum*. *Pharmaceutical Biology*, 55(1), 1638-1645. <https://doi.org/10.1080/13880209.2017.1316746>

Aguñiga-Sánchez, I., Soto-Hernández, M., Cadena-Íñiguez, J., Ruíz-Posadas, L. del M., Cadena-Zamudio, J. D., González-Ugarte, A. K., Steider, B. W., & Santiago-Osorio, E. (2015). Fruit extract from a *Sechium edule* hybrid induce apoptosis in leukaemic cell lines but not in normal cells. *Nutrition and Cancer*, 67(2), 250-257. <https://doi.org/10.1080/01635581.2015.989370>

Aung, L. H., Ball, A., & Kushad, M. (1990). Developmental and Nutritional Aspects of Chayote (*Sechium edule*, Cucurbitaceae). *Economic Botany*, 44(2), 157-164. JSTOR.

- Avendaño-Arrazate, C. H., Cadena-Iñiguez, J., Arévalo-Galarza, M. L. C., Cisneros-Solano, V. M., Morales-Flores, F. J., & Ruiz-Posadas, L. M. (2014). Mejoramiento genético participativo en chayote. *AgroProductividad*, 7, 30-39.
- Avendaño-Arrazate, C. H., Cadena-Iñiguez, J., Cisneros-Solano, V. M., Ramírez-Rodas, Y. C., & Mejía-Montoya, G. B. (2017). Variación morfológica a nivel inter a infraespecífico en *Sechium edule* spp. *AgroProductividad*, 10(9), 58-63.
- Avendaño-Arrazate, Carlos Hugo, Cadena-Iñiguez, J., Arévalo-Galarza, M. L., Cisneros-Solano, V. M., Aguirre-Medina, J. F., Moreno-Pérez, E. del C., Cortés-Cruz, M., Castillo-Martínez, C. R., & Ramírez-Vallejo, P. (2012). Genetic variation of an infraspecific chayote complex evaluated by isoenzymatic systems. *Pesquisa Agropecuária Brasileira*, 47(2), 244-252. <https://doi.org/10.1590/S0100-204X2012000200013>
- Cadena Iñiguez, J., Soto Hernández, M., Arévalo Galarza, M. de L., Avendaño Arrazate, C. H., Aguirre Medina, J. F., & Ruiz Posadas, L. del M. (2011). Biochemical characterization of domesticated varieties of chayote *Sechium edule* (Jacq.) Sw. Fruits compared to wild relatives. *Revista Chapingo Serie Horticultura*, 17(2), 45-55. <https://doi.org/10.5154/r.rchsh.2011.17.044>
- Cadena-Iñiguez, J., Arévalo-Galarza, L., Avendaño-Arrazate, C. H., Soto-Hernández, M., Ruiz-Posadas, L. M., Santiago-Osorio, E., Acosta-Ramos, M., Cisneros-Solano, V. M., Aguirre-Medina, J. F., & Ochoa-Martínez, D. (2007). Production, Genetics, Postharvest Management and Pharmacological Characteristics of *Sechium edule* (Jacq.) Sw. *Fresh Produce*, 1(1), 41-53.

- Cadena-Iñiguez, J., & Arévalo-Galarza, M. L. C. (2011). Las variedades de Chayote (*Sechium edule* (Jacq.) Sw.) y su comercio mundial (1 st ed.). Biblioteca Básica de Agricultura. 98 p.
- Cadena-Iñiguez, J., Soto-Hernández, M., Arévalo-Galarza, M. L. C., Avendaño-Arrazate, C. H., & Aguirre-Medina, J. F. (2013a). Modelos de mejoramiento genético participativo en chayote (*Sechium spp*) (1 st ed.). Editorial del Colegio de Postgraduados.
- Cadena-Iñiguez, J., Soto-Hernández, M., Torres-Salas, A., Aguiñiga-Sánchez, I., Ruíz-Posadas, L., Rivera-Martínez, A. R., Avendaño-Arrazate, C. H., & Santiago-Osorio, E. (2013b). The antiproliferative effect of chayote varieties (*Sechium edule* (Jacq.) Sw.) on tumour cell lines. *Journal of Medicinal Plants Research*, 7(8), 455-460. <https://doi.org/10.5897/JMPR12.866>
- Castro-Rodríguez, J. M., Toledo-Díaz, A. M., Rodríguez-Galdón, B., Perdomo-Molina, A., Rodríguez-Rodríguez, E. M., & Díaz-Romero, C. (2015). Caracterización morfológica y composición química de chayotas (*Sechium edule*) cultivadas en las Islas Canarias (España). *Archivos Latinoamericanos de Nutrición*, 65(4), 2-17.
- Chen, Z., Feng, K., Grover, C. E., Li, P., Liu, F., Wang, Y., Xu, Q., Shang, M., Zhou, Z., Cai, X., Wang, X., Wendel, J. F., Wang, K., & Hua, J. (2016). Chloroplast DNA Structural Variation, Phylogeny, and Age of Divergence among Diploid Cotton Species. *PLOS ONE*, 11(6), e0157183. <https://doi.org/10.1371/journal.pone.0157183>
- Chomicki, G., Schaefer, H., & Renner, S. S. (2019). Origin and domestication of Cucurbitaceae crops: Insights from phylogenies, genomics and archaeology. *New Phytologist*, 226(5), 1240-1255. <https://doi.org/10.1111/nph.16015>

- Cross, H., Lira, S. R., & Motley, T. J. (2006). Origin and diversification of chayote. En T. J. Motley, N. Zerega, & H. Cross (Eds.), *Darwin's Harvest: New Approaches to the Origins, Evolution, and Conservation of Crops* (1^o, pp. 171-194). Columbia University Press.
- De Donato, M., & Cequea, H. (1994). A Cytogenetic Study of Six Cultivars of the Chayote, *Sechium edule* Sw. (Cucurbitaceae). *Journal of Heredity*, 85(3), 238-241. <https://doi.org/10.1093/oxfordjournals.jhered.a111444>
- De Storme, N., & Mason, A. (2014). Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology*, 1, 10-33. <https://doi.org/10.1016/j.cpb.2014.09.002>
- Doebley, J., Goodman, M. M., & Stuber, C. W. (1987). Patterns of isozyme variation between maize and Mexican annual teosinte. *Economic Botany*, 41(2), 234-246. <https://doi.org/10.1007/BF02858971>
- EL-Bakatoushi, R., & Ahmed, D. G. A. (2018). Evaluation of genetic diversity in wild populations of *Peganum harmala* L., a medicinal plant. *Journal of Genetic Engineering and Biotechnology*, 16(1), 143-151. <https://doi.org/10.1016/j.jgeb.2017.11.007>
- Fauziah, N. A., Hidajati, K., & Soejoenoes, A. (2019). The effect of chayote Extract (*Sechium edule*) on blood pressure in pregnant women with hypertension. *Indonesian Journal of Medicine*, 4(3), 266-277. <https://doi.org/10.26911/theijmed.2019.04.03.10>
- Gao, Y., Yin, S., Yang, H., Wu, L., & Yan, Y. (2018). Genetic diversity and phylogenetic relationships of seven *Amorphophallus* species in southwestern China revealed by chloroplast DNA sequences. *mitochondrial DNA Part A*, 29(5), 679-686. <https://doi.org/10.1080/24701394.2017.1350855>

- GISeM. (2015). Diversidad de chayotes en México *Sechium* spp. Distribución de accesiones conservadas en el Banco de Germoplasma de *Sechium edule* (BANGESE) en Huatusco, Veracruz, México. [Mapa].
- Guevara, H. F., Rodríguez, L. L., Gómez, C. H., Pinto, R. R., Rodríguez, G. G., & Perezgrovas, G. R. (2014). Local criteria for chayote (*Sechium edule* Jacq. Sw.) seed selection in rural areas of Chiapas, Mexico. *Acta Agronómica*, 64(2), 178-185. <https://doi.org/10.15446/acag.v64n2.39776>
- Ikeda, H., Carlsen, T., Fujii, N., Brochmann, C., & Setoguchi, H. (2012). Pleistocene climatic oscillations and the speciation history of an alpine endemic and a widespread arctic-alpine plant. *New Phytologist*, 194(2), 583-594. <https://doi.org/10.1111/j.1469-8137.2012.04061.x>
- Islam, S., Kumar, A., Kumar, D. K., & Alom, S. (2018). Physicochemical analysis and nutritional properties of fresh, osmo-dehydrated and dried chayote (*Sechium edule* L.). *Journal of Postharvest Technology*, 6(2), 49-56.
- Jain, J. R., Timsina, B., Satyan, K. B., & Manohar, S. H. (2017). A comparative assessment of morphological and molecular diversity among *Sechium edule* (Jacq.) Sw. Accessions in India. *Biotech*, 7(2), 106. <https://doi.org/10.1007/s13205-017-0726-5>
- Jeffrey, C. (1966). On the Classification of the Cucurbitaceae. *Kew Bulletin*, 20(3), 417-426. JSTOR. <https://doi.org/10.2307/4108235>
- Jeffrey, C. (1978). Further Notes on Cucurbitaceae: IV: Some New-World Taxa. *Kew Bulletin*, 33(2), 347-380. JSTOR. <https://doi.org/10.2307/4109590>
- Kalpana, K., Dhananjaya, K., Ravikumar, K. R., & Mallesha, H. (2013). Antibacterial activity of *Sechium edule* (Jacq.) Swartz against gram negative food borne bacteria. *Advances in Applied Science Research*, 4(2), 259-261.

- Kück, P., & Wägele, J. W. (2016). Plesiomorphic character states cause systematic errors in molecular phylogenetic analyses: A simulation study. *Cladistics*, 32(4), 461-478. <https://doi.org/10.1111/cla.12132>
- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, 34(7), 1812-1819. <https://doi.org/10.1093/molbev/msx116>
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874. <https://doi.org/10.1093/molbev/msw054>
- Lalthansanga, J., & Samanta, A. K. (2015). Effect of feeding chayote (*Sechium edule*) meal on growth performance and nutrient utilization in indigenous pig (Zovawk) of Mizoram. *Veterinary World*, 8(7), 918-923. <https://doi.org/10.14202/vetworld.2015.918-923>
- Lira, R. (1996). Chayote. *Sechium edule* (Jacq.) Sw., promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute., Rome, Italy.
- Lira, R., Alvarado, J. L., & Castrejón, J. (1994). Nota sobre el polen de *Sechium chinantlense* Lira & Chiang y *Parasicyos dieterleae* Lira & Torres (Cucurbitaceae). *Boletín de la Sociedad Botánica de México*, 54, 275-280. <https://doi.org/10.17129/botsoci.1435>
- Lira, R., Caballero, J., & Dávila, P. (1997a). A contribution to the generic delimitation of *Sechium* (Cucurbitaceae, Sicyinae). *Taxon*, 46(2), 269-282. <https://doi.org/10.2307/1224097>

- Lira, R., Eguiarte, L. E., & Montes-Hernández, S. (2009). Proyecto recopilación y análisis de la información existente de las especies de los géneros *Cucurbita* y *Sechium* que crecen y/o se cultivan en México. 107 p.
- Lira, Rafael, Castrejón, J., Zamudio, S., & Zenteno, C. R. (1999). Propuesta de ubicación taxonómica para los chayotes silvestres (*Sechium edule*, Cucurbitaceae) de México. *Acta Botánica Mexicana*, 49, 47-61. <https://doi.org/10.21829/abm49.1999.838>
- Lira, Rafael, & Chiang, F. (1992). Two New Combinations in *Sechium* (Cucurbitaceae) from Central America, and a New Species from Oaxaca, Mexico. *Novon*, 2(3), 227-231. JSTOR. <https://doi.org/10.2307/3391556>
- Lira, Rafael, & Nee, M. (1999). A new species of *Sechium* sect. *Frantzia* (Cucurbitaceae, Sicyeae, Sicyinae) from Mexico. *Brittonia*, 51(2), 204-209. <https://doi.org/10.2307/2666628>
- Lira, Rafael, Villaseñor, J. L., & Dávila, P. D. (1997b). A cladistic analysis of the Subtribe Sicyinae (Cucurbitaceae). *Systematic Botany*, 22(3), 415-425. JSTOR. <https://doi.org/10.2307/2419817>
- Lira, S. R. (1995). Estudios taxonómicos en el género *Sechium* P. Br. Cucurbitaceae [Tesis Doctoral]. Universidad Nacional Autónoma de México.
- Lombardo-Earl, G., Roman-Ramos, R., Zamilpa, A., Herrera-Ruiz, M., Rosas-Salgado, G., Tortoriello, J., & Jiménez-Ferrer, E. (2014). Extracts and fractions from edible roots of *Sechium edule* (Jacq.) Sw. with antihypertensive activity. *Evidence-Based Complementary and Alternative Medicine*, 2014, 594326. <https://doi.org/10.1155/2014/594326>
- Machida-Hirano, R., Cortés-Cruz, M., González, B. A. A., Cadena-Iñiguez, J., Shirata, K., & Watanabe, K. N. (2015). Isolation and characterization of

novel microsatellite markers in chayote [*Sechium edule* (Jacq.) Sw.]. *American Journal of Plant Sciences*, 6(13), 720-726. <https://doi.org/10.4236/ajps.2015.613203>

Maity, S., Firdous, S. M., & Debnath, R. (2013). Evaluation of antidiabetic activity of ethanolic extract of *Sechium edule* fruits in alloxan-induced diabetic rats. *World Journal of Pharmacy and Pharmaceutical Sciences*, 2, 3612-3621.

Martiar, J. S. (2017). Effect of Antihypoglycemic *Sechium edule* Jacq. Swartz. ethanol extract on histopathologic changes in hyperglycemic *Mus musculus* L. *Journal of Medicine*, 2(2), 86-93. <https://doi.org/10.26911/theijmed.2017.02.02.02>

Mercado, P., Lira, R., & Castrejón, J. (1993). Estudios cromosómicos de *Sechium edule* P. Br. y *Sicana naudin* (Cucurbitaceae). 176 p.

Moncayo-Pérez, C. D., Magaña-Lira, N., Peña-Lomelí, A., Leyva-Mir, S.G., Sahagún-Castellanos, J., & Pérez-Grajales, M. (2020). White smut (*Entyloma australe* Speg.) resistance in tomatillo (*Physalis spp.*) germplasm. *Revista Chapingo. Serie Horticultura*, 26(1), 65-77. <https://doi.org/10.5154/r.rchsh.2019.09.016>

Monge, P. J. E., & Loría, M. (2017). Caracterización de frutos de cinco genotipos de tacaco [*Sechium tacaco* (Pittier) C. Jeffrey] en Costa Rica. *Tecnología en Marcha*, 30(3), 71-84. <https://doi.org/10.18845/tm.v30i3.3274>

Monroy-Vázquez, M. E., Soto-Hernández, M., Cadena-Iñiguez, J., Santiago-Osorio, E., Ruiz-Posadas, L. del M., & Rosas-Acevedo, H. (2009). Estudio biodirigido de un extracto alcohólico de frutos de *Sechium edule* (Jacq.) Swartz. *Agrociencia*, 43(8), 777-790.

Morales, A. J. (1994). Morfología general del tacaco, *Sechium tacaco* (Cucurbitaceae). *Revista de Biología Tropical*, 42, 59-71.

- Neeraja, K., Debnath, R., & Firdous, S. M. (2015). Cardioprotective activity of fruits of *Sechium edule*. *Bangladesh Journal of Pharmacology*, 10(1), 125-130. <https://doi.org/10.3329/bjp.v10i1.21329>
- Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *The New Phytologist*, 219(2), 779-793. <https://doi.org/10.1111/nph.15243>
- Newstrom, L. (1986). Studies in the origin and evolution of chayote, *Sechium edule* (Jacq.) Sw. (Cucurbitaceae) [Tesis Doctoral]. University of California.
- Newstrom, L. (1990). Origin and evolution of chayote, *Sechium edule*. In D. M. Bates, R. W. Robinson, & C. Jeffrey (Eds.), *Biology and Utilization of the Cucurbitaceae* (1º, pp. 141-149). Cornell University Press.
- Newstrom, L. (1991). Evidence for the origin of chayote, *Sechium edule* (Cucurbitaceae). *Economic Botany*, 45(3), 410-428. <https://doi.org/10.1007/BF02887082>
- Okamoto, Y., Nomura, A., Yamanaka, A., & Maruyama, S. (2008). The extract of *Sechium edule* fruit reduces endothelin-1 release in cultured cells. *Japan Journal of Food Engineering*, 8, 117-122.
- Olguín-Hernández, G., Valdovinos-Ponce, G., Cadena-Iñiguez, J., & Arévalo-Galarza, M. L. C. (2013). Etiología de la marchitez de plantas de chayote (*Sechium edule*) en el Estado de Veracruz. *Revista mexicana de fitopatología*, 31(2), 161-169.
- Olvera-Vásquez, S. G., Cadena-Iñiguez, J., Gilani, S. A., & Watanabe, K. N. (2019). The cytological studies on neglected and underutilized *Cucurbit* species with special reference to chayote, an Under-Exploited Species.

American Journal of Plant Sciences, 10(8), 1261-1279.
<https://doi.org/10.4236/ajps.2019.108091>

Palacios, R. (1987). Estudio exploratorio del número cromosómico del chayote, *Sechium edule Sw* [Tesis de Licenciatura]. Universidad Veracruzana.

Patwardhan, A., Ray, S., & Roy, A. (2014). Molecular markers in phylogenetic studies-a review. *Journal of Phylogenetics and Evolutionary Biology*, 2(2), 1-9. <https://doi.org/10.4172/2329-9002.1000131>

Ren, Y., Zhang, Z., Liu, J., Staub, J. E., Han, Y., Cheng, Z., Li, X., Lu, J., Miao, H., Kang, H., Xie, B., Gu, X., Wang, X., Du, Y., Jin, W., & Huang, S. (2009). An integrated genetic and cytogenetic map of the *Cucumber* genome. *PLOS ONE*, 4(6), e5795. <https://doi.org/10.1371/journal.pone.0005795>

Riviello-Flores, M. D. la L., Arévalo-Galarza, M. D. L., Cadena-Iñiguez, J., Soto-Hernández, R. M., Ruiz-Posadas, L. D. M., & Gómez-Merino, F. C. (2018). Nutraceutical characteristics of the extracts and juice of chayote (*Sechium edule* (Jacq.) Sw.) Fruits. *Beverages*, 4(2), 37. <https://doi.org/10.3390/beverages4020037>

Rosado-Pérez, J., Aguiñiga-Sánchez, I., Santiago-Osorio, E., & Mendoza-Núñez, V. M. (2019). Effect of *Sechium edule* var. *nigrum spinosum* (Chayote) on oxidative stress and pro-Inflammatory markers in older adults with metabolic syndrome: an exploratory study. *Antioxidants (Basel, Switzerland)*, 8(5), 146. <https://doi.org/10.3390/antiox8050146>

Salazar-Aguilar, S., Ruiz-Posadas, L. D. M., Cadena-Iñiguez, J., Soto-Hernández, M., Santiago-Osorio, E., Aguiñiga-Sánchez, I., Rivera-Martínez, A. R., & Aguirre-Medina, J. F. (2017). *Sechium edule* (Jacq.) Swartz, a new cultivar with antiproliferative potential in a human cervical cancer HeLa Cell Line. *Nutrients*, 9(8), E798. <https://doi.org/10.3390/nu9080798>

- Sanjur, O. I., Piperno, D. R., Andres, T. C., & Wessel-Beaver, L. (2002). Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *PNAS*, *99*(1), 535-540. <https://doi.org/10.1073/pnas.012577299>
- Sateesh, G., Fathullah, H. S., Santhosh, K. G., & Samba, S. R. V. (2012). Anti-ulcer activity of *Sechium edule* ethanolic fruit extract. *The Pharma Innovation*, *1*, 77-81.
- Schaefer, H., Heibl, C., & Renner, S. S. (2009). Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1658), 843-851. <https://doi.org/10.1098/rspb.2008.1447>
- Sebastian, P., Schaefer, H., Lira, R., Telford, I. R. H., & Renner, S. S. (2012). Radiation following long-distance dispersal: The contributions of time, opportunity and diaspore morphology in *Sicyos* (Cucurbitaceae). *Journal of Biogeography*, *39*(8), 1427-1438. <https://doi.org/10.1111/j.1365-2699.2012.02695.x>
- Shukla, S., Sinha, S., Khan, S., Kumar, S., Singh, K., Mitra, K., Maurya, R., & Meeran, S. M. (2016). Cucurbitacin B inhibits the stemness and metastatic abilities of NSCLC via downregulation of canonical Wnt/ β -catenin signaling axis. *Scientific Reports*, *6*, 21860. <https://doi.org/10.1038/srep21860>
- Smýkal, P., Coyne, C. J., Ambrose, M. J., Maxted, N., Schaefer, H., Blair, M. W., Berger, J., Greene, S. L., Nelson, M. N., Besharat, N., Vymyslický, T., Toker, C., Saxena, R. K., Roorkiwal, M., Pandey, M. K., Hu, J., Li, Y. H., Wang, L. X., Guo, Y., ... Varshney, R. K. (2015). Legume crops phylogeny and genetic diversity for science and breeding. *Critical Reviews in Plant Sciences*, *34*(1-3), 43-104. <https://doi.org/10.1080/07352689.2014.897904>

The Plant List. (2010). *Version 1*. <http://www.theplantlist.org/>

Verma, V. K., Pandey, A., Jha, A. K., & Ngachan, S. V. (2017). Genetic characterization of chayote [*Sechium edule* (Jacq.) Swartz.] landraces of northeastern hills of India and conservation measure. *Physiology and Molecular Biology of Plants*, 23(4), 911-924. <https://doi.org/10.1007/s12298-017-0478-z>

Wunderlin, R. P. (1976). Two New Species and a New Combination in *Frantzia* (Cucurbitaceae). *Brittonia*, 28(2), 239-244. JSTOR. <https://doi.org/10.2307/2805833>

Yuan, Z., Ma, Y., Zeng, J., Duan, L., Xue, X., Wang, H., Lin, T., Liu, Z., Zeng, K., Zhong, Y., Zhang, S., Hu, Q., Liu, M., Zhang, H., Reed, J., Moses, T., Liu, X., Huang, P., Qing, Z., ... Huang, S. (2016). Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. *Nature Plants*, 2, E16183. <https://doi.org/10.1038/nplants.2016.183>

Zhang, S., Miao, H., Sun, R., Wang, X., Huang, S., Wehner, T. C., & Gu, X. (2013). Localization of a new gene for bitterness in *Cucumber*. *Journal of Heredity*, 104(1), 134-139. <https://doi.org/10.1093/jhered/ess075>

Zhang, X., Xu, Z., Pei, H., Chen, Z., Tan, X., Hu, J., Yang, B., & Sun, J. (2017). Intraspecific variation and phylogenetic relationships are revealed by ITS1 secondary structure analysis and single-nucleotide polymorphism in *Ganoderma lucidum*. *PLOS ONE*, 12(1), e0169042. <https://doi.org/10.1371/journal.pone.0169042>

Zhou, J., Liu, M., Chen, Y., Xu, S., Guo, Y., & Zhao, L. (2019). Cucurbitacin B suppresses proliferation of pancreatic cancer cells by ceRNA: Effect of miR-146b-5p and lncRNA-AFAP1-AS1. *Journal of Cellular Physiology*, 234(4), 4655-4667. <https://doi.org/10.1002/jcp.27264>

CAPÍTULO III. RELACIONES FILOGENÉTICAS ENTRE ESPECIES MEXICANAS DEL GÉNERO *Sechium* (CUCURBITACEA)

PHYLOGENETIC RELATIONSHIPS AMONG MEXICAN SPECIES OF THE GENUS *Sechium* (CUCURBITACEAE)

RESUMEN

El género *Sechium* incluye 11 especies altamente relacionadas con el género *Sicyos*. *Sechium* se originó en Mesoamérica, pero solamente *S. edule* y *S. tacaco* han sido domesticadas. *Sechium edule* es la especie más aprovechada de su género debido a su amplia diversidad morfogénica, producto de la domesticación y la riqueza ecogeográfica de México. Se evaluaron 68 muestras de las especies *S. chinantlense*, *S. compositum*, *S. hintonii* y *S. edule*, y de ésta última especie se incluyeron sus complejos varietales (grupos *albus*, *nigrum* y *virens*). Se amplificaron los *loci* de ADN mitocondrial (*CoxIII*), ADN cloroplástico (*rbcL*) y la región ribosomal nuclear ITS1-5.8S-ITS2, esto con el objetivo de establecer relaciones filogenéticas que ayuden a comprender procesos evolutivos, los cuales contribuirán a la taxonomía del clado mexicano de *Sechium* para su aprovechamiento en programas de mejoramiento genético. Los árboles filogenéticos obtenidos mediante Inferencia Bayesiana y Máxima Parsimonia mostraron que *S. chinantlense*, *S. compositum* y *S. edule* forman un clúster con alto grado de parentesco filogenético. Posiblemente *S. chinantlense* y *S. compositum* sean poblaciones o variedades derivadas de *S. edule*.

Palabras clave: Evolución, domesticación, taxonomía, *loci*, clado mexicano.

ABSTRACT

The genus *Sechium* includes 11 species highly related to the genus *Sicyos*. *Sechium* is native to Mesoamerica, but only *S. edule* and *S. tacaco* have been domesticated. *Sechium edule* is the most exploited species of the genus due to its broad phenotypic variability. 68 samples from the species *S. chinantlense*, *S. compositum*, *S. hintonii* and *S. edule* were evaluated, as well as their respective varietal complexes (*albus*, *nigrum* and *virens* groups), thereby amplifying the *loci* of mitochondrial DNA (*CoxIII*), chloroplast DNA (*rbcL*) and the nuclear ribosomal ITS1-5.8S-ITS2 region. This was done with the aim of establishing phylogenetic relationships that help to understand evolutionary processes, which, in turn, will contribute to the taxonomy of the Mexican clade of *Sechium* and its exploitation in breeding. Phylogenetic trees obtained from Bayesian Inference and the Maximum Parsimony show that *S. chinantlense*, *S. compositum*, and *S. edule* form a cluster with a high phylogenetic kinship. *S. chinantlense* and *S. compositum* could possibly be populations or varieties derived from *S. edule*.

Key words: evolution, domestication, taxonomy, *loci*, Mexican clade.

3.1. INTRODUCCIÓN

La familia de las cucurbitáceas se originó en Asia (Schaefer *et al.*, 2009). The Plant List registra 11 especies aceptadas del género *Sechium* P. Brown, todas ellas tienen una semilla por fruto, los cuales son carnosos y fibrosos. En un comienzo, *S. panamensis*, *S. tacaco*, *S. talamancensis* y *S. villosum* fueron descritas dentro del género *Frantzia* (Wunderlin, 1976), pero posteriormente se anexaron a *Sechium* (Lira *et al.*, 1999), el cual ya se ha incluido a su vez dentro del género *Sicyos*, al tener alta afinidad filogenética (Sebastian *et al.*, 2012). Lira (1995) describió cuatro especies mexicanas acorde a su morfología, fenología y ecogeografía: *S. chinantlense* Lira & F. Chiang, *S. compositum* (Donn. Sm.) C. Jeffrey, *S. edule* (Jacq.) Swartz, y *S. hintonii* (Paul G. Wilson) C. Jeffrey. El resto del género está formado por *S. mexicanum* Lira & M. Nee, *S. panamense* (Wunderlin) Lira & F. Chiang, *S. pittieri* (Cogn.) C. Jeffrey, *S. talamancense* (Wunderlin) C. Jeffrey, *S. venosum* (LD Gómez) Lira & F. Chiang, *S. villosum* (Wunderlin) C. Jeffrey, y *S. tacaco* (Pittier) C. Jeffrey.

Las especies del género *Sechium* están ampliamente distribuidas en Mesoamérica, cinco especies en México (*S. edule*, *S. compositum*, *S. chinantlense*, *S. hintonii* y *S. mexicanum*) y seis en Centroamérica (*S. panamense*, *S. talamancensis*, *S. venosum*, *S. villosum*, *S. pittieri* y *S. tacaco*) (Wunderlin, 1976; Lira y Chiang, 1992; Lira y Nee, 1999). *Sechium edule* se originó y se domesticó por primera vez en el suroeste de México (Newstrom, 1990, 1991; Cross *et al.*, 2006); junto con *S. tacaco* (sinónimo: *Frantzia tacaco*), son las especies más explotadas dentro del género debido a su variabilidad fenotípica, el cual es producto de la domesticación (Cadena-Iñiguez y Arévalo-Galarza, 2011; Monge y Loría, 2017).

El género *Sechium* fue descrito por P. Brown para dar cabida a la especie *S. edule* (Jeffrey, 1966). El género *Sechium* se ha agrupado junto con los géneros *Microsechium*, *Parasicyos*, *Sechiopsis*, *Sicyos* y *Sicyosperma* en la subtribu *Sicyinae* (Lira *et al.*, 1997). Los géneros *Sechium* y *Sechiopsis* comparten la presencia de nectarios florales en la base del receptáculo (Lira *et al.*, 1997). Los

estudios del espaciador transcrito interno (ITS) de Cross *et al.* (2006) indicaron que *S. edule*, *S. chinantlense* y *S. compositum* son especies genéticamente muy similares y que el género es parafilético, las especies centroamericanas de *Sechium* se agrupan en otro clado alejado de las especies mexicanas. Por lo tanto, estos autores dividieron el género en dos géneros separados (*Sechium* y *Frantzia*). Filogenéticamente estos dos géneros también están distantes, entre ellos, las especies de *Frantzia* son de distribución restringida a Panamá y Costa Rica (Sebastian *et al.*, 2012).

Cross *et al.* (2006) plantearon la hipótesis de que *S. edule* podría haberse originado a partir de hibridaciones interespecíficas. Varios autores han informado el número de cromosomas de *S. edule* y los resultados difieren considerablemente. Mercado *et al.* (1993) reportaron $n = 13$ para *S. edule* de Oaxaca; De Donato y Cequea (1994) informaron $n = 14$ en seis cultivares de *S. edule*. Estas diferencias pueden deberse a un proceso conocido como disploidía, donde la cantidad de cromosomas cambia debido a diversas mutaciones, pero sin pérdida o ganancia de ADN, lo que puede conducir a eventos de especiación desde una perspectiva filogenética (De Storme y Mason, 2014).

Lira *et al.* (1999) sugirieron que *S. edule* incluye dos subespecies (*S. edule* subsp. *edule* y *S. edule* subsp. *sylvestre*), donde la subespecie silvestre (*S. edule* subsp. *sylvestre*) dio origen a la subespecie cultivada. *Sechium edule* comparte características morfológicas en flores y frutos con *S. chinantlense* y *S. compositum*; la estructura estaminal de *S. compositum* es parecida a *S. edule*, pero las ramificaciones de sus filamentos son más delgadas y rectas, por otra parte, sus frutos no tienen la hendidura apical distintiva de los frutos de *S. edule*, además su número cromosómico es $n = 14$. *Sechium chinantlense* ($n = 15$) posee hendidura apical al igual que *S. edule*, pero difieren morfológicamente en la estructura estaminal (Lira *et al.*, 1999) y algunos rasgos palinológicos como el eje polar, forma del grano de polen y el grosor de la exina (Lira *et al.*, 1994). *Sechium compositum* tiene una mayor afinidad con *S. edule* (Newstrom, 1991). *Sechium hintonii* parece ser la especie más lejana del clado mexicano (Cross *et al.*, 2006);

en comparación con las otras especies del género, sus inflorescencias son colgantes y no erectas, y sus frutos son pequeños (2.5–3.5 cm de largo y 2–2.5 cm de ancho) y de color marrón (Lira, 1995).

En el mejoramiento genético, es vital tener conocimiento de los parientes silvestres de los cultivos, que puedan conferir características agronómicas de interés para las especies cultivadas (Dempewolf *et al.*, 2017). Por lo tanto, las poblaciones silvestres de *S. chinantlense* y *S. compositum* relacionadas con *S. edule* pueden albergar características deseables que pueden usarse para mejorar el chayote cultivado (Lira *et al.*, 1999). Los diferentes tipos de cruces constituyen a menudo uno de los métodos clásicos de mejora genética, cuyo éxito dependerá en gran medida de la proximidad filogenética entre especies. En *S. edule* las poblaciones silvestres tienen frutos pequeños de color verde oscuro y sabor amargo; en contraparte, las poblaciones domesticadas pueden tener frutos más grandes con colores desde verde claro a amarillo y con sabores de neutro a dulce. El contraste morfológico entre poblaciones silvestres y domesticadas permite realizar estudios filogenéticos que nos brinden pistas sobre su evolución, estas a su vez, deben complementarse con estudios moleculares de ADN de mitocondria, cloroplasto y ribosomal (Patwardhan *et al.*, 2014), los cuales evolucionan de forma distinta. El ADN mitocondrial generalmente evoluciona más lentamente en un tercio, en comparación con el ADN cloroplástico, que evoluciona un 50 % más lento en comparación con el ADN nuclear (Wolfe *et al.*, 1987).

El objetivo de este estudio fue analizar secuencias genómicas de ADN mitocondrial, de cloroplasto y la región nuclear ITS1-5.8S-ITS2 en 68 muestras de cuatro especies mexicanas de *Sechium*, incluyendo los complejos varietales de *S. edule* para establecer su relación filogenética y comprender su evolución y diversificación. Esto a su vez contribuirá al conocimiento sobre el origen y formación de *S. chinantlense* y *S. compositum*, así como maximizar los recursos filogenéticos de *Sechium* mediante la formulación de estrategias de conservación y mejoramiento.

3.2. MATERIALES Y MÉTODOS

3.2.1. Material genético

Se estudiaron 19 poblaciones de diferentes localidades correspondientes a *S. chinantlense*, *S. compositum*, *S. hintonii* y los complejos varietales de *S. edule* descritos por Cadena-Iñiguez *et al.* (2008, 2011). El número de individuos varió en cada población (Cuadro 2). Se anexaron del GenBank las accesiones DQ535843 y JN560649 del *locus rbcL* correspondientes a *S. mexicanum* y *S. hintonii*. H-387-GISem es producto de hibridación entre *S. edule amarus sylvestris* x *S. edule var. virens levis* 290. Además, “Perla negra” se obtuvo de *S. edule var. nigrum minor* x *S. edule var. amarus sylvestris* (Cadena-Iñiguez *et al.*, 2013). Los datos de pasaporte se muestran en el Cuadro 2.

Cuadro 2. Datos de pasaporte de los materiales genéticos estudiados.

Material genético	Número de colección	Nombres de colectores	Identificador	Herbario
<i>S. edule</i> var. <i>albus dulcis</i>	274-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
	275-05	VMCS		
	285-05			
	286-05			
<i>S. edule</i> var. <i>albus levis</i> (grande)	287-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
	291-05	VMCS		
	366-06			
<i>S. edule</i> var. <i>albus levis</i> (normal)	295-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
	355-06	VMCS		
	293-05			
<i>S. edule</i> var. <i>albus minor</i>	261-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
	262-05	VMCS		
	294-05			
<i>S. edule</i> var. <i>albus spinosum</i>	522-09	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
	283-05	VMCS		
	284-05			
	522-09			

<i>S. edule</i> var. <i>nigrum</i>	331-06	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>conus</i>	348-06	VMCS.		
	351-06			
<i>S. edule</i> var. <i>nigrum</i>	378-07	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>levis</i>	392-07	VMCS		
	467-09			
	263-05			
<i>S. edule</i> var. <i>nigrum</i>	319-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>maxima</i>	372-06	VMCS		
	375-06			
	391-07			
<i>S. edule</i> var. <i>nigrum</i>	396-08	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>minor</i>	397-08	VMCS		
	327-06			
<i>S. edule</i> var. <i>nigrum</i>	354-06	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>spinosum</i>	357-06	VMCS		
	358-06			
	359-06			
<i>S. edule</i> var. <i>nigrum</i>	370-06	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>xalapensis</i>	371-06	VMCS.		
	379-07			
<i>Sechium edule</i> var.	631-12	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
Perla negra	631-13	VMCS.		
	631-14			
<i>Sechium edule</i> H-	387-07	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
387-GISem	387-09	VMCS		
	543-10			
<i>Sechium edule</i> var.	550-10	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México.
<i>virens levis</i>	551-10	VMCS		
	552-10			
<i>Sechium edule</i> var.	546-10	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México
<i>virens levis</i> (amargo)	547-10	VMCS		
	548-10			
	549-10			
<i>Sechium edule</i> wild	273-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de Posgraduados, Montecillo, Texcoco, México
	303-05	VMCS		
	360-06			

	361-06				
	273-05				
<i>Sechium chinantlense</i>	299-05	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de	
	385-07	VMCS		Posgraduados, Montecillo, Texcoco, México	
	386-07				
	390-07				
<i>Sechium compositum</i>	401-08	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de	
	405-09	VMCS		Posgraduados, Montecillo, Texcoco, México	
	406-09				
	544-10				
<i>Sechium hintonii</i>	418-09	JCI, CHAA,	JCI	Hortorio Herbarium, Colegio de	
	418-10	VMCS		Posgraduados, Montecillo, Texcoco, México	
	418-08				
	418-07				

JCI = Jorge Cadena-Iñiguez; CHAA = Carlos Hugo Avendaño-Arrazate; VMCS = Víctor Manuel Cisneros Solano.

3.2.2. Extracción y calidad de ADN

Se utilizaron brotes foliares jóvenes de cada individuo para realizar la extracción de ADN total mediante el protocolo CTAB (Doyle y Doyle, 1987). La calidad del ADN se determinó mediante electroforesis en geles de agarosa al 1.2 % con el amortiguador Tris-Acetato-EDTA (TAE 0.5 X) (Tris-Base, ácido acético y EDTA 0.5 M (pH 8.0)). El ADN se tiñó en bromuro de etidio a 1 mg L⁻¹ y se visualizó en fotodocumentador Kodak Electroforesis Documentación y Análisis (EDAS) 290 (Kodak, Ltd. Rochester, Nueva York, EE.UU.). El ADN se cuantificó con el espectrofotómetro GenesysTM 10 uv (Thermo Scientific, Waltham Massachusetts, EE. UU.). Posteriormente las muestras se diluyeron a una concentración de 10 ng µL⁻¹ para realizar las PCR.

3.2.3. Condiciones de PCR y amplificación de fragmentos

Se emplearon cuatro pares de iniciadores en las PCR: un par para la subunidad III del citocromo c oxidasa (*CoxIII*) del ADN mitocondrial (ADNmt); dos pares para DNA de cloroplasto (ADNcp) correspondientes a ribulosa bifosfato carboxilasa de

cadena larga (*rbcL*) y la región nuclear ITS1-5.8S-ITS2 amplificada por los iniciadores ITS5-ITS4 (Cuadro 3). Las reacciones de PCR se realizaron en el termociclador Techne TC-512 (Bibby Scientific, Vernon Hills, Illinois, EE. UU.). La amplificación se realizó en un volumen de 25 μL : 6 μL de ADN (10 ng μL^{-1}), 2.5 μL de amortiguador (10X), 10 μL de la mezcla de dNTP's (500 μM), 2 μL de MgCl_2 (50 mM), 2 μL de cada cebador (forward y reverse) (10 pM μL^{-1}) y 0.5 μL de ADN Taq Polimerasa (5U μL^{-1}) (Thermo Fisher Scientific, EE. UU.). Las condiciones de la PCR fueron las siguientes: a) un ciclo de desnaturalización inicial a 94 °C durante 3 min; b) 30 ciclos a 94 °C durante 1 min, temperatura de alineamiento (T_m °C) en función de los iniciadores (Cuadro 3) durante 1 min, 72 °C durante 2 min, y c) un ciclo de extensión final a 72 °C durante 10 min.

Cuadro 3. Características de los pares de iniciadores y temperatura de metilación (T_m).

Locus	Iniciador	Secuencia 5'-3'	T_m (°C)	Referencia
<i>CoxIII</i>	Cox3r	F: CTCCCCACCAATAGATAGAG	52	Duminil <i>et al.</i> (2002)
	Cox3f	R: CCGTAGGAGGTGTGATGT	50	Duminil <i>et al.</i> (2002)
<i>rbcL</i>	rbcL1f	F: ATGTCACCACAAACAGAAAC	48	Olmstead <i>et al.</i> (1992)
	rbcL724r	R: TCGCATGTACCTGCAGTAGC	54	Fay <i>et al.</i> (1997)
ITS1-5.8S-ITS2	ITS4	R: TCCTCCGCTTATTGATATGC	50	White <i>et al.</i> (1990)
	ITS5	F: GGAAGTAAAAGTCGTAACAAGG	51	White <i>et al.</i> (1990)

3.2.4. Electroforesis y purificación de fragmentos

Cada fragmento de cada gen por separado (*CoxIII*, *rbcL* y la región nuclear ITS1-5.8S-ITS2) fueron amplificados y separados de forma independiente en geles de agarosa al 1.2 %, se obtuvo una sola banda correspondiente a cada gen a 120 V durante una hora. Para teñir y visualizar las bandas, se siguió el mismo

procedimiento utilizado para corroborar la calidad del ADN. El tamaño de las bandas se determinó con el marcador de peso molecular GeneRuler™ 1 Kb DNA Ladder (Thermo Fisher Scientific, EE. UU.). De acuerdo con las instrucciones del proveedor, los fragmentos de gel se purificaron utilizando el kit de recuperación de ADN Zymoclean (Zymo Research, EE. UU.).

3.2.5. Cuantificación y secuenciación

Después de purificar los productos de PCR, se cuantificaron con un espectrofotómetro NanoDrop™ Lite (Thermo Scientific, Wilmington, NC, EE. UU.) y posteriormente se enviaron a Macrogen, Maryland (www.macrogenusa.com) para la secuenciación de Sanger. Las secuencias obtenidas se resguardaron en el GenBank.

3.2.6. Alineamiento y análisis estadístico

Las secuencias del GenBank de los géneros *Microsechium*, *Sicyos*, *Parasicyos*, *Sechiopsis*, *Cucumis*, *Citrullus* y *Cucurbita* se consideraron como grupos externos. También se incorporaron las secuencias del GenBank reportadas de *Sechium*. Los números de accesión de las secuencias del GenBank de los grupos externos y de *Sechium* aparecen en los dendrogramas (Figuras 18 y 19). Para cada *locus* se realizaron múltiples alineamientos de secuencias utilizando el algoritmo MUSCLE del software MEGA 7.0 (Kumar *et al.*, 2016). Se obtuvo una secuencia consenso con el software SeaView (Gouy *et al.*, 2010) de individuos muestreados de cada población, esto al observarse que las secuencias para el mismo *locus* eran prácticamente idénticas, de esta forma se consideraron para cada material genético representativas y se enviaron al GenBank para su resguardo (Cuadro 4).

La historia evolutiva se infirió utilizando la máxima parsimonia (MP) (Nei y Kumar, 2000). El árbol de consenso bootstrap fue inferido a partir de 1000 réplicas y se utilizó para representar la historia evolutiva de los taxones analizados; el árbol de MP se obtuvo utilizando el algoritmo *Subtree-Pruning-and-Regrafting* (SPR) con MEGA 7.0 (Kumar *et al.*, 2016). Se usó el software MrBayes versión 3.1.2

(Ronquist y Huelsenbeck, 2003) para el análisis de Inferencia Bayesiana (IB) basado en el cálculo de la probabilidad posterior de un árbol. El IB se construyó utilizando el modelo de Tiempo General Reversible con Distribución Gamma más Sitios Invariantes (GTR + G + I), ya que es uno de los modelos más complejos en condiciones estacionarias, reversibles y homogéneas, utilizando un parámetro diferente para modelar cada una de las sustituciones entre nucleótidos y sus frecuencias (Nei y Kumar, 2000). Se utilizó el algoritmo Markov Chain Monte Carlo (MCMC), el cual tiene cuatro cadenas (una cadena fría y tres calientes), las cadenas calientes elevan la probabilidad posterior de los árboles a una potencia entre 0 y 1, lo que favorece el salto entre los picos locales de dicha probabilidad, durante la búsqueda ocurre un intercambio de cadenas, de esta forma la cadena de mayor probabilidad se comporta como cadena fría y las demás como calientes; esto conlleva una búsqueda exhaustiva para un conjunto de árboles posibles. Para conocer la divergencia de las cadenas se analizaron las desviaciones estándar promedio y las diferencias entre ellas a un valor cercano o menor a 0.01. El análisis efectuó 100,000 generaciones con árboles muestreados cada 100 generaciones (Ronquist y Huelsenbeck, 2003).

Cuadro 4. Origen de los materiales genéticos y números de accesión del GenBank.

Material	Origen	Número de accession del GenBank		
		<i>CoxIII</i>	<i>rbcL</i>	ITS1-5.8S-ITS2
<i>S. edule</i> var. <i>albus dulcis</i>	Zaachila, Oaxaca, Méx.	MN984661	MT000109	MT112256
<i>S. edule</i> var. <i>albus levis</i> (grande)	Zaachila, Oaxaca, Méx.	MN984662	MT000110	MT112257
<i>S. edule</i> var. <i>albus levis</i> (normal)	Zaachila, Oaxaca, Méx.	MN984663	MT000111	MT112258
<i>S. edule</i> var. <i>albus minor</i> (3)	Huatusco, Veracruz, Méx.	MN984664	MT000112	MT112259
<i>S. edule</i> var. <i>albus spinosum</i>	Huatusco, Veracruz, Méx.	---	---	MT112260

<i>S. edule</i> var. <i>nigrum</i> <i>conus</i>	Zimatlán, Oaxaca, Méx.	MN984666	MT000114	MT112262
<i>S. edule</i> var. <i>nigrum</i> <i>levis</i> (4)	Huatusco, Veracruz, Méx.	---	MT000115	MT112263
<i>S. edule</i> var. <i>nigrum</i> <i>maxima</i>	Huatusco, Veracruz, Méx.	MN984667	---	MT112264
<i>S. edule</i> var. <i>nigrum</i> <i>minor</i>	Huatusco, Veracruz, Méx.	MN984668	MT000116	MT112265
<i>S. edule</i> var. <i>nigrum</i> <i>spinosum</i>	Zimatlán, Oaxaca, Méx.	MN984669	MT000117	MT112266
<i>S. edule</i> var. <i>nigrum</i> <i>xalapensis</i>	Zaachila, Oaxaca, Méx.	MN984670	MT000118	MT112267
<i>S. edule</i> var. Perla negra	Huatusco, Veracruz, Méx.	---	MT000119	MT112268
<i>S. edule</i> H-387-GISeM	Huatusco, Veracruz, Méx.	MN984665	MT000113	MT112261
<i>S. edule</i> var. <i>virens</i> <i>levis</i>	Coscomatepec , Veracruz, Méx.	MN984673	MT000122	MT112271
<i>S. edule</i> var. <i>virens</i> <i>levis</i> (amargo)	Huatusco, Veracruz, Méx.	MN984676	MT000123	MT112273
<i>S. edule</i> (wild)	Huatusco, Veracruz, Méx.	MN984674	---	MT112272
<i>S. chinantlense</i>	Huatusco, Veracruz, Méx.	MN984671	MT000120	MT112269
<i>S. compositum</i>	Tapachula, Chiapas, Méx.	MN984672	MT000121	MT112270
<i>S. hintonii</i>	Huatusco, Veracruz, Méx.	MN984675	JN560649	MT112274

3.3. RESULTADOS Y DISCUSIÓN

Se obtuvo un total de 179 secuencias para los *loci* *CoxIII* (57), *rbcL* (55) e ITS1-5.8S-ITS2 (67). Sin embargo, se generó una secuencia consenso de los individuos por población, determinado así por un alto grado de parentesco entre las secuencias y para poder visualizarlas en los árboles filogenéticos junto con los grupos externos. Con excepción de *coxIII*, el resto de *loci* tienen secuencias conservadas en más del 50 %; aun así, *CoxIII* mostró un mayor número de sitios parsimonio-informativos (32 %) (Cuadro 5).

Cuadro 5. Resumen general de los parámetros genéticos para cada locus.

Locus	ALPB	SC (%)	SV (%)	SPI (%)
<i>CoxIII</i>	1223	305 (25)	544 (44)	394 (32)
<i>rbcL</i>	1559	909 (58)	516 (33)	175 (11)
ITS1-5.8S-ITS2	737	378 (51)	320 (43)	172 (23)

ALPB = alineamiento longitudinal en pares de bases; SC = sitios conservados; SV = sitios variables; SPI = sitios parsimonio-informativos.

3.3.1. Análisis filogenético *CoxIII*

No hubo grupo externo debido a la ausencia de secuencias en el GenBank para especies de *Sechium* o géneros relacionados, por lo que no se pudo enraizar el árbol filogenético y solo se representó la relación entre las accesiones independientemente de su línea evolutiva. El árbol filogenético con más del 90 % de soporte se obtuvo mediante el método MP (Figura 17), y el árbol IB mostró baja probabilidad posterior en sus ramas, lo cual es inconsistente con el árbol MP. *Sechium edule* var. *nigrum spinosum* está distante del resto de las accesiones; sin embargo, la accesión de *S. hintonii* está relacionada con la accesión silvestre de *S. edule*. Los complejos varietales del grupo *albus* de *S. edule* se encuentran en el grupo remarcado de color rojo junto con *S. compositum* y *S. chinantlense* (Figura 17); el resto de los grupos varietales *nigrum* y la accesión silvestre de *S. edule* están en un segundo clúster; en este segundo grupo están las accesiones que tienen características morfológicas y químicas respectivas a poblaciones silvestres de chayote: frutos color verde oscuro y sabor amargo. Por otra parte, también se encuentra *S. hintonii*, especie del clado mexicano de *Sechium*, pero poco relacionado con *S. edule*.

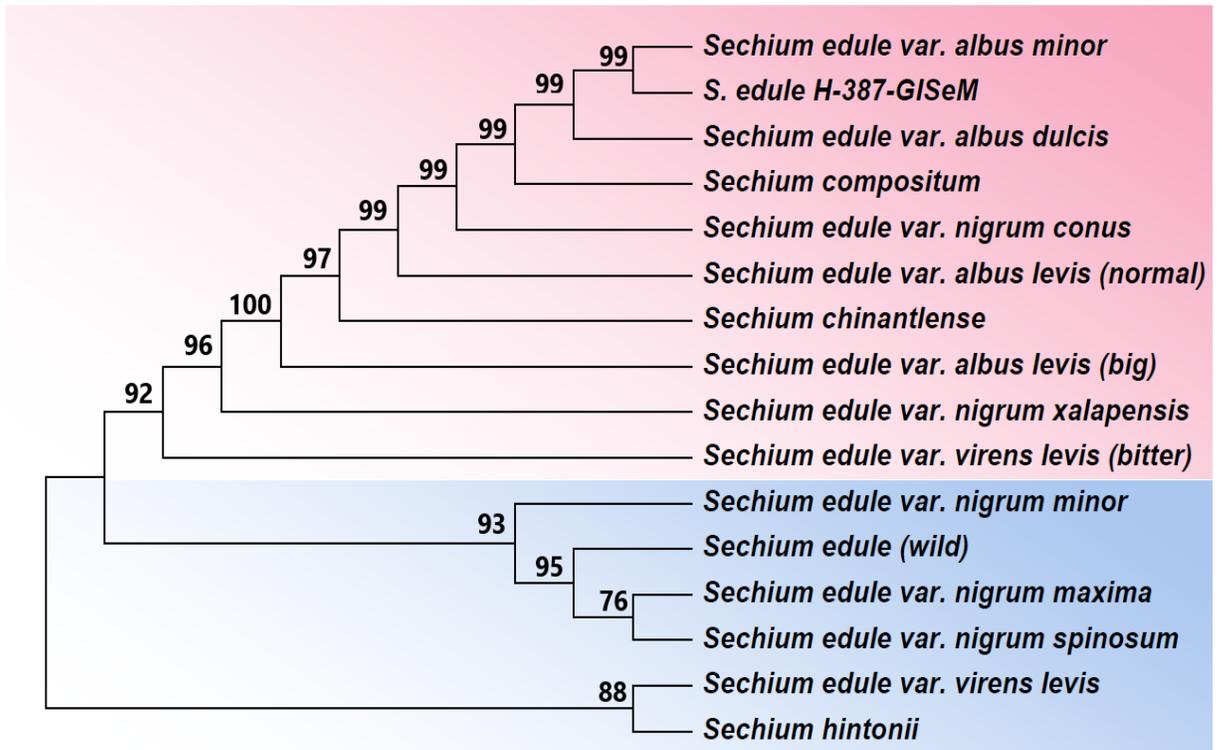


Figura 17. Árbol filogenético resultante de secuencias *CoxIII*. El árbol de consenso de arranque fue inferido a partir de 1000 réplicas y se utiliza para representar la historia evolutiva de los taxones analizados. El porcentaje de árboles replicados, en los que los taxones asociados se agrupan en la prueba de arranque, se muestra junto a las ramas mostrando los 50 árboles más parsimoniosos (longitud = 1502); el índice de consistencia es 0.7 con un índice de retención de 0.6.

3.3.2. Análisis filogenético *rbcl*

Los géneros *Sicyos* y *Sechiopsis* están muy cercanos entre sí con *Sechium*, no así para los géneros *Cucurbita* y *Cucumis*, que se separaron considerablemente, pero con valores bajos de bootstrapping en las ramas del árbol MP (Figura 18). En las especies de *Sechium* y sus complejos varietales se obtuvieron valores superiores al 90 % de soporte en las ramas correspondientes. El árbol filogenético de IB mostró politomías en sus ramas (árbol no mostrado), lo que hace difícil resolver las relaciones evolutivas en las especies mexicanas de *Sechium*. Esta afirmación también se sostiene con los trabajos moleculares de

Cross *et al.* (2006), donde con datos moleculares ITS también se observaron politomías. Las especies *S. chinantlense*, *S. compositum* y algunos complejos varietales del grupo *albus* se agruparon cercanamente; sin embargo, *S. edule* var. *albus levis* (grande) se agrupó con las secuencias de *Sicyos*.

3.3.3. Análisis filogenético de la región ITS

Los árboles IB y MP mostraron valores de soporte superiores a 0.5 para los grupos externos correspondientes a las especies *Sicyos*, *Microsechium*, *Parasicyos* y *Sechiopsis*, pero no para las especies *Sechium* y sus complejos varietales, obteniendo bajos valores de soporte para algunos nodos (Figura 20). Sin embargo, la región nuclear ITS1-5.8S-ITS2 se mantuvo parcialmente consistente con los *loci* *CoxIII* y *rbcL*, es decir, las especies mexicanas de *Sechium* y los complejos varietales de *S. edule* forman un grupo bien diferenciado de los grupos externos, además de *S. chinantlense* y *S. compositum* se relacionaron filogenéticamente con los complejos varietales de *S. edule*. Finalmente, la accesión de *S. hintonii* es la más alejada del clado *Sechium* (Figura 19).

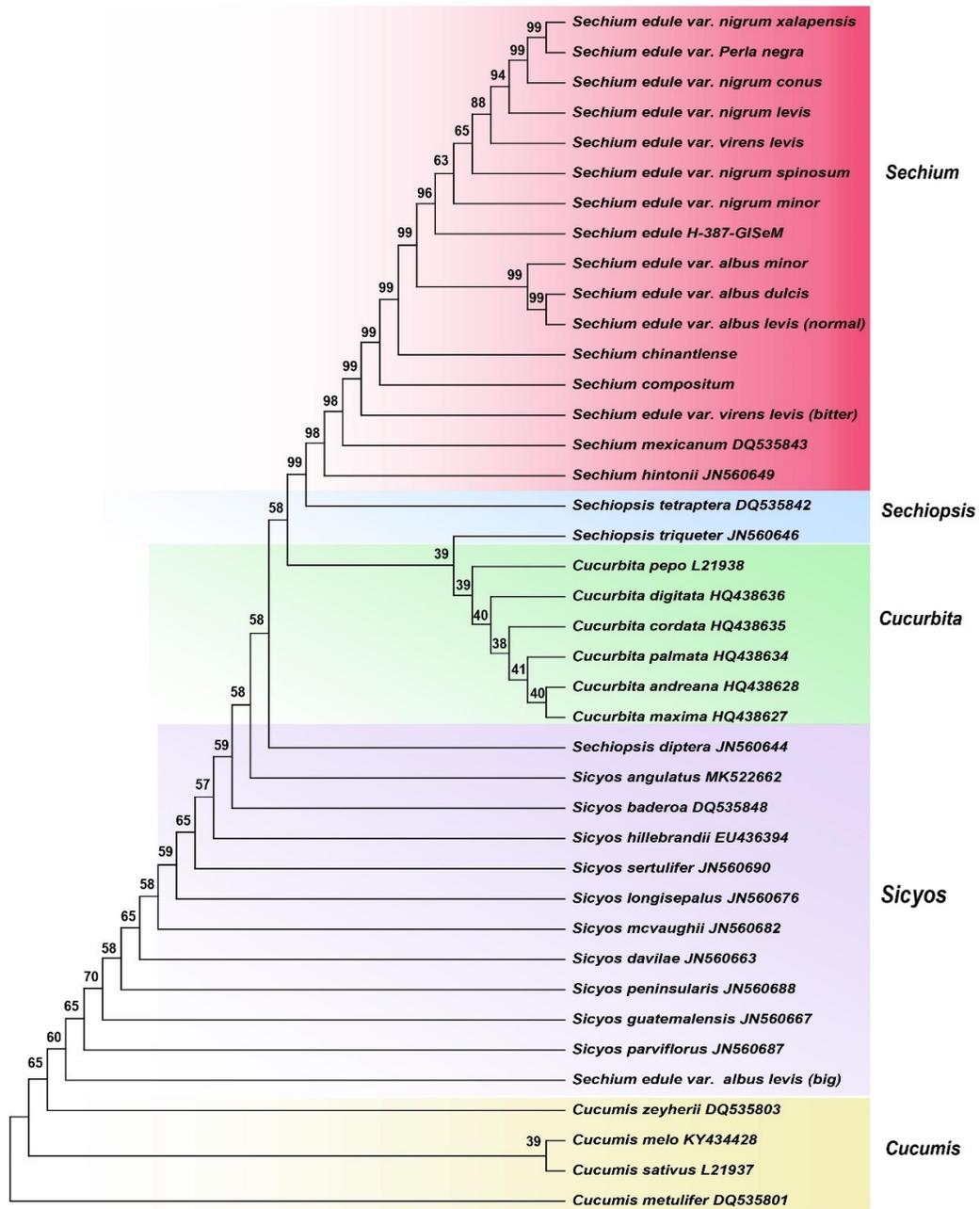
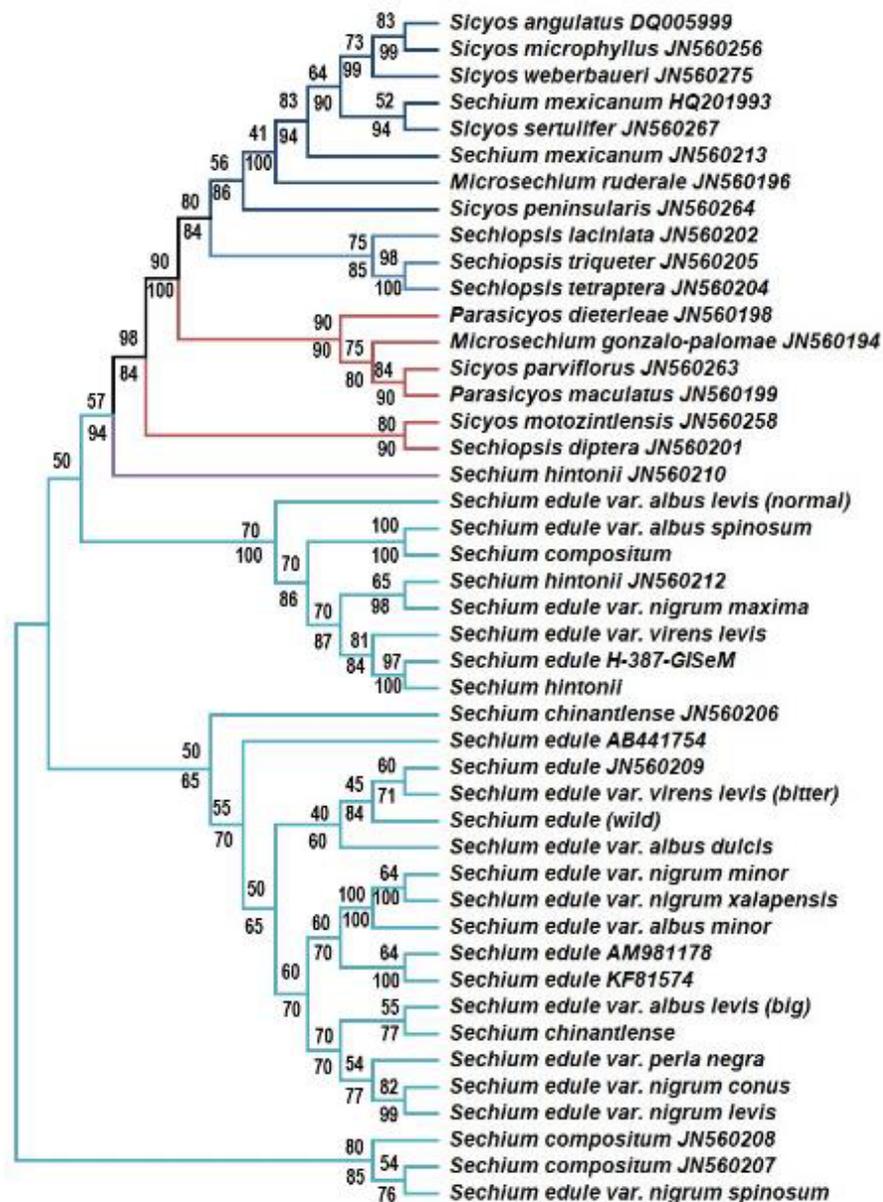


Figura 18. Árbol filogenético resultante de secuencias *rbcL*. El árbol de consenso de arranque inferido de 1000 réplicas se utiliza para representar la historia evolutiva de los taxones analizados. El porcentaje de árboles replicados, en los que los taxones asociados se agrupan en la prueba de arranque, se muestra junto a las ramas. El análisis de la parsimonia produjo 194 árboles más parsimoniosos con una longitud de 1602; el índice de consistencia es de 0.75 con un índice de retención de 0.78.



Sechium

Figura 19. Árbol filogenético IB resultante de secuencias ITS1-5.8S-ITS2 obtenidas de 100.000 generaciones. Las probabilidades posteriores se indican debajo de los nodos, y los valores máximos de soporte de arranque de parsimonia aparecen encima de los nodos. El análisis de parsimonia produjo 83 árboles más parsimoniosos con una longitud de 518; el índice de consistencia es de 0.65 con un índice de retención de 0.79.

3.3.4. Relaciones filogenéticas de especies mexicanas de *Sechium*

La filogenia de *CoxIII* (Figura 17) muestra que los complejos varietales de *S. edule* del grupo *albus* están relacionados con *S. chinantlense* y *S. compositum*. El grupo *albus* corresponde a chayotes distantes genéticamente del ancestro silvestre y se consideran de divergencia reciente con características evolutivas, como el desarrollo de la pubescencia en tallos y hojas, lo que les permite adaptarse a condiciones extremas de luminosidad y temperatura (Cadena-Iñiguez *et al.*, 2008). Asimismo, el híbrido H-387-GISeM es de formación reciente, donde uno de sus padres es *S. edule* var. *virens levis* y el otro padre es *S. edule amarus sylvestris* (no incluido en este estudio).

Sechium compositum y *S. chinantlense* se consideran especies diferentes de *S. edule*. Sin embargo, estudios filogenéticos han indicado que estas tres especies están próximas entre sí y evolucionaron recientemente dentro del clado mexicano de *Sechium* (Cross *et al.*, 2006; Sebastian *et al.*, 2012). El estudio molecular con AFLP de Cross *et al.* (2006) señala que *S. compositum*, *S. chinantlense* y *S. edule* son distintas a nivel de especie, pero con cierto grado de parentesco, es decir, las poblaciones de *S. compositum* se asemejan a las poblaciones de *S. edule* de Veracruz, y las poblaciones de *S. chinantlense* con las poblaciones de *S. edule* de Chiapas y Guerrero. El ADN nuclear es empleado para obtener distintos marcadores moleculares, entre ellos los AFLP, sin embargo, el ADN nuclear evoluciona más rápido en comparación con el ADN extranuclear, por lo que son poco recomendados para estudiar filogenia, aunque pueden dar ciertas evidencias de hibridaciones. Las primeras inferencias filogenéticas de las especies de *Sechium* se hicieron sobre marcadores morfológicos, los cuales son susceptibles a procesos como la domesticación y los efectos ambientales, por otra parte, es posible tener distintas perspectivas en la taxonomía de *Sechium* en función de la región geográfica de donde se analizaron los especímenes, ya que también existe el riesgo de que algunas de ellas sean producto de hibridación. Los complejos varietales del grupo *nigrum* están más próximos a las poblaciones silvestres de *S. edule*, y estos complejos fueron considerados los más cercanos

a los parientes silvestres en cuanto a características morfológicas (frutos de color verde oscuro, pequeños, amargos y con espinas) y fitoquímicos (clorofila y contenido de cucurbitacinas) (Cadena-Iñiguez *et al.*, 2008, 2011). *Sechium hintonii* y *S. edule* wild están filogenéticamente cercanos, lo que respalda los resultados de Cross *et al.* (2006).

La filogenia de *rbcL* (Figura 18) muestra que la mayoría de las accesiones de *Sechium* se agruparon en un solo grupo bien diferenciado de *Sicyos* y *Sechiopsis*. Sin embargo, la accesión de *S. mexicanum* se agrupó con *Sicyos*, *Sechiopsis* y *Microsechium*. Cross *et al.* (2006) y Sebastian *et al.* (2012) también obtuvieron los mismos resultados y estos últimos concluyeron que *S. mexicanum* debe agruparse en *Sicyos*. *Sechium hintonii* y *S. mexicanum* son los más antiguos dentro del clado mexicano de *Sechium*, con estimaciones de tiempo de divergencia de reloj molecular para *S. hintonii* de 5 millones de años, y de 15 millones de años para *S. mexicanum* (Sebastian *et al.*, 2012).

En la filogenia ITS1-5.8S-ITS2 (Figura 19) todas las accesiones de *Sechium* formaron un clado bien diferenciado de los grupos externos, a excepción de *S. hintonii* (JN560210), que comparte cierta similitud con los géneros externos. *Sechium hintonii* se distribuye únicamente en Guerrero y Estado de México, sus frutos son pequeños comparados con *S. edule* y carecen de hendidura apical (Lira *et al.* 1999); también muestra diferencias con especies mexicanas de *Sechium* con respecto a la morfología del polen: eje polar, eje ecuatorial y largo de espinas. En un principio fue clasificado como *Microsechium hintonii* Paul G. Wilson.

En la filogenia de *rbcL* se muestra un clado de *Sechium* bien diferenciado, pero con la inclusión de *Sicyos angulatus*. Ambos árboles filogenéticos (MP e IB) no tienen buen soporte para bootstrap en algunas ramas. Las discrepancias entre estos agrupamientos se pueden deber a diferentes tasas de mutación entre genes, ya que el ADNmt generalmente evoluciona más lentamente en un tercio, en comparación con el ADNcp, que evoluciona un 50 % más lento en

comparación con el ADN nuclear. Es decir, un ADN altamente conservado a través del tiempo o que ha sufrido pocas mutaciones, puede ayudar a hacer inferencias filogenéticas de las especies y no se vería tan afectado por los cambios morfológicos que las estructuras botánicas sí pudieran llegar a sufrir (Wolfe *et al.*, 1987). Sin embargo, tanto las tasas de mutación en el ADN nuclear como en el extranuclear pueden variar en función de los diferentes linajes de plantas (Stenøien, 2008; Yao *et al.*, 2019).

Las especies *S. mexicanum* y *S. hintonii* fueron más distantes en comparación con *S. chinantlense*, *S. compositum* y *S. edule*; sin embargo, algunas accesiones de *S. edule* silvestre estudiadas con ITS muestran cierta similitud con *S. hintonii*, posiblemente porque las accesiones de dichas especies fueron recolectadas en Guerrero, México y tal vez ocurrieron cruzamientos espontáneos que expliquen su similitud (Cross *et al.*, 2006). Estas cuatro especies forman un clado bien diferenciado en comparación con todas las especies del clado centroamericano de *Sechium*. Las primeras poblaciones de *S. chinantlense* y *S. compositum* podrían haber sido diferentes en sus inicios, a pesar de la dificultad que existe para demostrar la hibridación natural de estas dos especies con *S. edule*, Cadena-Iñiguez *et al.*, (2013) han realizado hibridaciones artificiales, obteniendo así nuevas variantes fenotípicas.

Los números de cromosomas de *S. compositum*, *S. chinantlense* y *S. edule* varían considerablemente; $n = 14$ para *S. compositum*, $n = 15$ para *S. chinantlense* (Lira, 1999), y existen reportes de $n = 13$ y 14 para *S. edule*, que dependen de la zona geográfica de las muestras recolectadas (Mercado *et al.*, 1993; De Donato y Cequea, 1994). Aunque estas condiciones pueden ser una barrera para la deriva genética y la reproducción entre estas especies, en la disploidía, el reordenamiento del ADN debido a irregularidades en la división meiótica puede alterar el número de cromosomas, conservando así la misma carga genética (De Storme y Mason, 2014). Esto asegurará de alguna manera el éxito de la hibridación, dando como resultado genotipos estables y mejorados (Cadena-Iñiguez *et al.*, 2013).

3.4. CONCLUSIONES

Los complejos varietales de *S. edule* se han diversificado a lo largo del tiempo a través de los efectos de la especialización adaptativa, así como de los procesos de selección natural y artificial. Los complejos varietales del grupo *nigrum* son los más antiguos y cercanos a un ancestro silvestre; sin embargo, debido a las múltiples variantes fenotípicas que existen en el estado de Veracruz, México, es necesario evaluar aún más poblaciones e individuos que ayuden a comprender la historia evolutiva de las poblaciones silvestres de *S. edule*. Mientras tanto, los complejos varietales del grupo *albus* son los más recientes en términos evolutivos. Las especies *S. chinantlense*, *S. compositum*, *S. hintonii* y *S. edule* formaron un grupo dentro del clado mexicano de *Sechium*. Las poblaciones de *S. edule* y sus complejos varietales podrían haberse originado o participado en la formación de *S. chinantlense* y *S. compositum*, y estas a su vez podrían incluirse como variantes de *S. edule*. La afinidad filogenética detectada entre estas tres especies (*S. chinantlense*, *S. compositum* y *S. edule*) sugiere la posibilidad de realizar mejoramiento genético entre ellas.

3.5. LITERATURA CITADA

- Avendaño-Arrazate CH, Cadena-Iñiguez J, Arévalo-Galarza MLC, Cisneros-Solano VM, Aguirre-Medina J *et al.* (2012). Genetic variation of an infraspecific chayote complex evaluated by isoenzymatic systems. *Pesquisa Agropecuária Brasileira* 47 (2): 244-252. doi:10.1590/S0100-204X2012000200013
- Avendaño-Arrazate CH, Cadena-Iñiguez J, Cisneros-Solano VM, Ramírez-Rodas YC, Mejía-Montoya GB (2017). Variación morfológica a nivel inter e infraespecífico en *Sechium* spp. *Agroproductividad* 10 (9): 58-63.
- Cadena-Iñiguez J, Arévalo-Galarza MLC (2011). Las variedades de Chayote (*Sechium edule* (Jacq.) Sw.) y su comercio mundial. 1st ed. Colegio de Postgraduados, Montecillo, Texcoco: bba. 98 p.

- Cadena-Iñiguez J, Arévalo-Galarza MLC, Avendaño-Arrazate CH, Soto-Hernández M, Ruiz-Posadas LM *et al.* (2007). Production, genetics, postharvest management and pharmacological characteristics of *Sechium edule* (Jacq.) Sw. *Fresh Produce* 1 (1): 41-53.
- Cadena-Iñiguez J, Avendaño-Arrazate CH, Soto-Hernández M, Ruiz-Posadas LM, Aguirre Medina J *et al.* (2008). Intraspecific variation of *Sechium edule* (Jacq.) Sw. in the state of Veracruz, México. *Genetic Resources and Crop Evolution* 55: 835-847. doi:10.1007/s10722-007-9288-4
- Cadena-Iñiguez J, Avendaño-Arrazate CH, Cisneros-Solano VM, Arévalo-Galarza MLC, Aguirre-Medina JF (2013). Modelos de mejoramiento genético participativo en chayote (*Sechium* spp). 1st ed. México: Editorial del Colegio de Postgraduados.
- Cadena-Iñiguez J, Soto-Hernández M, Arévalo-Galarza MLC, Avendaño-Arrazate CH, Aguirre-Medina JF *et al.* (2011). Biochemical characterization of domesticated varieties of chayote *Sechium edule* (Jacq.) Sw. fruits compared to wild relatives. *Revista Chapingo Serie Horticultura* 17 (2): 45-55.
- Cross H, Lira SR, Motley TJ (2006). Origin and diversification of chayote. In: Motley TJ, Zerega N, Cross H (editors). *Darwin's Harvest: new approaches to the origins, evolution, and conservation of crops*. 1st ed. New York, USA: Columbia University Press, pp.171-194.
- De Donato M, Cequea H (1994). A cytogenetic study of six cultivars of the chayote, *Sechium edule* Sw (Cucurbitaceae). *The Journal of Heredity* 85 (3): 238-241.
- De Storme N, Mason A (2014). Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology* 1: 10-33. doi: 10.1016/j.cpb.2014.09.002
- Dempewolf H, Baute G, Anderson J, Kilian B, Smith C *et al.* (2017). Past and future use of wild relatives in crop breeding. *Crop Science* 57 (3): 1070-1082. doi: 10.2135/cropsci2016.10.0885

- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19 (1): 11-15.
- Duminil J, Pemonge MH, Petit RJ (2002). A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. *Molecular Ecology Notes* 2 (4): 428-430. doi: 10.1046/j.1471-8286.2002.00263.x
- Fay MF, Cameron KM, Prance GT, Lledó MD, Chase MW (1997). Familial relationships of *Rhabdodendron* (Rhabdodendraceae): Plastid *rbcL* Sequences Indicate a Caryophyllid Placement. *Kew Bulletin* 52 (4): 923-932. doi: 10.2307/4117819
- Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* 27 (2): 221-223.
- Jeffrey C (1966). On the classification of the Cucurbitaceae. *Kew Bulletin* 20 (3): 417-426. doi: 10.2307/4108235
- Kumar S, Stecher G, Tamura K (2016). MEGA 7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33 (7): 1870-1874.
- Lira R, Caballero J, Dávila P (1997). A contribution to the generic delimitation of *Sechium* (Cucurbitaceae, Sicyinae). *Taxon* 46 (2): 269-282. doi:10.2307/1224097
- Lira R, Chiang F (1992). Two new combinations in *Sechium* (Cucurbitaceae) from Central America, and a new species from Oaxaca, Mexico. *Novon* 2 (3): 227-231. doi: 10.2307/3391556
- Lira R, Nee M (1999). A new species of *Sechium* Sect. *Frantzia* (Cucurbitaceae, Sicyeae, Sicyinae) from Mexico. *Brittonia* 51 (2): 204-209. doi: 10.2307/2666628
- Lira SR (1995). Estudios taxonómicos en el género *Sechium* P. Br. Cucurbitaceae. PhD, Universidad Nacional Autónoma de México, D.F. México.

- Lira SR, Castrejón J, Zamudio S, Rojas ZC (1999). Propuesta de ubicación taxonómica para los chayotes silvestres (*Sechium edule*, Cucurbitaceae) de México. *Acta Botánica Mexicana* 49: 47-61.
- Machida-Hirano R, Cortés-Cruz M, González BAA, Cadena-Iñiguez J, Shirata K *et al.* (2015). Isolation and characterization of novel microsatellite markers in chayote [*Sechium edule* (Jacq.) Sw.]. *American Journal of Plant Sciences* 6 (13): 2033-2041. doi: 10.4236/ajps.2015.613203
- Mercado P, Lira R, Castrejón J (1993). Estudios cromosómicos de *Sechium edule* P. Br. y *Sicana naudin* (Cucurbitaceae). In: 12th Congreso Mexicano de Botánica; Yucatán, México. p. 176.
- Monge PJ, Loría CM (2017). Caracterización de frutos de cinco genotipos de tacaco [*Sechium tacaco* (Pittier) C. Jeffrey] en Costa Rica. *Tecnología en Marcha* 30 (3): 71-84.
- Nei M, Kumar S (2000). *Molecular evolution and phylogenetics*. 1st ed. New York: USA: Oxford University Press.
- Newstrom LE (1991). Evidence for the origin of chayote, *Sechium edule* (Cucurbitaceae). *Economic Botany* 45 (3): 410-428. doi: 10.1007/BF02887082
- Newstrom LE (1990). Origin and evolution of chayote, *Sechium edule*. In: Bates DM, Robinson RW, Jeffrey C (editors). *Biology and Utilization of the Cucurbitaceae*. 1st ed. New York, USA: Cornell University Press, pp 141-149.
- Olmstead RG, Michaels HJ, Scott KM, Palmer JD (1992). Monophyly of the Asteridae and Identification of their major lineages inferred from DNA Sequences of *rbcL*. *Annals of the Missouri Botanical Garden* 79 (2): 249-265. doi:10.2307/2399768
- Patwardhan A, Ray S, Roy A (2014). Molecular markers in phylogenetic studies- a review. *Journal of Phylogenetics and Evolutionary Biology* 2 (2): 1-9. doi: 10.4172/2329-9002.1000131
- Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19 (12): 1572-1574.

- Schaefer H, Heibl C, Renner SS (2009). Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society B: Biological Sciences* 276 (1658): 843-851. doi: 10.1098/rspb.2008.1447
- Sebastian P, Schaefer H, Lira R, Telford IRH, Renner SR (2012) Radiation following long-distance dispersal: the contributions of time, opportunity and diaspore morphology in *Sicyos* (Cucurbitaceae). *Journal of Biogeography* 39 (8): 1427-1438. doi: 10.1111/j.1365-2699.2012.02695.x
- Stenøien HK (2008) Slow molecular evolution in 18S rDNA, *rbcL* and *nad5* genes of mosses compared with higher plants. *Journal of Evolutionary Biology* 21 (2): 566-571. doi: 10.1111/j.1420-9101.2007.01479.x
- White TJ, Bruns T, Lee S, Taylor J (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: White TJ, Bruns T, Lee S, Taylor J, Innis MA, Gelfand DH, Sninsky J (editors). *PCR protocols: a guide to methods and applications*. 1st ed. Massachusetts, USA: Academic Press, Inc, pp. 315-322.
- Wolfe KH, Li WH, Sharp PM (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *National Academy of Sciences* 84 (24): 9054-9058. doi: 10.1073/pnas.84.24.9054
- Wunderlin RP (1976) Two new species and a new combination in *Frantzia* (Cucurbitaceae). *Brittonia* 28 (2): 239-244.
- Yao X, Tan Y, Yang J, Wang Y, Corlett RT *et al.* (2019). Exceptionally high rates of positive selection on the *rbcL* gene in the genus *Ilex* (*Aquifoliaceae*). *BMC Evolutionary Biology* 19: 1-13. doi:10.1186/s12862-019-1521-1

CAPÍTULO IV. MODELOS DE DISTRIBUCIÓN POTENCIAL DE *Sechium tacaco* (PITTIER) C. JEFFREY EN COSTA RICA

POTENTIAL DISTRIBUTION MODELS OF *Sechium tacaco* (PITTIER) C. JEFFREY IN COSTA RICA

RESUMEN

Sechium tacaco (Pittier) C. Jeffrey es una cucurbitácea endémica a las regiones montañosas de Costa Rica. El objetivo de este estudio fue determinar la distribución actual y potencial de *S. tacaco* en Costa Rica con siete modelos de aprendizaje automático, con la finalidad de optimizar el manejo de los recursos fitogenéticos de *S. tacaco*, encaminado a la identificación de patrones de distribución geográfica y posibles adaptaciones climáticas que permitan tener perspectivas sobre su conservación y mejoramiento genético. Se emplearon 21 puntos de ocurrencia junto con 19 variables bioclimáticas y la variable altitud para evaluar siete modelos de aprendizaje automático y un ensamble de estos. Se utilizaron librerías de código abierto ejecutadas en Rstudio. Los modelos de distribución se infirieron por las variables bio1, bio2, bio3, bio4, bio12, bio13, bio14, bio18 y bio19. El modelo aditivo generalizado obtuvo los valores de SUC (0.96) y TSS (0.90) más elevados, sin embargo, los siete modelos probados y el ensamble mostraron rendimiento adecuados (AUC > 0.5 y TSS > 0.4). Las variables bioclimáticas relacionadas a la temperatura fueron las de mayor contribución a los modelos y las principales limitantes en la distribución de *S. tacaco*. Las zonas con alta idoneidad para la distribución potencial de *S. tacaco* se encuentran en los valles centrales de Costa Rica, cubriendo regiones de las provincias de Alajuela, Cartago, San José y Puntarenas. Estas pueden ser fuentes de germoplasma para estudios futuros de conservación y mejoramiento.

Palabras clave: Aprendizaje automático, germoplasma, conservación, mejoramiento.

ABSTRACT

To determine the current and potential distribution of *S. tacaco* in Costa Rica with seven machine learning models, to optimize the management of phylogenetic resources of *S. tacaco*, aimed at identifying patterns of geographic distribution and possible climatic adaptations that allow to have perspectives on their conservation and breeding. Twenty-one occurrence records together with 19 bioclimatic variables and the altitude variable were used to evaluate seven machine learning models and an assembly of them. Open-source libraries executed in Rstudio were used. The distribution models were inferred by the variables bio1, bio2, bio3, bio4, bio12, bio13, bio14, bio18 and bio19. The GAM model obtained the highest AUC (0.96) and TSS (0.9) values, however, the seven evaluated models and the assembly showed adequate performance (AUC > 0.5 and TSS > 0.4). The bioclimatic variables related to temperature turned out to be the ones with the greatest contribution to the models and the main limitations in the distribution of *S. tacaco*. A greater number of occurrence records may be required to evaluate distribution models. The areas with high suitability for the potential distribution of *S. tacaco* are found in the central valleys of Costa Rica, covering regions of the provinces of Alajuela, Cartago, San José and Puntarenas. These can be sources of germplasm for future conservation and breeding studies.

Key words: Machine learning, germplasm, conservation, breeding.

4.1. INTRODUCCIÓN

Sechium tacaco (Pittier) C. Jeffrey es una especie endémica a las regiones montañosas de Costa Rica, donde localmente es conocida como “tacaco” (Wunderlin, 1976). En este país también se distribuye su posible ancestro silvestre, *S. talamancensis* (Wunderlin) C. Jeffrey. Los frutos del tacaco son representativos de la cultura y gastronomía costarricense. Mediante selección se ha logrado una gran variación fenotípica de los frutos del tacaco, en función de su peso, ancho, grosor, número de espinas y suturas longitudinales (Monge y Loría, 2017). Generalmente, las plantaciones de tacaco se encuentran en rangos altitudinales de 500-1700 msnm (Wunderlin, 1976; Monge y Loría, 2017) e inclusive puede llegar hasta los 2000 msnm (Lira, 1995). *Sechium tacaco* es una especie poco representada en los bancos de germoplasma y con problemas de erosión genética, debido al desconocimiento del cultivo, perturbación de su hábitat y por la introducción de cultivos (Lira, 1995).

Los modelos de distribución de especies (MDE) son herramientas que se apoyan en sistemas de información geográfica (SIG) y datos de presencias reales para predecir zonas de idoneidad para las especies, esto en función de sus características ambientales (Mateo *et al.*, 2011). En forma general, la metodología empleada por los MDE consiste en compilar ubicaciones geográficas de la especie de interés; posteriormente se obtienen datos espaciales de variables edafoclimáticas acordes a los puntos de ocurrencia de la especie. Los datos espaciales obtenidos son procesados mediante técnicas estadísticas que pueden predecir zonas idóneas para la distribución de las especies (Hijmans y Elith, 2013). Los MDE pueden clasificarse en métodos descriptivos, tales como Bioclim, Dominio, distancia de Mahalanobis y Anuclim, que solo necesitan datos de presencia para poder ser modelados (Mateo *et al.*, 2011). Existen también técnicas discriminantes como los modelos lineales generalizados con sus respectivas variantes, por ejemplo, el modelo aditivo generalizado y el modelo potenciado generalizado. Dentro de este grupo también se encuentran los métodos de aprendizaje automatizado como árboles

aleatorios, máxima entropía y máquinas de vectores de soporte, los cuales son empleados para computar regresiones numéricas para tareas de predicción. Las técnicas discriminantes basadas en aprendizaje automático son muy flexibles en términos de cálculos computacionales, pueden procesar gran cantidad de información y sus resultados suelen ser más consistentes en comparación con las técnicas descriptivas (Mateo *et al.*, 2011; Hijmans y Elith, 2013; Schmitt *et al.*, 2017). El modelo de máxima entropía (Maxent) es el más aplicado en modelos de distribución de especies debido a una interfaz sencilla en su programación y por arrojar resultados adecuados (Phillips *et al.*, 2006). Sin embargo, es conveniente optimizar la configuración de Maxent para obtener el mejor modelo (Muscarella *et al.*, 2014), y, sobre todo, compararlo con otros algoritmos para tener perspectivas múltiples sobre la especie en cuestión.

Los MDE pueden tener diversas aplicaciones en la agricultura. Por ejemplo, tienen utilidad para determinar los efectos del cambio climático sobre la distribución de las especies (Beck, 2012); monitorear la presencia de plagas, especies invasoras (Lantschner *et al.*, 2018) y agentes polinizadores (Polce *et al.*, 2013), así como para detectar endemismos vegetales, hábitats amenazados, patrones de diversidad y estudios de conservación, solo por mencionar algunos ejemplos (Mateo *et al.*, 2011). Todas estas aplicaciones son útiles para la formulación de estrategias que pretendan maximizar los recursos fitogenéticos, ya sea para mitigar los daños ocasionados por otras especies y para priorizar áreas vulnerables y ricas en germoplasma (Flores-Tolentino *et al.*, 2019).

Respecto al género *Sechium*, existen pocos estudios sobre modelos de distribución de especies. *Sechium edule* y *S. tacaco* son las únicas especies cultivadas de su género. González-Santos *et al.* (2017) pronosticaron con el modelo de Maxent que para el año 2050 algunos complejos varietales de *S. edule* podrían perder más de la mitad de su distribución actual, lo cual es un panorama alarmante debido a que también en las zonas de alto riesgo existen poblaciones silvestres de *S. edule* que son importantes para conocer los procesos filogenéticos de las especies mexicanas de *Sechium*. Por otra parte, los estudios

de *S. tacaco* están limitados a su morfología (Monge and Loría, 2017) y su filogenia (Sebastian *et al.*, 2012).

El objetivo de este estudio fue determinar la distribución actual y potencial de *S. tacaco* en Costa Rica con siete modelos de aprendizaje automático, con la finalidad de optimizar el manejo de los recursos fitogenéticos de esta especie, encaminado a la identificación de patrones de distribución geográfica y posibles adaptaciones climáticas que permitan tener perspectivas sobre su conservación y mejoramiento genético.

4.2. MATERIALES Y MÉTODOS

4.2.1. Datos de ocurrencia e información ambiental

Se obtuvieron 21 puntos de ocurrencia (latitud y longitud) de *S. tacaco* de las bases de datos de Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>) y Monge y Loría (2017). Se verificó que los puntos de ocurrencia no fueran atípicos y repetidos. Se emplearon las 19 variables bioclimáticas de WorldClim versión 2.1 del período 1970-2000) (Cuadro 6) con resolución espacial $\sim 1 \text{ km}^2$ (Fick and Hijmans, 2017). Se añadieron los modelos ráster con resolución espacial $\sim 1 \text{ km}^2$ para datos de elevación en metros (Fick and Hijmans, 2017); clasificación climática de Köppen-Geiger (Beck *et al.*, 2018) y tipos de suelo de Harmonized World Soil Database version 1.2 (Fischer *et al.*, 2008). Los valores ráster para las 22 capas se obtuvieron con Point Sampling Tool de QGIS versión 3.16.2 (QGIS Development Team, 2020).

Cuadro 6. Lista de variables utilizadas para el desarrollo de modelos de distribución de *S. tacaco* en Costa Rica.

Variable	Descripción	Unidad
Bio1*	Temperatura media anual	°C
Bio2*	Rango medio diario	°C
Bio3*	Isotermalidad	°C
Bio4*	Estacionalidad de la temperatura	°C
Bio5	Temperatura máxima del mes más cálido	°C
Bio6	Temperatura mínima del mes más frío	°C
Bio7	Rango anual de temperatura	°C
Bio8	Temperatura media del trimestre más lluvioso	°C
Bio9	Temperatura media del trimestre más seco	°C
Bio10	Temperatura media del trimestre más cálido	°C
Bio11	Temperatura media del trimestre más frío	°C
Bio12*	Precipitación anual	mm
Bio13*	Precipitación del mes más lluvioso	mm
Bio14*	Precipitación del mes más seco	mm
Bio15	Estacionalidad de la precipitación	mm
Bio16	Precipitación del trimestre más lluvioso	mm
Bio17	Precipitación del trimestre más seco	mm
Bio18*	Precipitación del trimestre más cálido	mm
Bio19*	Precipitación del trimestre más frío	mm
Elevación	Modelo de elevación digital	msnm

* Variables seleccionadas en los modelos de distribución de *S. tacaco* en Costa Rica.

4.2.2. Modelos de distribución

Todos los paquetes mencionados en el escrito se ejecutaron en Rstudio (R Core Team, 2020). Para los datos ráster de las 20 variables ambientales (19 variables de WorldClim y la variable elevación) se calculó la correlación de Pearson y se eliminaron aquellas variables con correlaciones ≥ 0.8 , esto para prevenir colinealidad y evitar la afectación de los modelos (Feng *et al.*, 2019). Las variables de tipos de clima y suelo se usaron como información descriptiva y se omitieron en las predicciones de los modelos de distribución. En el caso del clima, este se relaciona con algunas variables como la precipitación media (bio12) y temperatura media anual (bio1). En el caso del suelo, se carece de información referente a su calidad, sin embargo, con la exclusión de estas variables se pueden tener rendimientos sólidos en los modelos de distribución (Evans *et al.*, 2010).

Se emplearon técnicas de aprendizaje automático ampliamente usadas según la literatura debido a su alto rendimiento y resultados óptimos (Mateo *et al.*, 2011), incluyéndose así los siguientes algoritmos: Modelo Lineal Generalizado (MLG), Modelo Aditivo Generalizado (MAG), Modelo de Regresión Potenciado Generalizado (GBM), Análisis de Árboles de Clasificación (AAC), Maxent, Bosque aleatorio (RF) y Máquinas de vectores de soporte (MVS). Para la ejecución de estos siete modelos se empleó el paquete SSDM (Schmitt *et al.*, 2017) con su configuración predeterminada y las dependencias gam (Wood, 2017), stats (R Core Team, 2020), maxent (Hijmans *et al.*, 2017), rpart (Therneau y Atkinson, 2019), gbm (Greenwell *et al.*, 2020), randomForest (Liaw y Wiener, 2002) y e1071 (Meyer *et al.*, 2019). Sin embargo, para optimizar el modelo Maxent, se empleó el paquete ENMeval (Muscarella *et al.*, 2014) con la siguiente configuración: se desactivaron las funciones de umbral y bisagra para evitar el sobreajuste de las curvas de la respuesta; también se desactivaron las funciones cuadráticas y de producto y el multiplicador de regularización fue de 1.25 con una función lineal. La evaluación de los modelos se cuantificó con el AUC, donde valores > 0.5 indican modelos adecuados. Sin embargo, para eliminar el sesgo

de clasificación espacial estudiado por Lobo *et al.* (2007), se realizó un muestreo de distancia puntual, es decir, se calculó la diferencia de distancias para presencias y ausencias en los datos de entrenamiento (75 %) y prueba (25 %) (Hijmans y Elith, 2013). Para verificar el rendimiento de los modelos se calcularon los estadísticos kappa y TSS. Kappa cuantifica la proporción de puntos predichos correctamente después de eliminar la probabilidad de coincidencia aleatoria, su valor oscila entre -1 y 1; valores cercanos a 1 indican rendimiento excelente del modelo y valores cercanos a -1 pobre rendimiento. Para corregir la dependencia de la prevalencia de kappa, se calculó el TSS ((Especificidad + Sensibilidad) -1), el cual optimiza y corrige los problemas de kappa. Los criterios de rendimiento de la TSS son los mismos que kappa (Allouche *et al.*, 2006).

Adicionalmente se efectuó la prueba de Jackknife para observar la contribución de cada variable ambiental a los modelos de distribución, que junto con los estadísticos AUC, kappa, TSS, sensibilidad, especificidad, proporción de predicciones correctas (PPC) y un ensamble de los siete algoritmos se calcularon con el mismo paquete SSDM (Schmitt *et al.*, 2017). Todos estos valores se promediaron de los resultados para los datos de prueba y entrenamiento. Los formatos ráster producto de los siete modelos y el ensamble se exportaron con el paquete raster (Hijmans, 2020) a QGIS Development Team versión 3.16.2 (2020).

4.3. RESULTADOS Y DISCUSIÓN

En los puntos de ocurrencia de *S. tacaco* se encontraron cinco tipos de climas: Am (tropical monzónico, 38.1 %), Af (tropical ecuatorial, 23.8 %), Aw (tropical de sabana, 19 %), Cwb (invierno templado y seco, verano cálido, 14.3 %) y Cfb (templado, sin estación seca y verano cálido 4.76 %). Los tipos de suelo hallados fueron andosol (38.1 %), cambisol (33.3 %) y alisol (28.6 %). Las variables bioclimáticas para evaluar los modelos de distribución fueron bio1, bio2, bio3, bio4, bio12, bio13, bio14, bio18 y bio19, las cuales resultaron del análisis de correlación. Las variables relacionadas a la temperatura mostraron muy poca variación, por ejemplo, bio1 obtuvo un rango de 17.65-21.3 °C, obteniéndose un

promedio de ~ 20 °C. Para las variables relacionadas a la precipitación se obtuvieron rangos más amplios, por ejemplo, para bio12 hubo registros con rango de 2200–4000 mm de precipitación.

El MAG obtuvo el AUC (0.96) y TSS (0.90) más elevados. De forma general el AUC para los siete algoritmos fue > 0.8, indicando así modelos adecuados. Por otra parte, el TSS estuvo dentro del rango de rendimiento óptimo de los modelos $0.4 > TSS < 0.7$ (Allouche *et al.*, 2006). El ensamble de los siete modelos fue adecuado en términos de AUC, kappa y TSS. La PPC obtuvo valores > 0.8 y fue considerada como aceptable (Cuadro 7).

Cuadro 7. Parámetros estadísticos para los modelos de distribución de *S. tacaco* en Costa Rica.

Modelo	AUC	PPC	Sensibilidad	Especificidad	kappa	TSS
MLG	0.89	0.88	0.91	0.88	0.22	0.79
MAG	0.96	0.93	0.98	0.92	0.37	0.90
GBM	0.83	0.81	0.83	0.83	0.66	0.66
ACC	0.83	0.83	0.91	0.75	0.66	0.57
RF	0.85	0.83	0.83	0.83	0.66	0.66
Maxent	0.92	0.88	0.83	0.88	0.023	0.71
MVS	0.91	0.92	0.91	0.91	0.83	0.82
Ensamble	0.83	0.83	0.83	0.84	0.42	0.67

Respecto a la prueba Jackknife, las variables con mayor contribución en los modelos fueron bio1, bio2, bio3 y bio14. En el ensamble de los siete modelos la variable con mayor contribución fue bio1. El modelo AAC asignó la misma contribución para todas las variables predictoras (Cuadro 8).

Cuadro 8. Contribución (%) de las variables ambientales a los modelos de distribución de *S. tacaco* en Costa Rica.

Modelo	Bio1	Bio2	Bio3	Bio4	Bio12	Bio13	Bio14	Bio18	Bio19
MLG	0.22	4.01	8.01	0.1	6.54	12.36	4.76	37.99	26.07
MAG	23.92	0.29	40.19	14.8	1.68	4.47	1.24	1.65	11.73
GBM	26.63	30.42	2.11	18.04	1.66	0.64	0.21	1.05	18.85
AAC	11.11	11.11	11.11	11.11	11.11	11.11	11.11	11.11	11.11
RF	53.74	7.85	6.61	7.9	3.25	5.03	6.85	1.99	6.74
Maxent	16.8	8.4	43.7	3.5	1.6	1.5	0.3	16	8.3
MVS	24.52	3.78	2.79	11.83	2.17	6.87	30.68	13.06	4.24
Ensamble	27.48	10.01	10.54	10.10	6.07	8.57	4.72	10.41	12.04

Los modelos MLG y MAG mostraron grandes extensiones territoriales para una idoneidad de 0.25 (Figura 20). El GBM indicó idoneidad con probabilidad de 1 para los Valles Centrales de Costa Rica, peculiarmente el modelo ACC mostró la máxima idoneidad en un eje longitudinal de territorio costarricense que cubre parte de las provincias Alajuela, San José, Cartago y Puntarenas.

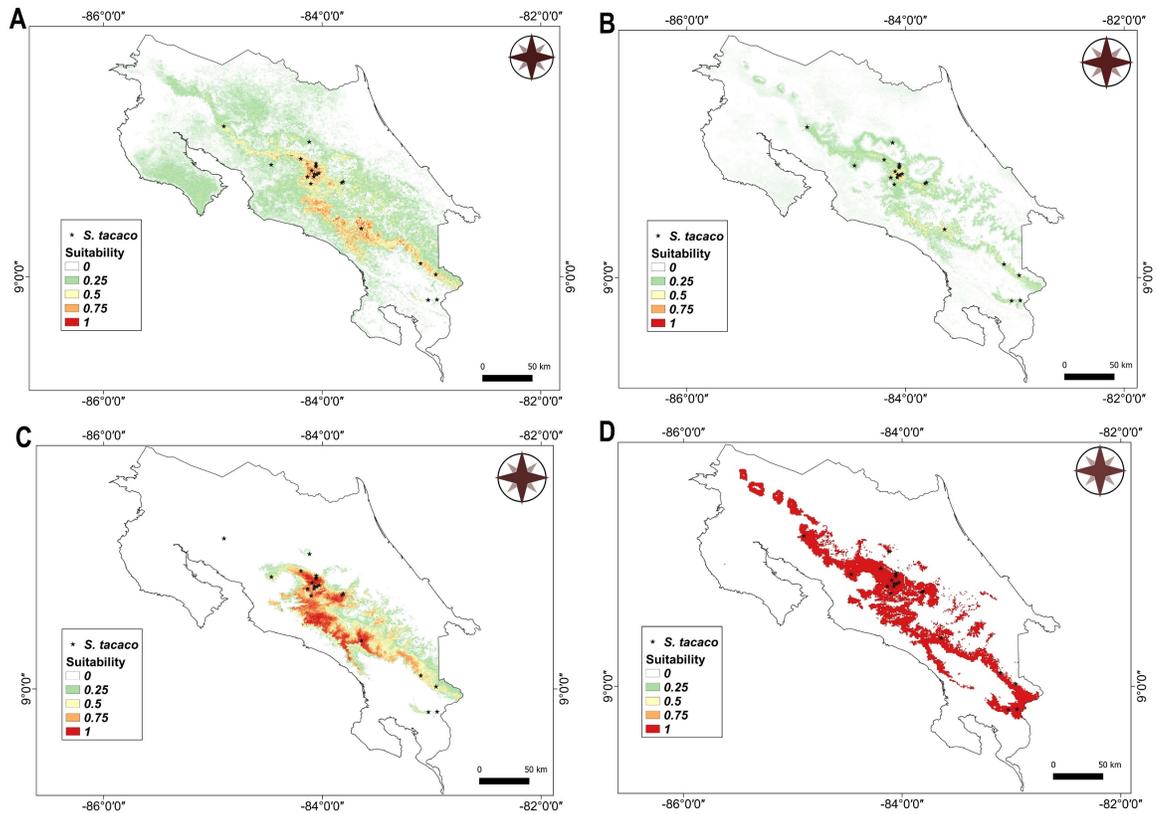


Figura 20. Modelos de distribución para *S. tacaco* en Costa Rica. **A.** MLG. **B.** MAG. **C.** GBM. **D.** ACC.

Los modelos de RF, Maxent y MVS mostraron idoneidad de 1 en regiones muy parecidas (Figura 21 A-C), sin embargo, RF tuvo mayor extensión territorial en Valles Centrales y en el Pacífico Sur donde se ubica la cordillera de Talamanca. El ensamble de los siete modelos (Figura 21D) presentó una proyección cartográfica muy parecida a la de Maxent.

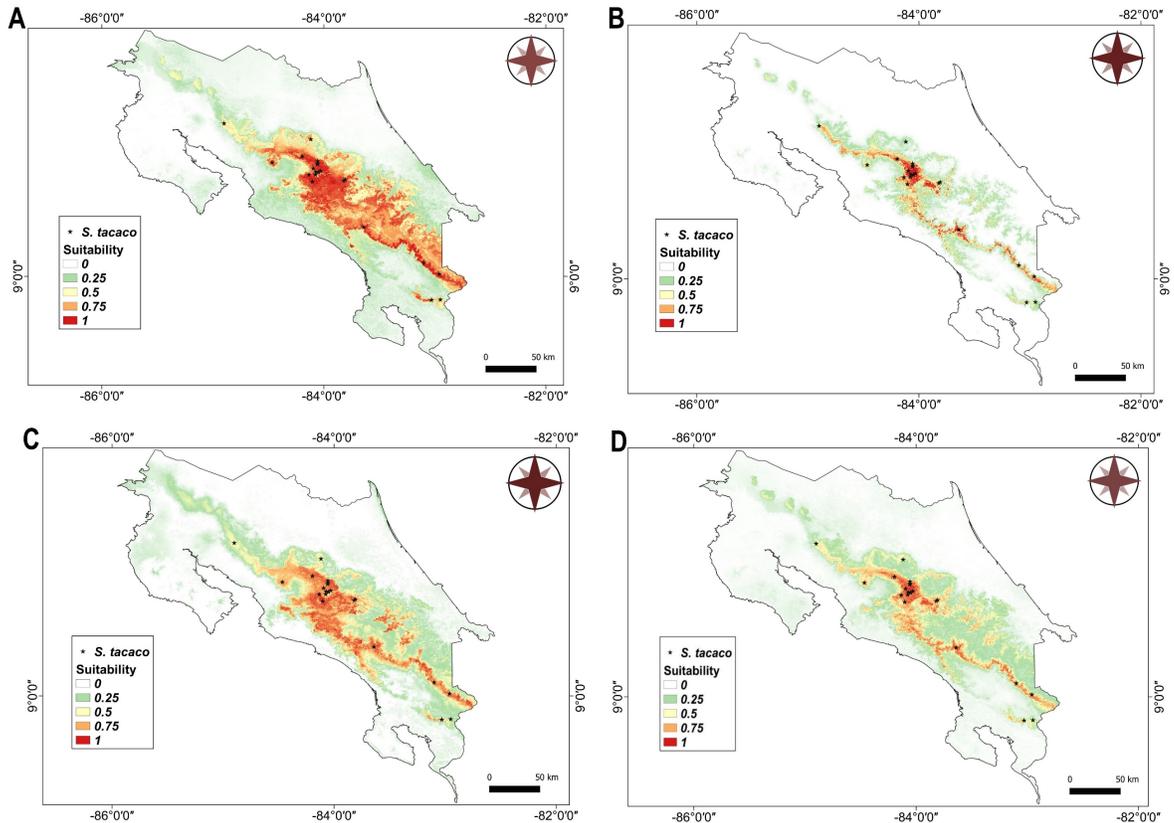


Figura 21. Modelos de distribución de *S. tacaco* en Costa Rica. Técnicas basadas en aprendizaje automático, **A.** RF, **B.** Maxent, **C.** MVS y **D.** Ensemble de los siete modelos.

El 80 % del territorio costarricense tiene un clima tipo A, razón por la cual no se esperaría que el tipo de clima sea una condicional para determinar la distribución de *S. tacaco*, considerando también que el clima está fuertemente relacionado con las variables bioclimáticas bio1 y bio12.

Los valores de AUC para entrenamiento y prueba (0.97 y 0.93) determinaron un buen modelamiento en la distribución de *S. tacaco*, valores muy cercanos a 1 suelen ser indicativos de la distribución restringida de las especies (Phillips *et al.*, 2006), tal y como lo es *S. tacaco*. Lira *et al.* (2018) encontraron valores AUC superiores a 0.95 en poblaciones silvestres de *S. edule*, las cuales únicamente se distribuyen en los estados de Oaxaca y Veracruz, México. Generalmente las

especies endémicas suelen tener bajos niveles de diversidad genética debido a los reducidos tamaños de sus poblaciones, sin embargo, es necesario evaluarlas con marcadores moleculares, ya que en algunos casos se ha encontrado que dichas especies pueden tener niveles moderados o altos de diversidad genética (Forrest *et al.*, 2017).

La temperatura media anual (bio1) juega un papel importante en la distribución de *S. tacaco*, si bien es una especie semi-cultivada, debido a la escasa investigación agronómica en *S. tacaco* es difícil determinar los requerimientos óptimos de temperatura y riego del cultivo. Las Figuras 20 y 21, en concordancia con Monge y Loría (2017), muestran una posible tendencia de que *S. tacaco* pueda introducirse o distribuirse en Panamá, sobre todo con los modelos ACC, RF, Maxent y MVS. Esto constituye un aspecto importante para investigar la respuesta agronómica de *S. tacaco* en otras condiciones edafoclimáticas.

Los suelos andosol y cambisol que predominan en 2/3 partes de los puntos de ocurrencia, se caracterizan por ser de origen volcánico y tener alta capacidad de intercambio catiónico, la cual se correlaciona positivamente con la cantidad de materia orgánica (Dai *et al.*, 2018).

Las especies del género *Sechium* se caracterizan por tener semilla recalcitrante (Lira, 1995), lo que es una limitante para su conservación en bancos de germoplasma *ex situ*, ya que las semillas no soportan las condiciones de temperatura y humedad y pierden su viabilidad. La forma óptima de conservar este tipo de especies es en su hábitat natural (bancos de germoplasma *in situ*).

Si bien los modelos de distribución son predicciones de la idoneidad de las especies, nos brindan información relevante sobre las regiones geográficas que contienen germoplasma de interés, lo cual sin duda es importante para estudios de conservación y mejoramiento.

4.4. CONCLUSIONES

El MAG resultó tener los mejores índices de AUC y TSS, sin embargo, los siete modelos y el ensamble mostraron rendimiento adecuado. Las variables bioclimáticas relacionadas a la temperatura resultaron ser las de mayor contribución a los modelos y las principales limitantes en la distribución de *S. tacaco*, que, al ser una especie de distribución restringida, requiere condiciones climáticas muy específicas. Las zonas con alta idoneidad se encuentran en los valles centrales de Costa Rica, cubriendo regiones de las provincias de Alajuela, Cartago, San José y Puntarenas, mismas que pueden ser fuentes de germoplasma para estudios futuros de conservación y mejoramiento.

4.5. LITERATURA CITADA

- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223-1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1), 1-12. <https://doi.org/10.1038/sdata.2018.214>
- Beck, J. (2012). Predicting climate change effects on agriculture from ecological niche modeling: Who profits, who loses? *Climatic Change*, 116. <https://doi.org/10.1007/s10584-012-0481-x>
- Evans, J. M., Fletcher, R. J., & Alavalapati, J. (2010). Using species distribution models to identify suitable areas for biofuel feedstock production. *GCB Bioenergy*, 2(2), 63-78. <https://doi.org/10.1111/j.1757-1707.2010.01040.x>

- Feng, X., Park, D. S., Liang, Y., Pandey, R., & Papeş, M. (2019). Collinearity in ecological niche modeling: Confusions and challenges. *Ecology and Evolution*, 9(18), 10365-10376. <https://doi.org/10.1002/ece3.5555>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315. <https://doi.org/10.1002/joc.5086>
- Fischer, G., Nachtergaele, F., Prieler, S., van Velthuisen, H. T., Verelst, D., & Wiberg, D. (2008). *Global Agro-ecological Zones Assessment for Agriculture*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.
- Flores-Tolentino, M., Ortiz, E., Villaseñor, J. L., Flores-Tolentino, M., Ortiz, E., & Villaseñor, J. L. (2019). Ecological niche models as a tool for estimating the distribution of plant communities. *Revista Mexicana de Biodiversidad*, 90. <https://doi.org/10.22201/ib.20078706e.2019.90.2829>
- González-Santos, R., Cadena-Íñiguez, J., Morales-Flores, F. J., Ruiz-Vera, V. M., & Pimentel-López, J. (2017). Prediction of the effects of climate change on *Sechium edule* (Jacq.) Swartz varietal groups in Mexico. *Genetic Resources and Crop Evolution*, 64(4), 791-804. <https://doi.org/10.1007/s10722-016-0401-4>
- Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers. (2020). *gbm: Generalized boosted regression models* (R package version 2.1.8) [Computer software]. <https://CRAN.R-project.org/package=gbm>
- Hijmans, R. J. (2020). *raster: Geographic Data Analysis and Modeling* (R package version 3.3-13) [Computer software]. <https://CRAN.R-project.org/package=raster>
- Hijmans, R. J., & Elith, J. (2013). *Species distribution modeling with R*. 77 p.

- Hijmans, R. J., Phillips, S. J., Leatwick, J., & Elith, J. (2017). Dismo: Species Distribution Modeling. R package version 1.1-4. <https://CRAN.R-project.org/package=dismo>
- Lantschner, V., Vega, G., & Corley, J. (2018). Predicting the distribution of harmful species and their natural enemies in agricultural, livestock and forestry systems: An overview. *International Journal of Pest Management*, 65, 1-17. <https://doi.org/10.1080/09670874.2018.1533664>
- Liaw, M., & Wiener, M. (2002). Classification and regression by RandomForest. *R News*, 2(3), 18-22.
- Lira, S. R. (1995). *Estudios taxonómicos en el género Sechium P. Br. Cucurbitaceae* [Tesis Doctoral]. Universidad Nacional Autónoma de México.
- Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, 145-151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
- Mateo, R. G., Felicísimo, A. M., & Muñoz, J. (2011). Species distributions models: A synthetic revision. *Revista Chilena de Historia Natural*, 84, 217-240. <http://dx.doi.org/10.4067/S0716-078X2011000200008>
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2019). *E1071: Misc Functions of the Department of Statistics (E1071), Probability Theory Group (Formerly: E1071) TU Wien.* (R package version 1.7-3) [Computer software]. <https://CRAN.R-project.org/package=e1071>
- Monge, P. J. E., & Loría, M. (2017). Caracterización de frutos de cinco genotipos de tacaco [*Sechium tacaco* (Pittier) C. Jeffrey] en Costa Rica. *Tecnología en Marcha*, 30(3), 71-84. <https://doi.org/10.18845/tm.v30i3.3274>

- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution*, 5(11), 1198-1205. <https://doi.org/10.1111/2041-210X.12261>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Polce, C., Termansen, M., Aguirre-Gutiérrez, J., Boatman, N., Budge, G., Crowe, A., Garratt, M., Pietravalle, S., Potts, S., Ramirez, J., Somerwill, K., & Biesmeijer, J. (2013). Species distribution models for crop pollination: A modelling framework applied to Great Britain. *PloS one*, 8, e76308. <https://doi.org/10.1371/journal.pone.0076308>
- QGIS Development Team. (2020). QGIS Geographic Information System. Open-Source Geospatial Foundation Project (3.16.2) [Computer software]. <https://www.qgis.org/en/site/>
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing (1.3.1093) [Computer software]. <https://www.R-project.org/>
- Schmitt, S., Pouteau, R., Justeau, D., Boissieu, F. de, & Birnbaum, P. (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, 8(12), 1795-1803. <https://doi.org/10.1111/2041-210X.12841>
- Sebastian, P., Schaefer, H., Lira, R., Telford, I. R. H., & Renner, S. S. (2012). Radiation following long-distance dispersal: The contributions of time, opportunity and diaspore morphology in *Sicyos* (Cucurbitaceae). *Journal*

of *Biogeography*, 39(8), 1427-1438. <https://doi.org/10.1111/j.1365-2699.2012.02695.x>

Therneau, T., & Atkinson, B. (2019). rpart: Recursive partitioning and regression trees (R package version 4.1-15) [Computer software]. <https://CRAN.R-project.org/package=rpart>

Wood, S. N. (2017). Generalized additive models (Second). Chapman and Hall CRC Press.

Wunderlin, R. P. (1976). Two new species and a new combination in *Frantzia* (Cucurbitaceae). *Brittonia*, 28(2), 239-244. JSTOR. <https://doi.org/10.2307/2805833>

CAPÍTULO V. MODELOS DE DISTRIBUCIÓN POTENCIAL DE *Sechium edule* (CUCURBITACEAE) EN MÉXICO

POTENTIAL DISTRIBUTION MODELS OF *Sechium edule* (CUCURBITACEAE) IN MEXICO

RESUMEN

México es el centro de origen del chayote (*Sechium edule* (Jacq.) Sw), una planta importante en la alimentación humana y en la farmacéutica. El objetivo de este estudio fue determinar la distribución actual y potencial de *S. edule* en México utilizando siete modelos de distribución de especies, con el fin de gestionar eficientemente los recursos de *S. edule* y ayudar a su conservación y mejoramiento mediante la identificación de patrones de distribución geográfica. Se utilizaron 162 puntos de ocurrencia de la GBIF y nueve capas en formato raster para evaluar algoritmos de modelos de distribución (Bioclim, Domain, Mahalanobis, MLG, Maxent, Bosque Aleatorio y MVS). Para evaluar la confiabilidad y el rendimiento de los modelos se utilizaron los estadísticos AUC y TSS. Los tipos de clima predominantes fueron Cwb (33.3 %) y Aw (17.9 %); los tipos de suelo predominantes fueron leptosol (33.3 %) y phaozem (16.7 %). Los siete modelos mostraron áreas de alta idoneidad (> 0.75) para los estados de Chiapas, Guerrero, Oaxaca, Veracruz, Tabasco, Puebla e Hidalgo. Los valores de AUC para los siete modelos fueron > 0.8 y su rendimiento fue adecuado ($0.4 > TSS < 0.7$). Se encontró que MLG fue el mejor modelo medido por AUC-TSS, sin embargo, los siete modelos fueron adecuados para explicar la distribución de *S. edule* en México. La adaptabilidad climática de *S. edule* también le permite distribuirse hacia la Península de Yucatán, occidente y norte de México.

Palabras clave: modelos de distribución, conservación, mejoramiento genético, *Sechium edule*, AUC, TSS.

ABSTRACT

Mexico is the native country of the chayote (*Sechium edule* (Jacq.) Sw), an important plant in human food and pharmaceuticals. The aim of this study was to determine the current and potential distribution of *S. edule* in Mexico using seven species distribution models, in order to efficiently manage the resources of *S. edule* and aid its conservation and breeding, identify geographical distribution patterns. 162 occurrence records from the GBIF and nine environmental images in raster format were used to evaluate algorithms of the distribution models (Bioclim, Domain, Mahalanobis, GLM, Maxent, Random Forest and SVM). To evaluate the reliability and performance of the models, the AUC and TSS statistics were used. The predominant climate types were Cwb (33.3%) and Aw (17.9%); the predominant soil types were leptosol (33.3%) and phaozem (16.7%). The seven models showed areas of high suitability (> 0.75) for the states of Chiapas, Guerrero, Oaxaca, Veracruz, Tabasco, Puebla, and Hidalgo. The AUC values for the seven models were > 0.8 and their performance was adequate ($0.4 > \text{TSS} < 0.7$). GLM was found to be the best model as measured by AUC-TSS, however, all seven models were adequate to explain the distribution of *S. edule* in Mexico. The climatic adaptability of *S. edule* also allows it to be distributed towards the Yucatan Peninsula, western and northern Mexico.

Keywords: Distribution models, conservation, breeding, *Sechium edule*, AUC, TSS.

5.1. INTRODUCCIÓN

El chayote [*Sechium edule* (Jacq.) Swartz] es una especie originaria de las regiones montañosas de México. Actualmente se encuentra en estado silvestre en los estados de Veracruz, Guerrero y Oaxaca (Lira *et al.*, 1999), donde también se practica su cultivo, y en menor medida en el occidente y centro de México. Debido a que las poblaciones silvestres de chayote y las especies estrechamente relacionadas como *S. compositum* y *S. chinantlense* solo se han encontrado en este país, México es considerado el centro de origen de *S. edule* (Newstrom, 1990; Lira, 1995; Cross *et al.*, 2006). Las poblaciones cultivadas y silvestres de chayote en México muestran gran variabilidad fenotípica y fitoquímica en las características de los frutos, hojas y flores (Cadena-Iñiguez *et al.*, 2008, 2011). México ocupa el primer lugar en el mundo en producción y exportación de chayote. México produce alrededor de 2972 h con un rendimiento promedio de 66 t h⁻¹, aportando el estado de Veracruz más del 90 % de la producción (SIACON, 2020). El fruto del chayote se utiliza en la gastronomía nacional y sus extractos alcohólicos se están estudiando como potencial para el tratamiento de diversos tipos de cáncer (Cadena-Iñiguez *et al.*, 2013; Aguiñiga-Sánchez *et al.*, 2015, 2017; Salazar- Aguilar *et al.*, 2017).

La distribución de una especie está en función de las condiciones ecológicas donde se desarrolla y se reproduce. Los modelos de distribución de especies (MDE) se basan en procedimientos estadísticos y sistemas de información geográfica. Usando datos de presencia-ausencia, estos modelos permiten la inferencia de sitios adecuados con características ambientales óptimas para la especie. Los distintos algoritmos para su cálculo se agrupan en técnicas descriptivas, discriminatorias y mixtas (Mateo *et al.*, 2011). Las técnicas descriptivas como Bioclim incluyen métodos envolventes para caracterizar los sitios que se encuentran dentro del espacio ambiental ocupado por la especie en cuestión. En este grupo también existen algoritmos basados en distancias (Domain y Mahalanobis) que evalúan sitios en términos de similitud ambiental con datos de presencias (Elith *et al.*, 2006). Las técnicas discriminantes como el

modelo lineal generalizado (MLG) requieren datos de presencia-ausencia y tienen la capacidad de modelar relaciones ecológicas realistas. Dentro de estas técnicas se encuentran las que utilizan el aprendizaje automático, como la máxima entropía (Maxent), el bosque aleatorio y las máquinas de vectores de soporte (MVS). Estos son ampliamente utilizados debido a su capacidad para procesar grandes cantidades de información y construir grupos para su clasificación, o para generar regresiones numéricas para tareas de predicción (Gobeyn *et al.*, 2019). También existen modelos geográficos nulos como distancia geográfica, cascos convexos, círculos y presencia-ausencia que utilizan la ubicación geográfica de las presencias, pero no se basan en variables predictivas y no son de uso común ya que su desempeño y confiabilidad aún no se han evaluado (Hijmans y Elith, 2013). El modelo de Maxent se usa con más frecuencia debido a que su cálculo computacional es rápido y fácil (Phillips *et al.*, 2006; Zhang *et al.*, 2019), sin embargo, es necesario evaluar más modelos para obtener una visión más completa.

Los MDE tienen varias aplicaciones en la agricultura. En especies cultivadas como *Zea mays* y *Panicum virgatum*, utilizadas como biocombustibles en los EUA, se han utilizado los modelos Maxent y MVS para estimar la idoneidad de las regiones geográficas en las que estas especies pueden prosperar (Evans *et al.*, 2010). La información de los MDE también puede proporcionar referencias filogeográficas e inferir eventos históricos que intentan explicar el origen y la estructura genética de las poblaciones (Carstens y Richards, 2007), aunque esto también debe corroborarse con datos moleculares. Los MDE se utilizan para monitorear la disponibilidad de insectos polinizadores, así como para proponer estrategias para hacer frente a cualquier cambio adverso en el medio ambiente que los afecte (Polce *et al.*, 2013). También es posible predecir la idoneidad de especies nocivas (malezas e insectos), pronosticar su ubicación, distribución y formular estrategias para mitigar el daño socioeconómico en áreas vulnerables a la invasión (Lantschner *et al.*, 2018; Manzoor *et al.*, 2018). Otras aplicaciones a destacar son la evaluación de riqueza de especies, endemismo, hábitats

amenazados, protección y conservación de especies amenazadas, patrones de diversidad, estudios de conservación, respuesta al cambio climático, reintroducción de especies amenazadas y el estudio de efectos antrópicos en la distribución de las especies. Aunque muchas de estas aplicaciones se utilizan en especies animales, también pueden extrapolarse a especies vegetales (Mateo *et al.*, 2011).

El AUC es el estadístico más comúnmente utilizado para medir la confiabilidad de los MDE. Este valor corresponde a la probabilidad de que, al tomar datos de presencia/ausencia al azar, el modelo asigne valores de idoneidad más altos a los datos de presencia (Mateo *et al.*, 2011). El AUC tiene algunas limitaciones en su uso, especialmente en especies con un alto rango de adaptabilidad (Lobo *et al.*, 2007). Sin embargo, a través de ajustes en los algoritmos como la validación cruzada, el análisis de correlación, kappa de Cohen, la TSS y el aumento del número de presencias (> 20), su consistencia y rendimiento pueden aumentar considerablemente (Hijmans y Elith, 2013).

Respecto a *S. edule*, González-Santos *et al.* (2017) evaluaron 106 puntos de ocurrencia pertenecientes a cinco grupos varietales de chayote, los cuales fueron modelados utilizando las 19 capas bioclimáticas de WorldClim y el modelo Maxent para pronosticar el efecto del cambio climático en la distribución de estos grupos varietales para el año 2050. Algunos complejos como *nigrum spinosum* perderá alrededor del 50 % de su distribución actual, mientras que otros como *virens levis* mantendrán y aumentarán levemente su distribución, ya que este complejo es muy valioso y se usa comúnmente en el mejoramiento (Cadena-Iñiguez y Arévalo-Galarza, 2011; Cadena-Iñiguez *et al.*, 2013).

Lira *et al.* (2018) analizaron 34 puntos de ocurrencia de poblaciones silvestres de *S. edule* en México utilizando el modelo de Maxent propuesto por Phillips *et al.* (2006), con las 19 capas bioclimáticas y con una resolución ~1 km². Los valores de AUC para datos de entrenamiento y prueba fueron 0.96 y 0.98, respectivamente, lo que indica que el modelo es bastante bueno. A pesar de esto,

los valores altos de AUC son indicativos de especies con rangos de distribución restringidos (Phillips, 2009), ya que solo se han encontrado poblaciones silvestres de *S. edule* en Oaxaca, Guerrero y Veracruz (Lira *et al.*, 1999). La mayor idoneidad para el potencial de distribución de *S. edule* tuvo un rango de probabilidad de 50-60 %. En Lira *et al.* (2018), se omitieron ciertos análisis previos, como la eliminación de variables altamente correlacionadas que inciden en la complejidad del modelo (Phillips *et al.*, 2006; Zhang *et al.*, 2019). Además, no se calcularon estadísticos como la TSS para evaluar el desempeño del modelo.

El objetivo de este estudio fue determinar la distribución actual y potencial de *S. edule* en México utilizando siete modelos de distribución de especies. La obtención de diferentes resultados generará múltiples perspectivas sobre la especie, con el fin de gestionar eficientemente los recursos de *S. edule*, ayudando así a su conservación y mejoramiento.

5.2. MATERIALES Y MÉTODOS

5.2.1. Datos de ocurrencia e información ambiental

Los puntos de ocurrencia de *S. edule* se obtuvieron del Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>). A nivel mundial existen 2327 puntos de ocurrencia, de los cuales 983 son de México. Los datos de estos registros se filtraron para eliminar puntos con coordenadas repetidas, ausencia de coordenadas, errores de georreferenciación y exclusión de especímenes silvestres, dejando 162 puntos de ocurrencia.

Se utilizaron las 19 variables bioclimáticas de WorldClim versión 2.1 del período 1970-2000, con una resolución espacial de $\sim 1 \text{ km}^2$ (Fick y Hijmans, 2017). Se agregaron modelos ráster con una resolución espacial de $\sim 1 \text{ km}^2$ para los datos de elevación en metros (Fick y Hijmans, 2017) (Cuadro 6, del capítulo IV). La

clasificación climática de Köppen-Geiger (Beck *et al.*, 2018) y los tipos de suelo de Harmonized World Soil Database versión 1.2 (Fischer *et al.*, 2008) se utilizaron para agrupar los registros de ocurrencia. Los valores ráster para las 22 capas se obtuvieron utilizando Point Sampling Tool de QGIS versión 3.16.2 (QGIS Development Team, 2020). Las ubicaciones específicas de los puntos de ocurrencia se encontraron utilizando las Subprovincias Fisiográficas del Mapa Digital de México, elaborado por el Instituto Nacional de Estadística y Geográfica (INEGI) (<https://www.inegi.org.mx/temas/fisiografia>).

5.2.2. Modelos de distribución.

Todas las bibliotecas mencionadas en este documento se ejecutaron en RStudio (R Core Team, 2020). Las variables cualitativas de clima y tipo de suelo se procesaron utilizando el paquete tidyverse (Wickham *et al.*, 2019) con datos representados en gráficos de barras. No se incluyeron en los MDE ya que el tipo de clima se correlaciona con variables bioclimáticas como la precipitación media anual (Evans *et al.*, 2010), y el tipo de suelo en general, no parece restringir la distribución debido a la adaptabilidad del cultivo (Lira *et al.*, 2009). Se calculó la correlación de Pearson para los valores raster de las 19 variables bioclimáticas de WorldClim y la variable de elevación. Las variables con valores de correlación > 0.8 se eliminaron para evitar la afectación de los modelos (Phillips *et al.*, 2006) y proporcionar una explicación biológica de los resultados (Graham, 2003).

Se utilizaron siete algoritmos comunes de la literatura: Bioclim (Busby, 1986), Domain (Carpenter *et al.*, 1993) y Mahalanobis (Hirzel *et al.*, 2002) como métodos descriptivos; Modelo lineal generalizado (MLG) (Hijmans y Elith, 2013), Maxent (Phillips *et al.*, 2006), Bosque aleatorio (Breiman, 2001) y Máquina de vectores de soporte (MVS) (Vapnik, 1998) como métodos de aprendizaje automático o técnicas de discriminación. Para la ejecución de los modelos se siguió el protocolo de Hijmans y Elith (2013). El rendimiento de los modelos se midió utilizando el AUC. Sin embargo, el AUC puede presentar sesgo y variación en sus valores con la extensión espacial para seleccionar puntos de fondo (Lobo *et al.*, 2007). Para corregir este problema, se utilizó un muestreo de distancia

puntual. Esto implicó calcular la distancia entre los puntos de ocurrencia de presencia para los conjuntos de datos de entrenamiento y prueba, así como la distancia desde los datos de ausencia para los conjuntos de datos de entrenamiento y prueba (25 % datos de prueba y 75 % datos de entrenamiento) (Hijmans, 2012; Hijmans y Elith, 2013). Para evaluar el desempeño de los modelos se utilizó el TSS, donde $TSS > 0.75$ indica desempeño excelente, $0.4 < TSS < 0.75$ desempeño óptimo y $TSS < 0.4$ desempeño pobre (Eskildsen *et al.*, 2013). La ejecución de los modelos y el cálculo de AUC y TSS se realizó utilizando las librerías *dismo* (Hijmans *et al.*, 2017), *raster* (Hijmans, 2020), *sp* (Pebesma y Bivand, 2005), *maptools* (Bivand y Lewin-Koh, 2020), *SDMtune* (Vignali *et al.*, 2020), *RandomForest* (Liaw y Wiener, 2002) y *PresenceAbsence* (Freeman y Moisen, 2008). Las capas ráster se exportaron en formato GRD a QGIS Development Team 3.16.2 (2020) para la edición final de los mapas del modelo de distribución.

5.3. RESULTADOS Y DISCUSIÓN

5.3.1. Información edafoclimática

Se registraron 21 subprovincias fisiográficas para 162 puntos de ocurrencia, la mayoría de las cuales corresponden a las subprovincias fisiográficas de Sierra Oriental (20 %), Volcanes Lagos y Anáhuac (13 %), Chiconquiaco (10.5 %), Sierra de Chiapas (9.5 %) y Sierras del sur de Chiapas (8.5 %) (Figura 22).

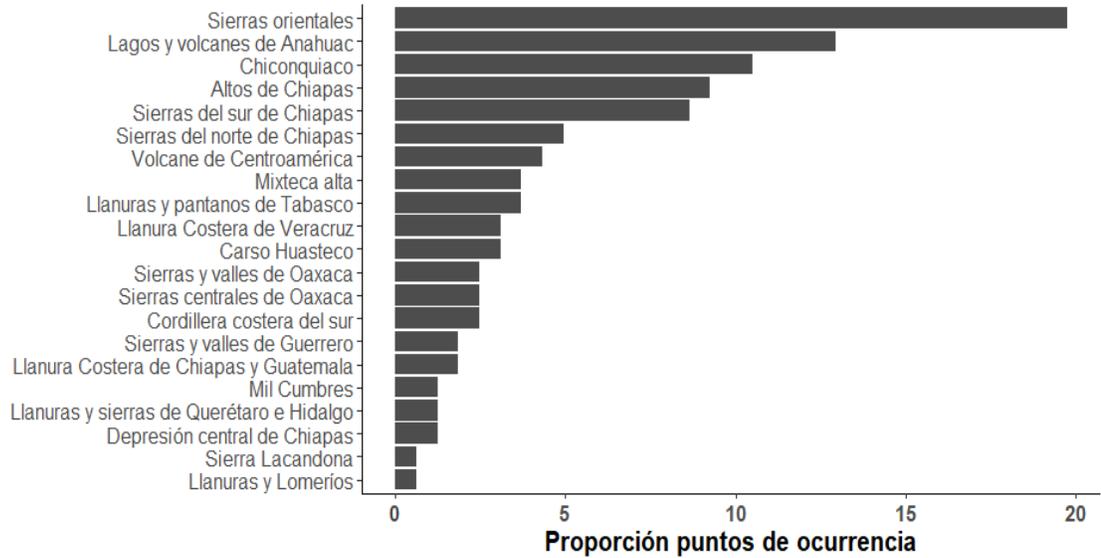


Figura 22. Proporción de puntos de ocurrencia para las subprovincias fisiográficas de México.

Se encontraron tres grupos climáticos presentes en los registros de ocurrencia de *S. edule* según la clasificación de Köppen: tropical (A), seco (B) y templado (C). El tipo de clima predominante fue templado con invierno seco y verano cálido (Cwb), seguido del clima tropical con lluvias de verano (Aw) y tropical con lluvias monzónicas (Am). En las subprovincias fisiográficas de Volcanes y Lagos de Anáhuac y Sierras Orientales, se observaron hasta siete y ocho tipos diferentes de climas, respectivamente. Llanuras y Lomeríos y Sierra Lacandona, con pocos puntos de ocurrencia, predominaron los climas Cwb y Am, respectivamente (Figura 23A). Ocho grupos de suelos predominantes están presentes en los puntos de ocurrencia de *S. edule*: leptosol 33 %, phaozem 16 %, luvisol y regosol con 13 %. En algunas subprovincias fisiográficas, solo predominó un tipo de suelo (Figura 23B)

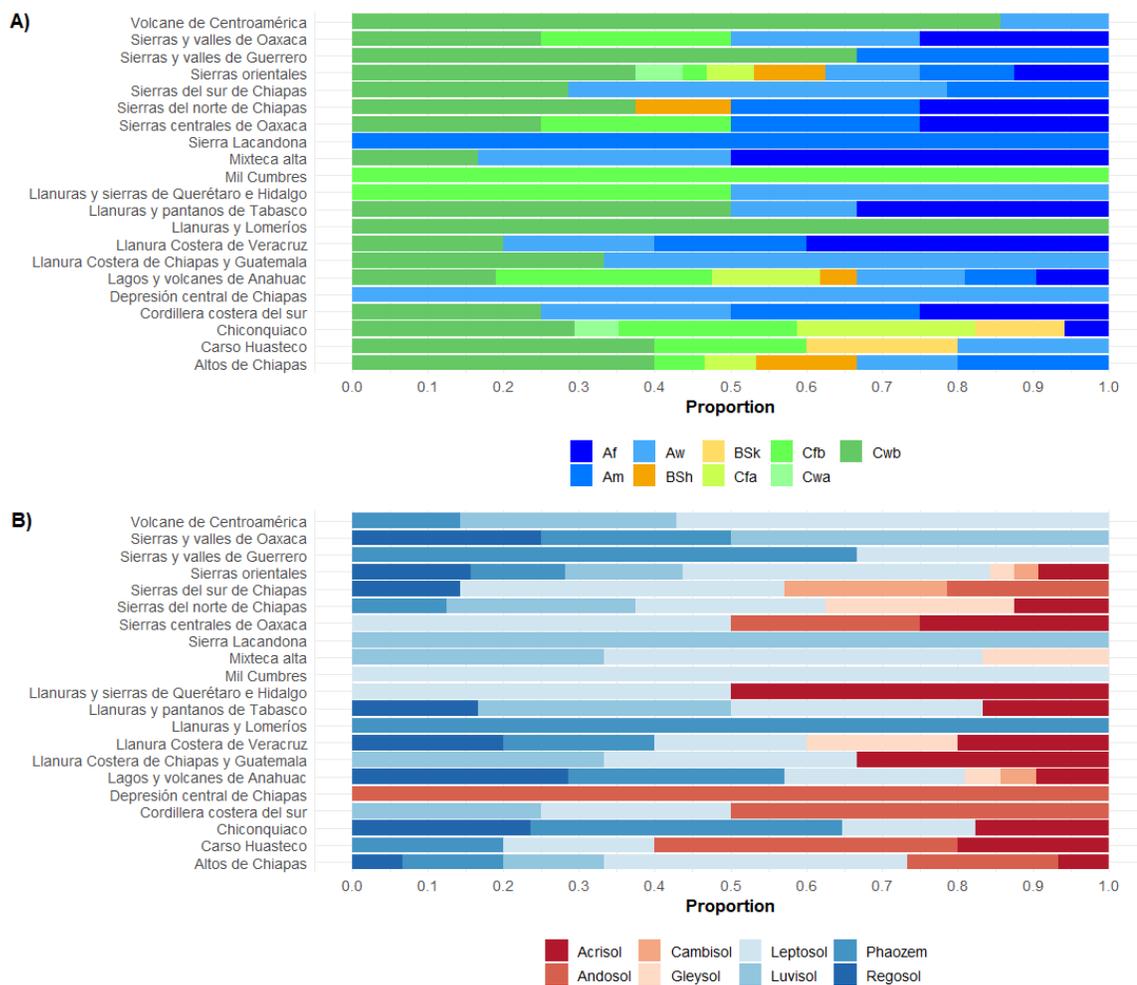


Figura 23. Proporción de **A)** tipos de climas y **B)** tipos de suelo en las subprovincias fisiográficas de México respecto a *S. edule*.

5.3.2. Modelos de distribución

Los modelos de distribución de *S. edule* en México fueron inferidos por las capas bio1, bio2, bio3, bio4, bio12, bio14, bio15 y bio18 y elevación. Se promediaron los valores de AUC y TSS para datos de entrenamiento y prueba. El modelo de bosque aleatorio obtuvo la puntuación AUC más alta (0.95) pero tuvo el rendimiento más bajo (TSS = 0.35). En general, el AUC para los siete modelos fue > 0.5 y mostraron un desempeño adecuado (0.4 < TSS > 0.7) (Cuadro 9).

Cuadro 9. Evaluación y rendimiento de los modelos empleando los estadísticos AUC y TSS.

Model	AUC	TSS
Bioclim	0.88	0.53
Domain	0.86	0.60
Mahalanobis	0.92	0.59
MLG	0.90	0.64
Maxent	0.93	0.62
Random Forest	0.95	0.35
MVS	0.86	0.55

El modelo Bioclim (Figura 24A) mostró la máxima idoneidad para el estado de Chiapas. El modelo de Dominio (Figura 24B) mostró alta idoneidad en el centro y sur de México, en los estados de Chiapas, Guerrero y Veracruz. Se encontró idoneidad moderada en pequeñas regiones del norte del país, donde el clima es generalmente más seco y con pocas precipitaciones. Esto puede explicarse por la coincidencia de los tipos climáticos BSh y BSk en ciertos puntos de ocurrencia. La distancia de Mahalanobis (Figura 24C) mostró resultados muy similares al modelo de Dominio, solo que en este caso la idoneidad se encuentra en el centro y sur de México, particularmente en el estado de Tabasco. El MLG (Figura 24D) mostró muy poca idoneidad para los estados de Chihuahua y Baja California Norte. Este modelo mostró el mejor desempeño y, a diferencia de los demás, su idoneidad cubre gran parte del territorio nacional.

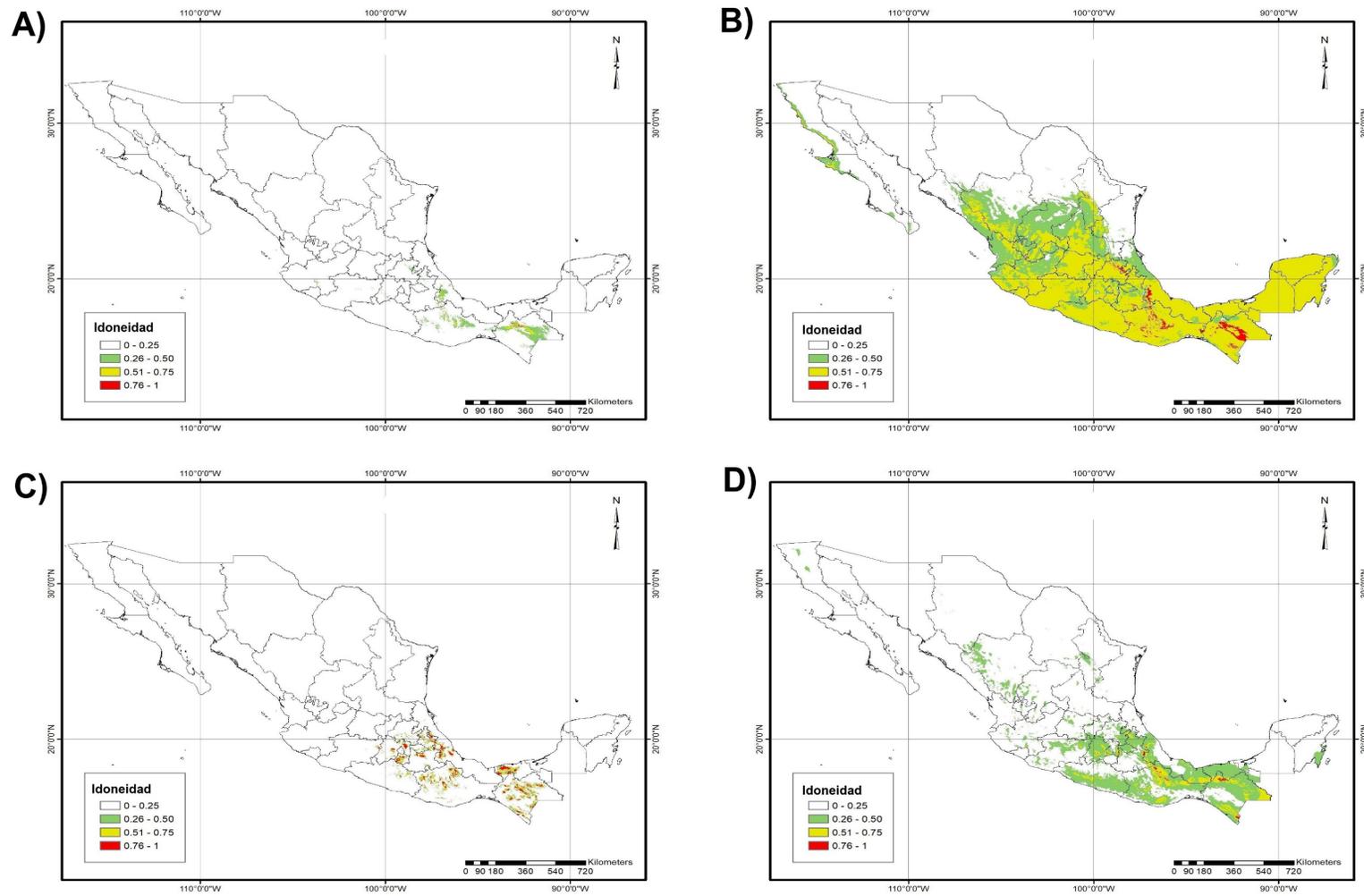


Figura 24. Modelos de distribución para *S. edule* en México. **A)** Bioclim, **B)** Dominio, **C)** Mahalanobis y **D)** MLG.

El modelo de Maxent (Figura 25A) tiene zonas de alta idoneidad similares a los modelos Mahalanobis y Domain, sin embargo, la alta idoneidad se presentó en Veracruz, Guerrero y Chiapas. Los modelos MVS y Bosque aleatorio (Figuras 25 B-C) mostraron mapas muy similares, con áreas de alta idoneidad similares a las observadas en el modelo de Maxent.

Los modelos descriptivos como Domain y Mahalanobis mostraron una mayor área de idoneidad concentrada en el centro y sur de México. Bioclim y MVS dieron resultados muy similares, siendo su idoneidad muy limitada y concentrada en los estados de Veracruz, Guerrero y Oaxaca. Entre los modelos de aprendizaje automático, Maxent mostró áreas más grandes de alta idoneidad en comparación con Bosque aleatorio y MVS. Los modelos Domain y MLG fueron los únicos que mostraron baja idoneidad en el norte y sur de Baja California. En general, las áreas de idoneidad baja a moderada se encontraron en la Península de Yucatán, oeste y norte de México, y las áreas de idoneidad alta se encontraron en los estados de Chiapas, Guerrero, Oaxaca, Puebla, Hidalgo, Veracruz y Tabasco (Figura 25D).

La temperatura media anual juega un papel importante en la distribución de *S. edule*, con un promedio de 20 °C en los 162 puntos de ocurrencia. El rango de temperatura óptimo es 13-21 °C, ya que temperaturas por debajo de esta causan daños en los frutos, mientras que las superiores a 28 °C provocan problemas en la retención de flores y frutos (Lira, 1996). Las poblaciones silvestres de *S. edule* generalmente tienden a distribuirse en los climas Am y Af; sin embargo, en condiciones de crecimiento, *S. edule* puede prosperar en una amplia gama de climas secos y templados (Aung *et al.*, 1990).

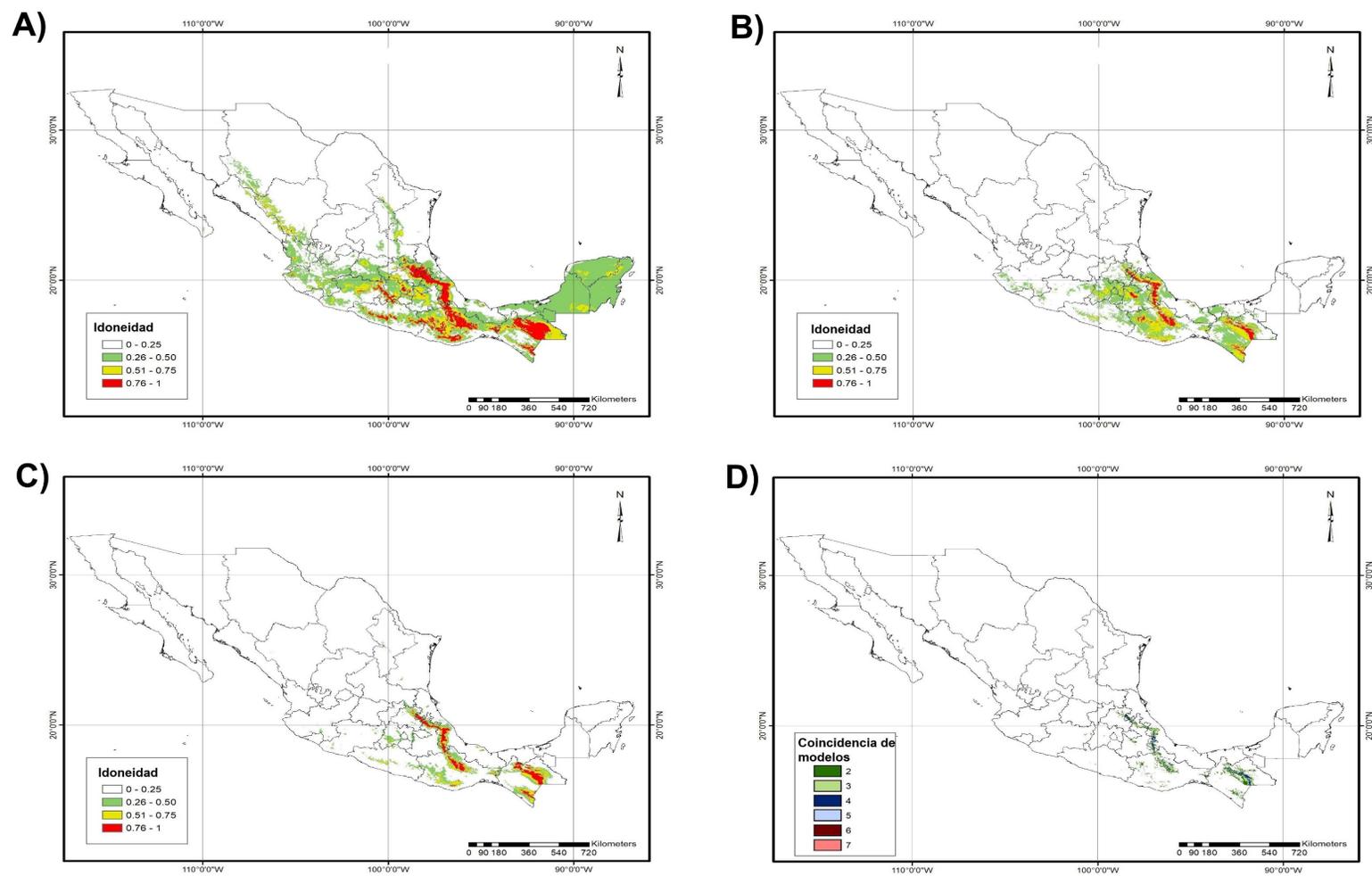


Figura 25. Modelos de distribución para *S. edule* en México. **A)** Maxent, **B)** MVS, **C)** Bosque aleatorio y **D)** Ensemble de coincidencia para los siete modelos.

Los componentes calcáreos de los leptosoles pueden inmovilizar los nutrientes minerales en el suelo y, en el caso del chayote, pueden reducir la productividad y calidad de los frutos (Cadena-Iñiguez, 2005). Sin embargo, algunos leptosoles pueden tener pequeñas capas ricas en materia orgánica, que es un requisito edáfico para las plantaciones de chayote (Lira, 1996). Phaozem, presente en el 50 % de las subprovincias fisiográficas, son suelos considerados ideales para la agricultura por su alto contenido en materia orgánica. Los suelos ideales para la producción de chayote deben estar bien drenados, ricos en materia orgánica y tener un pH de 4.5-6.5 (Lira, 1996), sin embargo, las diferentes variantes fenotípicas del chayote permiten un alto nivel de adaptabilidad a diferentes tipos de suelo (Lira *et al.*, 2009).

Las zonas de alta idoneidad (> 0.8) mostradas en los siete modelos coincidieron con la presencia de dos especies que están estrechamente relacionadas con *S. edule*: *S. chinantlense* y *S. compositum* (Cross *et al.*, 2006). Los modelos de distancia (Domain y Mahalanobis) tienen una extensión territorial de mayor proporción en comparación con los otros modelos, ya que es en estas regiones donde se han tomado el mayor número de muestreos y registros de *S. edule*. Los complejos varietales *virens levis*, *nigrum spinosum*, *nigrum xalapensis*, *albus dulcis* y *albus minor* también se encuentran distribuidos en estas regiones (González-Santos *et al.*, 2017). Sin embargo, *S. edule* es vulnerable en estas mismas áreas. El riesgo de erosión genética en estos lugares se debe principalmente a la presencia de nuevas variedades mejoradas que están desplazando a las poblaciones semidomesticadas, que generalmente tienen frutos con características morfológicas no aptas para el mercado, como la presencia de espinas, tamaño pequeño y sabor amargo. Además, la introducción de cultivos como el café provoca una disminución de las poblaciones de *S. edule* (Cadena-Iñiguez, 2005). En 2007, para contrarrestar los efectos de la erosión genética de *Sechium* spp., se creó el banco de germoplasma en Huatusco, Veracruz, México, donde se protegen *in situ* variantes domesticadas y silvestres

de distintas partes de México y Centroamérica (Cadena-Iñiguez y Arévalo-Galarza, 2011).

La recalcitrancia de la semilla de *S. edule* no permite su resguardo en bancos de germoplasma *ex situ*, por lo que se deben encontrar sitios adecuados para mantener y estudiar las accesiones de *S. edule*. Actualmente se trabaja en la creación de nuevos bancos de germoplasma para facilitar el acceso a la diversidad genética a los productores de chayote, así como para incentivar el intercambio de material genético entre los agricultores (Cadena-Iñiguez *et al.*, 2013).

En este estudio, solo se consideraron los puntos de ocurrencia pertenecientes a *S. edule* cultivado. Sin embargo, Lira *et al.* (2018) emplearon el modelo Maxent y solo utilizaron puntos de ocurrencia de *S. edule silvestre*, sus resultados son muy similares a los de este estudio utilizando el mismo modelo. Esto se debe al hecho de que los requisitos ecológicos de *S. edule* cultivada y silvestre son similares. La historia evolutiva de *S. edule* puede complementar el análisis de los modelos de distribución. Las primeras poblaciones silvestres de *S. edule* se originaron en los estados de Guerrero, Chiapas y Veracruz. A excepción del modelo Bioclim, todos los modelos mostraron áreas de alta idoneidad en estos estados. La hibridación entre poblaciones y variedades de *S. edule*, así como con especies de *S. chinantlense* y *S. compositum*, junto con la especialización adaptativa, han dado lugar a nuevas variantes fenotípicas adaptadas a diferentes condiciones ambientales. *Sechium edule* se considera una especie que puede prosperar en diferentes condiciones climáticas. Si bien los complejos varietales de chayote difieren morfológicamente, las diferencias no son muy notorias cuando se comparan los valores de las variables bioclimáticas (González-Santos *et al.*, 2017). Por tanto, los modelos utilizados pueden ser representativos de todas las poblaciones de *S. edule* (Mateo *et al.*, 2011).

La distribución de especies también puede estar limitada por otros factores que deben considerarse, pero resulta difícil incorporarlos en los modelos de

distribución. Por ejemplo, la competencia con otras especies, la polinización y la influencia humana pueden afectar la distribución. Es necesario evaluar diferentes modelos para obtener una perspectiva amplia sobre la distribución de especies.

5.4. CONCLUSIONES

Las poblaciones cultivadas de *S. edule* prosperan en una amplia gama de tipos de clima y suelo. Se encontró que el MLG es el mejor según los estadísticos AUC y TSS. Sin embargo, los siete modelos fueron relativamente adecuados para explicar la distribución de *S. edule* en México. Las áreas de alta idoneidad para el desarrollo de *S. edule* se ubican en regiones de los estados de Chiapas, Guerrero, Oaxaca, Veracruz, Tabasco, Puebla e Hidalgo, las cuales deben ser consideradas como áreas de alta prioridad para la conservación. La adaptabilidad climática de *S. edule* también permite que se distribuya hacia el oeste y norte de México.

5.5. LITERATURA CITADA

- Aguñiga-Sánchez, I., Cadena-Iñiguez, J., Santiago-Osorio, E., Gómez-García, G., Mendoza-Núñez, V. M., Rosado-Pérez, J., Ruíz-Ramos, M., Cisneros-Solano, V. M., Ledesma-Martínez, E., Delgado-Bordonave, A. de J., & Soto-Hernández, R. M. (2017). Chemical analyses and in vitro and in vivo toxicity of fruit methanol extract of *Sechium edule* var. *nigrum spinosum*. *Pharmaceutical Biology*, 55(1), 1638-1645. <https://doi.org/10.1080/13880209.2017.1316746>
- Aguñiga-Sánchez, I., Soto-Hernández, M., Cadena-Iñiguez, J., Ruíz-Posadas, L. del M., Cadena-Zamudio, J. D., González-Ugarte, A. K., Steider, B. W., & Santiago-Osorio, E. (2015). Fruit extract from a *Sechium edule* hybrid induce apoptosis in leukaemic cell lines but not in normal cells. *Nutrition and Cancer*, 67(2), 250-257.

- Aung, L. H., Ball, A., & Kushad, M. (1990). Developmental and nutritional aspects of chayote (*Sechium edule*, Cucurbitaceae). *Economic Botany*, 44(2), 157-164. JSTOR.
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1), 1-12. <https://doi.org/10.1038/sdata.2018.214>
- Bivand, R. S., & Lewin-Koh, N. (2020). Maptools: Tools for Handling Spatial Objects. R package version 1.0-2. <https://CRAN.R-project.org/package=maptools>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Busby, J. R. (Australian N. P. and W. S. (1991). BIOCLIM - a bioclimate analysis and prediction system. *Plant Protection Quarterly (Australia)*. <https://agris.fao.org/agris-search/search.do?recordID=AU9103158>
- Cadena Iñiguez, J., Soto Hernández, M., Arévalo Galarza, M. de L., Avendaño Arrazate, C. H., Aguirre Medina, J. F., & Ruiz Posadas, L. del M. (2011). Biochemical characterization of domesticated varieties of chayote *Sechium edule* (Jacq.) Sw. Fruits compared to wild relatives. *Revista Chapingo Serie Horticultura*, 17(2), 45-55. <https://doi.org/10.5154/r.rchsh.2011.17.044>
- Cadena-Iñiguez, J. (2005). Caracterización morfoestructural, fisiológica, química y genética de diferentes tipos de chayote (*Sechium edule*) [Doctoral]. Colegio de Postgraduados, Campus Montecillo.

- Cadena-Iñiguez, J., & Arévalo-Galarza, M. L. C. (2011). Las variedades de Chayote (*Sechium edule* (Jacq.) Sw.) y su comercio mundial (1 st ed.). bba. 98 p.
- Cadena-Iñiguez, J., Avendaño-Arrazate, C. H., Soto-Hernández, M., Ruiz-Posadas, L. M., Aguirre-Medina, J. F., & Arévalo-Galarza, L. (2008). Intraspecific variation of *Sechium edule* (Jacq.) Sw. In the state of Veracruz, Mexico. *Genetic Resources and Crop Evolution*, 55(6), 835-847. <https://doi.org/10.1007/s10722-007-9288-4>
- Cadena-Iñiguez, J., Soto-Hernández, M., Arévalo-Galarza, M. L. C., Avendaño-Arrazate, C. H., & Aguirre-Medina, J. F. (2013). Modelos de mejoramiento genético participativo en chayote (*Sechium spp*) (1 st ed.). Editorial del Colegio de Postgraduados.
- Cadena-Iñiguez, J., Soto-Hernández, M., Torres-Salas, A., Aguiñiga-Sánchez, I., Ruíz-Posadas, L., Rivera-Martínez, A. R., Avendaño-Arrazate, C. H., & Santiago-Osorio, E. (2013). The antiproliferative effect of chayote varieties (*Sechium edule* (Jacq.) Sw.) on tumour cell lines. *Journal of Medicinal Plants Research*, 7(8), 455-460. <https://doi.org/10.5897/JMPR12.866>
- Carpenter, G., Gillison, A. N., & Winter, J. (1993). DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity & Conservation*, 2(6), 667-680. <https://doi.org/10.1007/BF00051966>
- Carstens, B. C., & Richards, C. L. (2007). Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, 61(6), 1439-1454. <https://doi.org/10.1111/j.1558-5646.2007.00117.x>
- Cross, H., Lira, S. R., & Motley, T. J. (2006). Origin and diversification of chayote. In T. J. Motley, N. Zerega, & H. Cross (Eds.), *Darwin's Harvest: New*

Approaches to the Origins, Evolution, and Conservation of Crops (1^o, pp. 171-194). Columbia University Press.

Elith, J., Graham*, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. M., Peterson, A. T., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129-151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Eskildsen, A., Roux, P. C. le, Heikkinen, R. K., Høye, T. T., Kissling, W. D., Pöyry, J., Wisz, M. S., & Luoto, M. (2013). Testing species distribution models across space and time: High latitude butterflies and recent warming. *Global Ecology and Biogeography*, 22(12), 1293-1303. <https://doi.org/10.1111/geb.12078>

Evans, J. M., Fletcher, R. J., & Alavalapati, J. (2010). Using species distribution models to identify suitable areas for biofuel feedstock production. *GCB Bioenergy*, 2(2), 63-78. <https://doi.org/10.1111/j.1757-1707.2010.01040.x>

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315. <https://doi.org/10.1002/joc.5086>

Fischer, G., Nachtergaele, F., Prieler, S., van Velthuisen, H. T., Verelst, D., & Wiberg, D. (2008). *Global Agro-ecological Zones Assessment for Agriculture*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Freeman, E. A., & Moisen, G. (2008). PresenceAbsence: An R Package for Presence-Absence Model Analysis. *Journal of Statistical Software*, 23(11), 1-31. <https://doi.org/10.18637/jss.v023.i11>

- Gobeyn, S., Mouton, A. M., Cord, A. F., Kaim, A., Volk, M., & Goethals, P. L. M. (2019). Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling*, 392, 179-195. <https://doi.org/10.1016/j.ecolmodel.2018.11.013>
- González-Santos, R., Cadena-Íñiguez, J., Morales-Flores, F. J., Ruiz-Vera, V. M., & Pimentel-López, J. (2017). Prediction of the effects of climate change on *Sechium edule* (Jacq.) Swartz varietal groups in Mexico. *Genetic Resources and Crop Evolution*, 64(4), 791-804. <https://doi.org/10.1007/s10722-016-0401-4>
- Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology*, 84, 2809-2815. <https://doi.org/10.1890/02-3114>
- Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null-model. *Ecology*, 93, 679-688. <https://doi.org/10.1890/11-0826.1>
- Hijmans, R. J. (2020). raster: Geographic Data Analysis and Modeling (R package version 3.3-13) [Computer software]. <https://CRAN.R-project.org/package=raster>
- Hijmans, R. J., & Elith, J. (2013). Species distribution modeling with R. 77 p.
- Hirzel, A., Hausser, J., Chessel, D., & Perrin, N. (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, 83, 2027-2036. <https://doi.org/10.2307/3071784>
- Lantschner, V., Vega, G., & Corley, J. (2018). Predicting the distribution of harmful species and their natural enemies in agricultural, livestock and forestry systems: An overview. *International Journal of Pest Management*, 65, 1-17. <https://doi.org/10.1080/09670874.2018.1533664>

- Liaw, M., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18-22.
- Lira, R. (1996). Chayote. *Sechium edule* (Jacq.) Sw., promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute., Rome, Italy. 54 p.
- Lira, R., Eguiarte, L. E., & Montes-Hernández, S. (2009). Proyecto recopilación y análisis de la información existente de las especies de los géneros *Cucurbita* y *Sechium* que crecen y/o se cultivan en México. 107 p.
- Lira, Rafael, Castrejón, J., Zamudio, S., & Zenteno, C. R. (1999). Propuesta de ubicación taxonómica para los chayotes silvestres (*Sechium edule*, Cucurbitaceae) de México. *Acta Botánica Mexicana*, 49, 47-61. <https://doi.org/10.21829/abm49.1999.838>
- Lira, S. R. (1995). Estudios taxonómicos en el género *Sechium* P. Br. Cucurbitaceae [Tesis Doctoral]. Universidad Nacional Autónoma de México. 267 p.
- Lira, S. R., Sosa-Guzmán, L. J., Cabral-Soto, I. V., & Téllez, O. V. (2018). Distribución potencial de las especies silvestres de la familia Cucurbitaceae de México (Informe final SNIB-CONABIO Proyecto No. JM003; p. 18). Universidad Nacional Autónoma de México. Facultad de Estudios Superiores Iztacala.
- Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, 145-151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>

- Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size – finer may not always be better. *Scientific Reports*, 8(1), 7168. <https://doi.org/10.1038/s41598-018-25437-1>
- Mateo, R. G., Felicísimo, A. M., & Muñoz, J. (2011). Species distributions models: A synthetic revision. *Revista Chilena de Historia Natural*, 84, 217-240. <http://dx.doi.org/10.4067/S0716-078X2011000200008>
- Newstrom, L. (1990). Origin and evolution of chayote, *Sechium edule*. En D. M. Bates, R. W. Robinson, & C. Jeffrey (Eds.), *Biology and Utilization of the Cucurbitaceae* (1º, pp. 141-149). Cornell University Press.
- Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5(2). <https://cran.r-project.org/doc/Rnews/>
- Phillips, S.J. (2009). A Brief tutorial on Maxent. Network of conservation educators and practitioners, Center for Biodiversity and Conservation, American Museum of Natural History. Lessons in Conservation.
- Phillips, Steven J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Polce, C., Termansen, M., Aguirre-Gutiérrez, J., Boatman, N., Budge, G., Crowe, A., Garratt, M., Pietravalle, S., Potts, S., Ramirez, J., Somerwill, K., & Biesmeijer, J. (2013). Species distribution models for crop pollination: A modelling framework applied to Great Britain. *PloS one*, 8, e76308. <https://doi.org/10.1371/journal.pone.0076308>
- QGIS Development Team. (2020). QGIS Geographic Information System. Open-Source Geospatial Foundation Project (3.16.2) [Computer software]. <https://www.qgis.org/en/site/>

- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing (1.3.1093) [Computer software]. <https://www.R-project.org/>
- Salazar-Aguilar, S., Ruiz-Posadas, L. D. M., Cadena-Iñiguez, J., Soto-Hernández, M., Santiago-Osorio, E., Aguiñiga-Sánchez, I., Rivera-Martínez, A. R., & Aguirre-Medina, J. F. (2017). *Sechium edule* (Jacq.) Swartz, a new cultivar with antiproliferative potential in a human cervical cancer HeLa cell line. *Nutrients*, 9(8), E798. <https://doi.org/10.3390/nu9080798>
- SIACON. (2020). Sistema de Información Agroalimentaria de Consulta Nueva Generación. SIAP. <https://www.gob.mx/siap>
- Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley. 768 p.
- Vignali, S., Barras, A., Arlettaz, R., & Braunisch, V. (2020). SDMtune: An R package to tune and evaluate species distribution models. *Ecology and Evolution*, 10. <https://doi.org/10.1002/ece3.6786>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open-Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Zhang, J., Jiang, F., Li, G., Qin, W., Li, S., Gao, H., Cai, Z., Lin, G., & Zhang, T. (2019). Maxent modeling for predicting the spatial distribution of three raptors in the Sanjiangyuan National Park, China. *Ecology and Evolution*, 9(11), 6643-6654. <https://doi.org/10.1002/ece3.5243>

CAPÍTULO VI. DESCRIPTORES EDAFOCLIMÁTICOS Y MODELOS DE DISTRIBUCIÓN DEL CLADO CENTROAMERICANO DE *Sechium* P. BROWN

EDAPHOCLIMATIC DESCRIPTORS AND DISTRIBUTION MODELS OF THE CENTRAL AMERICAN CLADO OF *Sechium* P. BROWN

RESUMEN

Las especies centroamericanas de *Sechium* son endémicas a Costa Rica y Panamá. Estas especies han sido descritas morfológicamente, sin embargo, la información sobre su ecogeografía se encuentra limitada. Mediante puntos de ocurrencia y sistemas de información geográfica, el objetivo de este estudio fue determinar las condiciones edafoclimáticas bajo las cuales prosperan estas especies, esto con la finalidad de contribuir y enriquecer las bases de datos de estas, en aras de conservarlas y aprovecharlas para mejorar las características de especies cultivadas. Se emplearon 22 variables bioclimáticas con resolución espacial de $\sim 1 \text{ Km}^2$ y se extrajeron los valores ráster para los respectivos análisis estadísticos. En el análisis clúster de k-medioídes se obtuvieron tres grupos donde todos los puntos de ocurrencia aparecieron mezclados unos con otros, y no se lograron diferenciar por especies. En el ACP los dos primeros componentes contribuyeron con el 60.3 % de la variación total, el primer componente aportó el 36.7 % de la variación y estuvo integrado por las variables bio1, bio2, bio4, bio14, bio15 y bio19; el segundo componente principal aportó el 23.6 % de la variación y estuvo conformado por las variables bio12, bio13 y bio16. Los climas predominantes fueron Af (29.2 %) Cfb (27.3 %), Am (20.6 %), Cwb (18.7 %), Aw (3.83 %) y Cwc (0.47 %). Los tipos de suelo predominantes fueron andosol (38.3 %), cambisol (30.6 %), arenosol (16.7 %) y nitisol (7.18 %). La temperatura media anual para las seis especies del clado centroamericano tiene un promedio de 18.9 °C. La precipitación en esta zona geográfica tiene un promedio de 3000 mm.

Palabras clave: endemismo, ecogeografía, condiciones edafoclimáticas, k-medioídes, humedad, bajas temperaturas.

ABSTRACT

The Central American species of *Sechium* are endemic to Costa Rica and Panama. These species have been morphologically described, however, information on their ecogeography is limited. Through occurrence records and geographic information systems, the objective of this study was to determine the edaphoclimatic conditions of these species under which they thrive, this in order to contribute and enrich the databases of these species, in order to conserve and take advantage of them to improve the characteristics of cultivated species. 22 bioclimatic variables with spatial resolution of $\sim 1 \text{ Km}^2$ were used and raster values were extracted for the respective statistical analyzes. In the cluster analysis of k-medoids, three groups were obtained where all the occurrence records appeared mixed with each other, they were not able to differentiate by species. In the PCA, the first two components contributed 60.3% of the total variation, the first component contributed 36.7% of the variation and was made up of the variables bio1, bio2, bio4, bio14, bio15 and bio19; the second main component contributed 23.6% of the variation and was made up of the variables bio12, bio13 and bio16. The predominant climates were Af (29.2%) Cfb (27.3%), Am (20.6%), Cwb (18.7%), Aw (3.83%) and Cwc (0.47%). The predominant soil types were andosol (38.3%), cambisol (30.6%), arenosol (16.7%) and nitisol (7.18%). The mean annual temperature for the six species of the Central American clade has an average of 18.9 °C. Precipitation in this geographical area has an average of 3000 mm.

Key words: endemism, ecogeography, edaphoclimatic conditions, k-medoids, humidity, low temperatures.

6.1. INTRODUCCIÓN

El género *Sechium* P. Brown pertenece a la familia Cucurbitaceae y está integrado por 11 especies. Seis especies de este género son endémicas a Costa Rica y Panamá: *S. panamense* (Wunderlin) Lira & Chiang, *S. pittieri* (Cogn.) C. Jeffrey, *S. tacaco* (Pittier) C. Jeffrey, *S. talamancense* (Wunderlin) C. Jeffrey, *S. venosum* (L.D. Gómez) Lira & Chiang y *S. villosum* (Wunderlin) C. Jeffrey; estas seis especies están en el clado centroamericano de *Sechium*. Las especies restantes *S. chinantlense* Lira & Chiang, *S. compositum* (Donn. Sm.) C. Jeffrey, *S. edule* (Jacq.) Swartz, *S. hintonii* (P.G. Wilson) C. Jeffrey y *S. mexicanum* Lira & Nee se distribuyen en las zonas montañosas de México (Newstrom, 1990, 1991; Lira, 1995)

Sechium panamense se restringe a Panamá y tiene un intervalo de adaptación a la altitud entre los 1800-3000 msnm, prosperando así en climas fríos (Lira y Chiang, 1992; Lira, 1995). *Sechium pittieri* tiene una distribución en Costa Rica, Panamá y peculiarmente en Nicaragua, por lo que su adaptabilidad es mucho mayor comparada con las especies del clado centroamericano (Lira, 1995). *Sechium tacaco* es la especie más aprovechada dentro del clado centroamericano y constituye formas semidomesticadas en San José, Costa Rica, donde la utilizan ampliamente en la gastronomía (Lira, 1995). Al estar sometida a cierto grado de domesticación, las poblaciones de *S. tacaco* han experimentado cambios fenotípicos en la morfología de sus frutos (Monge y Loría, 2017). *Sechium talamancense*, *S. venosum* y *S. villosum* se distribuyen en Costa Rica en altitudes que pueden sobrepasar los 2000 msnm (Lira, 1995).

El estudio de Lira (1995) aportó información de las especies centroamericanas y su posible contribución en la mejora hacia las especies cultivadas (*S. edule* y *S. tacaco*). Sin embargo, la información edafoclimática de las especies centroamericanas es limitada. Actualmente los Sistemas de Información Geográfica (SIG) nos brindan información sobre las variables climáticas y ecológicas presentes en una porción específica de territorio; por otra parte, los

GIS son de importancia para realizar estudios de conservación, localización y distribución de las especies de interés (Mateo *et al.*, 2011).

El objetivo de este estudio fue determinar las condiciones edafoclimáticas y la distribución potencial de las especies centroamericanas de *Sechium*, con la finalidad de contribuir y enriquecer las bases de datos relacionadas a estas especies, en aras de conservarlas y aprovecharlas en un futuro para mejorar las características de las especies cultivadas.

6.2. MATERIALES Y MÉTODOS

6.2.1. Puntos de ocurrencia y variables bioclimáticas

Se obtuvieron 209 puntos de ocurrencia para las seis especies centroamericanas de *Sechium* de GBIF (<https://www.gbif.org/>) (nueve de *S. panamense*, 147 de *S. pittieri*, 16 de *S. tacaco*, 13 de *S. talamancense*, 15 de *S. venosum* y nueve de *S. villosum*). Se emplearon las 19 variables bioclimáticas de WorldClim 2.0 (Fick y Hijmans, 2017); se incorporaron también modelos ráster para elevación (Fick y Hijmans, 2017), tipos de clima de Köppen-Geiger (Beck *et al.*, 2018) y tipos de suelo según la clasificación de World Soil Database version 1.2 (Fischer *et al.*, 2008). Las 22 capas empleadas tienen una resolución espacial de ~ 1 km². Los valores ráster se obtuvieron con Point Sampling Tool de QGIS versión 3.16.2 (2020).

6.2.2. Análisis descriptivo

Todos los paquetes por mencionar se ejecutaron en Rstudio (R Core Team, 2020). De las 19 variables bioclimáticas de WorldClim y la variable elevación se obtuvieron los valores primer y tercer cuartiles, mínimos, máximos y media; los valores mínimos y máximos representaron los límites a partir de los cuales cada especie puede tolerar según la variable bioclimática en cuestión. Las variables

categorías de tipos de clima y suelo se graficaron con tidyverse (Wickham *et al.*, 2019) para observar la proporción en función de las especies.

6.2.3. Análisis de agrupamiento

Para verificar si existe tendencia al agrupamiento de los valores ráster se calculó el estadístico de Hopkins (H) con clustertend (Yilan y Rutong, 2015). El mejor algoritmo de agrupación (jerárquico o de partición) se determinó con clValid (Brock *et al.*, 2008). De esta forma se escogió el método de partición k-medioides para determinar similitudes edafoclimáticas en los puntos de ocurrencia de las especies con la librería factoextra (Kassambara y Mundt, 2020). K-medioides necesita especificar el número óptimo de grupos, el cual se determinó con NbClust (Charrad *et al.*, 2014), donde previamente los valores ráster deben ser normalizados.

6.2.4. Modelos de distribución

En las variables cuantitativas se aplicó un análisis de correlación de Pearson y se eliminaron aquellas con valores > 0.8 para evitar la afectación a los modelos (Phillips *et al.*, 2006). Se corrieron los siguientes algoritmos de aprendizaje automático: Modelo lineal generalizado (MLG), Modelo aditivo generalizado (MAG), Esplines adaptativos de regresión multivariante (SARM), Modelo de regresión potenciado generalizado (MRPG), Análisis de árboles de clasificación (AAC), Bosque aleatorio (RF), Maxent, Red neuronal artificial (RNA) y Máquinas de vectores de soporte (MVS). Estos ocho algoritmos se ensamblaron para observar un solo mapa de distribución de cada especie. Se efectuó la prueba de Jackknife para cuantificar la contribución de cada variable bioclimática a los modelos de distribución, por otra parte, se calcularon los estadísticos de AUC, PPC, kappa y TSS para verificar la fiabilidad y rendimiento de cada modelo. Todas las funciones mencionadas en este párrafo se ejecutaron con SSDM (Schmitt *et al.*, 2017).

6.3. RESULTADOS Y DISCUSIÓN

Los valores correspondientes a primer y tercer cuartiles, mínimos, máximos y medias se encuentran en el Apéndice 1. El correlograma (Figura 26) mostró como la variable elevación correlaciona de forma inversa con variables relacionadas a la temperatura, es decir, a mayores altitudes la temperatura tiende a bajar.

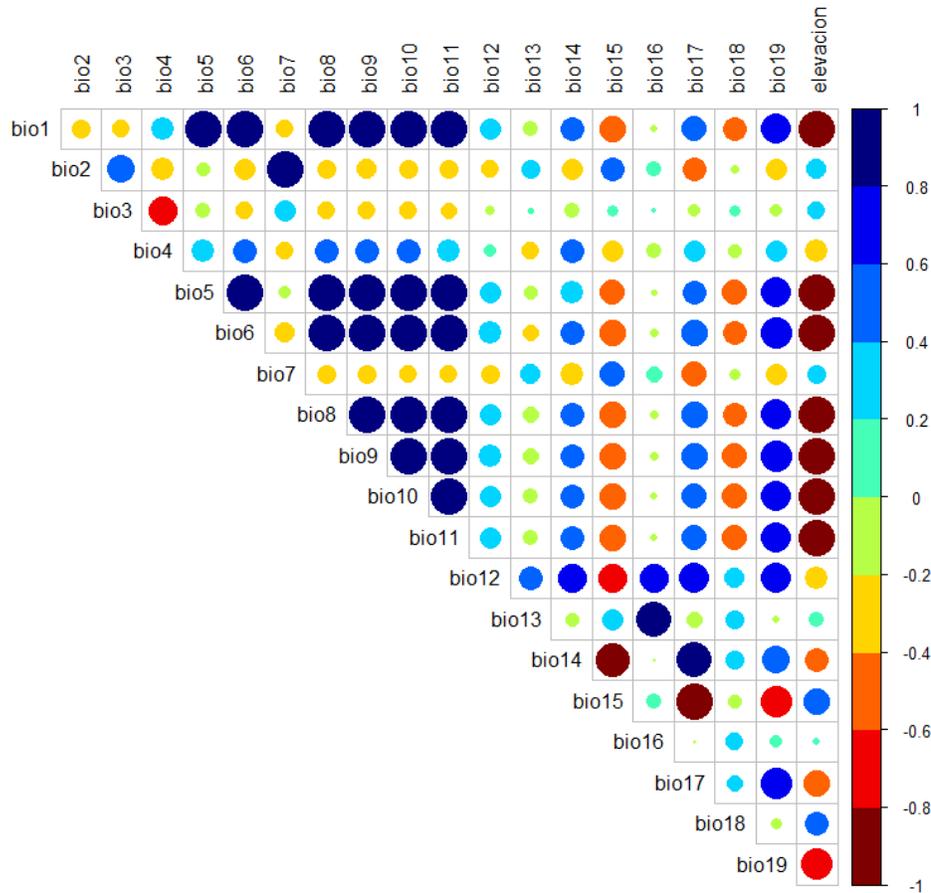


Figura 26. Correlograma de las variables cuantitativas para las especies centroamericanas de *Sechium*

6.3.1. Tipos de clima y suelo

El clima predominante fue el tropical ecuatorial (Af) presentándose en el 29.2 % de los 209 puntos de ocurrencia; una proporción de todas las especies se desarrollan bajo este tipo de clima. Le siguen los tipos de climas Cfb 27.3 %, Am

20.6 %, Cwb 18.7 %, Aw 3.83 % y Cwc 0.47 %, éste último únicamente presente en una accesión de *S. pittieri*. La proporción de las accesiones de cada especie en función a tipos de clima y suelo se muestran en la Figura 27. Las accesiones de *S. villosum* y *S. panamense* únicamente se encuentran en suelos tipo andosol. Las accesiones de *S. pittieri* están presentes en todos los tipos de suelo: cambisol 38.1 %, andosol 25.9 %, arenosol 18.4 %, nitisol 7.48 %, kastanozem 5.44 %, regosol 4.08 % y leptosol 0.68 % (Figura 27 A-B). El andosol es el suelo predominante en las especies centroamericanas de *Sechium* y se caracterizan por tener alto contenido de materia orgánica; son suelos de origen volcánico, tienen gran capacidad de retención de agua e intercambio catiónico.

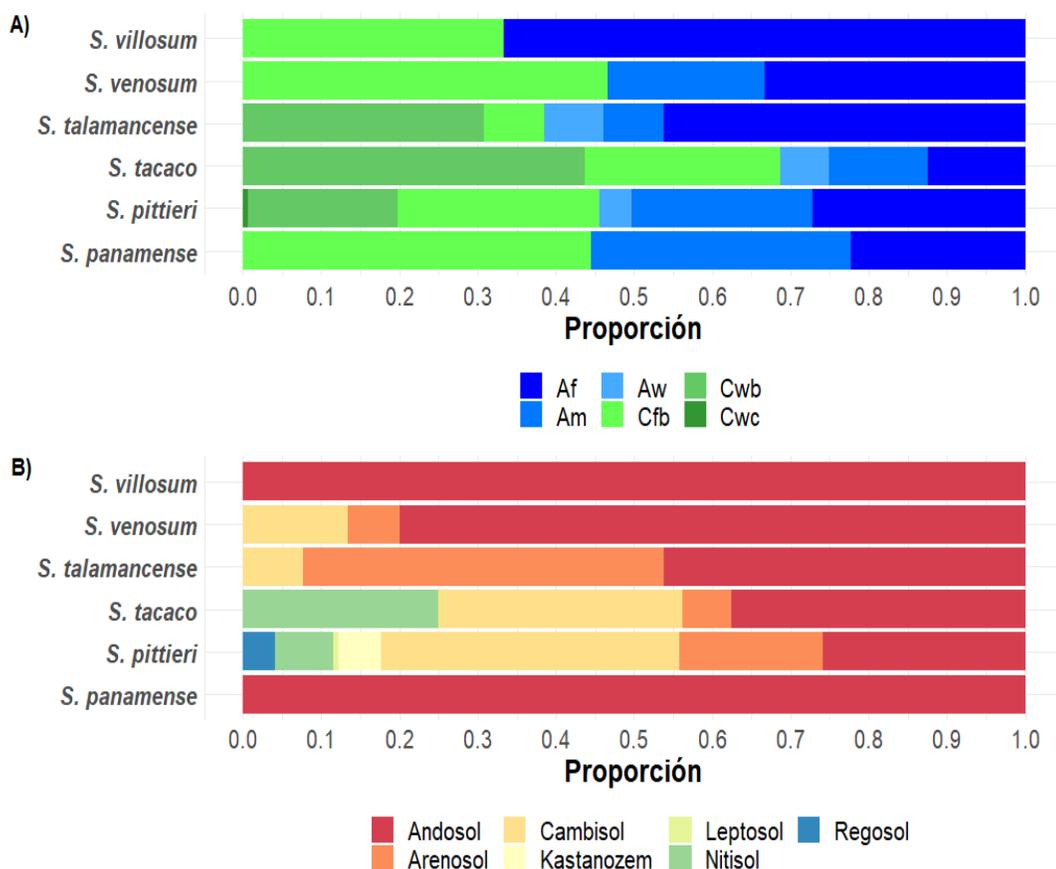


Figura 27. Proporción de **A)** tipos de clima y **B)** tipos de suelo en relación con los puntos de ocurrencia para cada especie.

S. panamense

Se desarrolla a una temperatura media anual de 19 °C, pero puede prosperar en temperaturas mínimas de 10 °C y máximas de 31 °C; tolera precipitaciones que oscilan los 2600-3000 mm. Se le encuentra en climas tipo Am, Af y Cfb (Figura 27A). El suelo donde mejor se adapta es el andosol, con orígenes volcánicos, permeables, ligeros y con una ligera capa de materia orgánica. Su intervalo de altitud es de 1300-2000 msnm, sin embargo, Lira (1995) reportó que puede extender sus límites hasta los 3000 msnm.

S. pittieri

Se desarrolla a una temperatura media anual de 19 °C; puede prosperar en temperaturas mínimas de 2 °C y máximas de 32 °C; se adapta a lugares con precipitaciones promedio de 2600-3200 mm, sin embargo, se registró un punto con hasta 4000 mm (Apéndice 1, sección *S. pittieri*, bio12). Su rango de adaptabilidad climática le permite establecerse en los seis tipos de climas encontrados (Af, Am, Aw, Cfb, Cwb y Cwc) (Figura 27A) y los siete tipos de suelo encontrados en este estudio (Figura 27B). Lira (1995) argumentó que esta especie puede distribuirse desde Costa Rica hasta Nicaragua, en donde se consume como verdura. Su intervalo de altitud es de 900-2900 msnm, guardando concordancia con el estudio de Lira (1995), aunque se encontró un punto de *S. pittieri* en una altitud de 16 msnm (Apéndice 1, sección *S. pittieri*, elevación). Su adaptabilidad puede sugerir una amplia variabilidad genética, la cual todavía no ha sido explorada y debe ser estudiada con marcadores morfológicos y moleculares.

S. tacaco

Su requerimiento medio anual de temperatura es de 18 °C, pudiendo soportar temperaturas mínimas de 9 °C y máximas de 30 °C. Se desarrolla en lugares con precipitaciones de 2200-2800 mm; su intervalo de altitud está entre los 1400-2000 msnm (Apéndice 1). Los puntos de ocurrencia coincidieron en lugares con climas tropical (Aw, Af y Am) y templado (Cwb y Cfb) (Figura 27). A pesar de

adaptarse a una amplia gama de suelo, se desarrolla bien en condiciones de cultivo en andosoles y nitisoles, debido a su alta estabilidad estructural, profundidad y baja saturación de bases.

S. talamancense

Tiene un requerimiento medio anual de temperatura de 20 °C, aun así, puede tolerar temperaturas mínimas de 9 °C y máximas de 30 °C. La especie se desarrolla en condiciones húmedas con registros de precipitación pluvial de 3300-3800 mm en un intervalo de altitud de 500-2200 m (Apéndice 1), sin embargo, su límite puede extenderse hasta los 3200 m de altitud (Lira, 1995). Se le encuentra en los tipos de clima Af, Am, Aw, Cfb y Cwb; debido a sus altos requerimientos hídricos prefiere suelos arenosos (arenosol) pero también puede desarrollarse en cambisoles y andosoles (Figura 27).

S. venosum

Su requerimiento medio anual de temperatura es de 17 °C, tolerando temperaturas mínimas de 6 °C y máximas de 30 °C. Al igual que *S. talamancense*, prospera en ambientes muy húmedos de 2600-3700 mm; su intervalo de altitud oscila los 1100-2500 m. Se adapta en climas Cfb, Am y Af; prefiere suelos tipo andosol, arenosol y cambisol (Figura 27 A-B). La estructura genética de *S. venosum* puede ser una fuente de germoplasma en la búsqueda de genes resistentes al ataque de fitopatógenos (Lira, 1995).

S. villosum

Tiene un requerimiento medio anual de 19 °C, sin embargo, tolera temperaturas mínimas de 6 °C y máximas de 29 °C. Al igual que las demás especies de su clado, se desarrolla en ambientes húmedos con 2900-3500 mm; su intervalo de altitud oscila los 800-900 msnm (Apéndice 1), pero puede sobrepasar los 2700 msnm. Lira (1995) reportó intervalos de altitud para esta especie entre los 1500-2000 m.

6.3.2. Análisis ACP y clúster k-medioides

Con los 209 puntos de ocurrencia se formaron tres clusters ($k = 3$), que fue el número óptimo calculado por NbClust. El coeficiente H fue de 0.24 y está por debajo del umbral (0.5), es decir, los datos son candidatos para análisis de agrupamiento. Las dos primeras dimensiones contribuyeron con el 60.3 % de la variación total (Figura 28 A-B). Los puntos de ocurrencia de las especies quedaron distribuidos en distintos clusters, es decir, en un cluster pueden agruparse varios puntos de ocurrencia de distintas especies, ya que se distribuyen en nichos ecológicos similares debido a su distribución restringida en Costa Rica y Panamá. (Figura 28 A-C). El 36.8 % de los puntos de ocurrencia se agruparon en el clúster 1, el 35.4 % en el clúster 2 y el 27.8 % en el clúster 3.

El clúster 1 quedó conformado por puntos de ocurrencia de *S. venosum*, *S. talamancense*, *S. tacaco* y *S. pittieri*. Por ejemplo, para *S. pittieri*, el 32.7 % de sus accesiones están en el clúster 1, el 34 % en el clúster 2 y 33.3 % en el clúster 3; para *S. villosum* el 22.2 % de sus accesiones están en el grupo 2 y el 77.8 % en el clúster 3 (Figura 28)

Los dos primeros componentes principales contribuyeron con el 60.3 % de la variación total (Figura 28B). El primer componente aportó el 36.7 % de la variación total y estuvo integrado por las variables bio1, bio2, bio4, bio14, bio15 y bio19; el segundo componente principal aportó el 23.6 % de la variación total y estuvo conformado por las variables bio12, bio13 y bio16.

Al igual que el análisis clúster de k-medioides, no se muestran grupos bien definidos de las especies, esto debido a que todas ellas se distribuyen de forma restringida en Costa Rica y Panamá, y por ende las condiciones edafoclimáticas son muy similares, pudiendo ser inclusive que muchas de ellas interactúen en los mismos nichos ecológicos.

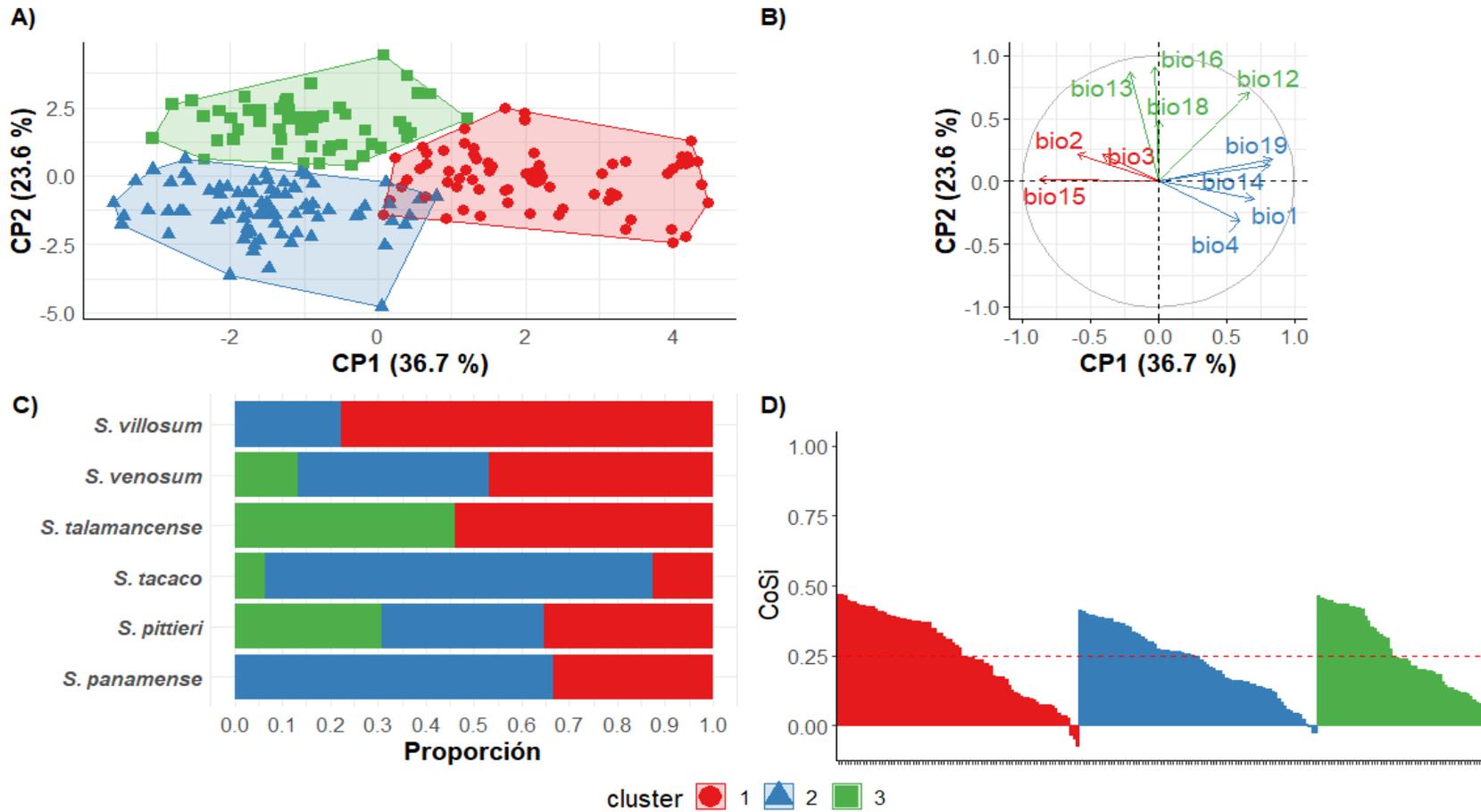


Figura 28. **A)** Agrupación k-medioides, **B)** principales variables bioclimáticas que contribuyen en la distribución de las especies centroamericanas de *Sechium*, **C)** proporción de los puntos de ocurrencia de las especies en cada clúster y **D)** consistencia del agrupamiento mediante el coeficiente de la Silueta.

6.3.3. Modelos de distribución

S. panamense

Todos los modelos de distribución presentaron AUC > 0.8 con excepción de MPG (0.75); los estadísticos de Kappa y TSS estuvieron dentro de los rangos óptimos (Cuadro 10). Respecto a la prueba Jackknife, la variable bio1 (temperatura media anual) fue la que más contribuyó en los modelos de distribución, es decir, esta variable es la responsable de permitir la distribución de *S. panamense* (Cuadro 11). Las zonas de alta idoneidad se encuentran en la región que divide Bocas del Río y Chiriquí (Figura 29).

Cuadro 10. Parámetros estadísticos para los nueve modelos de distribución y ensamble para *S. panamense*.

Modelo	AUC	PPC	Sensibilidad	Especificidad	Kappa	TSS
MLG	0.82	0.81	0.83	0.81	0.21	0.64
MAG	0.82	0.81	0.83	0.81	0.27	0.64
MPG	0.75	0.75	0.66	0.83	0.5	0.49
MARS	0.90	0.97	0.83	0.97	0.35	0.80
AAC	0.91	0.91	0.83	0.99	0.83	0.82
MAXENT	0.99	0.98	0.99	0.98	0.12	0.97
RNA	0.91	0.91	0.83	0.99	0.83	0.82
RF	0.83	0.80	0.83	0.83	0.66	0.66
MVS	0.83	0.82	0.83	0.85	0.66	0.68
Ensamble	0.89	0.88	0.88	0.89	0.58	0.77

Cuadro 11. Prueba Jackknife para evaluar la contribución de las variables a los modelos en *S. panamense*.

Modelo	Bio1	Bio2	Bio3	Bio12	Bio13
MLG	93.43	0.04	2.53	3.91	0.09
MAG	94.34	0.13	0.09	4.36	1.09
MPG	85.76	5.17	7.99	0.72	0.36
MARS	61.73	4.25	15.57	9.62	8.83
AAC	20	20	20	20	20
MAXENT	11.95	21.67	22.14	22.14	22.11
RNA	73.07	0	26.93	0	0
RF	25.7	22.12	9.73	6.3	36.15
MSV	66.02	5.54	5.21	21.71	1.52
Ensamble	55.15	12.08	11.53	10.03	11.21

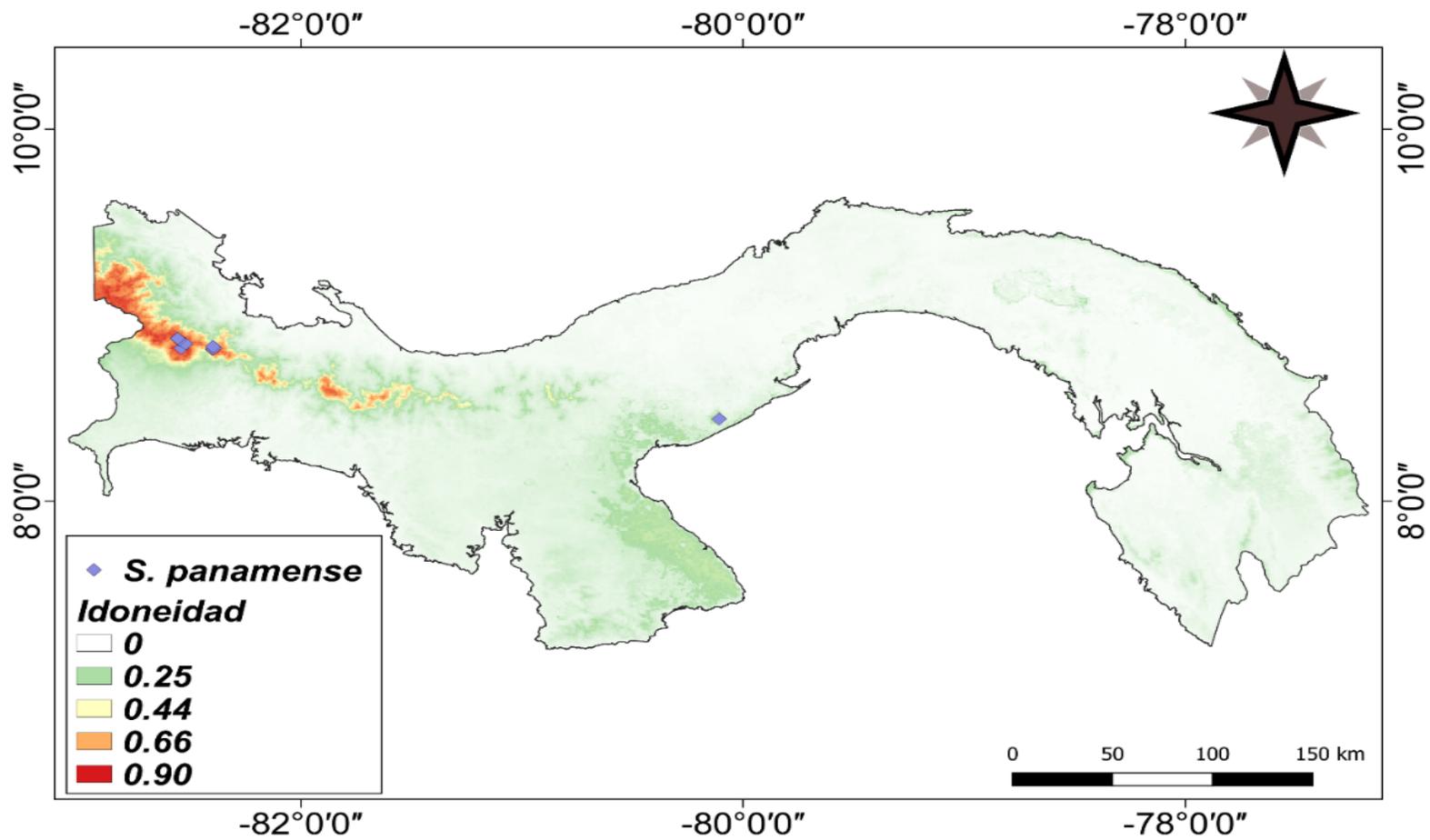


Figura 29. Ensamble de los nueve modelos de distribución de especies para *S. panamense* en Panamá.

S. pittieri

Los nueve modelos para *S. pittieri* fueron adecuados en términos de AUC, kappa y TSS. El mejor modelo en relación a esto tres estadísticos fue MARS (Cuadro 12). En la prueba Jackknife la variable con mayor contribución fue bio1, aunque la variación está distribuida en forma equitativa (Cuadro 13). Las regiones de alta idoneidad (> 0.65) se encuentran en Panamá (Bocas del Río y Chiriquí) y en gran parte del territorio central de Costa Rica (Limón, Cartago, Puntarenas, San José, Heredia y Alajuela) (Figura 30).

Cuadro 12. Parámetros estadísticos para los nueve modelos de distribución y ensamble para *S. pittieri*.

Modelo	AUC	PPC	Sensibilidad	Especificidad	Kappa	TSS
MLG	0.84	0.84	0.84	0.84	0.56	0.68
MAG	0.81	0.81	0.82	0.81	0.50	0.63
MPG	0.80	0.80	0.80	0.79	0.60	0.59
MARS	0.84	0.84	0.84	0.84	0.56	0.68
AAC	0.81	0.81	0.81	0.81	0.62	0.62
MAXENT	0.83	0.83	0.83	0.83	0.14	0.66
RNA	0.79	0.79	0.79	0.79	0.59	0.58
RF	0.80	0.80	0.80	0.80	0.60	0.60
SVM	0.82	0.82	0.82	0.82	0.65	0.64
Ensamble	0.83	0.83	0.84	0.84	0.57	0.68

Cuadro 13. Prueba Jackknife para evaluar la contribución de las variables a los modelos de distribución de *S. pittieri*.

Modelo	Bio1	Bio2	Bio3	Bio4	Bio12	Bio13	Bio14	Bio18	Bio19	Altitud
MLG	0.86	0.25	7.01	8.31	15.96	7.01	0.41	17.5	37.9	4.78
MAG	20.69	0.1	5.3	6.99	3.33	0.51	4.45	21.85	26.61	10.16
MPG	27.17	50.57	0.93	1.35	3.31	1.09	10.78	1.75	0.82	2.22
MARS	16.35	7.8	21.57	0	7.06	5.37	17.12	13.37	6.19	5.16
AAC	19.53	3.83	7.76	0	13.15	16.18	13.38	2.73	23.44	0
MAXENT	10.03	10.05	10.06	9.99	10.02	10.06	9.87	10.04	9.9	9.98
RNA	11.36	13.97	4.48	8.72	12.19	11.36	9.7	6.8	8.57	12.84
RF	8.37	12.29	6.68	5.82	11.38	9.68	10.9	11.4	15.79	7.68
MSV	2.46	9.25	10.65	10.83	12.65	4.83	23.88	3.07	19.74	2.63
Ensamble	13.01	12.37	9.32	8.69	9.95	5.71	10.63	10.67	12.59	7.06

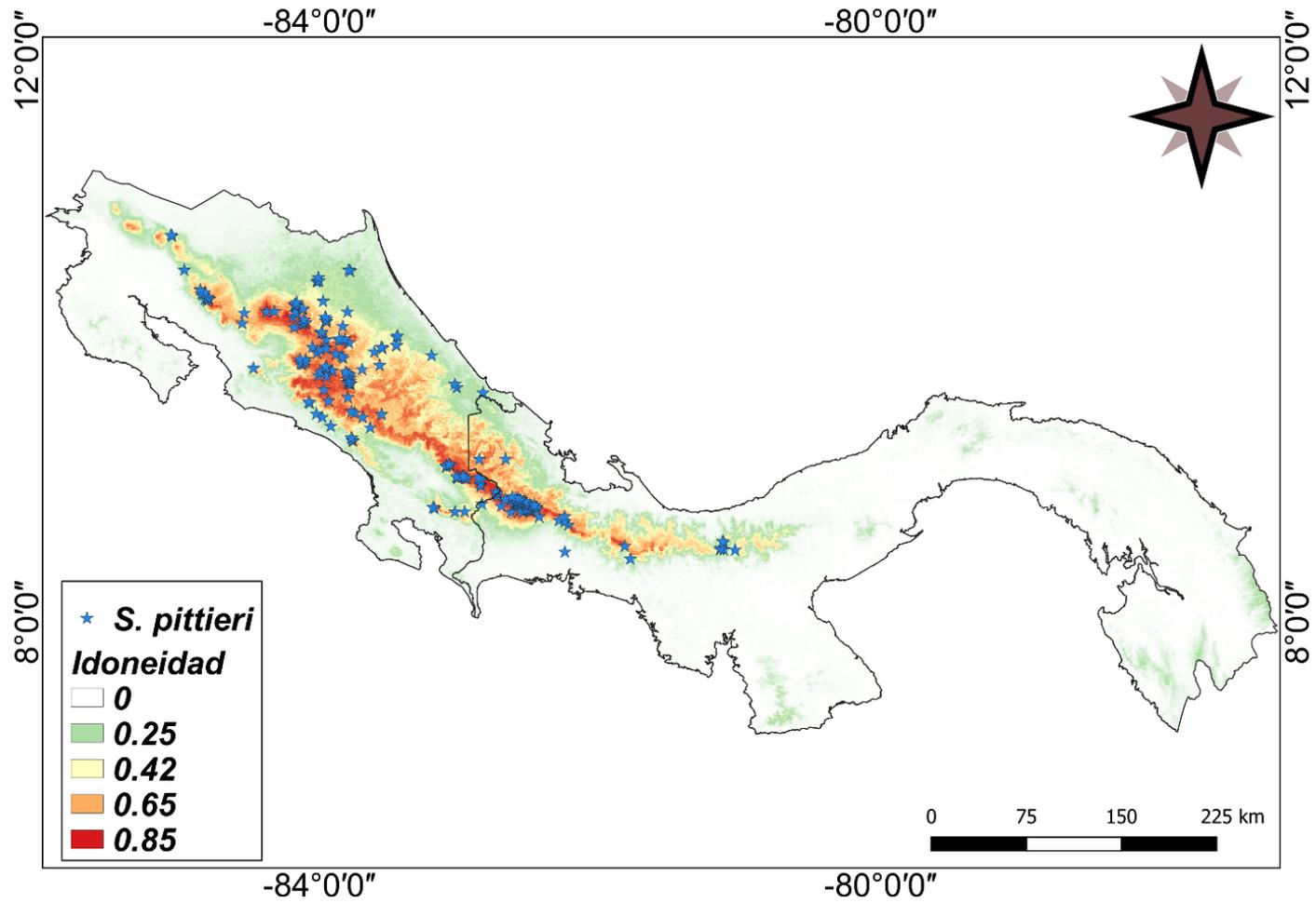


Figura 30. Ensamble de los nueve modelos de distribución de especies para *S. pittieri* en Costa Rica-Panamá

S. talamancense

Se obtuvieron valores de fiabilidad y rendimiento pobres para los modelos MAG, MARS y Maxent en relación con los estadísticos AUC y kappa (Cuadro 14). La variable bio4 (estacionalidad de la temperatura) resultó ser la de mayor contribución para explicar los modelos de distribución (Cuadro 15). Las zonas con alta idoneidad se encuentran en Costa Rica, en las provincias de Limón, Cartago y San José (Figura 31).

Cuadro 14. Parámetros estadísticos para los nueve modelos de distribución y ensamble para *S. talamancense*.

Modelo	AUC	PPC	Sensibilidad	Especificidad	Kappa	TSS
MLG	0.86	0.84	0.87	0.84	0.15	0.71
MAG	0.62	0.62	0.62	0.62	0.02	0.24
MPG	0.87	0.87	0.87	0.87	0.75	0.74
MARS	0.59	0.68	0.5	0.69	0.03	0.19
ACC	0.62	0.62	0.5	0.75	0.25	0.25
MAXENT	0.79	0.83	0.75	0.83	0.009	0.58
RNA	0.81	0.81	0.75	0.87	0.62	0.62
RF	0.75	0.75	0.75	0.75	0.5	0.5
MVS	0.75	0.75	0.75	0.75	0.5	0.5
Ensamble	0.84	0.84	0.81	0.85	0.45	0.66

Cuadro 15. Prueba Jacknife para evaluar la contribución de las variables a los modelos de distribución de *S. talamancense*.

Modelo	Bio1	Bio3	Bio4	Bio12	Bio13	Bio14	Bio18	Elevación
MLG	26.46	1.86	10.35	15.39	15.83	5.21	5.19	19.72
MAG	13.75	0.02	76.13	0.28	2.02	2.9	3.64	1.27
MPG	2.92	3.48	87.23	1.03	1.15	0.48	2.79	0.9
ACC	0	100	0	0	0	0	0	0
MAXENT	12.65	12.48	12.57	12.66	12.47	11.85	12.66	12.66
RNA	5.38	17.63	5.38	17.63	5.38	14.11	5.38	29.11
RF	8.92	13.38	50.35	2.74	4.47	3.51	8.17	8.46
MVS	6.89	9.45	13.84	2.52	44.42	2.66	15.15	5.1
Ensamble	11.29	15.29	24.82	7.72	10.23	9.98	10.28	10.39

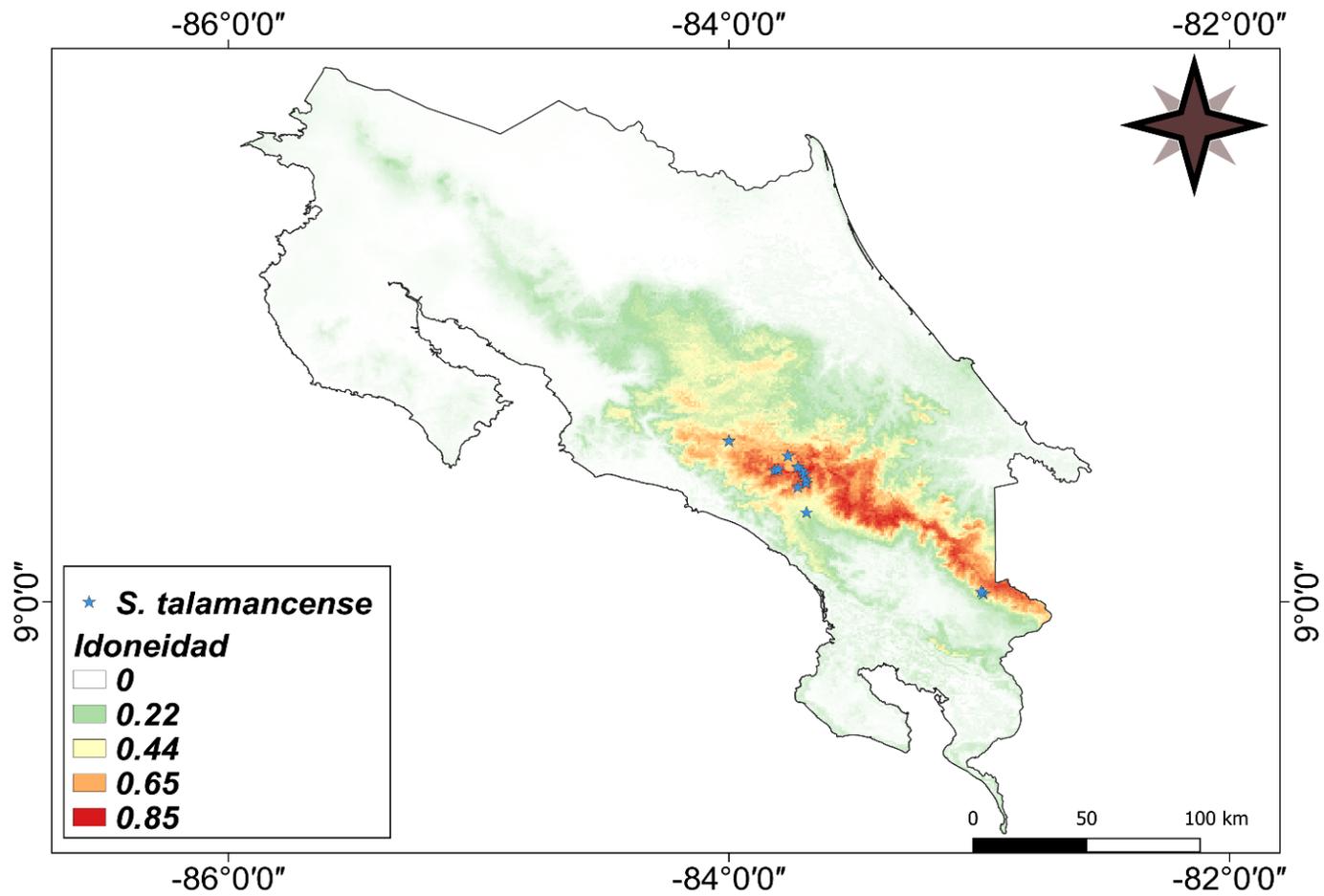


Figura 31. Ensamble de los nueve modelos de distribución de especies para *S. talamancense* en Costa Rica.

S. venosum

El modelo de bosque aleatorio obtuvo los valores más altos de AUC (0.90) y TSS (0.90). En general todos los modelos mostraron buenos rendimientos en términos de AUC y TSS (Cuadro 16). Respecto a la prueba Jackknife, bio14 obtuvo la mayor contribución para explicar dichos modelos (Cuadro 17). Las zonas de alta idoneidad se presentaron en Costa Rica (Limón y Cartago) y Panamá (Bocas del Toro) (Figura 32).

Cuadro 16. Parámetros estadísticos para los nueve modelos de distribución y ensamble para *S. venosum*.

Modelo	AUC	PPC	Sensibilidad	Especificidad	Kappa	TSS
MLG	0.79	0.78	0.80	0.78	0.08	0.58
MAG	0.81	0.83	0.80	0.83	0.11	0.66
MPG	0.70	0.70	0.70	0.70	0.40	0.4
MARS	0.76	0.82	0.70	0.83	0.09	0.53
CTA	0.65	0.65	0.70	0.60	0.30	0.3
MAXENT	0.78	0.76	0.80	0.76	0.008	0.56
RNA	0.75	0.75	1.00	0.50	0.50	0.50
RF	0.90	0.90	0.90	0.90	0.80	0.80
SVM	0.80	0.80	0.80	0.80	0.60	0.60
Ensamble	0.78	0.79	0.77	0.79	0.36	0.56

Cuadro 17. Prueba Jackknife para evaluar la contribución de las variables a los modelos de distribución de *S. venosum*

Modelo	Bio1	Bio3	Bio12	Bio13	Bio14	Bio15	Bio18	Bio19
MLG	0	7.32	17.64	9.28	56.27	8.27	0.39	0.83
MAG	3.13	2.29	6.66	3.11	38.25	24.24	1.23	21.09
MPG	3.02	9.6	6.71	6.87	42.53	12.74	13.98	4.55
MARS	10.76	7.25	15.19	13.25	23.42	13.98	0	16.14
AAC	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
MAXENT	12.71	12.5	12.71	12.71	11.72	12.70	12.59	12.38
RNA	16.93	2.99	12.55	23.92	6.75	2.99	16.93	16.93
RF	15.16	6.83	4.15	3.08	25.57	12	19.27	13.94
MVS	15.79	7.61	0.58	18.61	17.17	3.68	2.32	34.23
Ensamble	10.63	8.12	11.05	6.80	26.96	11.86	12.88	11.7

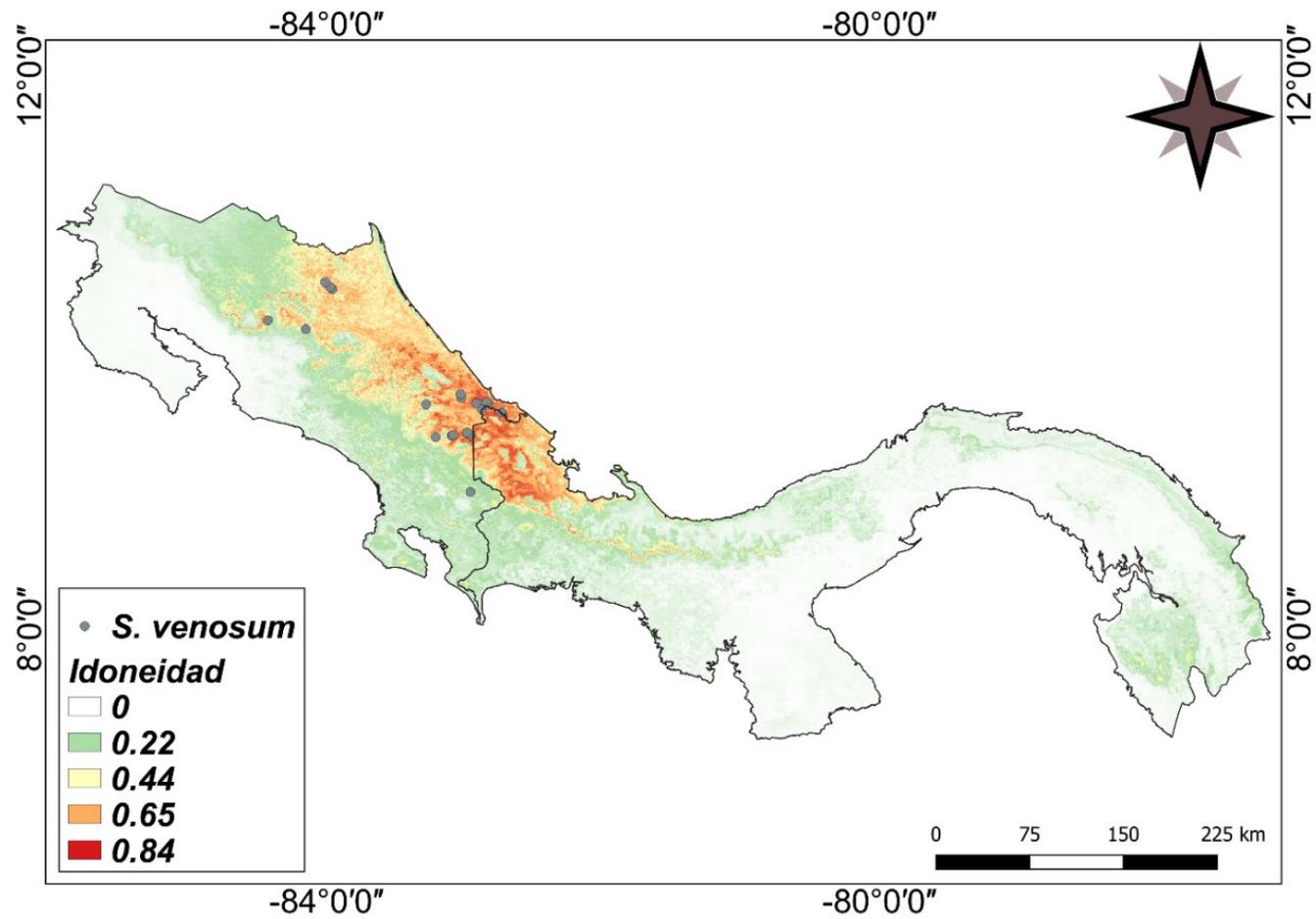


Figura 32. Ensamble de los nueve modelos de distribución de especies para *S. venosum* en Panamá-Costa Rica.

S. villosum

Los modelos MPG, AAC y Maxent mostraron valores > 0.9 para AUC. El modelo RNA mostró rendimientos pobres en términos de AUC (0.66) y TSS (0.33) (Cuadro 18). Respecto a la prueba Jackknife, las variables relacionadas a la temperatura (bio1, bio2, bio3 y bio7) fueron las más importantes para explicar la distribución de *S. villosum* (Cuadro 19). Las zonas de alta idoneidad se presentaron en el eje Central de Costa Rica (Figura 33).

Cuadro 18. Parámetros estadísticos para los nueve modelos de distribución y ensamble para *S. villosum*

Modelo	AUC	PPC	Sensibilidad	Especificidad	Kappa	TSS
MLG	0.83	0.83	0.83	0.83	0.07	0.66
MAG	0.80	0.78	0.83	0.78	0.08	0.61
MPG	0.91	0.91	0.83	0.99	0.82	0.82
MARS	0.80	0.76	0.86	0.76	0.06	0.62
AAC	0.91	0.91	0.99	0.83	0.83	0.82
MAXENT	0.94	0.89	0.99	0.89	0.01	0.88
RNA	0.66	0.66	0.66	0.66	0.33	0.33
RF	0.83	0.83	0.83	0.83	0.66	0.66
MVS	0.83	0.83	0.83	0.83	0.66	0.66
Ensamble	0.84	0.82	0.81	0.84	0.40	0.65

Cuadro 19. Prueba Jackknife para evaluar la contribución de las variables a los modelos de distribución de *S. villosum*.

Modelo	Bio1	Bio2	Bio3	Bio7	Bio12	Bio13	Bio14	Bio18	Elevación
MLG	13.95	6.52	5.17	5.63	16.8	2.5	28.2	0	21.22
MAG	11.17	7.78	17.82	1.98	1.83	10.29	15.35	27.82	5.98
MPG	23.33	2.25	4.42	1.85	8.84	30.24	20.81	6.18	2.09
MARS	27.89	0	35.15	6.12	0	0	0	14.95	15.9
AAC	11.11	11.11	11.11	11.11	11.11	11.11	11.11	11.11	11.11
MAXENT	11.22	11.2	11.22	11.2	11.12	11.22	10.3	11.22	11.19
RNA	16.6	0	16.68	0	16.67	16.68	16.74	16.63	0
RF	11.94	36.87	1.78	12.84	1.65	2.86	6.88	1.21	23.94
MVS	1.52	6.73	12.83	1.25	4.92	15.13	52.35	3.61	1.65
Ensamble	14.47	9.24	13.71	5.75	8.06	8.31	17.09	10.01	13.37

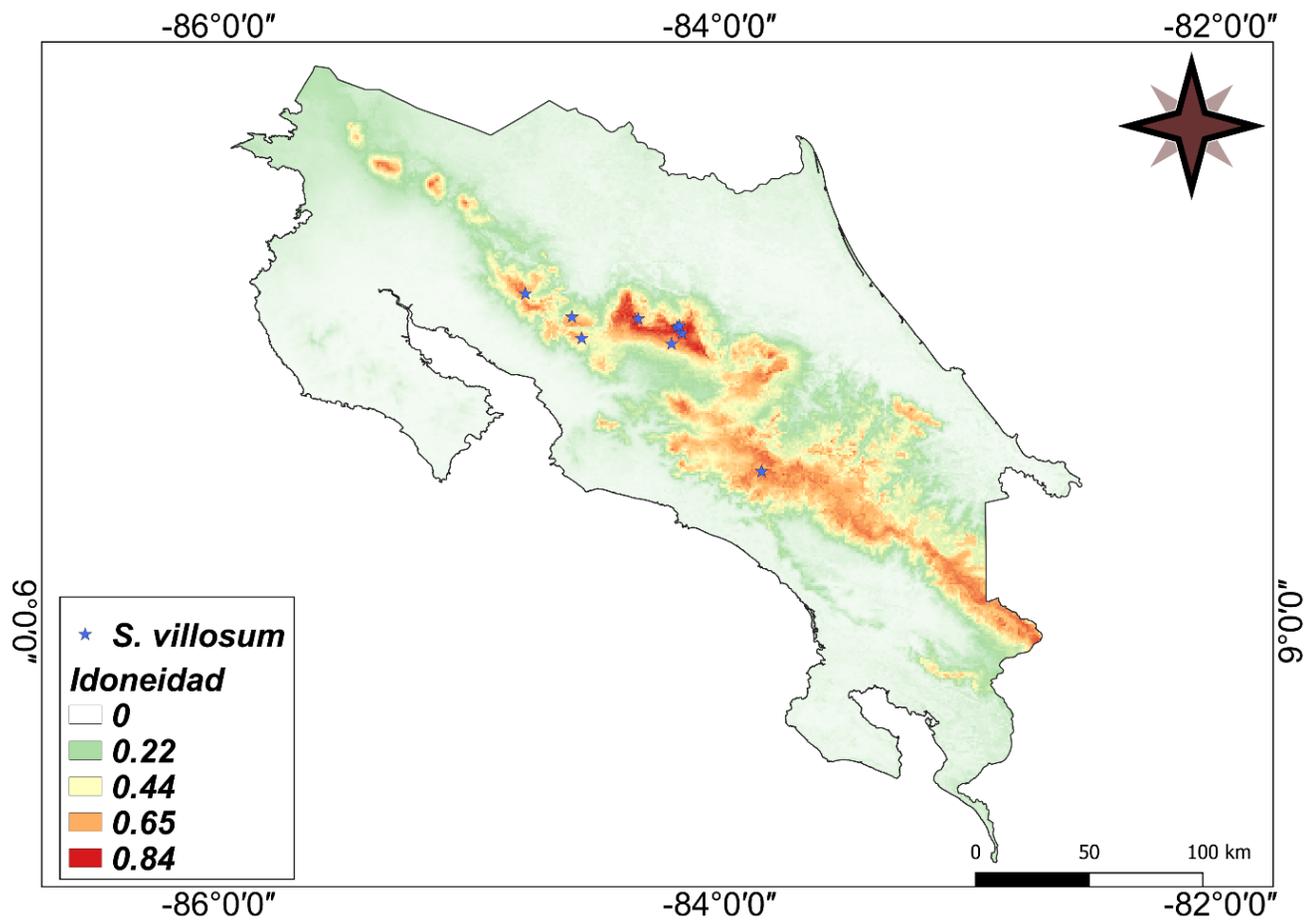


Figura 33. Ensamble de los nueve modelos de distribución de especies de *S. villosum* en Costa Rica.

6.4. CONCLUSIONES

Las especies centroamericanas de *Sechium* se distribuyen en Costa Rica y Panamá donde prosperan en condiciones de temperatura muy similares, más no así para las variables bioclimáticas de precipitación. Estas especies pueden tolerar y/o resistir altas condiciones de humedad y bajas temperaturas, características que pudieran ser introducidas para mejorar las especies cultivadas como *S. edule*.

6.5. LITERATURA CITADA

- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1), 1-12. <https://doi.org/10.1038/sdata.2018.214>
- Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). cIValid: An R Package for Cluster Validation. *Journal of Statistical Software*, 25(4), 1-22.
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, 61(6), 1-36.
- de Mendiburu, F. (2020). agricolae: Statistical Procedures for Agricultural Research (R package version 1.3-3) [Computer software].
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315. <https://doi.org/10.1002/joc.5086>
- Fischer, G., Nachtergaele, F., Prieler, S., van Velthuisen, H. T., Verelst, D., & Wiberg, D. (2008). Global agro-ecological zones assessment for agriculture. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

- Kassambara, A., & Mundt, F. (2020). factoextra: extract and visualize the results of multivariate data analyses (R package version 1.0.7) [Computer software]. <https://CRAN.R-project.org/package=factoextra>
- Lira, R., & Chiang, F. (1992). Two new combinations in *Sechium* (Cucurbitaceae) from Central America, and a new species from Oaxaca, Mexico. *Novon*, 2(3), 227-231. JSTOR. <https://doi.org/10.2307/3391556>
- Lira, S. R. (1995). Estudios taxonómicos en el género *Sechium* P. Br. *Cucurbitaceae* [Tesis Doctoral]. Universidad Nacional Autónoma de México.
- Mateo, R. G., Felicísimo, A. M., & Muñoz, J. (2011). Species distributions models: A synthetic revision. *Revista Chilena de Historia Natural*, 84, 217-240.
- Monge, P. J. E., & Loría, M. (2017). Caracterización de frutos de cinco genotipos de tacaco [*Sechium tacaco* (Pittier) C. Jeffrey] en Costa Rica | Revista Tecnología en Marcha. *Tecnología en Marcha*, 30(3), 71-84. <https://doi.org/10.18845/tm.v30i3.3274>
- Newstrom, L. (1990). Origin and evolution of chayote, *Sechium edule*. In D. M. Bates, R. W. Robinson, & C. Jeffrey (Eds.), *Biology and Utilization of the Cucurbitaceae* (1º, pp. 141-149). Cornell University Press.
- Newstrom, L. (1991). Evidence for the origin of chayote, *Sechium edule* (Cucurbitaceae). *Economic Botany*, 45(3), 410-428. <https://doi.org/10.1007/BF02887082>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>

- QGIS Development Team. (2020). QGIS Geographic Information System. Open-Source Geospatial Foundation Project (3.16.2) [Computer software]. <https://www.qgis.org/en/site/>
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing (1.3.1093) [Computer software]. <https://www.R-project.org/>
- Schmitt, S., Pouteau, R., Justeau, D., Boissieu, F. de, & Birnbaum, P. (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, 8(12), 1795-1803. <https://doi.org/10.1111/2041-210X.12841>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open-Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Yilan, L., & Rutong, Z. (2015). clustertend: Check the Clustering Tendency (R package version 1.4.) [Computer software]. <https://CRAN.R-project.org/package=clustertend>

Apéndice 1. Estadísticos descriptivos para las variables bioclimáticas de las especies centroamericanas de *Sechium*.

Variable	<i>S. panamense</i>	<i>S. pittieri</i>	<i>S. tacaco</i>	<i>S. talamancense</i>	<i>S. venosum</i>	<i>S. villosum</i>
Bio1						
Min	15.72	9.971	14.86	14.44	11.51	11.51
1 st Qu.	17.67	16.148	16.08	18.5	12.58	16
Media	18.78	19.058	17.9	20.16	17.44	18.97
3 ^{er} Qu.	19.32	21.456	18.66	23.66	20.38	21.8
Max	25.09	26.992	24.22	24.28	24.29	23.08
Bio2						
Min	8.817	8.117	8.45	8.408	8.508	8.692
1 st Qu.	8.975	9	9.867	8.958	8.929	8.708
Media	9.021	9.494	9.959	9.539	9.068	8.819
3 ^{er} Qu.	9.042	9.887	10.254	10.108	9.175	8.9
Max	9.283	12.4	10.367	10.883	9.75	9.083
Bio3						
Min	74.66	74.63	78.97	77.85	78.06	78.3
1 st Qu.	75	78.35	80.14	79.28	78.74	78.76
Media	76.65	79.51	80.43	80.23	79.51	79.22
3 ^{er} Qu.	78.01	81.16	80.9	81.34	79.94	79.62
Max	79.43	83.27	81.3	82.45	81.34	79.89
Bio4						
Min	66.62	44.03	60.39	58.83	60.73	62.89
1 st Qu.	73.19	57.34	62.31	61.1	65.25	65.03
Media	75.24	61.77	63.93	66.47	69.05	69.07
3 ^{er} Qu.	78.52	65.4	64.61	72.27	73.35	71.06
Max	80.51	80.45	69.23	74.93	78.53	78.53
Bio5						
Min	21.9	15.5	21	20	16.9	16.9
1 st Qu.	23.8	22.2	22.43	25.1	18.3	21.9
Media	24.9	25.17	24.32	26.34	23.25	24.69
3 ^{er} Qu.	25.7	28.1	25.15	29.4	26.15	27.6
Max	31.2	32.5	29.8	30	30	28.8
Bio6						
Min	10	4	8.9	8.7	5.6	5.6
1 st Qu.	11.9	10.2	10.1	12.5	6.7	10.3
Media	13.12	13.23	11.94	14.46	11.84	13.56
3 ^{er} Qu.	13.6	15.9	12.68	18.3	15.1	16.6
Max	19.6	22.3	19.1	19.1	19.1	17.9
Bio7						
Min	11.1	10.1	10.7	10.7	10.7	10.9
1 st Qu.	11.6	11.5	12.3	11.3	11.25	11
Media	11.78	11.93	12.38	11.88	11.41	11.13
3 ^{er} Qu.	12	12.5	12.7	12.6	11.6	11.3
Max	12.3	15.1	13	13.2	12.2	11.6
Bio8						
Min	15.78	10.27	14.6	14.62	11.87	11.88
1 st Qu.	17.62	16	16.05	18.12	12.93	15.73
Media	18.64	18.91	17.87	20.17	17.51	19.03
3 ^{er} Qu.	19.12	21.18	18.65	23.85	20.33	21.98
Max	24.78	26.42	24.68	24.27	24.27	23.27
Bio9						
Min	15.32	9.417	15.02	13.93	11.4	11.4
1 st Qu.	18.02	16	15.76	18.3	11.95	15.55
Media	19.1	19.07	17.92	20.34	17.39	19.23
3 ^{er} Qu.	19.8	21.667	18.98	24.22	20.73	22.32
Max	25.8	27.317	24.77	24.92	24.93	23.6
Bio10						
Min	16.5	10.53	15.52	15.13	12.27	12.27
1 st Qu.	18.6	16.79	16.75	19.27	13.3	16.75
Media	19.71	19.83	18.63	21	18.23	19.81
3 ^{er} Qu.	20.37	22.3	19.4	24.65	21.16	22.7

Max	26.2	27.95	25.15	25.33	25.35	24.03
Bio11						
Min	14.77	9.133	14.03	13.5	10.43	10.43
1 st Qu.	16.75	15.475	15.27	17.73	11.58	15.07
Media	17.89	18.305	17.08	19.35	16.57	18.13
3 ^{er} Qu.	18.47	20.925	17.88	22.82	19.62	21.08
Max	24.18	26.417	23.4	23.42	23.42	22.3
Bio12						
Min	2349	1720	1945	2970	2235	2303
1 st Qu.	2649	2596	2262	3310	2596	2922
Media	2908	2973	2508	3378	3074	3157
3 ^{er} Qu.	3023	3286	2813	3442	3516	3506
Max	3954	4052	3103	3863	3660	3775
Bio13						
Min	388	305	317	349	380	380
1 st Qu.	416	406	359	399	425	422
Media	449.6	463.8	383.8	464.7	447.1	429
3 ^{er} Qu.	484	518	394.8	506	475	450
Max	510	623	521	606	573	457
Bio14						
Min	17	4	18	27	14	48
1 st Qu.	21	23	28	58	21	94
Media	58.11	56.55	48.06	92.46	53.4	80.67
3 ^{er} Qu.	82	85.5	52.25	109	81.5	108
Max	156	190	144	190	122	127
Bio15						
Min	35.3	18.98	26	23.52	34.68	33.39
1 st Qu.	40.55	43.1	56.65	34.63	37.93	39.17
Media	56	56.1	56.91	44.91	52.25	46.16
3 ^{er} Qu.	62.09	69.23	63.83	64.1	65.01	57.16
Max	82.4	86.84	71.15	68.88	76.53	63.79
Bio16						
Min	1030	779	805	970	941	941
1 st Qu.	1109	1065	953.5	1109	1076	1146
Media	1168	1178	1016.9	1230	1181	1144
3 ^{er} Qu.	1223	1274	1066.5	1334	1303	1190
Max	1320	1601	1271	1513	1436	1217
Bio17						
Min	72	42	89	139	77	140
1 st Qu.	154	120	144.5	247	122.5	186
Media	239.7	235.5	200	360.3	285.5	307.8
3 ^{er} Qu.	348	379.5	205	428	411	363
Max	498	582	492	587	475	471
Bio18						
Min	315	248	279	425	371	441
1 st Qu.	349	506	411.5	516	469.5	488
Media	433.4	598.5	501.6	642.4	627.4	570.4
3 ^{er} Qu.	491	681	614.8	727	783	617
Max	726	1106	652	1104	929	819
Bio19						
Min	250	113	296	492	162	276
1 st Qu.	472	394	358.8	717	297.5	811
Media	618	609.8	487.5	777.2	589.2	781.9
3 ^{er} Qu.	802	838	614.2	885	845	932
Max	1020	1183	870	957	934	1032
Elevación						
Min	275	16	412	364	360	580
1 st Qu.	1337	937	1468	504	1096	847
Media	1437	1353	1632	1194	1650	1370
3 ^{er} Qu.	1639	1846	2001	1509	2544	1938
Max	1991	2894	2223	2194	2731	2731

CAPÍTULO VII. DESCRIPTORES EDAFOCLIMÁTICOS Y MODELOS DE DISTRIBUCIÓN DEL CLADO MEXICANO DE *Sechium* P. BROWN

EDAPHOCLIMATIC DESCRIPTORS AND DISTRIBUTION MODELS OF THE MEXICAN CLADO OF *Sechium* P. BROWN

RESUMEN

Los sistemas de información geográfica analizan la ubicación espacial de puntos georreferenciados para obtener información descriptiva de los mismos y aplicar algoritmos que permitan ver la distribución de las especies. Mediante el uso de 280 puntos de ocurrencia y sistemas de información geográfica, el objetivo de este estudio fue determinar las condiciones edafoclimáticas bajo las cuales prosperan las especies del clado mexicano de *Sechium*, esto con la finalidad de contribuir y enriquecer las bases de datos de estas especies, en aras de conservarlas y aprovecharlas para mejorar las características de especies cultivadas. Se emplearon 22 variables bioclimáticas con resolución espacial de ~1 Km² y se extrajeron los valores ráster para los respectivos análisis estadísticos y la evaluación del modelo Maxent. En el análisis clúster de k-medioides se obtuvieron cuatro grupos donde todos los puntos de ocurrencia aparecieron mezclados unos con otros. No se lograron diferenciar por especies. En el ACP los dos primeros componentes contribuyeron con el 72.3 % de la variación total, el primer componente aportó el 42.5 % de la variación y estuvo integrado por las variables bio1, bio2, bio12 y bio18; el segundo componente principal aportó el 29.8 % de la variación y estuvo conformado por las variables bio3 y bio4. Los climas predominantes fueron Cwb (29.3 %), Aw (16.1 %) y Am (15 %). Los suelos predominantes fueron leptosol (29.3 %), phaeozem (15.4 %) y andosol (13.9 %). Las variables relacionadas a la temperatura son las responsables de la distribución de las especies mexicanas de *Sechium*.

Palabras clave: edafoclimáticas, k-medioides, puntos de ocurrencia, temperatura.

ABSTRACT

Geographic information systems analyze the spatial location of georeferenced points to obtain descriptive information about them and apply algorithms that allow the distribution of species to be seen. Through 280 occurrence records and geographic information systems, the objective of this study was to determine the edaphoclimatic conditions of these species under which they thrive, this in order to contribute and enrich the databases of these species, in order to conserve and take advantage of them to improve the characteristics of cultivated species. 22 bioclimatic variables with spatial resolution of $\sim 1 \text{ Km}^2$ were used and raster values were extracted for the respective statistical analyzes and the evaluation of the Maxent model. In the cluster analysis of k-medoids, four groups were obtained where all the points of occurrence appeared mixed with each other, they were not able to differentiate by species. In the PCA, the first two components contributed 72.3% of the total variation, the first component contributed 42.5% of the variation and was made up of the variables bio1, bio2, bio12 and bio18; the second main component contributed 29.8% of the variation and was made up of the bio3 and bio4 variables. The predominant climates were Cwb (29.3%), Aw (16.1%) and Am (15%). The predominant soils were leptosol (29.3%), phaeozem (15.4%) and andosol (13.9%). Variables related to temperature are responsible for the distribution of the Mexican species of *Sechium*.

Key words: edaphoclimatic, k-medoids, points of occurrence, temperature.

7.1. INTRODUCCIÓN

La distribución de una especie está en función de las condiciones ecológicas donde se desarrolla y se reproduce. Los modelos de distribución de especies (MDE) se basan en procedimientos estadísticos y sistemas de información geográfica. Usando datos de presencia-ausencia, estos modelos permiten la inferencia de sitios adecuados con características ambientales óptimas para la especie. Los distintos algoritmos para su cálculo se agrupan en técnicas descriptivas, discriminatorias y mixtas (Mateo *et al.*, 2011). Las técnicas descriptivas como Bioclim incluyen métodos envolventes para caracterizar los sitios que se encuentran dentro del espacio ambiental ocupado por la especie en cuestión. En este grupo también existen algoritmos basados en distancias (Domain y Mahalanobis) que evalúan sitios en términos de similitud ambiental con datos de presencias (Elith *et al.*, 2006). Las técnicas discriminantes como el modelo lineal generalizado (MLG) requieren datos de presencia-ausencia y tienen la capacidad de modelar relaciones ecológicas realistas. Dentro de estas técnicas se encuentran las que utilizan el aprendizaje automático, como la máxima entropía (Maxent), el bosque aleatorio y las máquinas de vectores de soporte (MVS). Estos son ampliamente utilizados debido a su capacidad para procesar grandes cantidades de información y construir grupos para su clasificación, o para generar regresiones numéricas para tareas de predicción (Gobeyn *et al.*, 2019). También existen modelos geográficos nulos como distancia geográfica, cascos convexos, círculos y presencia-ausencia que utilizan la ubicación geográfica de las presencias, pero no se basan en variables predictivas y no son de uso común ya que su desempeño y confiabilidad aún no se han evaluado (Hijmans y Elith, 2013). El modelo de Maxent se usa con más frecuencia debido a que su cálculo computacional es rápido y fácil (Phillips *et al.*, 2006; Zhang *et al.*, 2019), sin embargo, es necesario evaluar más modelos para obtener una visión más completa.

Los MDE tienen varias aplicaciones en la agricultura. En especies cultivadas como *Zea mays* y *Panicum virgatum*, utilizadas como biocombustibles en los

EUA, se han utilizado los modelos Maxent y MVS para estimar la idoneidad de las regiones geográficas en las que estas especies pueden prosperar (Evans *et al.*, 2010). La información de los MDE también puede proporcionar referencias filogeográficas e inferir eventos históricos que intentan explicar el origen y la estructura genética de las poblaciones (Carstens y Richards, 2007), aunque esto también debe corroborarse con datos moleculares. Los MDE se utilizan para monitorear la disponibilidad de insectos polinizadores, así como para proponer estrategias para hacer frente a cualquier cambio adverso en el medio ambiente que los afecte (Polce *et al.*, 2013). También es posible predecir la idoneidad de especies nocivas (malezas e insectos), pronosticar su ubicación, distribución y formular estrategias para mitigar el daño socioeconómico en áreas vulnerables a la invasión (Lantschner *et al.*, 2018; Manzoor *et al.*, 2018). Otras aplicaciones para destacar son la evaluación de riqueza de especies, endemismo, hábitats amenazados, protección y conservación de especies amenazadas, patrones de diversidad, estudios de conservación, respuesta al cambio climático, reintroducción de especies amenazadas y el estudio de efectos antrópicos en la distribución de las especies. Aunque muchas de estas aplicaciones se utilizan en especies animales, también pueden extrapolarse a especies vegetales (Mateo *et al.*, 2011).

los valores altos de AUC son indicativos de especies con rangos de distribución restringidos (Phillips, 2009), ya que solo se han encontrado poblaciones silvestres de *S. edule* en Oaxaca, Guerrero y Veracruz (Lira *et al.*, 1999). La mayor idoneidad para el potencial de distribución de *S. edule* tuvo un rango de probabilidad de 50-60 %. En Lira *et al.* (2018), se omitieron ciertos análisis previos, como la eliminación de variables altamente correlacionadas que inciden en la complejidad del modelo (Phillips *et al.*, 2006; Zhang *et al.*, 2019). Además, no se calcularon estadísticos como la TSS para evaluar el desempeño del modelo.

El objetivo de este estudio fue determinar las condiciones edafoclimáticas y la distribución potencial de las especies mexicanas de *Sechium*, esto con la

finalidad de contribuir y enriquecer las bases de datos relacionadas a estas especies, en aras de conservarlas y aprovecharlas en un futuro para mejorar las características de las especies cultivadas.

7.2. MATERIALES Y MÉTODOS

7.2.1. Puntos de ocurrencia y variables bioclimáticas

Se obtuvieron 280 puntos de ocurrencia para las seis especies mexicanas de *Sechium* de GBIF (<https://www.gbif.org/>) (162 de *S. edule*, 47 de *S. compositum*, 16 de *S. chinantlense*, 23 de *S. hintonii* y 32 de *S. mexicanum*). Se emplearon las 19 variables bioclimáticas de WorldClim 2.0 (Fick y Hijmans, 2017); se incorporaron también modelos ráster para elevación (Fick y Hijmans, 2017), tipos de clima de Köppen-Geiger (Beck *et al.*, 2018) y tipos de suelo según la clasificación de World Soil Database version 1.2 (Fischer *et al.*, 2008). Las 22 capas empleadas tienen una resolución espacial de ~ 1 km². Los valores ráster se obtuvieron con Point Sampling Tool de QGIS versión 3.16.2 (2020).

7.2.2. Análisis descriptivo

Todos los paquetes por mencionar se ejecutaron en Rstudio (R Core Team, 2020). De las 19 variables bioclimáticas de WorldClim y la variable elevación se obtuvieron los valores primer y tercer cuartiles, mínimos, máximos y media; los valores mínimos y máximos representaron los límites a partir de los cuales cada especie puede tolerar según la variable bioclimática en cuestión (Apéndice 2). Las variables categóricas de tipos de clima y suelo se graficaron con tidyverse (Wickham *et al.*, 2019) para observar la proporción en función de las especies.

7.2.3. Análisis de agrupamiento

Para verificar si existe tendencia al agrupamiento de los valores ráster se computó el estadístico de Hopkins (H) con clustertend (Yilan y Rutong, 2015). El mejor algoritmo de agrupación (jerárquico o de partición) se determinó con cIValid

(Brock *et al.*, 2008), de esta forma se escogió el método de partición k-medoides para determinar similitudes edafo-climáticas en los puntos de ocurrencia de las especies con la librería factextra (Kassambara y Mundt, 2020). K-medoides necesita especificar el número óptimo de grupos, el cual se determinó con NbClust (Charrad *et al.*, 2014), donde previamente los valores ráster deben ser normalizados. Adicionalmente se realizó un análisis de agrupamiento para todas las especies de *Sechium* con todas variables ambientales donde se obtuvo la mediana por variable y por especie.

7.2.4. Modelo Maxent

En las variables cuantitativas se aplicó un análisis de correlación de Pearson y se eliminaron aquellas con valores > 0.8 para evitar la afectación a los modelos (Phillips *et al.*, 2006). En el capítulo IV se mostró que Maxent muestra muy buenos rendimientos en comparación con los demás algoritmos; por otra parte, el mapa de los modelos ensamblados se asemejaba al modelo Maxent. Sin embargo, Maxent fue optimizado con el paquete ENMeval (Muscarella *et al.*, 2014) con la siguiente configuración: se desactivaron las funciones de umbral y bisagra para evitar el sobreajuste de las curvas de la respuesta, también se desactivaron las funciones cuadráticas y de producto y el multiplicador de regularización fue de 1.50 con una función lineal. El rendimiento del modelo se cuantificó con la AUC, Kappa y TSS, tal y como se describe a detalle en el capítulo IV. La prueba de Jackknife fue ejecutada con el paquete SSDM (Schmitt *et al.*, 2017). Los formatos ráster producto de los siete modelos y el ensamble se exportaron con el paquete raster (Hijmans, 2020) a QGIS Development Team versión 3.16.2 (2020).

7.3. RESULTADOS Y DISCUSIÓN

7.3.1. Tipos de clima y suelo

El tipo de suelo predominante en las especies mexicanas de *Sechium* fue leptosol con el 29.3 %, seguido por phaeozem (15.4 %) y andosol (13.9 %). *Sechium edule* abarcó la mayor diversidad de tipos de suelo en sus colectas, con nueve

de los 11 tipos de suelos presentes. Por ejemplo, registró 31.5 % para leptosol, 17.3 % andosol y luvisol 13.6 %. El resto de los tipos de suelos representó poco menos del 10 % (Figura 34A). Para los tipos de clima, *S. edule* se desarrolla en todos los climas registrados; algunas especies como *S. hintonii* solo se desarrollan en climas tipo A (Figura 34B). Los componentes calcáreos de los leptosoles pueden inmovilizar los nutrientes minerales en el suelo y, en el caso del chayote, pueden reducir la productividad y calidad de los frutos (Cadena-Iñiguez, 2005). Sin embargo, algunos leptosoles pueden tener pequeñas capas ricas en materia orgánica, que es un requisito edáfico para las plantaciones de chayote (Lira, 1996). Phaozem, presente en el 50 % de las subprovincias fisiográficas, son suelos considerados ideales para la agricultura por su alto contenido en materia orgánica. Los suelos ideales para la producción de chayote deben estar bien drenados, ricos en materia orgánica y tener un pH de 4.5-6.5 (Lira, 1996), sin embargo, las diferentes variantes fenotípicas del chayote permiten un alto nivel de adaptabilidad a diferentes tipos de suelo (Lira *et al.*, 2009).

Tradicionalmente el chayote se cultiva a una altitud que oscila los 300-2000 m. Requiere una alta humedad relativa de 80-85 % con una precipitación media anual de 1500-2000 mm distribuidos uniformemente en el año. Temperaturas por debajo de 13 °C reducen la producción ya que afectan el crecimiento de los frutos y temperaturas superiores a los 28 °C estimulan la caída de flores (Lira, 1996). Sin embargo, en el Apéndice 2, algunas accesiones de chayote se desarrollan en temperaturas mínimas de 10 °C y máximas de 27 °C; algunas de ellas también crecen con condiciones mínimas y máximas de precipitación (528 mm y 4650 mm, respectivamente).

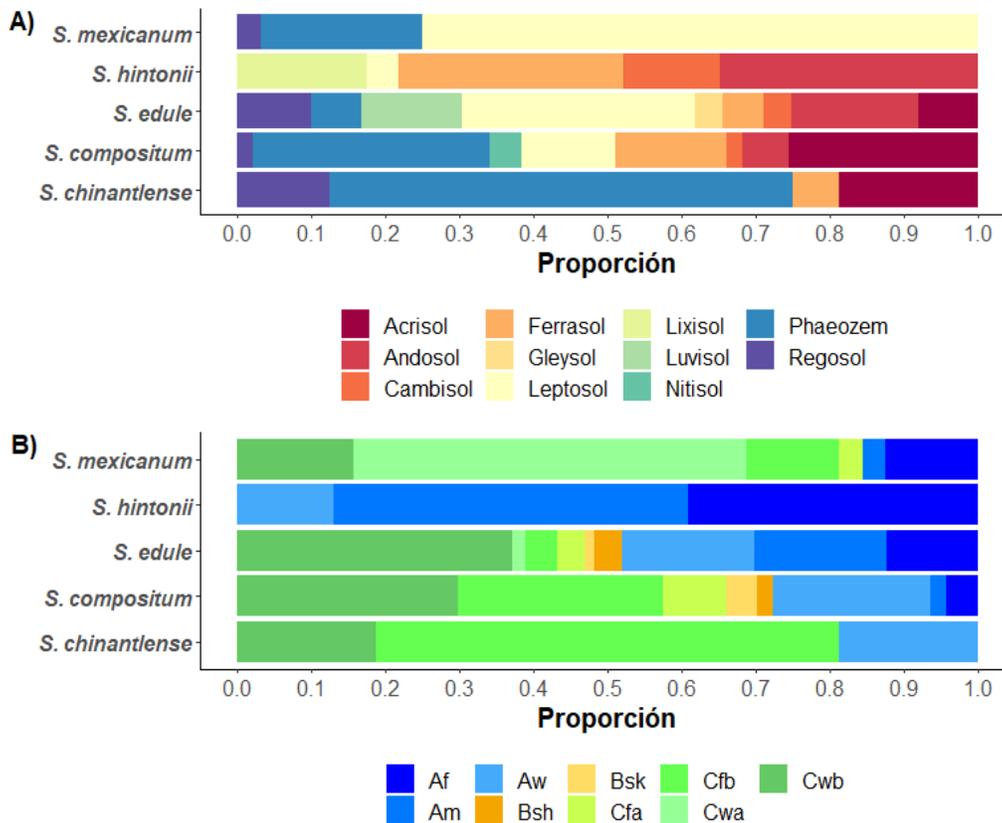


Figura 34. Proporción de **A)** tipos de suelo y **B)** tipos de clima en relación con los puntos de ocurrencia para cada especie.

7.3.2. Análisis ACP y clúster k-medioides

En el ACP, los dos primeros componentes explicaron el 72.3 % de variación total (CP1= 42.5 % y CP2 = 29.8 %). El CP1 fue representado por las variables bio1, bio2, bio12, bio14 y bio18; mientras el CP2 fue explicado por bio3 y bio4 (Figura 35 A-B). Con los 280 puntos de ocurrencia se formaron cuatro clusters (k = 4), que fue el número óptimo calculado por NbClust. El índice H fue 0.12 e indica alta tendencia al agrupamiento. El paquete cIValid infirió el método k-medioides como el mejor método para agrupar los datos mediante los estadísticos de conectividad, Dunn y la Silueta. El 18.5 % de los puntos de ocurrencia se agruparon en el clúster 1, el 35 % en el clúster 2, el 25.5 % en el clúster 3 y 21 % en el clúster 4. La Figura 35C muestra la proporción de las colectas por especies que integraron cada uno de los clusters.

La Figura 35D indica un promedio para el coeficiente de la silueta de 0.34. En el clúster 4 se muestra que todos sus elementos están perfectamente agrupados. En los grupos restantes se observa que algunos individuos tienen cierta afinidad por otros grupos, sin embargo, solo representan un pequeño porcentaje que no afectó dicho análisis (Figuras 35A, 35D).

La Figura 36 muestra una variedad de gráficos que permiten hacer inferencias sobre las principales variables bioclimáticas (bio1, bio4, bio12, bio15 y elevación). Por ejemplo, en el gráfico de dispersión, para las variables bio1 y elevación, se muestra que a una mayor altitud existirá un descenso de temperatura. Estas mismas variables también muestran una correlación altamente significativa ($p < 0.001$). Por otra parte, las cajas de bigotes muestran que existen algunos datos atípicos para las variables evaluadas, esto puede explicar que el paquete `clValid` haya encontrado al algoritmo k-medioides como la mejor opción, ya que trata de minimizar el efecto que tiene los datos atípicos.

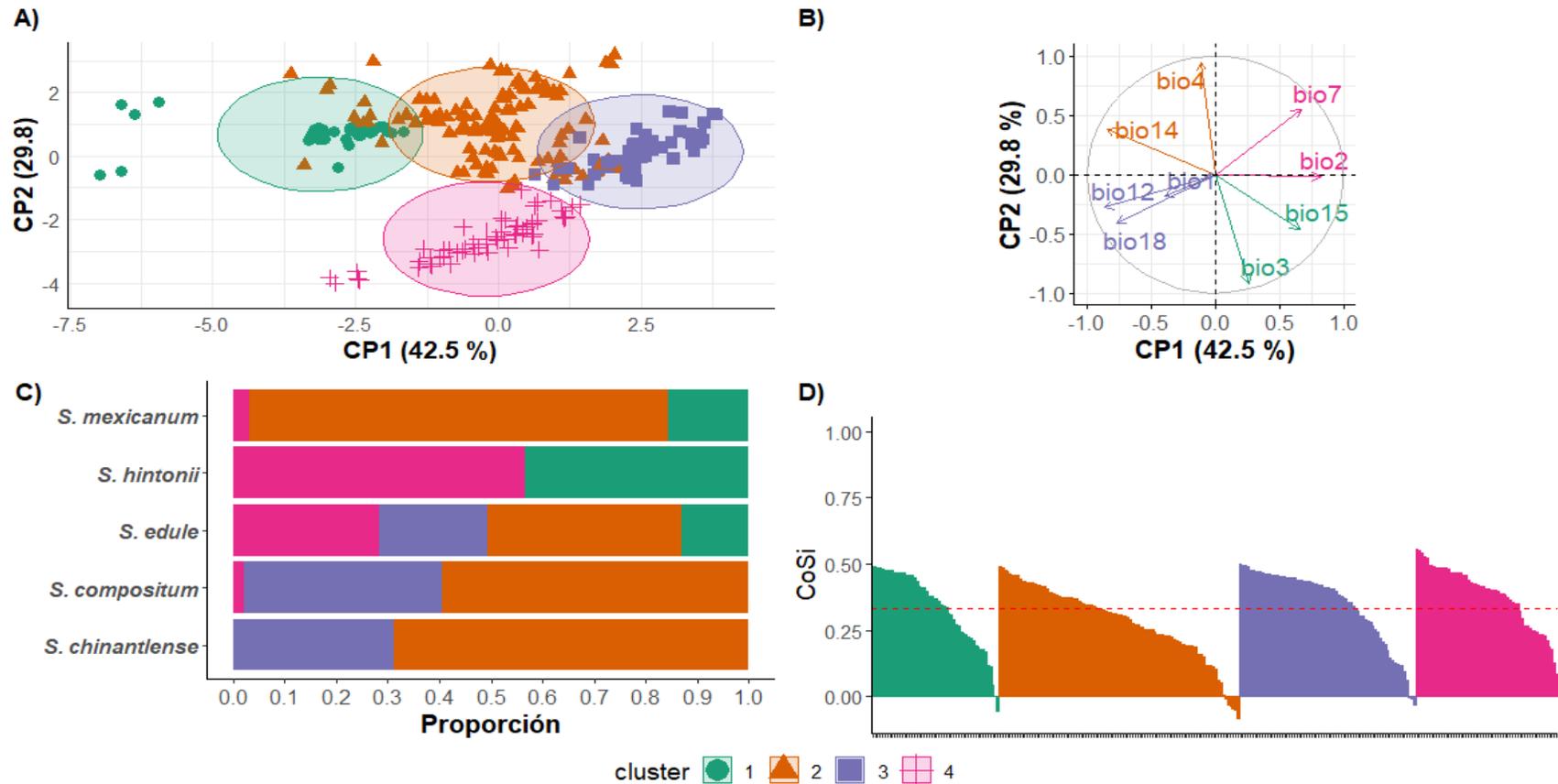


Figura 35. A) Agrupación k-medioides, **B)** principales variables bioclimáticas que contribuyen en la distribución de las especies mexicanas de *Sechium*, **C)** proporción de los puntos de ocurrencia de las especies en cada clúster y **D)** consistencia del agrupamiento mediante el coeficiente de la Silueta

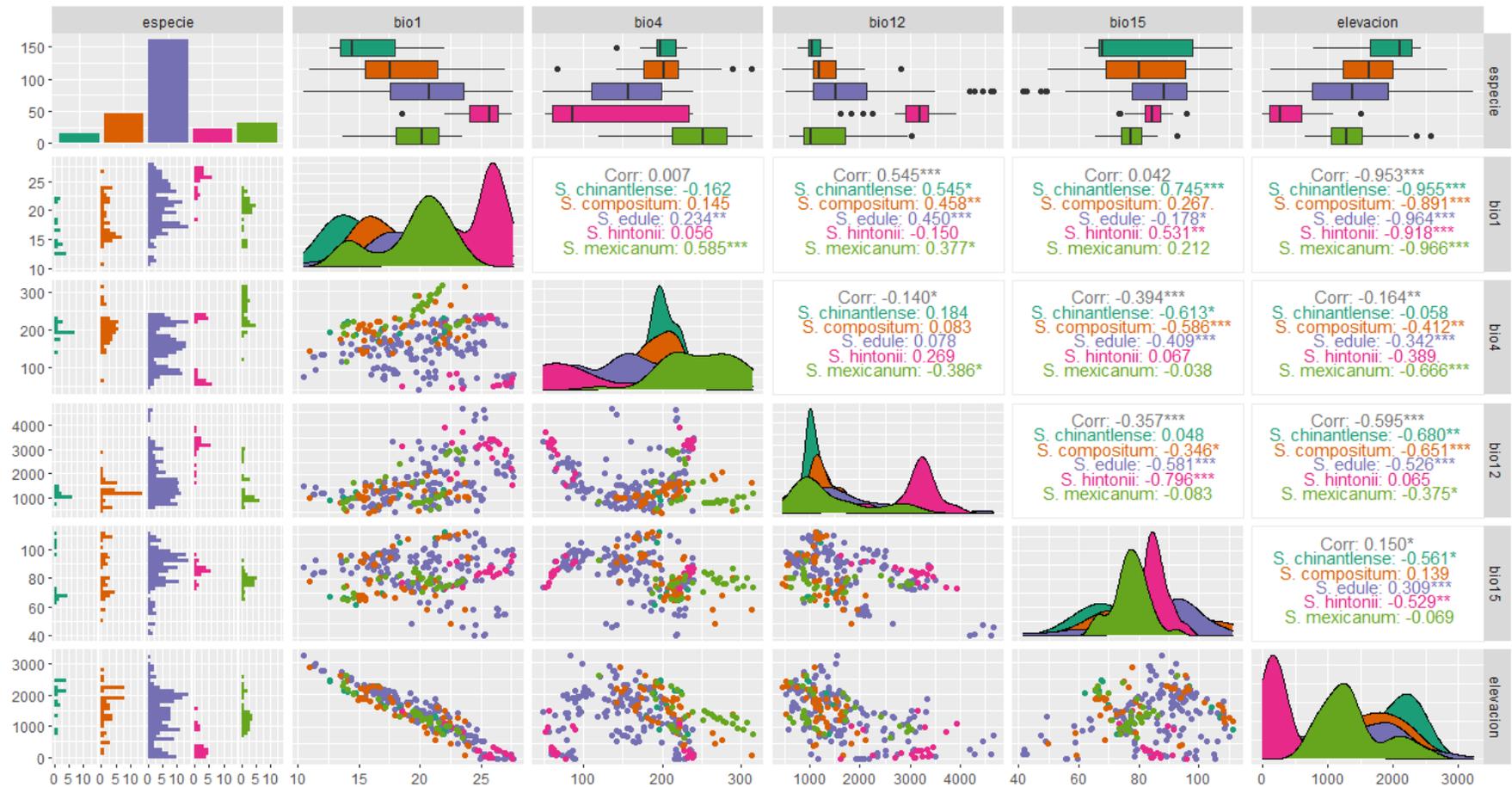


Figura 36. Histogramas de frecuencias, cajas de bigotes, gráficos de densidad y correlaciones para las variables bio1, bio4, bio12, bio15 y elevación de las especies mexicanas de *Sechium*.

7.3.3. Análisis de agrupamiento para el género *Sechium*

La Figura 37 muestra que el agrupamiento con base en variables bioclimáticas tiene cierta relación con los estudios moleculares citados en este escrito. Se logró una separación entre las especies centroamericanas y mexicanas (Figura 37A), aunque con $k = 3$, *S. hintonii* se agrupa en forma solitaria (Figura 37B).

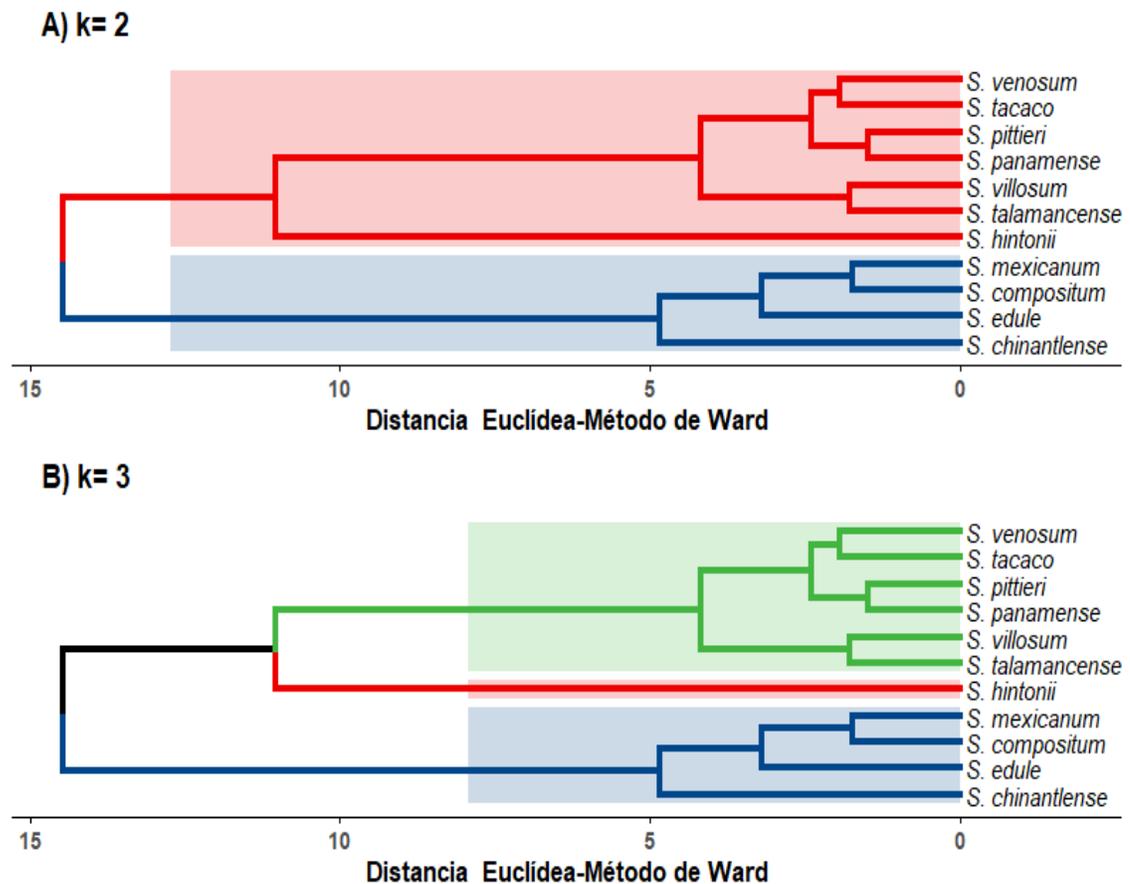


Figura 37. Dendrogramas calculados a partir de las medianas de las 19 variables bioclimáticas y la variable elevación para las especies de *Sechium*. **A)** $k = 2$, índice de consistencia promedio de 0.48 y **B)** $k = 3$ con índice de consistencia de 0.55.

7.3.4. Modelos de distribución

El modelo Maxent mostró tener buenos rendimientos en términos de AUC y TSS. *Sechium hintonii* mostró el AUC y TSS más bajos, sin embargo, estos rendimientos se consideran buenos según la literatura (Cuadro 20). Los mapas inferidos para *S. compositum* y *S. chinantlense* muestran regiones de alta idoneidad en los estados de Veracruz, Chiapas, Guerrero y Oaxaca (Figuras 38 y 39). En el capítulo V (Figuras 24 y 25) pueden observarse los mapas de idoneidad para *S. edule*.

Cuadro 20. Estadísticos para el modelo Maxent de las especies mexicanas de *Sechium*.

Especie	AUC	PPC	Sensibilidad	Especificidad	TSS
<i>S. chinantlense</i>	0.94	0.89	0.90	0.72	0.62
<i>S. compositum</i>	0.97	0.94	0.87	0.67	0.54
<i>S. edule</i>	0.93	0.85	0.98	0.64	0.62
<i>S. hintonii</i>	0.84	0.81	0.94	0.47	0.41
<i>S. mexicanum</i>	0.91	0.94	0.80	0.70	0.50

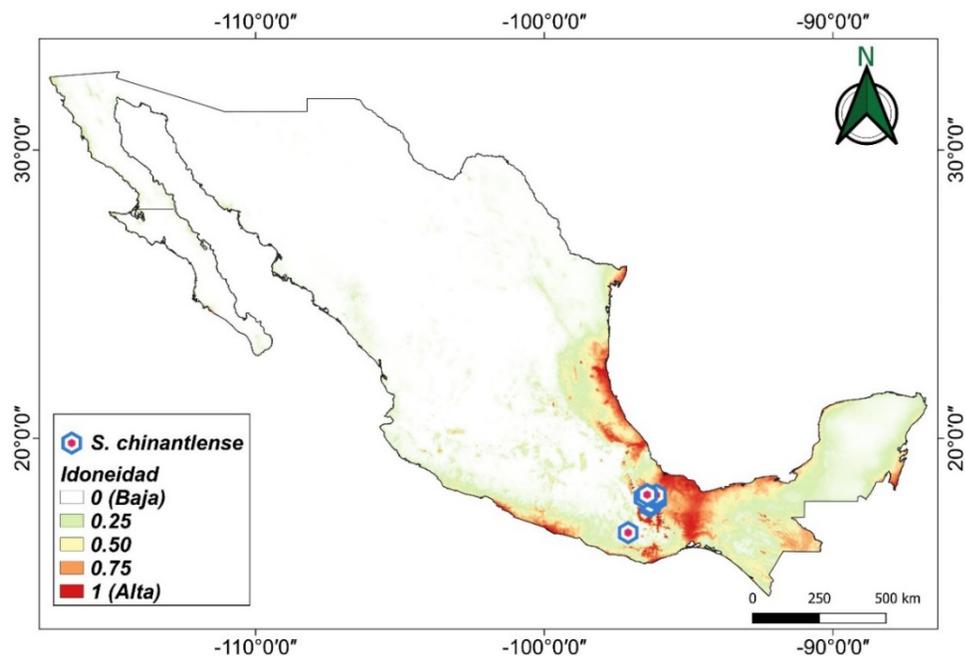


Figura 38. Modelo de distribución Maxent de *S. chinantlense* en México.

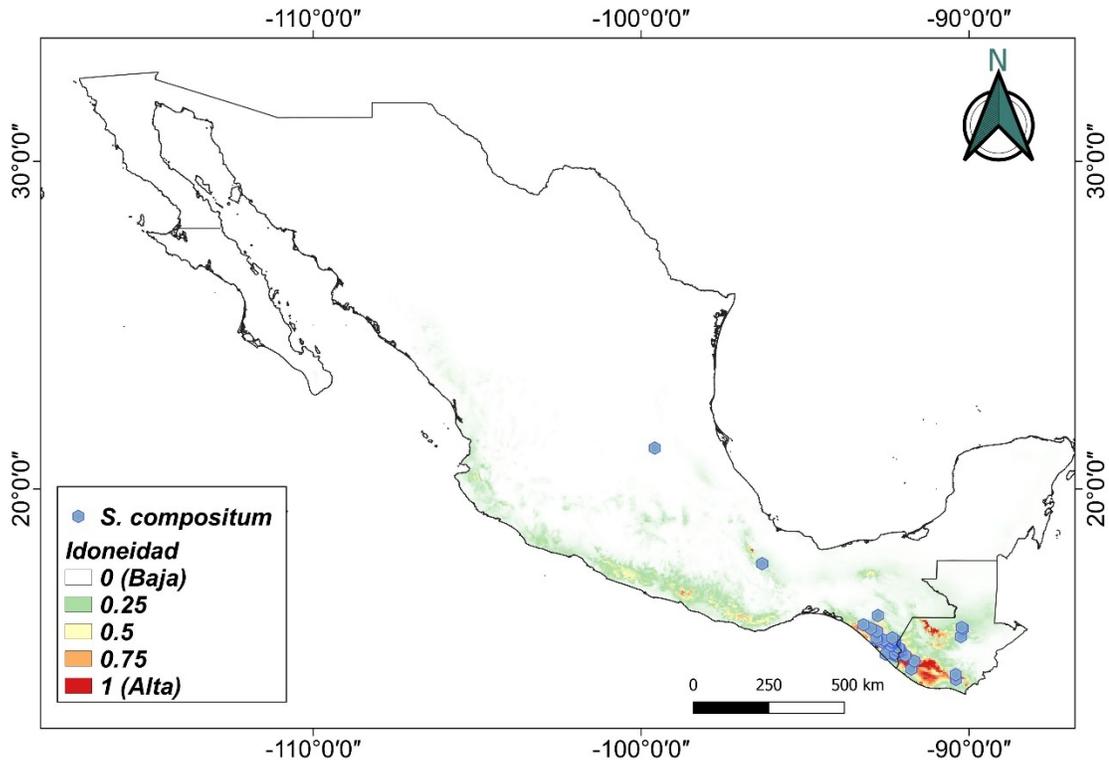


Figura 39. Modelo de distribución Maxent de *S. compositum* para México-Guatemala.

Sechium hintonii mostró zonas de alta idoneidad con mayor dispersión sobre territorio mexicano en comparación con *S. compositum*, *S. chinantlense* y *S. mexicanum* (Figura 40). Lira (1995) reportó que *S. hintonii* únicamente se distribuía en el Estado de México y el sureste de Guerrero, sin embargo, en el mapa de la Figura 40 se muestra que existen puntos de ocurrencia en los estados de Veracruz, Hidalgo, Puebla, Michoacán, Jalisco y Guanajuato.

Sechium mexicanum tiene una distribución muy restringida en la región montañosa que divide a los estados de Veracruz, Hidalgo y Puebla (Figura 41). De forma general, podemos concluir que es una especie difícil de encontrar en México.

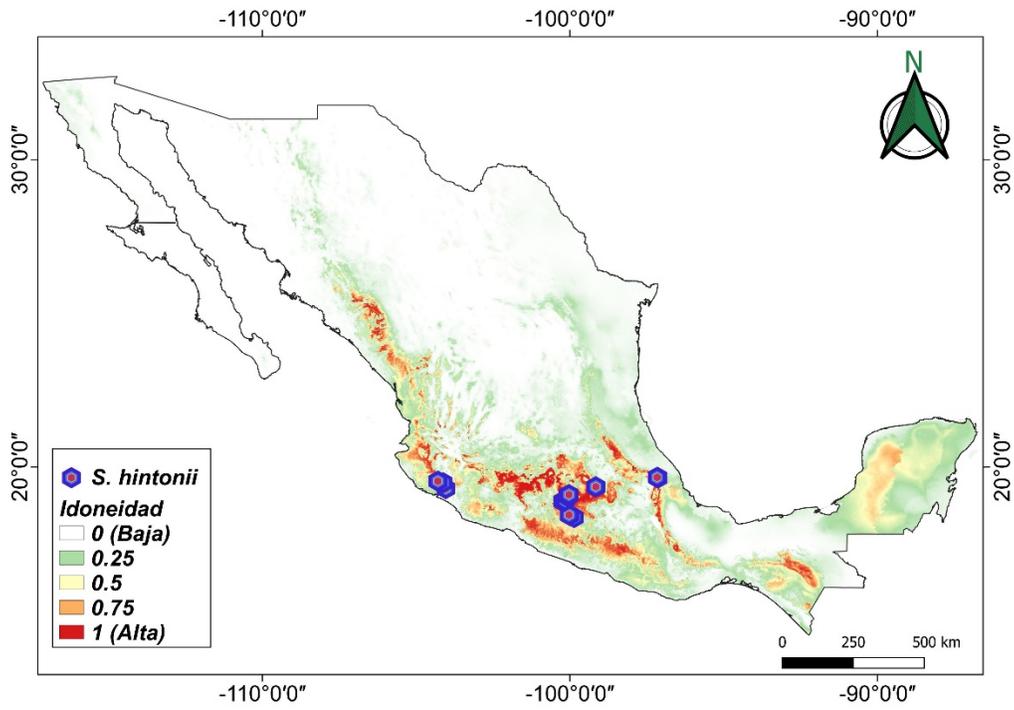


Figura 40. Modelo de distribución Maxent de *S. hintonii* para México.

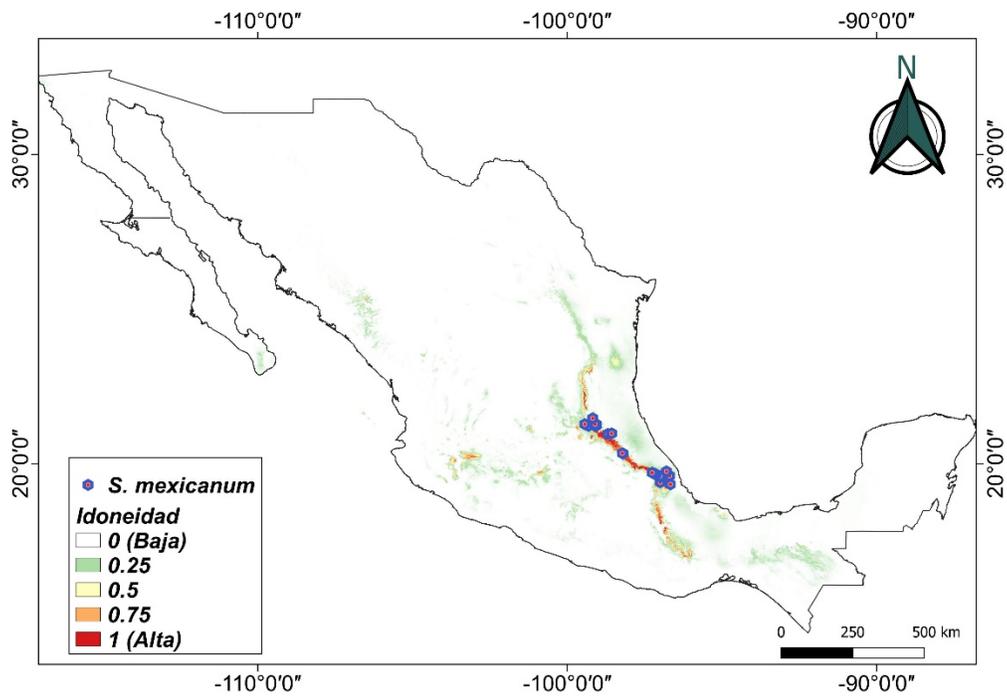


Figura 41. Modelo de distribución Maxent de *S. mexicanum* para México.

7.4. CONCLUSIONES

Las especies mexicanas de *Sechium* tiene gran potencial de idoneidad en los estados de Veracruz, Tabasco, Guerrero, Oaxaca y Chiapas; hacia El Bajío, Occidente de México y con moderada idoneidad hacia la Península de Yucatán. El modelo Maxent arrojó buenos rendimiento en términos de AUC y TSS. Las zonas mencionadas considerarse fuentes de germoplasma posibles de utilizar en el futuro. También, pueden aportar información referente al origen, domesticación y evolución de las especies mexicanas de *Sechium*.

7.5. LITERATURA CITADA

- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1), 1-12. <https://doi.org/10.1038/sdata.2018.214>
- Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: an R package for cluster validation. *Journal of Statistical Software*, 25(4), 1-22. <https://doi.org/10.18637/jss.v025.i04>
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, 61(6), 1-36. <https://doi.org/10.18637/jss.v061.i06>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315. <https://doi.org/10.1002/joc.5086>

- Fischer, G., Nachtergaele, F., Prieler, S., van Velthuisen, H. T., Verelst, D., & Wiberg, D. (2008). Global agro-ecological zones assessment for agriculture. IIASA, Laxenburg, Austria and FAO, Rome, Italy.
- Hijmans, R. J. (2020). *raster*: Geographic Data Analysis and Modeling (R package version 3.3-13) [Computer software]. <https://CRAN.R-project.org/package=raster>
- Kassambara, A., & Mundt, F. (2020). *factoextra*: extract and visualize the results of multivariate data analyses (R package version 1.0.7) [Computer software]. <https://CRAN.R-project.org/package=factoextra>
- Lira, S. R. (1995). Estudios taxonómicos en el género *Sechium* P. Br. *Cucurbitaceae* [Tesis Doctoral]. Universidad Nacional Autónoma de México.
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution*, 5(11), 1198-1205. <https://doi.org/10.1111/2041-210X.12261>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- QGIS Development Team. (2020). QGIS Geographic Information System. Open-Source Geospatial Foundation Project (3.16.2) [Computer software]. <https://www.qgis.org/en/site/>
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing (1.3.1093) [Computer software]. <https://www.R-project.org/>

- Schmitt, S., Pouteau, R., Justeau, D., Boissieu, F. de, & Birnbaum, P. (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, 8(12), 1795-1803. <https://doi.org/10.1111/2041-210X.12841>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open-Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Yilan, L., & Rutong, Z. (2015). clustertend: check the clustering tendency (R package version 1.4.) [Computer software]. <https://CRAN.R-project.org/package=clustertend>

Apéndice 2. Estadísticos descriptivos para las variables bioclimáticas de las especies mexicanas de *Sechium*.

Variable	<i>S. chinantlense</i>	<i>S. compositum</i>	<i>S. edule</i>	<i>S. hintonii</i>	<i>S. mexicanum</i>
Bio1					
Min	12.58	10.94	10.45	18.57	13.64
1 st Qu.	13.50	15.57	17.55	24.04	18.08
Media	16.01	18.38	20.52	25.02	19.43
3 ^{er} Qu.	17.93	21.41	23.56	26.38	21.52
Max	22.02	26.93	27.66	27.55	23.53
Bio2					
Min	9.97	8.44	8.95	9.875	9.49
1 st Qu.	11.717	10.73	11.51	10.463	11.55
Media	12.654	12.97	12.68	11.487	12.16
3 ^{er} Qu.	13.927	14.97	13.49	12.588	12.56
Max	15.217	17.65	18.49	13.117	14.51
Bio3					
Min	57.00	53.77	58.52	57.75	57.12
1 st Qu.	61.66	60.66	65.60	59	58.83
Media	63.65	64.04	69.25	70.61	60.50
3 ^{er} Qu.	65.41	68.84	73.37	79.13	61.69
Max	69.80	77.38	82.39	82.62	73.55
Bio4					
Min	142.20	66.51	48.73	51.06	119.60
1 st Qu.	192.30	176.81	110.03	60.75	212.90
Media	198.30	199.68	155.27	138.38	246.60
3 ^{er} Qu.	216.90	219.9	198.47	232.76	282
Max	232.00	313.46	238.87	238.87	315.20
Bio5					
Min	21.60	20.50	17.10	25.20	21.10
1 st Qu.	22.65	25.10	26.50	32.60	26.95
Media	25.59	28.20	29.53	33.20	29.20
3 ^{er} Qu.	27.93	30.75	32.98	35	31.70
Max	33.40	35.60	37.60	35.50	33.90
Bio6					
Min	2.70	0.50	1	11.10	2.70
1 st Qu.	3.20	4.95	8.32	16.10	8.12
Media	5.76	8.07	11.16	16.75	9.03
3 ^{er} Qu.	7.65	11.55	14.40	18.05	10.15
Max	13.30	18.30	19.10	19.40	14.30
Bio7					
Min	17.30	15.70	12.60	13.20	14.40
1 st Qu.	18.70	17.50	16.60	15.80	18.60
Media	19.82	20.13	18.36	16.45	20.17
3 ^{er} Qu.	20.85	22.50	19.80	17.55	21.32
Max	23.10	25.60	26.30	19.60	24.70
Bio8					
Min	13.75	11.33	11.23	18.73	14.67
1 st Qu.	14.78	16.83	18.47	24.62	19.50
Media	17.18	19.51	21.33	25.59	21.01
3 ^{er} Qu.	18.72	22.04	24.49	27.12	23.41
Max	23.75	26.77	28	27.58	24.77
Bio9					
Min	10.77	9.06	9.26	17.77	10.83
1 st Qu.	11.53	13.67	16.05	22.98	15.77
Media	14.58	16.79	19.29	24.12	17.34
3 ^{er} Qu.	17.27	19.84	22.37	25.73	18.60
Max	21.8	26.28	28.06	26.85	22.97
Bio10					
Min	14.72	12.78	11.27	19.08	15.42
1 st Qu.	15.75	17.59	19.07	25.68	20.40
Media	18.17	20.60	22.25	26.56	22.03
3 ^{er} Qu.	19.82	23.88	25.55	28.21	24.56

Max	24.4	27.90	29.15	28.48	26.03
Bio11					
Min	9.9	9.06	9.26	17.77	10.83
1 st Qu.	10.62	13.04	15.72	21.84	14.96
Media	13.35	15.72	18.42	23.18	16.04
3 ^{er} Qu.	15.89	18.56	21.42	25.73	17.36
Max	19.4	26.28	26.93	26.75	20.42
Bio12					
Min	766	474	528	1615	610
1 st Qu.	974	1086	1087	2922	884
Media	1101	1241	1722	3024	1414
3 ^{er} Qu.	1216	1522	2142	3356	1722
Max	1479	2824	4650	3921	3023
Bio13					
Min	162	85	111	346	117
1 st Qu.	194	205	224.20	558	196.80
Media	221	242.70	325.60	578.80	281.60
3 ^{er} Qu.	258	285.50	412.50	629.50	376.20
Max	275	547	694	693	555
Bio14					
Min	2	3	1	1	7
1 st Qu.	7.50	6	4.25	10.50	16.75
Media	24.56	24.15	24.14	34.91	31.91
3 ^{er} Qu.	32	39	36.75	60	46.25
Max	42	77	166	94	67
Bio15					
Min	62.09	49.67	41.39	73.25	65.41
1 st Qu.	66.66	68.97	77.65	82.27	74.31
Media	78.81	83.62	85.64	84.28	77.35
3 ^{er} Qu.	98.03	95.36	95.87	87.33	80.77
Max	111.32	111.43	110.31	95.82	92.60
Bio16					
Min	435	232	281	829	306
1 st Qu.	445.80	543.50	609.20	1446	456.20
Media	557.20	641.10	861.20	1523	709.60
3 ^{er} Qu.	684	747	1063	1704	843
Max	766	1385	1857	1973	1550
Bio17					
Min	18	13	4	14	29
1 st Qu.	25.50	23.50	19	50	55.75
Media	81.38	80.28	85.19	121.40	103.47
3 ^{er} Qu.	103.50	123	124.75	192.50	144.50
Max	143	243	577	298	218
Bio18					
Min	220	142	175	317	152
1 st Qu.	247	274.50	314.20	567.50	264.20
Media	327.30	347.80	473.90	673.20	376.20
3 ^{er} Qu.	317.80	420.50	616.20	761.50	445
Max	704	695	1064	1080	660
Bio19					
Min	24	22	4	14	38
1 st Qu.	75.25	33.50	21	70.50	65.75
Media	91.75	89.32	120.60	157.30	120.62
3 ^{er} Qu.	106.75	127	137.20	230.50	154.50
Max	173	276	976	397	281
Elevación					
Min	779	119	7	5	660
1 st Qu.	1660	1240	759.80	102	1045
Media	1909	1568	1337.80	401.20	1379
3 ^{er} Qu.	2292	2002	1924.50	607	1531
Max	2436	2832	3226	1509	2586

*Min = valor mínimo; Max = valor máximo; 1st y 3^{er} Qu: Primer y tercer cuartil.

CAPÍTULO VIII. ANÁLISIS DESCRIPTIVO Y MULTIVARIADO DE LA VARIACIÓN INFRAESPECÍFICA DE *S. edule* EN UN BANCO DE GERMOPLASMA *ex situ*

DESCRIPTIVE AND MULTIVARIATE ANALYSIS OF THE INFRASPECIFIC VARIATION OF *S. edule* IN AN *ex situ* GERMOPLASMA BANK

RESUMEN

El chayote (*Sechium edule*) es una cucurbitácea originaria de las regiones montañosas del suroeste mexicano, donde sus frutos son empleados como alimento. México es el primer productor y exportador de chayote a nivel mundial, y tiene establecido un banco de germoplasma *ex situ*. El objetivo de este estudio fue analizar mediante técnicas descriptivas y multivariadas las accesiones del banco de germoplasma, esto en aras de encontrar asociaciones y patrones de similitud entre las accesiones que permitan gestionar eficientemente los recursos fitogenéticos de *S. edule*. El 23 % de las colectas está representada por *virens levis* y el 47.6 % provienen de Veracruz. El 62.8 % de las accesiones provienen de traspacios. La mujer resguarda y conserva el 58.9 % de dichas colectas. El 70.8 % de las colectas se emplean para comercio y el resto para autoconsumo. En el análisis de correspondencia múltiple se obtuvo el 27.24 % de la variación para los dos primeros componentes, donde las variables con mayor contribución fueron las relacionadas al sabor del fruto (dulce y amargo), condición de las poblaciones (bosque, ruderal-silvestre y huerta), color del fruto (amarillo y verde claro), tamaño pequeño de fruto y forma aplanada. El estudio de las colectas en los bancos de germoplasma es de importancia para tener un consenso sobre los materiales y tomar decisiones encaminadas a su conservación y mejoramiento genético, estas mismas colectas deben ir acompañadas de una caracterización morfológica, química y molecular.

Palabras clave: banco de germoplasma, análisis descriptivo, análisis de correspondencia múltiple, conservación, mejoramiento genético.

ABSTRACT

The chayote (*Sechium edule*) is a cucurbit native to the mountainous regions of the Mexican southwest, where its fruits are used as food. Mexico is the world's leading producer and exporter of chayote, and has established a germplasm bank in situ. The objective of this study was to analyze the accessions of the germplasm bank using descriptive and multivariate techniques, this in order to find associations and patterns of similarity between the accessions that allow the efficient management of the plant genetic resources of *S. edule*. 23% of the accessions are represented by *virens levis* and 47.6% come from Veracruz. 62.8% of the accessions come from backyards. The woman protects and conserves 58.9% of these accessions. 70.8% of the accessions are used for commerce and the rest for self-consumption. In the multiple correspondence analysis, 27.24% of the variation was obtained for the first two components, where the variables with the greatest contribution were those related to the taste of the fruit (sweet and bitter), condition of the populations (forest, ruderal-wild and orchard), fruit color (yellow and light green), small fruit size and flattened shape. The study of the accessions in the germplasm banks is of importance to have a consensus on the materials and make decisions aimed at their conservation and genetic improvement, these same accessions must be accompanied by a morphological, chemical and molecular characterization.

Key words: germplasm bank, descriptive analysis, multiple correspondence analysis, conservation, genetic improvement.

8.1. INTRODUCCIÓN

El chayote [*Sechium edule* (Jacq) Swartz] es un cultivo nativo de Mesoamérica, específicamente originado y diversificado en el suroeste de México (Newstrom, 1990; Lira, 1995). A lo largo del tiempo, las poblaciones de *S. edule* se han diversificado por factores naturales y artificiales en los denominados complejos varietales, los cuales difieren significativamente en características físico-químicas de flores, hojas y frutos (Cadena-Iñiguez, 2005; Cadena-Iñiguez *et al.*, 2008, 2011). Actualmente algunos complejos varietales como *nigrum minor*, *albus dulcis* y *nigrum xalapensis* se encuentran en riesgo de erosión genética y se estima que sus áreas de distribución disminuyan alrededor del 50 % en las próximas décadas (González-Santos *et al.*, 2017). Para aminorar problemas de erosión genética, en 2007 se creó el Banco Nacional de Germoplasma de *Sechium* spp. (BANGESE) en Huatusco, Veracruz, México, donde se resguardan *in situ* colecciones silvestres y domesticadas de México y Centroamérica correspondientes a *S. edule*, *S. compositum*, *S. chinantlense* y *S. hintonii* (Cadena-Iñiguez and Arévalo-Galarza, 2011).

Los bancos de germoplasma son de importancia para la conservación a largo plazo de los recursos fitogenéticos, sin embargo, si no existe un análisis sobre la documentación de sus colectas, es difícil tomar decisiones encaminadas al ámbito de la investigación, específicamente en términos de conservación y mejoramiento genético (Weise *et al.*, 2020). El análisis de los datos de pasaporte, junto con estudios de caracterización morfológica y molecular, son herramientas que ayudan a gestionar eficientemente las accesiones dentro de un banco de germoplasma (Barrera-Guzmán *et al.*, 2020). Por ejemplo, respecto a chayote, mediante la gestión de las colectas del BANGESE se han implementado métodos de mejoramiento genético participativo y se han liberado variedades comerciales de *S. edule* aptas para el mercado nacional e internacional (Cadena-Iñiguez *et al.*, 2013a).

El objetivo de esta investigación fue analizar mediante técnicas descriptivas y multivariadas los datos de pasaporte de las accesiones del BANGESE, esto en

aras de encontrar asociaciones y patrones de similitud entre las accesiones que permitan gestionar eficientemente los recursos fitogenéticos de la variabilidad infraespecífica de *S. edule*.

8.2. MATERIALES Y MÉTODOS

8.2.1. Datos de pasaporte

Todas los paquetes estadísticos y gráficos por mencionar se ejecutaron en Rstudio (R Core Team, 2020). Se obtuvieron 231 accesiones pertenecientes a la variabilidad infraespecífica de *S. edule* del BANGESE, las cuales se graficaron con el paquete tidyverse (Wickham *et al.*, 2019) para observar la proporción de complejos varietales en función al lugar de su procedencia, tipo de establecimiento y propietario. Adicionalmente se evaluó la variable altitud con estadísticos de tendencia central y dispersión para observar algún patrón relacionado a los rangos de distribución bajo los cuales se desarrollan los complejos varietales de *S. edule*.

8.2.2. Análisis multivariado

Debido a que no todas las 231 accesiones tienen datos de pasaporte completos, únicamente se emplearon 129 accesiones las cuales tienen datos obtenidos con base en la guía de descriptores de los complejos varietales de *S. edule* (Cadena-Iñiguez *et al.*, 2017). Las variables cualitativas se estudiaron mediante un análisis de correspondencia múltiple (ACM) con los paquetes FactoMineR (Lê *et al.*, 2008) y factoextra (Kassambara and Mundt, 2020). Para brindar una explicación más sencilla, se decidió evaluar el ACM por gráficos separados que correspondían a los grupos de variables (condición, propietario, color del fruto, textura, tamaño de fruto, sabor y forma del fruto). Para el cálculo de los eigenvectores, eigenvalores y el gráfico referente a la contribución de cada variable a los componentes, se emplearon los mismos paquetes estadísticos para correr el ACM.

8.3. RESULTADOS Y DISCUSIÓN

8.3.1. Análisis descriptivo

Los complejos varietales de *S. edule* son *albus dulcis*, *albus levis*, *albus minor*, *albus spinosum*, *amarus silvestrys*, *nigrum conus*, *nigrum levis*, *nigrum maxima*, *nigrum minor*, *nigrum spinosum*, *nigrum xalapensis* y *virens levis*. En el BANGESE se encuentran 231 accesiones que representan la variabilidad infraespecífica de *S. edule*, adicionalmente, también se cuenta con cuatro accesiones de *S. chinantlense* y 12 de *S. compositum*. Los complejos varietales *virens levis* y *nigrum spinosum* son los que tienen mayor representatividad en el BANGESE (Figura 42A); en el estado de Veracruz se encuentran 110 accesiones (47.6 %) de la colección total de *S. edule* (Figura 42B). Algunas accesiones del BANGESE pertenecientes a *virens levis* y *nigrum spinosum* provienen de Costa Rica y Guatemala, sin embargo, la mayoría de ellas son originarias de Veracruz, México (Figura 43).

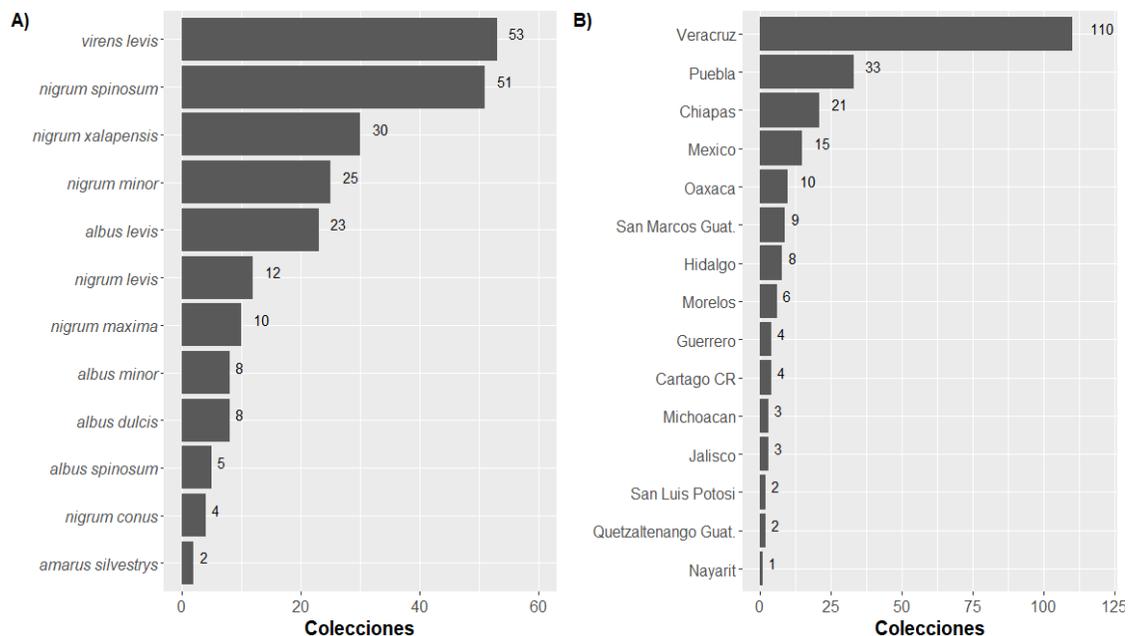


Figura 42. Colectas de *S. edule* en función a sus **A)** complejos varietales y **B)** lugar de procedencia.

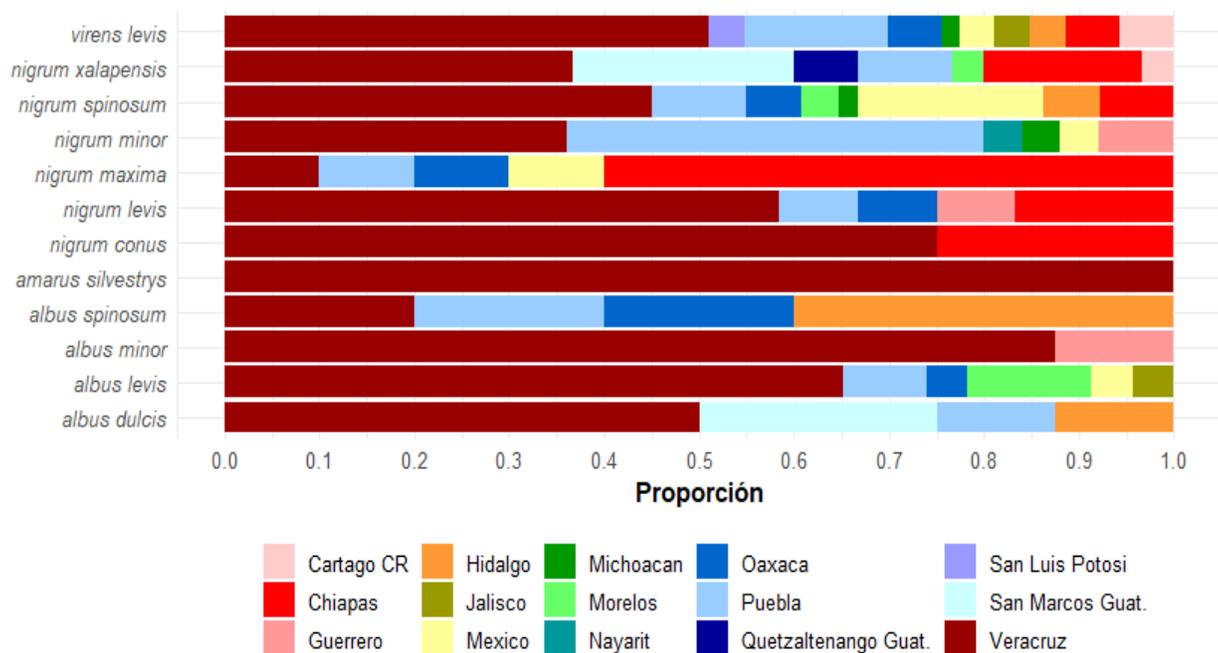


Figura 43. Proporciones y distribución de *Sechium* en función al lugar de procedencia.

De las 231 colectas en BANGESE, solo 129 cuenta con datos de pasaporte completos. De esta manera, el 62.8 % de las colectas se encuentran en traspatio y el 23.3 % en huertos comerciales. *Sechium compositum*, *S. chinantlense* y *amarus silvestrys* se encuentran en condiciones ruderales y silvestres (Figura 44).

La mujer maneja y conserva el 58.9 % de las colectas de *S. edule*, por lo que es un pilar fundamental en la conservación de estos recursos fitogenéticos; por otra parte, los complejos varietales destinados a agronegocios son *virens levis*, *nigrum spinosum*, *nigrum minor*, *nigrum maxima*, *albus minor* y *albus dulcis* (Figura 45). Las accesiones silvestres están fueran del alcance tanto de hombres como de mujeres, sin embargo, su conservación en bancos de germoplasma resulta muy importante. Los complejos varietales *virens levis*, *nigrum xalapensis*,

nigrum spinosum, *nigrum levis* y *albus dulcis* se establecen como hombre-empresa, ya que comúnmente se comercializan a gran escala.

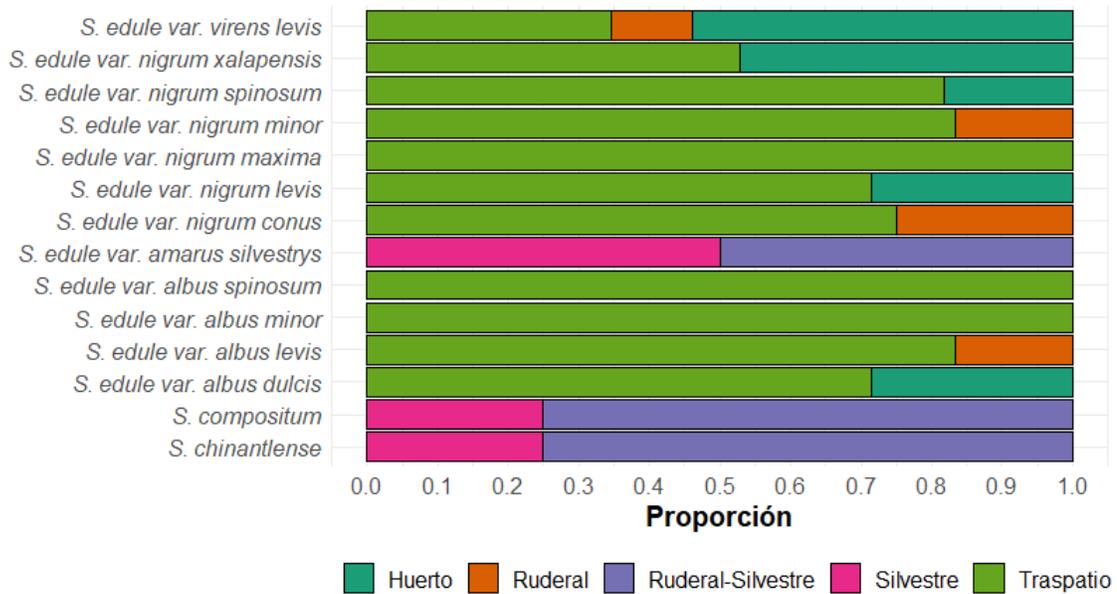


Figura 44. Establecimiento de la variación infraespecífica de *S. edule* y las especies *S. compositum* y *S. chinantlense*.

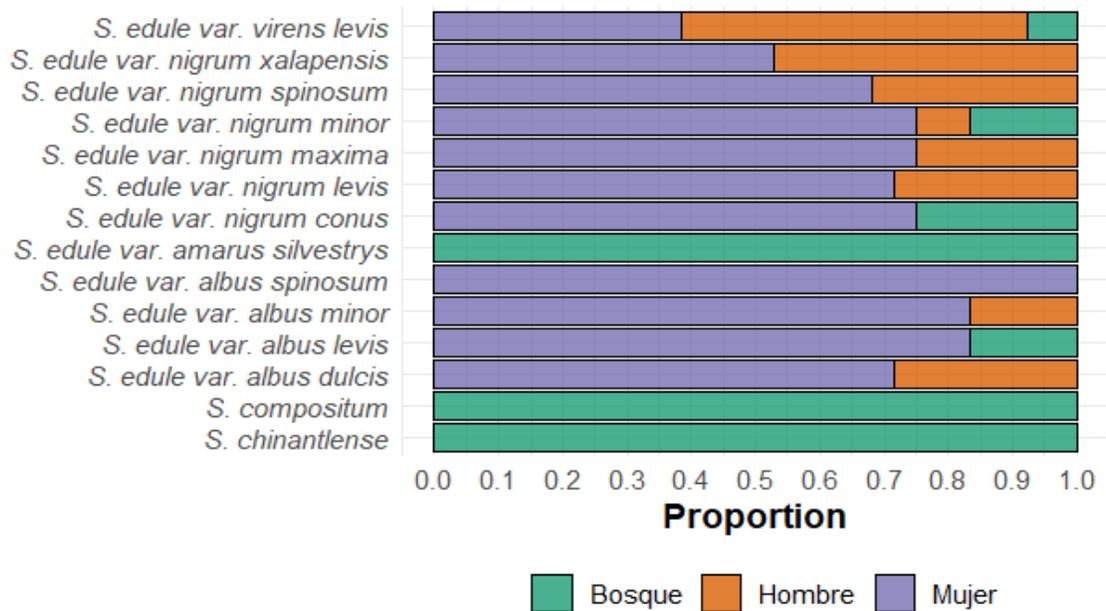


Figura 45. Propietario y manejo de los complejos varietales de *S. edule* y las especies *S. chinantlense* y *S. compositum*.

La Figura 46 muestra que algunos complejos varietales como *albus spinosum* se destinan exclusivamente para autoconsumo. Los grupos varietales *virens* y *nigrum* son empleados mayoritariamente para comercialización, mientras el grupo *albus* son destinados al autoconsumo. En la Figura 46 se observa que gran parte de los complejos varietales se usan para el comercio, a excepción de las accesiones silvestres. Algunos de ellos como *albus spinosum* y *albus minor* son exclusivos para autoconsumo. El 70.8 % de los complejos varietales es empleado para fines comerciales, mientras el 29.2 % para autoconsumo.

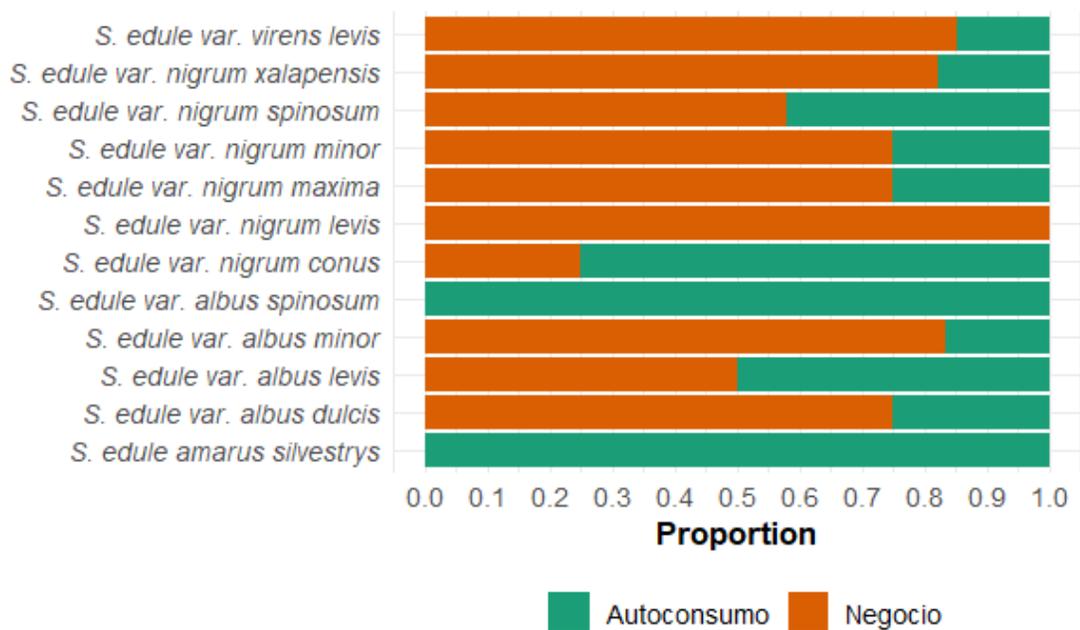


Figura 46. Destino final de la variación infraespecífica de *S. edule*.

En la Cuadro 21 se muestran estadísticos de tendencia central y dispersión para la variable altitud de las 129 colectas. El promedio de altitud más elevado fue para *nigrum spinosum* con 1993 msnm. La presencia de 'outliers' puede afectar algunas medidas de tendencia central como la media, algunas poblaciones de los complejos varietales parecen estar muy alejadas de los rangos óptimos para cada variedad en específico (Figura 47).

Cuadro 21. Estadísticos de tendencia central y dispersión para la variable altitud.

Varietal complex	Media	Mediana	Mínimo	Máximo	Rango	σ
<i>nigrum spinosum</i>	1993.70	1925.50	13.00	3376.00	3363.00	788.19
<i>albus minor</i>	1934.00	2223.00	1356.00	2223.00	867.00	500.56
<i>albus spinosum</i>	1790.00	1790.00	1357.00	2223.00	866.00	612.35
<i>nigrum minor</i>	1639.75	1672.00	1356.00	1859.00	503.00	210.54
<i>nigrum xalapensis</i>	1571.94	1486.00	424.00	2918.00	2494.00	558.11
<i>albus dulcis</i>	1498.57	1357.00	1164.00	2256.00	1092.00	383.64
<i>virens levis</i>	1437.89	1516.00	151.00	2918.00	2767.00	636.33
<i>albus levis</i>	1357.90	1356.00	753.00	2071.00	1318.00	393.04
<i>nigrum conus</i>	1351.25	1468.00	831.00	1638.00	807.00	367.53
<i>nigrum levis</i>	1323.00	1235.00	838.00	2148.00	1310.00	443.93
<i>nigrum maxima</i>	995.80	549.00	24.00	2918.00	2894.00	1150.76
<i>amarus silvestrys</i>	858.50	858.50	822.00	895.00	73.00	51.62
<i>S. chinantlense</i>	652.50	652.50	24.00	1281.00	1257.00	888.83
<i>S. compositum</i>	214.00	217.00	132.00	312.00	180.00	80.05

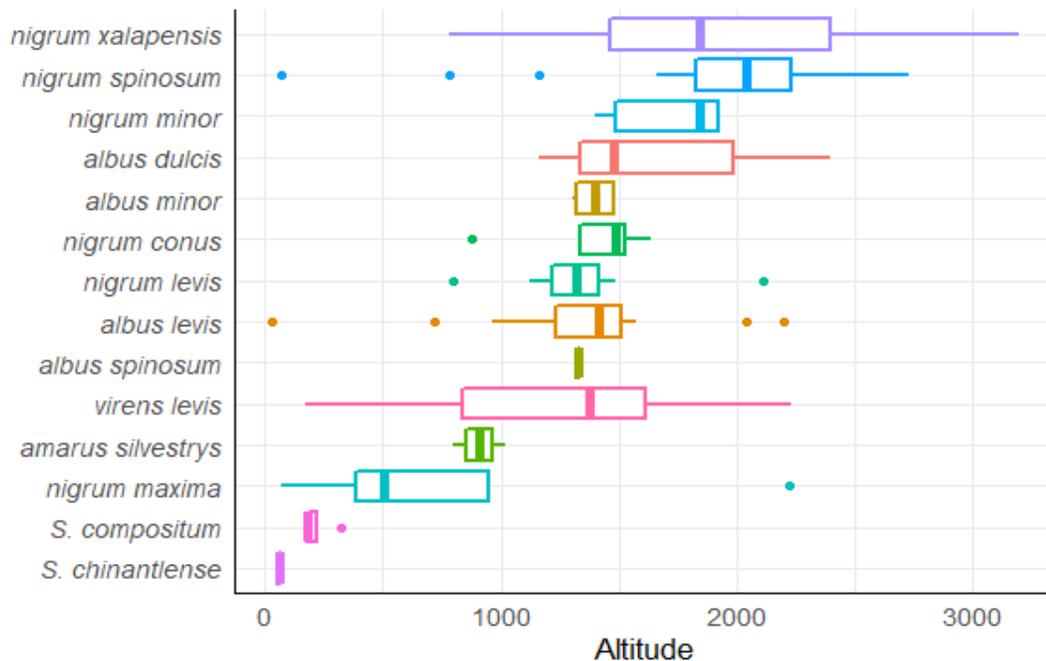


Figura 47. Cajas de bigotes para las colecciones de *Sechium* referente a la altitud.

8.3.2. Análisis de correspondencias múltiples

En los dos primeros componentes principales se obtuvo el 27.24 % de la variación total, los Cuadros 22 y 23 muestran los eigenvalores y eigenvectores respectivamente. Para los CP1 y CP2 las variables con mayor contribución fueron las relacionadas al sabor del fruto (dulce y amargo), condición de las poblaciones (bosque, ruderal-silvestre y huerta), color del fruto (amarillo y verde claro), tamaño pequeño de fruto y forma aplanada.

La Figura 48A presenta la distribución de las accesiones bajo el tipo de ambiente en el cual se desarrollan; de esta forma, las accesiones de *S. compositum* (Sco), *S. chinantlense* (Sci) y *amarus silvestrys* (AW) se encuentran en forma silvestre y ruderal-silvestre. El complejo *virens levis* (VL) está bajo condiciones de huerta, ya que es la forma comercial de *S. edule*; los complejos varietales de fruto color amarillo se encuentran generalmente en traspatio. Respecto a los propietarios, los hombres están asociados al complejo varietal comercial *virens levis*; mientras las mujeres resguardan complejos varietales de los grupos *albus* y *nigrum* (Figura 48B).

Cuadro 22. Eigenvalores y varianza explicada para los primeros 10 componentes del ACM

Componente	Eigenvalor	% varianza	% varianza acumulada
CP 1	0.49	15.50	15.50
CP 2	0.42	13.48	28.98
CP 3	0.30	9.48	38.46
CP 4	0.24	7.79	46.24
CP 5	0.20	6.49	52.74
CP 6	0.19	6.16	58.90
CP 7	0.18	5.82	64.72
CP 8	0.16	5.20	69.92
CP 9	0.15	4.82	74.73
CP 10	0.15	4.70	79.44

Cuadro 23. Vectores propios para las variables empleados en el ACM.

Variable	V1	V2	V3	V4	V5
Traspatio	-0.17	-0.86	-0.68	0.46	0.20
Huerto	-1.17	1.45	1.88	-0.63	0.25
Ruderal	1.03	0.25	-0.01	-4.08	-4.09
Ruderal-silvestre	4.12	2.43	-0.52	1.57	-0.62
Silvestre	3.83	2.34	0.89	1.15	4.38
Bosque	2.88	1.50	-0.03	-0.94	-1.14
Hombre	-1.07	1.21	1.43	-0.49	0.35
Mujer	-0.14	-0.91	-0.67	0.44	0.09
Verde oscuro	0.77	0.25	-1.14	-0.56	0.09
Verde-amarillo	-0.78	-0.05	-1.33	0.47	-2.00
Verde claro	-1.03	0.82	-0.37	0.33	-1.30
Verde muy oscuro	-0.64	1.11	1.30	0.14	3.44
Amarillo	0.51	-2.24	1.78	0.41	-0.30
Espinoso	-0.55	0.22	-0.75	1.88	0.12
Liso	0.15	-0.06	0.21	-0.52	-0.03
Grande	-0.68	-0.04	-1.36	1.07	-1.06
Mediano	-0.50	0.95	0.66	-0.30	-0.55
Pequeño	1.30	-1.09	0.97	1.37	-0.80
Muy grande	-1.02	0.23	-1.81	2.05	3.48
Muy pequeño	0.59	-1.11	-1.08	-3.00	1.99
Cónico	0.99	-0.71	-1.39	-0.79	-4.29
Aovado	0.43	-0.29	2.39	0.54	0.46
Elongado	-0.91	0.13	-2.72	1.52	2.30
Aplanado	4.00	2.22	-0.74	1.93	1.18
Piriforme	-0.65	0.37	-0.44	0.54	-0.56
Redondo	0.45	-1.10	-0.74	-2.50	1.37
Amargo	4.03	2.40	-0.09	1.44	0.88
Neutro	-0.59	0.40	-0.51	-0.28	-0.01
Dulce	0.51	-2.24	1.78	0.41	-0.30

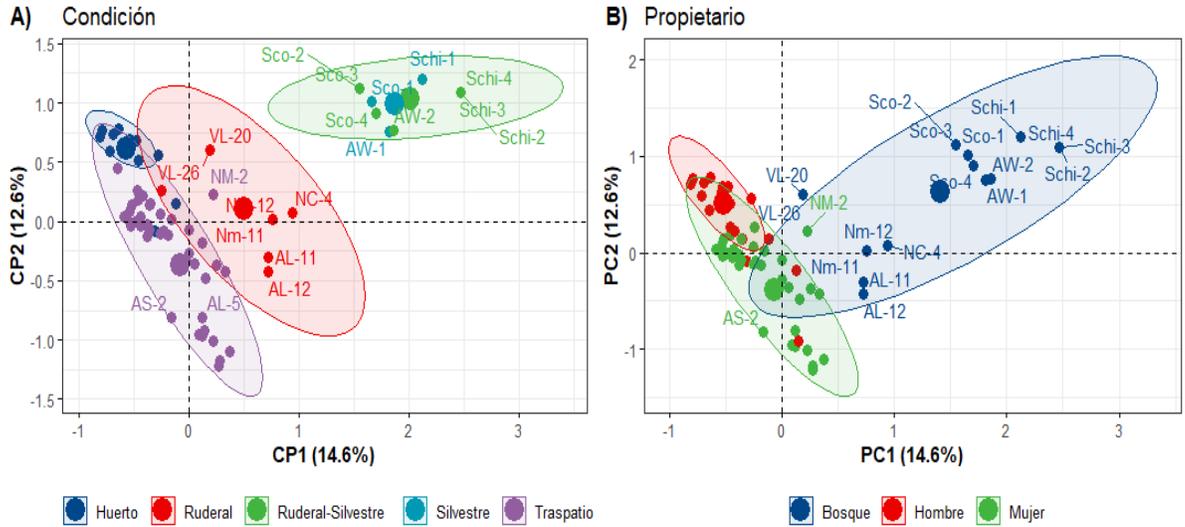


Figura 48. ACM para las variables **A)** Condición y **B)** Propietario.

Para color de frutos, la mayoría de ellos corresponde a diferentes tonalidades en color verde, a excepción de los del grupo *albus* (Figure 49A). En la Figura 49B se muestran que las accesiones de *amarus silvestrys* (Aw), *nigrum spinosum* (Ns) y *albus spinosum* (As) presentan espinas en sus frutos.

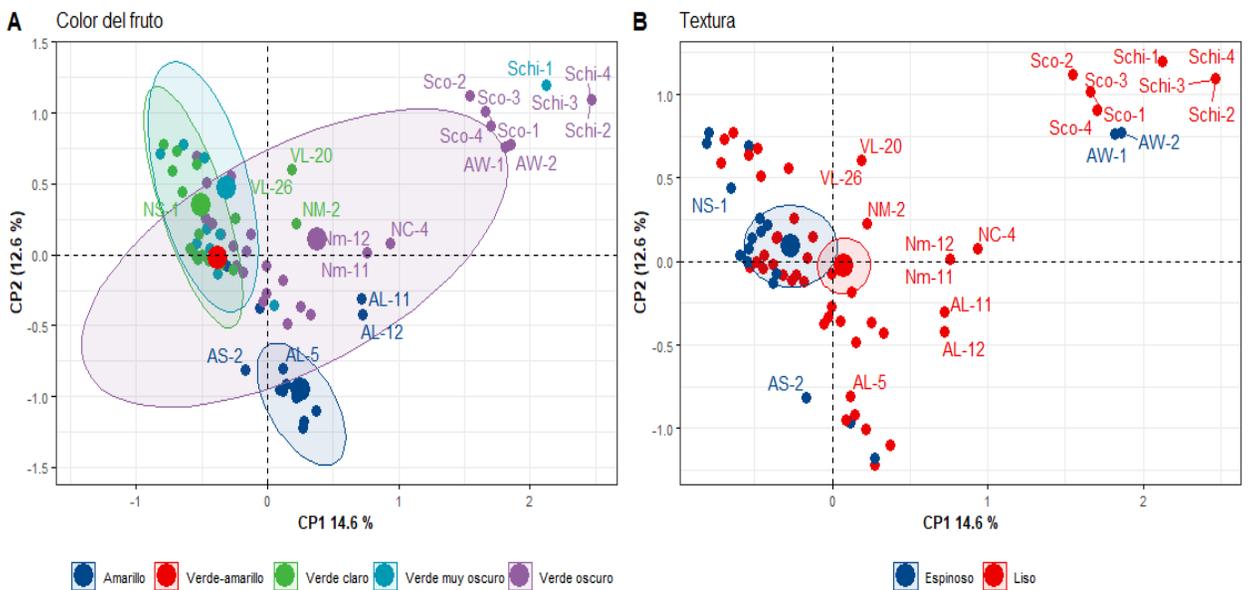


Figura 49. ACM para las variables **A)** Color del fruto y **B)** Textura del fruto.

El tamaño de los frutos de *Sechium* resulta ser muy variado. Los frutos de *amarus silvestrys* (AW), *nigrum minor* (Nm), *albus minor* (Am) y *S. chinantlense* (Schi) son pequeños. Frutos de tamaño grande y muy grande se presentaron en *nigrum maxima* (NM), *nigrum xalapensis* (Nx) y *albus spinosum*, principalmente (Figura 50A). Respecto al sabor, las elipses muestran una separación muy notoria: los frutos de *S. chinantlense*, *S. compositum* y *amarus silvestris* son amargos, los complejos del grupo *albus* tienen sabor dulce y los complejos del grupo *nigrum* y *virens* tienen sabor neutro (Figura 50B). La forma de los frutos es muy variada y resulta difícil establecer un patrón claro de agrupamiento (Figure 51).

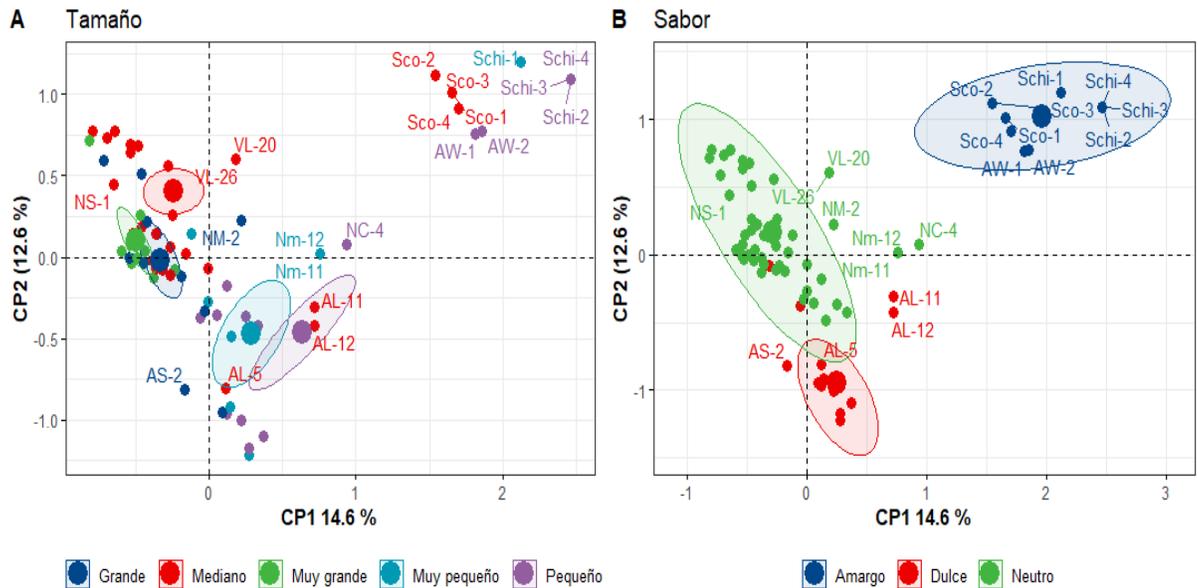


Figura 50. ACM para las variables A) Tamaño y B) Sabor de frutos.

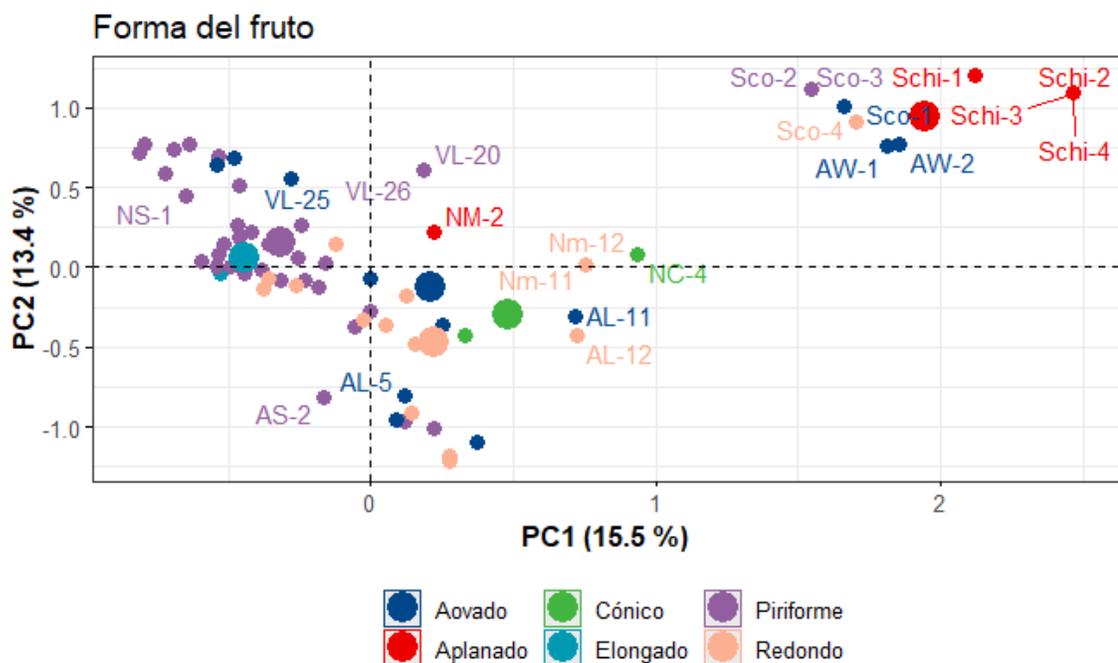


Figura 51. ACM para la variable forma de fruto.

La preferencia en el consumo de los grupos varietales de chayote radica en las características físicas y químicas de sus frutos. En México los grupos varietales con mayor presencia en el mercado son *virens levis*, *nigrum xalapensis* y *nigrum spinosum*. Estos suelen tener consistencia acuosa y poca fibra. En el ámbito de la exportación, se prefieren frutos del complejo varietal *virens levis*, ya que sus frutos son lisos, piriformes y de sabor neutro. Sin embargo, en los mercados regionales se opta por una gama más variada; los frutos dulces y almidonados también forman parte de la preferencia de los consumidores, ya que con ellos suelen prepararse ciertos platillos tradicionales, de esta forma también se conservan los recursos fitogenéticos de *S. edule*, sobre todo aquellos complejos varietales que están en peligro de erosión genética, pero son conservados en traspatios y/o huertos familiares.

El complejo *virens levis* es el utilizado por excelencia para la comercialización nacional e internacional. Los requisitos que debe cumplir son frutos piriformes,

color verde claro, liso sin espinas, sin surcos, 12-15 cm de longitud, 8-10 cm de ancho ecuatorial, 270-310 g, sabor neutro y libre de agentes patógenos. Sin embargo, el resto de los complejos varietales también tienen parámetros de estandarización para comercializarse (Cadena-Iñiguez y Arévalo-Galarza, 2011, p. 79). Respecto al autoconsumo, se prefiere que los frutos tengan alto contenido de agua, con sabor dulce o neutro, consistencia almidonada y libres de fibra.

En el bosque mesófilo se desarrollaron las primeras poblaciones silvestres de *S. edule*, caracterizándose por tener frutos color verde oscuro, espinosos y amargos; por otra parte, este tipo de vegetación es el adecuado para la variación infraespecífica de *S. edule* (Cadena-Iñiguez y Arévalo-Galarza, 2011). Los bosques mesófilos se distribuyen en las partes altas de la Sierra Madre Oriental, Sierra Norte de Chiapas y la Sierra Madre del Sur. La altitud varía entre 600-3100 msnm con una precipitación de 1000-3000 mm y con temperaturas de 12-23 °C.

Las rutas migratorias también son eventos, que, junto con las selecciones natural y artificial, provocaron la variación en el color y sabor del fruto (Cadena-Iñiguez y Arévalo-Galarza, 2011). Cuando las poblaciones silvestres de *S. edule* se trasladaron a la selva baja subperennifolia comenzó el cambio en el contenido de clorofila (tipo a) debido a una mayor irradiancia (Cadena Iñiguez *et al.*, 2011), dando origen a algunos complejos varietales como *virens levis* y aquellos del grupo *nigrum*. En los altiplanos y valles altos es común encontrar complejos varietales verdes espinosos. En los tipos de vegetación selva mediana perennifolia y baja caducifolia se comenzó a modificar el color y forma de las hojas y frutos, para posteriormente dar origen a los complejos varietales del grupo *albus*, donde sus frutos tienen bajo contenido de clorofilas y cucurbitacinas, pero con alto contenido de carotenoides y ácido ascórbico (Cadena Iñiguez *et al.*, 2011). La existencia de frutos amarillos también está fuertemente influenciada por la presión de selección (Valenta *et al.*, 2018).

El color amarillo de los frutos en chayote se debe a cambios bioquímicos producto de rutas de migración hacia otros tipos de vegetación y por los efectos de la

selección (Cadena-Iñiguez y Arévalo-Galarza, 2011). Sin embargo, Nevo *et al.* (2018) argumentaron que el color de la fruta desempeña un papel importante en la dispersión de semillas y puede ser considerada una señal evolutiva de las plantas. De esta manera, se busca tener un contraste entre el color de las hojas y los frutos para el éxito de la zoocoria. Los chayotes de color amarillo son considerados los más evolucionados dentro de la variación infraespecífica de *S. edule* (Cadena Iñiguez *et al.*, 2011)

El desempeño de la mujer en la conservación de los recursos fitogenéticos radica fundamentalmente en el manejo de los huertos familiares, empleando dichos recursos para la gastronomía y en ocasiones para la venta local. Por otra parte, el intercambio en los excedentes de la producción permite el flujo de diversidad genética a nivel regional (Watson y Eyzaguirre, 2001). El uso de recursos fitogenéticos destinados al consumo humano converge en una fuerte presión de selección, en el caso del chayote, se prefiere que los frutos tengan sabor neutro o dulce y con poca fibra, también se prefieren aquellos con consistencia almidonada (chayotes del grupo *albus*) (Cadena-Iñiguez, 2010).

El ACM referente al tamaño de los frutos, muestra que esta variable tiende a ser dinámica. Las poblaciones silvestres de *S. edule*, *S. chinantlense* y *S. compositum* exhiben frutos relativamente pequeños (< 10 cm) y con sabor amargo, ésta última característica se relaciona a los genes bi-1 y Bt-1, que confieren sabor amargo a hojas y frutos (Ren *et al.*, 2009), posteriormente con la selección esta característica ha ido en decremento. El sabor amargo está influenciado por la concentración de cucurbitacinas, la cual es mayor en las poblaciones silvestres y tienen aplicaciones en el área farmacéutica para el tratamiento de varios tipos de cáncer (Cadena-Iñiguez *et al.*, 2013b; Aguiñiga-Sánchez *et al.*, 2015, 2017; Salazar-Aguilar *et al.*, 2017; Rosado-Pérez *et al.*, 2019).

8.4. CONCLUSIONES

El complejo varietal *virens levis* es el más representado en el BANGESE y sus características morfológicas las hacen ideales para el mercado nacional e internacional. La mujer es la principal partícipe en la manejo, conservación y flujo de la diversidad genética de los recursos fitogenéticos de *S. edule*. El estudio de las colectas en los bancos de germoplasma es de importancia para tener un consenso sobre los materiales y tomar decisiones encaminadas a su conservación y mejoramiento genético, estas mismas accesiones deben ir acompañadas de una caracterización morfológica, química y molecular.

8.5. LITERATURA CITADA

- Aguñiga-Sánchez, I., Cadena-Íñiguez, J., Santiago-Osorio, E., Gómez-García, G., Mendoza-Núñez, V.M., Rosado-Pérez, J., Ruíz-Ramos, M., Cisneros-Solano, V.M., Ledesma-Martínez, E., Delgado-Bordonave, A. de J., Soto-Hernández, R.M., 2017. Chemical analyses and in vitro and in vivo toxicity of fruit methanol extract of *Sechium edule* var. *nigrum spinosum*. *Pharmaceutical Biology* 55, 1638–1645. <https://doi.org/10.1080/13880209.2017.1316746>
- Aguñiga-Sánchez, I., Soto-Hernández, M., Cadena-Íñiguez, J., Ruíz-Posadas, L. del M., Cadena-Zamudio, J.D., González-Ugarte, A.K., Steider, B.W., Santiago-Osorio, E., 2015. Fruit extract from a *Sechium edule* hybrid induce apoptosis in leukaemic cell lines but not in normal cells. *Nutrition and Cancer* 67, 250–257. <https://doi.org/10.1080/01635581.2015.989370>
- Barrera-Guzmán, L.A., Legaria-Solano, J.P., Ortega-Paczka, R., 2020. Diversidad genética en poblaciones de razas mexicanas de maíz. *Revista Fitotecnia Mexicana* 43, 121–125.

- Cadena Iñiguez, J., Soto Hernández, M., Arévalo Galarza, M. de L., Avendaño Arrazate, C.H., Aguirre Medina, J.F., Ruiz Posadas, L. del M., 2011. Biochemical characterization of domesticated varieties of chayote *Sechium edule* (Jacq.) Sw. fruits compared to wild relatives. *Revista Chapingo Serie horticultura* 17, 45–55. <https://doi.org/10.5154/r.rchsh.2011.17.044>
- Cadena-Iñiguez, J., 2010. El chayote (*Sechium edule* (Jacq.) Sw.), importante recurso fitogenético mesoamericano. *AgroProductividad* 3, 3–10.
- Cadena-Iñiguez, J., 2005. Caracterización morfoestructural, fisiológica, química y genética de diferentes tipos de chayote (*Sechium edule*) (Doctoral). Colegio de Postgraduados, Campus Montecillo, Texcoco, Estado de México.
- Cadena-Iñiguez, J., Arévalo-Galarza, M.L.C., 2011. Las variedades de Chayote (*Sechium edule* (Jacq.) Sw.) y su comercio mundial, 1 st ed. ed. bba, Montecillo, Texcoco. 98 p.
- Cadena-Iñiguez, J., Avendaño-Arrazate, C.H., Cisneros-Solano, V.M., Arévalo-Galarza, M.L.C., Ruíz-Posadas, L., Aguirre-Medina, J.F., Watanabe, K.N., 2017. Guía de descriptores varietales de *Sechium edule* (Jacq.) Sw. para la protección legal de su variación, 1º Edición. ed. Colegio de Postgraduados-GISeM, México.
- Cadena-Iñiguez, J., Avendaño-Arrazate, C.H., Soto-Hernández, M., Ruiz-Posadas, L.M., Aguirre-Medina, J.F., Arévalo-Galarza, L., 2008. Intraspecific variation of *Sechium edule* (Jacq.) Sw. in the state of Veracruz, Mexico. *Genetic resources and crop evolution* 55, 835–847. <https://doi.org/10.1007/s10722-007-9288-4>

- Cadena-Iñiguez, J., Soto-Hernández, M., Arévalo-Galarza, M.L.C., Avendaño-Arrazate, C.H., Aguirre-Medina, J.F., 2013a. Modelos de mejoramiento genético participativo en chayote (*Sechium* spp), 1 st ed. ed. Editorial del Colegio de Postgraduados, Montecillo, Texcoco.
- Cadena-Iñiguez, J., Soto-Hernández, M., Torres-Salas, A., Aguiñiga-Sánchez, I., Ruíz-Posadas, L., Rivera-Martínez, A.R., Avendaño-Arrazate, C.H., Santiago-Osorio, E., 2013b. The antiproliferative effect of chayote varieties (*Sechium edule* (Jacq.) Sw.) on tumour cell lines. *Journal of Medicinal Plant Research* 7, 455–460. <https://doi.org/10.5897/JMPR12.866>
- González-Santos, R., Cadena-Iñiguez, J., Morales-Flores, F.J., Ruiz-Vera, V.M., Pimentel-López, J., 2017. Prediction of the effects of climate change on *Sechium edule* (Jacq.) Swartz varietal groups in Mexico. *Genetic Resources and Crop Evolution* 64, 791–804. <https://doi.org/10.1007/s10722-016-0401-4>
- Kassambara, A., Mundt, F., 2020. factoextra: extract and visualize the results of multivariate data analyses. <https://cloud.r-project.org/package=factoextra>
- Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software* 25, 1–18. <https://doi.org/10.18637/jss.v025.i01>
- Lira, S.R., 1995. Estudios taxonómicos en el género *Sechium* P. Br. Cucurbitaceae (Tesis Doctoral). Universidad Nacional Autónoma de México, México, D.F.
- Nevo, O., Valenta, K., Razafimandimby, D., Melin, A.D., Ayasse, M., Chapman, C.A., 2018. Frugivores and the evolution of fruit color. *Biology Letters* 14, 20180377. <https://doi.org/10.1098/rsbl.2018.0377>

- Newstrom, L., 1990. Origin and evolution of chayote, *Sechium edule*, in: Bates, D.M., Robinson, R.W., Jeffrey, C. (Eds.), *Biology and Utilization of the Cucurbitaceae*. Cornell University Press, New York, USA, pp. 141–149.
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- Ren, Y., Zhang, Z., Liu, J., Staub, J.E., Han, Y., Cheng, Z., Li, X., Lu, J., Miao, H., Kang, H., Xie, B., Gu, X., Wang, X., Du, Y., Jin, W., Huang, S., 2009. An integrated genetic and cytogenetic map of the *Cucumber* genome. *PLOS ONE* 4, e5795. <https://doi.org/10.1371/journal.pone.0005795>
- Rosado-Pérez, J., Aguiñiga-Sánchez, I., Santiago-Osorio, E., Mendoza-Núñez, V.M., 2019. Effect of *Sechium edule* var. *nigrum spinosum* (Chayote) on oxidative stress and pro-inflammatory markers in older adults with metabolic syndrome: an exploratory study. *Antioxidants* 8, 146. <https://doi.org/10.3390/antiox8050146>
- Salazar-Aguilar, S., Ruiz-Posadas, L.D.M., Cadena-Iñiguez, J., Soto-Hernández, M., Santiago-Osorio, E., Aguiñiga-Sánchez, I., Rivera-Martínez, A.R., Aguirre-Medina, J.F., 2017. *Sechium edule* (Jacq.) Swartz, a new cultivar with antiproliferative potential in a human cervical cancer HeLa cell line. *Nutrients* 9, E798. <https://doi.org/10.3390/nu9080798>
- Valenta, K., Kalbitzer, U., Razafimandimby, D., Omeja, P., Ayasse, M., Chapman, C.A., Nevo, O., 2018. The evolution of fruit color: phylogeny, abiotic factors and the role of mutualists. *Scientific Reports* 8, 14302. <https://doi.org/10.1038/s41598-018-32604-x>
- Watson, J.W., Eyzaguirre, P.B., 2001. Home gardens and in situ conservation of plant genetic resources in farming systems. Proceedings of the Second International Home Gardens. Witzenhausen, Federal Republic of Germany.

Weise, S., Lohwasser, U., Oppermann, M., 2020. Document or Lose It-on the importance of information management for genetic resources conservation in gene banks. *Plants* 9, 1050. <https://doi.org/10.3390/plants9081050>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemond, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. *Journal of Open-Source Software* 4, 1686. <https://doi.org/10.21105/joss.01686>

CONCLUSIONES GENERALES

Las especies mexicanas y centroamericanas de *Sechium* están bien diferenciadas por distancias genéticas obtenidas a partir de marcadores moleculares y secuencias de ADN. *Sechium edule* y *S. tacaco* son las especies cultivadas de su género, el resto de las especies son silvestres, pero con posible utilidad para mejorar las especies cultivadas de *Sechium*.

Mediante secuencias de ADN extranuclear se determinó que las especies *S. edule*, *S. chinantlense* y *S. compositum* tienen un origen en común. Posiblemente las poblaciones silvestres de *S. compositum* y *S. chinantlense*, mediante hibridaciones y cruzamientos espontáneos dieron origen o participaron en la formación de las poblaciones cultivadas y silvestres de *S. edule*.

Los sistemas de información geográfica otorgaron información descriptiva referente a las condiciones edafoclimáticas bajo las cuales prosperan las especies de *Sechium*. Mediante algoritmos de aprendizaje automático se determinaron regiones con alta idoneidad de las especies. Para *S. edule* se determinó que las regiones con alta idoneidad se encuentran en regiones de los estados de Chiapas, Guerrero, Oaxaca, Puebla, Hidalgo y Veracruz, principalmente.

El estudio de las colectas en los bancos de germoplasma es de importancia para tener un consenso sobre los materiales y tomar decisiones encaminadas a su conservación y mejoramiento genético, estas mismas accesiones deben ir acompañadas de una caracterización morfológica, química y molecular. El complejo varietal *virens levis* es el más representado en el BANGESE y sus características morfológicas las hacen ideales para el mercado nacional e internacional. La mujer es la principal partícipe en el manejo, conservación y flujo de la diversidad genética de los recursos fitogenéticos de *S. edule*.

Apéndice 3. Variación infraespecífica de *S. edule*.



Figura 52. *Sechium edule* var. *nigrum xalapensis*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).

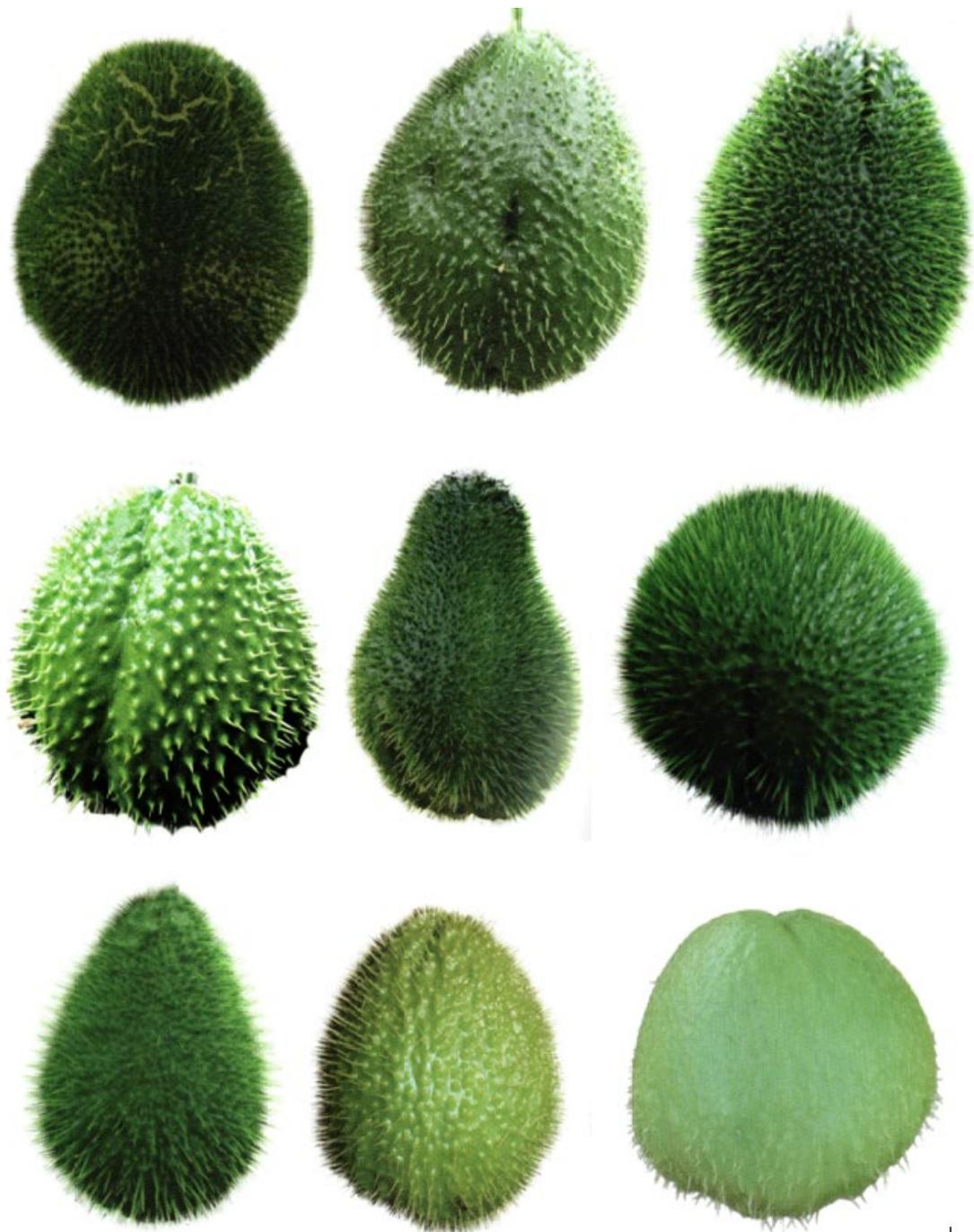


Figura 53. *Sechium edule* var. *nigrum spinosum*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 54. *Sechium edule* var. *nigrum levis*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 55. *Sechium edule* var. *nigrum minor*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 56. *Sechium edule* var. *nigrum maxima*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 57. *Sechium edule* var. *nigrum conus*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 58. *Sechium edule* var. *virens levis*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).

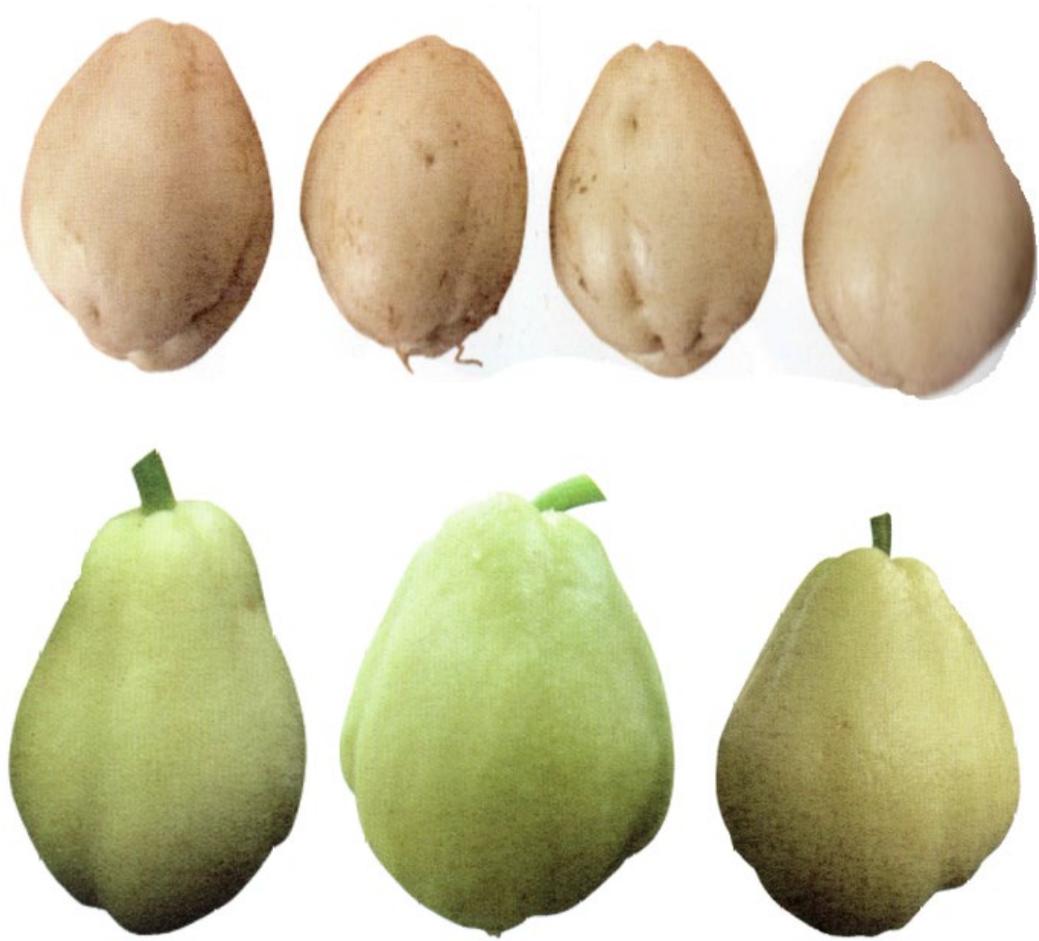


Figura 59. *Sechium edule* var. *albus dulcis*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).

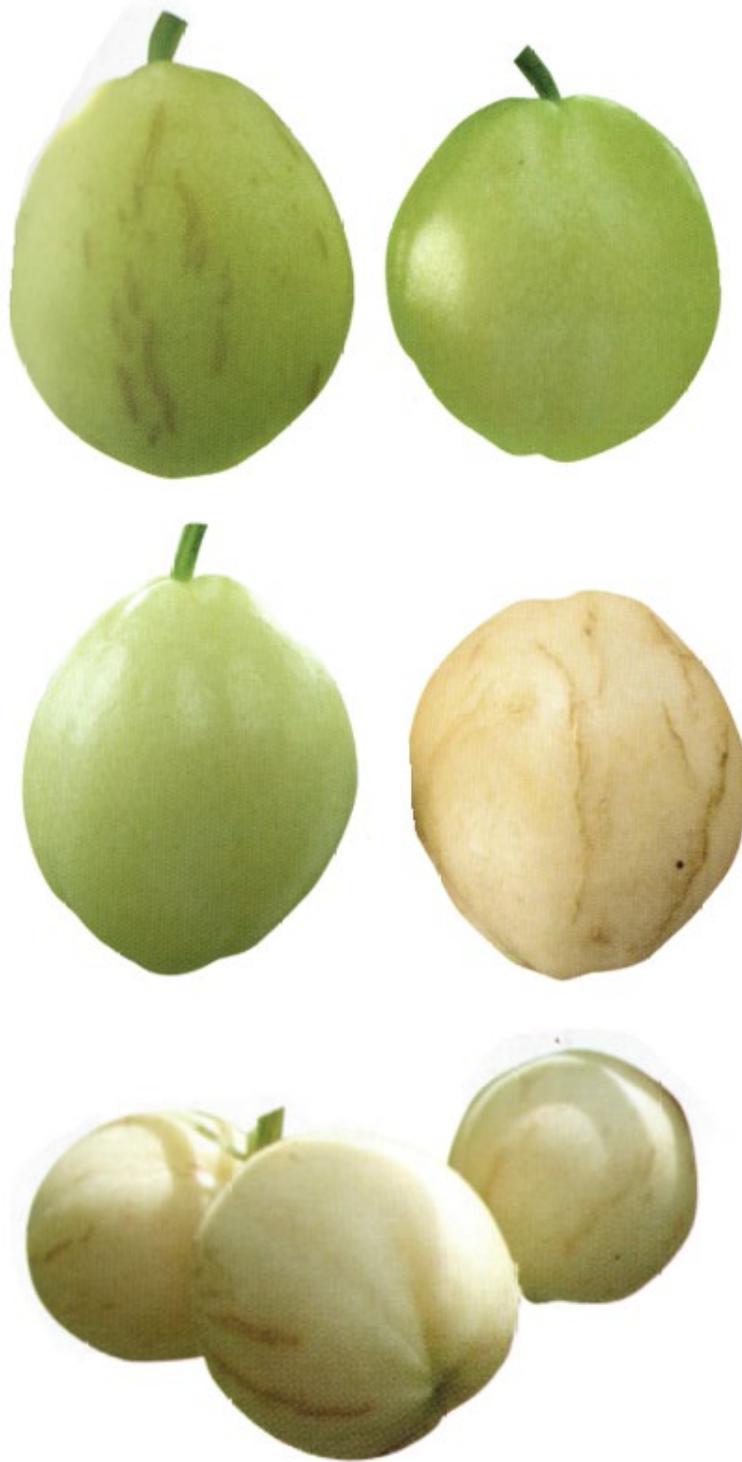


Figura 60. *Sechium edule* var. *albus minor*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).



Figura 61. *Sechium edule* var. *albus levis*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).

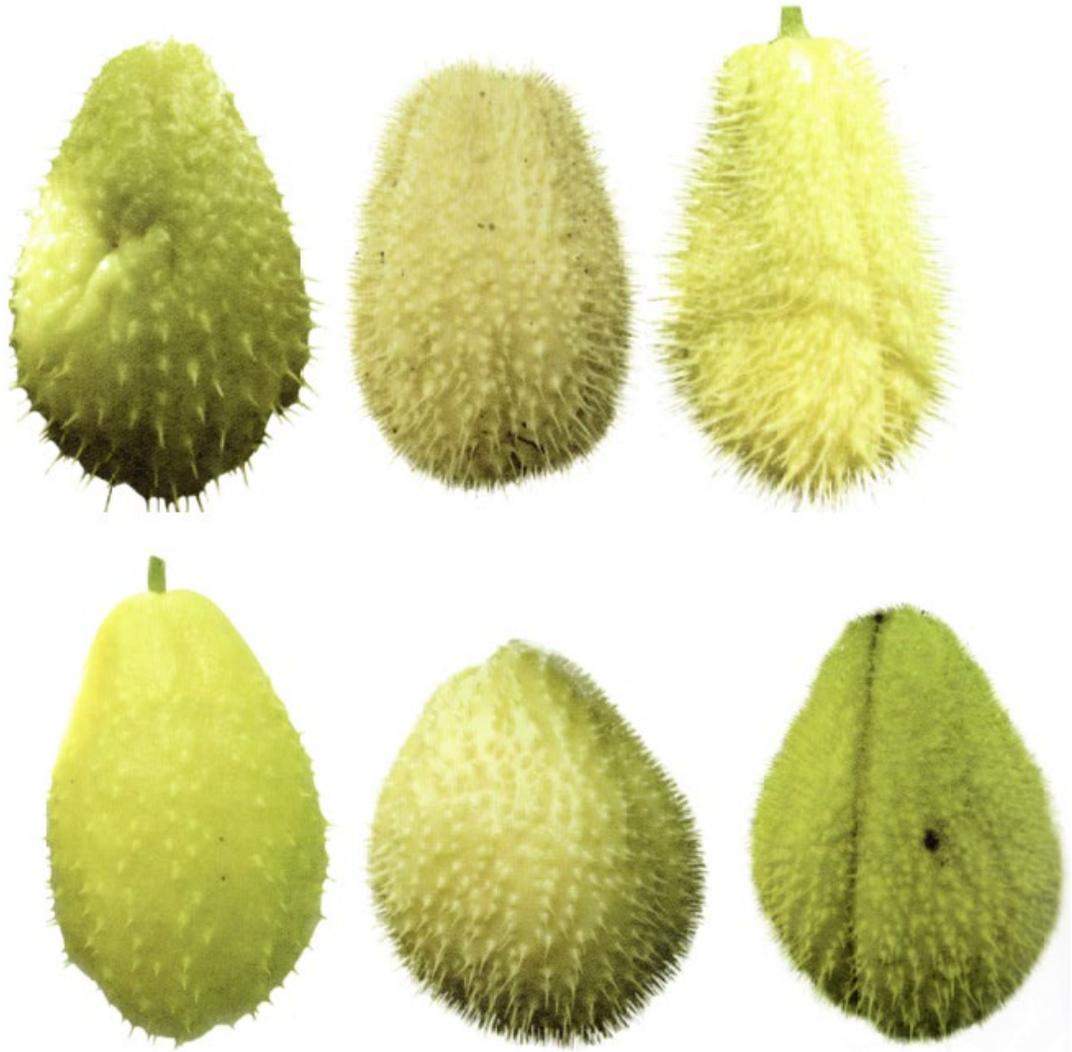


Figura 62. *Sechium edule* var. *albus spinosum*. Edición propia con fuente de Cadena-Iñiguez y Arévalo-Galarza (2011).