

Informe final* del Proyecto JM003
Distribución potencial de las especies silvestres de la familia Cucurbitaceae de México

Responsable: Dr. Rafael Lira Saade
Institución: Universidad Nacional Autónoma de México
Facultad de Estudios Superiores Iztacala
División de Investigación y Posgrado
Unidad de Biotecnología y Prototipos
Dirección: Av. de los Barrios # 1, Los Reyes Iztacala, Tlalnepantla, Mex, 54090 , México
Correo electrónico: rlira@servidor.unam.mx
Teléfono/Fax: 5623 1333 ext. 39780 Fax: 5623 1193
Fecha de inicio: Mayo 31, 2012.
Fecha de término: Mayo 15, 2018.
Principales resultados: Base de datos, cartografía, informe final.
Forma de citar el informe final y otros resultados:** Lira S., R., L.J. Sosa-Guzmán, I.V. Cabral-Soto, O. Téllez V. 2018. Distribución potencial de las especies silvestres de la familia Cucurbitaceae de México. Universidad Nacional Autónoma de México. Facultad de Estudios Superiores Iztacala. **Informe final SNIB-CONABIO, Proyecto No. JM003.** Ciudad de México.

Resumen:

Este proyecto tiene como objetivo llevar a cabo la modelación de la distribución geográfica de las especies arbóreas de las zonas áridas y semiáridas de México. Para este fin se definirá el polígono de ubicación de estas zonas, con base en las Ecoregiones Terrestres de la Conabio (2008) y se elaborará tanto una base de datos con información geográfica y taxonómica, como modelos de distribución de las especies arbóreas de estas regiones. Aproximadamente se incorporará a la base de datos información sobre 500 especies y categorías infraespecíficas, aunque posiblemente haya en estas zonas, cerca de 600. Se hará el modelaje geográfico de todas las especies que sea posible incluir, pues tomando en consideración que hay muchas especies que prácticamente no tienen información sobre su distribución, nos parece que podemos comprometernos a modelar la distribución geográfica del 70 % (350) de las especies y categorías infraespecíficas incluidas en la base de datos. Evidentemente, si en el desarrollo del proyecto existen más que puedan ser modeladas, éstas se incluirán en los resultados. Asimismo, se entregarán los perfiles bioclimáticos de todas las especies y de todas sus localidades, así como una sinopsis del análisis de vacío de información de las especies y fichas descriptivas de 100 especies arbóreas. La base de datos contendrá 19 campos que se construirán en el sistema Brahms, debido a la facilidad para manejar los datos. El modelaje bioclimático se hará por medio de MAXENT y los perfiles bioclimáticos con BIOCLIM.

-
- * El presente documento no necesariamente contiene los principales resultados del proyecto correspondiente o la descripción de los mismos. Los proyectos apoyados por la CONABIO así como información adicional sobre ellos, pueden consultarse en www.conabio.gob.mx
 - ** El usuario tiene la obligación, de conformidad con el artículo 57 de la LFDA, de citar a los autores de obras individuales, así como a los compiladores. De manera que deberán citarse todos los responsables de los proyectos, que proveyeron datos, así como a la CONABIO como depositaria, compiladora y proveedora de la información. En su caso, el usuario deberá obtener del proveedor la información complementaria sobre la autoría específica de los datos.

Informe final* del Proyecto JM003

Distribución potencial de las especies silvestres de la familia Cucurbitaceae de México

Responsable: Dr. Rafael Lira Saade

Institución: Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Iztacala

Unidad de Biotecnología y Prototipos

Dirección: Av. De Los Barrios 1, Los Reyes Iztacala, Tlalnepantla de Baz, Estado de México, 54090, México

Correo electrónico: rlira@unam.mx

Teléfono, fax Tel: 56221134

Principales resultados: base de datos DarwinCore, Modelos de distribución; Informe final

Forma de citar el informe final y otros resultados:**

Lira R., L.J. Sosa-Guzmán, I.V. Cabral-Soto, O. Téllez V. 2018.

Distribución potencial de las especies silvestres de la familia Cucurbitaceae de México.

Facultad de Estudios Superiores Iztacala

Informe final SNIB-CONABIO, proyecto JM003. México, D. F.

Febrero 2018

INDICE

Resumen.....	3
Introducción.....	4
Control de calidad de la base de datos.....	4
Proceso de modelación.....	8
Variables predictoras.....	8
Selección de la región de modelado 'M'.....	8
Modelación de MaxEnt.....	8
Desempeño del modelo.....	8
Umbral para definir la idoneidad.....	9
Conclusiones y recomendaciones.....-	9
Literatura consultada.....	10

ANEXO I Número de registros por especie en la base de datos del proyecto JM003

ANEXO II Valores de la prueba ROC (AUC) y umbral de corte

ANEXO III. Especies con valores inferiores de AUC a 0.800

Resumen

Se generaron modelos de distribución potencial de 96 especies, casi el 70% de las especies silvestres de la familia Cucurbitaceae que prosperan en México. Los modelos fueron generados, a partir de una base de datos del responsable, la cual contaba con cerca de 5,000 registros de más de 25 herbarios de México y otros países, así como también con la información complementaria brindada por el SNIB de CONABIO, la consulta de ejemplares del Herbario Nacional MEXU y literatura actualizada. La finalidad del proyecto era generar modelos de las 141 especies silvestres de México, sin embargo, varias de ellas contaron con registros insuficientes para la elaboración del modelo. Numerosos registros fueron excluidos del proceso de modelación debido a que estaban repetidos, mal determinados, o eran anómalos. Otros registros fueron excluidos debido a que las coordenadas estaban mal registradas y no pudieron ser corregidas. La base depurada contiene 5,340 registros correspondientes a 142 especies. Los modelos se generaron en forma de archivos digitales en formato ascii, acompañados de su respectivo metadato. Se cumplieron los objetivos de este proyecto, ya que logramos generar un conjunto de modelos, en su mayoría, con un buen desempeño (AUC mayores a 0.8). No obstante, por tratarse de especies ampliamente consumidas por el hombre, algunas distribuciones mostraron sesgos respecto a la distribución esperada de forma natural. La recomendación general que hacemos es explorar la modelación de las especies de cucurbitáceas sin definir una región de modelado 'M', que fue lo que se hizo en este proyecto, y en lugar de eso, hacer la modelación para todo el territorio mexicano.

Introducción

La familia Cucurbitaceae es un grupo de plantas rastreras o trepadoras, anuales o perennes, con flores unisexuales (con unos cuantos casos de bisexualidad) con ovario ínfero y frutos de diversos tipos, formas y dimensiones (Lira et al., 1998, 2002). A nivel mundial la familia incluye 118 géneros y 825 especies (Jeffrey, 1990). Es un grupo vegetal sumamente relevante desde muchos puntos de vista. Por ejemplo, muchas de sus especies están entre las plantas domesticadas de mayor interés para el hombre, ya que varias de ellas han sido parte fundamental de la dieta y otras facetas de la vida humana en todo el mundo (Lira 1995, 1996; Lira & Bye 1996; Lira et al., 2002); otras son especies silvestres útiles a nivel regional en el país, algunas son de distribución muy restringida y algunas de ellas, además, están cercanamente relacionadas con las plantas cultivadas de importancia económica (Cross et al., 2006; Lira 1995, 1996; Lira & Casas 1998; Lira & Caballero, 2002; Lira et al., 1998, 2009 a, b). De acuerdo con Lira et al. (2002), en la República Mexicana, la familia está representada por 34 géneros y 141 especies y taxa infraespecíficos, de los cuales cinco géneros y 70 especies o taxa infraespecíficos son endémicos al país.

Como antecedente de esta propuesta están los proyectos P097 y Q010 apoyados por CONABIO, los cuales tuvieron por objeto reunir la información necesaria para conformar una base de datos de la familia y llevar a cabo un análisis fitogeográfico cuantitativo y una propuesta de conservación preliminares del grupo en México (Lira et al., 2002). Otro proyecto también apoyado por CONABIO, derivó en el Catálogo de la Familia, y una colaboración con CONABIO permitió llevar a cabo un análisis de la distribución de las plantas domesticadas del género *Cucurbita* que se cultivan en México y sus parientes silvestres, con el objeto de evaluar el posible riesgo de liberación de transgénicos de ese género en el país (Arriaga et al., 2006). En los últimos años, además, también en este contexto y con apoyo de CIBIOGEM y CONABIO, se llevó a cabo una recopilación y análisis de la información existente de las especies de *Cucurbita* y *Sechium* que crecen y/o se cultivan en México y sus parientes silvestres (Lira et al., 2009b). Simultáneamente con esos trabajos, se llevó a cabo el análisis del efecto del cambio climático en los taxa silvestres más cercanamente emparentados con especies domesticadas de los géneros *Cucurbita* y *Sechium* (Lira et al., 2009 a).

No obstante a la gran cantidad de información que se ha logrado reunir sobre la familia Cucurbitaceae en México, no existe información sobre la distribución potencial ni las variables climáticas que influyen en la distribución de las especies silvestres que prosperan en el país. Este proyecto llena ese importante hueco de información, lo que permitirá plantear con mayor detalle estrategias de conservación de este importante grupo, por ejemplo, detectar de qué manera las áreas actualmente determinadas como reserva protegen a las especies de cucurbitáceas.

Control de calidad de la base de datos

Los registros de la familia Cucurbitaceae fueron obtenidos en su mayoría de la base de datos (BD) personal del responsable del proyecto, que originalmente contaba con 4947 registros. Se eliminaron los taxa no determinados a nivel de especie y se realizó la

georreferenciación de las localidades que no disponían de coordenadas o que contenían errores utilizando como herramientas principales Google-earth, cartografía de INEGI y recursos en línea. Se complementó esta información con la base de datos del SNIB proporcionada por CONABIO, la cual contaba originalmente con 5676 registros. Los registros fueron revisados minuciosamente por el Dr. Rafael Lira Saade, con el fin de detectar errores nomenclaturales, geográficos y anomalías. Los nombres de las especies brindados por el Sistema Nacional de Información sobre la Biodiversidad fueron revisados con relación a publicaciones especializadas, como revisiones y monografías, con relación a páginas en línea como www.tropicos.org y the plant list, herramientas con la cuales se actualizaron todos los nombres del proyecto JM003. La base fue revisada minuciosamente con el objetivo de mejorar la calidad de los datos, ya que contaba con registros con nombres científicos incorrectos, otros "antiguos" e incluso con errores ortográficos. Se realizó una consulta con el especialista en el grupo taxonómico de estudio, quien seleccionó los registros de especies cuyos determinadores cuentan con un amplio conocimiento o experiencia en la familia Cucurbitaceae. Posteriormente, los datos se cotejaron con la base de datos del Dr. Lira para encontrar registros duplicados. Esta base también se corrigió en cuanto a georreferencias. Los registros proporcionados por el responsable del proyecto, así como los proporcionados por CONABIO se unificaron en una sola base de datos en formato Access. Para algunas especies no se logró conjuntar un número suficientes de registros para la elaboración de modelos, por lo que se prosiguió a la consulta del Herbario Nacional MEXU, así como la consulta de literatura actualizada. De MEXU se obtuvieron 179 registros, de los cuales se georreferenciaron el 48%, y el 7.3 % fue corregido taxonómicamente, de acuerdo al conocimiento del especialista y a la consulta de bases de datos como Trópicos (<http://www.tropicos.org>) e IPNI (<http://www.ipni.org/index.html>). Para los géneros *Sicyos* se consultó la tesis de maestría de Rodríguez-Arévalo. Finalmente se revisó el artículo de Kearns (1994), de donde se tomaron 24 registros para el género *Ibervillea*. El 100% de estos datos tuvieron que georreferenciarse, ya que el artículo sólo proporcionó datos de la localidad sin coordenadas.

Con el fin de elaborar modelos de mejor calidad para cada especie, se identificaron los *outliers* o registros anómalos mediante modelos preliminares. Una vez identificados, se eliminaron o corrigieron en la base de datos de Darwin Core. Como ejemplo, mostramos los casos de las figuras 1-4. *Sicyos galeotti*, que es una especie que se distribuye principalmente en Veracruz y Oaxaca, con un registro anómalo en Chiapas. Otro caso es el de *Sechium edule*, que su distribución conocida es en Oaxaca, Veracruz, Querétaro y Puebla, presentando un registro anómalo en Campeche. De igual manera *Apodanthera aspera*, con distribución en Puebla y Oaxaca y con un registro anómalo en Durango y por último el caso de *Apodanthera undulata* que se distribuye del centro al Norte del país, con un registro anómalo en el Sur.

La base depurada contiene 5,340 registros correspondientes a 142 especies. El ANEXO I se puede consultar el número final de registros obtenidos para cada especie.

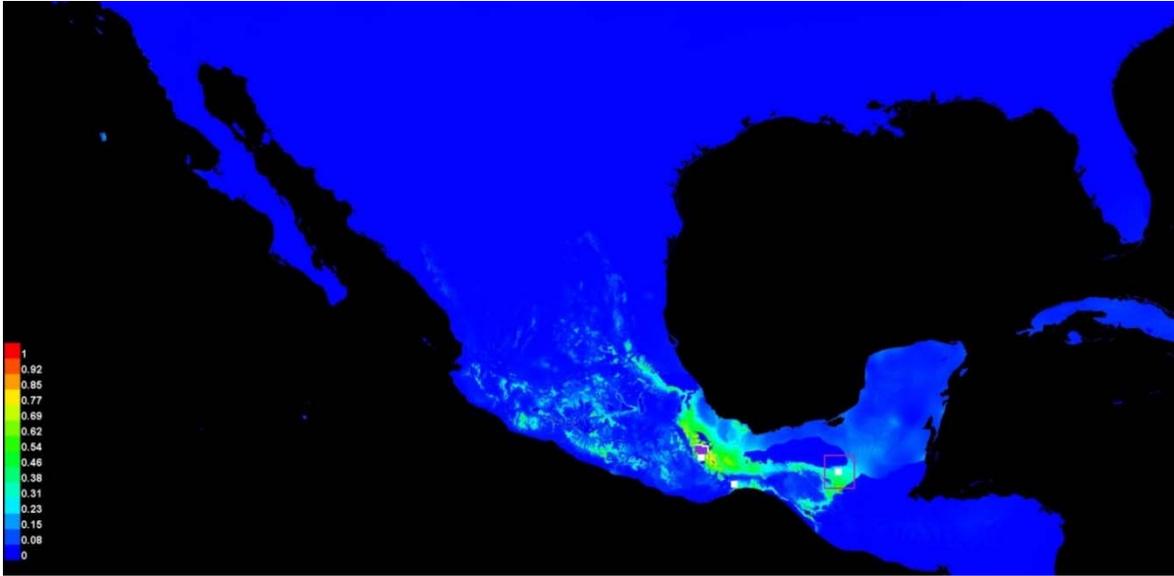


Figura 1. Modelo preliminar de *Sicyos galeotti*, con un registro anómalo en Chiapas.

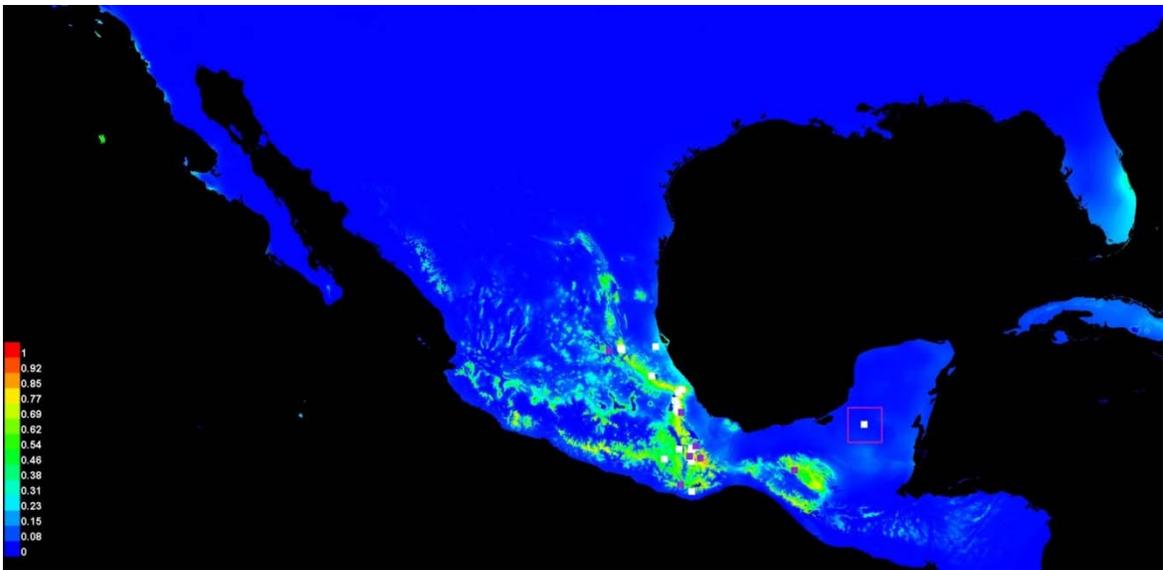


Figura 2. Modelo preliminar de *Sechium edule*, con un registro anómalo en Campeche.

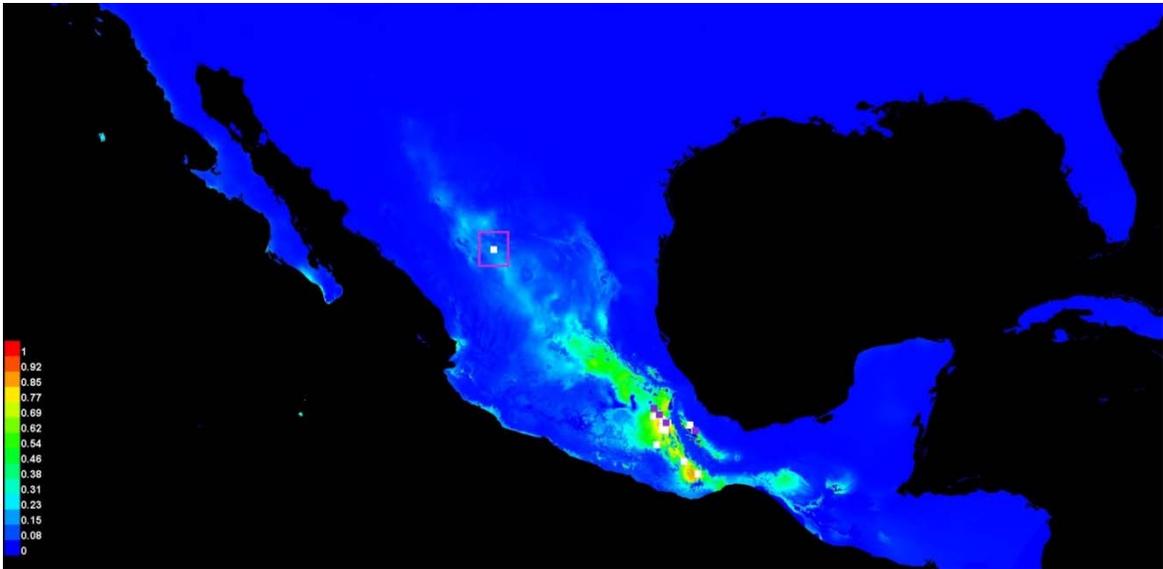


Figura 3. Modelo preliminar de *Apodanthera aspera*, con un registro anómalo en Durango.

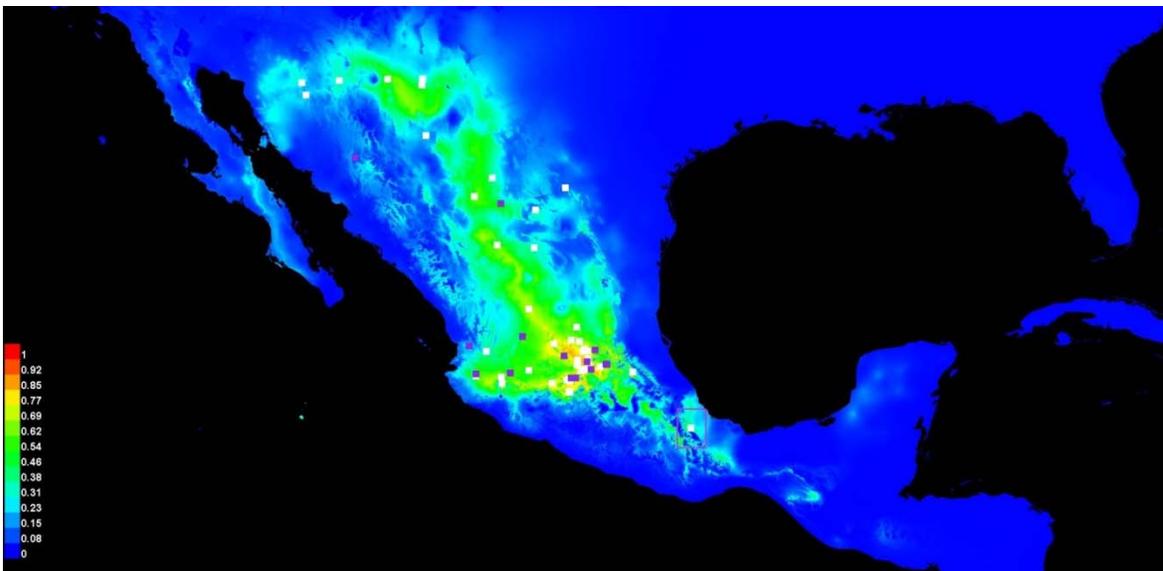


Figura 4. Modelo preliminar de *Apodanthera undulata*, con un registro anómalo en el Sur.

Con relación a los posibles errores geográficos, se revisaron las coordenadas de las localidades en que fueron recolectadas las especies, corrigiéndose aquellas que fueron no validadas geográficamente por CONABIO, y eliminándose las que se encontraban ubicadas en el mar. Finalmente, diversos registros que resultaron ser anómalos con respecto a la mayoría de los registros que mostraban la distribución conocida de la especie fueron eliminados. Al final del minucioso proceso de control de calidad de datos, la base de datos DarwinCore contiene 5,340 registros, correspondientes a 142 especies.

PROCESO DE MODELACIÓN

Variables predictoras

Como variables predictoras de la distribución de las especies, se utilizaron las 19 capas climáticas interpoladas por Cuervo-Robayo et al. (2013), las cuales cubren el periodo entre 1910-2009. Todas las capas empleadas tienen un formato ascii raster con una resolución espacial de 1 km²

Selección de la región de modelado ‘M’

Se definió una región a modelar ‘M’ consistente en un polígono formado por coordenadas extremas, basada en la regionalización de Provincias Bióticas (con énfasis en criterios morfotectónicos) Ferrusquía-Villafranca (1990). La M de cada especie fue seleccionada de acuerdo con la distribución de los registros de cada una de ellas. Se encontró que se formaron patrones entre grupos de especies y con la ayuda de un script en R fueron recortados. En algunos casos, de acuerdo con la localización de los sitios, las provincias fueron recortadas para ajustar la M.

Modelación en MaxEnt

Para generar el modelo de distribución potencial de cada especie se utilizó el programa MaxEnt BlackBox (Phillips et al., 2017), que es un algoritmo ampliamente usado para entender la relación entre las variables ambientales y la presencia de las especies (Elith et al. 2011).

La calibración por omisión del programa fue la siguiente: multiplicador de regularización fue 1, umbral de convergencia en 10^{-5} y el número máximo de punto de espacio ambiental de 1,000 (los cuales sirven de pseudoausencias para el análisis de MaxEnt). Se seleccionó un máximo de 1500 iteraciones (Phillips et al. 2006) y se omitieron las opciones de Extrapolate y Clamping, las cuales pueden ser empleadas para predicciones futuras.

Se utilizaron el 90% de registros para entrenamiento y el 10% para prueba como en otros trabajos (Londoño-Murcia & Sánchez-Cordero, 2011). Esto permitió a MaxEnt usar el porcentaje para generar una curva para cada modelo (Receiver Operating Characteristic – por sus siglas en inglés - ROC), con un valor asociado de AUC (area under the curve) para el mayor juego de datos (Phillips & Dudik, 2008; Marmion et al. 2009).

Las predicciones generadas con MaxEnt del nicho ecológico espacial tienen una escala de 0 (no adecuado) a 1 (muy adecuado) (Phillips et al. 2006). En todos los modelos se seleccionó el formato de salida Cloglog y el formato de ASCII para las predicciones, los cuales permitieron una interpretación más sencilla del modelo. Se seleccionaron las variables que como resultado brinda el análisis de Jackknife incluido en MaxEnt (Phillips & Dudik 2008)

Se produjeron los archivos en formato ascii raster para cada una de las especies que fue modelada de acuerdo con la calibración referida.

Desempeño del modelo

Para tener cierto grado de certidumbre del rendimiento del programa MaxEnt se utilizó el algoritmo de la curva ROC que está incorporado en el mismo programa como medida del rendimiento del programa y la fortaleza de los modelos generados. De acuerdo con Phillips

et al. (2006) todos aquellos modelos con valores superiores a 0.5 que es como si los modelos hubiesen sido generados aleatoriamente serían aceptables. Sin embargo, nosotros tomamos como límite para definir que los modelos fueron mejores cuando estos tuvieron valores superiores a 0.800 de AUC (Anexo III). Sin embargo, se generaron 11 modelos en los cuales los registros fuertemente anómalos entre sí, llevaron a manejar una calibración distinta en el sentido del espacio ambiental, el cual fue reducido hasta 50, con lo cual los modelos visualmente ajustaron muy razonablemente a las distribución conocida de las especies, sin embargo, los valores de AUC fueron inferiores a 0.700, incluso llegando hasta 0.200. Dado que los modelos muestran muy razonablemente la distribución de estas especies fueron aceptados como resultado.

Umbral para definir la idoneidad

Aunque existen diferentes alternativas para determinar el umbral óptimo para definir la idoneidad resultante (Freeman & Moisen 2008), no se han establecido reglas para determinar ese umbral óptimo (Hu & Jiang 2010). Nosotros optamos por seleccionar aquella que se denomina equal training sensitivity and specificity (Liu, 2005).

En su mayoría los modelos muestran estadísticas diagnósticas satisfactorias; los valores de entrenamiento superaron 0.800 al igual que los de prueba para la prueba AUC, por lo que consideramos que estos representan la distribución de las especies de una manera razonable, (ANEXO II). Algunos modelos tuvieron un desempeño menor a 0.8 (ANEXO III)

Conclusiones y recomendaciones

Definitivamente, la parte débil del proyecto ha sido la BDs, quizá como en el caso de cualquier grupo biológico. Todas las colectas de las especies de cucurbitáceas tienen un sesgo geográfico o ambiental. Mientras no exista un muestreo razonable de la especie bajo estudio en estos espacios la generación de modelos se verá con ciertos problemas y los resultados tendrán un grado de incertidumbre.

En las especies de Cucurbitaceae, por ser ampliamente consumidas por el hombre, algunas de ellas pueden mostrar una distribución altamente sesgada con respecto a la esperada de forma natural. Varias de las especies se encuentran ligadas a áreas alteradas e incluso basureros. Por lo anterior algunas de las áreas modeladas como parte del nicho fundamental en otras áreas del entorno natural que pudieran estar ocupando este tipo de especies son un poco dudosas, por el fuerte sesgo debido a la dispersión hecha por el hombre, a áreas en donde de forma natural no se distribuirían dichas especies.

Si bien la búsqueda de una "M" que refleje de una forma más natural la distribución de las especies es una idea excelente, los registros no permiten plasmar esta idea durante el modelado. Me parece que sería sano contrastar una M abierta versus esta empleada para recortar los modelos, ya que la disposición de los registros en el espacio geográfico en relación con las regiones o provincias de una regionalización pueden también sesgar fuertemente el modelado.

Literatura consultada

- Arriaga, L., E. Huerta, R. Lira, E. Moreno & J. Alarcón. 2006. Assessing the risk of releasing transgenic *Cucurbita* spp. in Mexico. *Agriculture, Ecosystems & Environments* 112: 291-299.
- Cross, H. B., R. Lira & T. J. Motley. 2006. Origins and diversification of Chayote. In: *Darwin's Harvest: New Approaches to the Origins, Evolution, and Conservation of Crops* (eds. T. J. Motley, N. J. C. Zerega & H. B. Cross). Columbia University Press, New York. Pp. 171-194.
- Cuervo-Robayo, A.P., Téllez-Valdés, Oswaldo, Gómez, Miguel, Venegas-Barrera, Crystian, Manjarrez, J. and E. Martínez-Meyer 2013. An update of high-resolution monthly climate surfaces for Mexico. *International Journal of Climatology*.
- Hutchinson, M.F. 2004. ANUSPLIN version 4.3 user guide. Centre for Resource and Environmental Studies, the Australian National University, Canberra, 54 p.
- Kearns, D. M. 1994a. The genus *Ibervillea* (Cucurbitaceae): An enumeration of the species and two new combinations *Madroño* 41: 13-22.
- Lira, R. & A. Casas. 1998. Uso y manejo en *Ibervillea millspaughii* (Cogn.) C. Jeffrey, *Melothria pendula* L. y otras especies silvestres de la familia Cucurbitaceae: Posibles procesos de domesticación incipiente. *Boletín Sociedad Botánica De México* 62: 77-89.
- Lira, R. & J. Caballero. 2002. Ethnobotany of the wild Mexican Cucurbitaceae. *Economic Botany*. 56: 380-398.
- Lira, R. 1995. Estudios taxonómicos y ecogeográficos de las cucurbitaceae latinoamericanas de importancia económica: *Cucurbita*, *Sechium*, *Sicana* y *Cyclanthera*. Systematic and ecogeographic studies on crop gene pools. 9. International Plant Genetic Resources Institute/Instituto de Biología, Universidad Nacional Autónoma de México. Roma. 281 pp.
- Lira, R. 1996. Chayote, *Sechium edule* (Jacq.) Sw. Promoting the conservation and use of underutilized and neglected crops. 8. Institute für Pflanzengenetik und Kulturpflanzenforschung Gatersleben/International Plant Genetic Resources Institute. Roma. 58 pp.
- Lira, R. y R. Bye. 1996. Las Cucurbitaceas en la alimentación de los dos mundos. In: J. Long-Solís (Ed.). *Conquista y comida. Consecuencias del encuentro de dos mundos*. Instituto de Investigaciones Históricas/Programa Universitario de Alimentos, Universidad Nacional Autónoma de México. México, D. F. pp. 199-226.
- Lira, R., C. Rodríguez, J.L. Alvarado, I. Rodríguez, J. Castrejón & A. Domínguez-Mariani. 1998. Diversidad e importancia de la familia Cucurbitaceae en México. *Acta Botánica Mexicana* 42: 43-77.
- Lira, R., J.L. Villaseñor & E. Ortíz. 2002. A proposal for the conservation of the family Cucurbitaceae in Mexico. *Biodiversity and Conservation* 11: 1699-1720.
- Lira, R., L. Eguiarte & S. Montes. 2009 b. Recopilación y análisis de la información existente de las especies de los géneros *Cucurbita* y *Sechium* que crecen y/o se cultivan en México. Informe Final de Proyecto CONABIO.
- Lira, R., O. Tellez & P. Dávila. 2009 a. The effects of climate change on the geographic distribution of wild Mexican Cucurbitaceae related to cultivated plants. *Genetic Resources and Crop Evolution* 56: 691-703.

- Peterson, A.T. & K.P. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecol. Mod.* 117: 159–164.
- Phillips S.J., R.P. Anderson & R.E. Schapire 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190: 231–259
- Rodríguez-Arévalo I. Revisión del género *Sicyos* L. (Cucurbitaceae) en México. Tesis de Maestría. Facultad de Ciencias, UNAM. México. 184 p
- Téllez, O., M.A. Hutchinson, H.A. Nix & P. Jones 2011. Desarrollo de coberturas digitales climáticas para México. In: Sánchez-Rojas, G., Ballesteros, B.C. & N. Pavón (Eds.) *Cambio Climático. Aproximaciones para el estudio de su efecto sobre la biodiversidad*. Universidad Autónoma de Hidalgo. Pachuca, Hidalgo. ISBN 978-607-482-152-9.
- Téllez, V.O. 1995. Flora, vegetación y fitogeografía de Nayarit. Tesis de Maestría. UNAM.
- Thuiller W, Albert C, Araújo MB, Berry PM, Cabeza M, Guisan A, Hickler T, Midgley GF, Paterson J, Schurr FM, Sykes MT, Zimmermann NE. 2008. Predicting global change impacts on plant species' distributions: Future challenges. *Persp Plant Ecol Evol Syst* 9: 137-152.

ANEXO I. Especies que conforman la base de datos DarwinCore, para las cuales se generó un modelo y número de registros obtenidos para cada especie

acceptedNameUsage	CuentaDecuenta
Apodanthera aspera	22
Apodanthera undulata	64
Brandegea bigelovii	16
Cayaponia attenuata	84
Cayaponia racemosa	81
Cionosicyos excisus	82
Cionosicyos macranthus	24
Citrullus lanatus	33
Cucumis anguria	48
Cucumis dipsaceus	39
Cucumis melo	62
Cucumis sativus	13
Cucurbita argyrosperma	172
Cucurbita cordata	46
Cucurbita digitata	34
Cucurbita foetidissima	138
Cucurbita lundelliana	40
Cucurbita moschata	83
Cucurbita palmata	21
Cucurbita pedatifolia	46
Cucurbita pepo	72
Cucurbita radicans	60
Cucurbita scabridifolia	11
Cyclanthera dissecta	157
Cyclanthera integrifoliola	42
Cyclanthera langaei	76
Cyclanthera minima	23
Cyclanthera multifoliolata	83
Cyclanthera ribiflora	80
Cyclanthera rostrata	15
Cyclanthera steyermarkii	7
Cyclanthera tamnoides	61
Chalema synanthera	15
Dieterlea fusiformis	32
Dieterlea maxima	29
Doyerea emetocathartica	47
Echinopepon cirrhopedunculatus	28

acceptedNameUsage	CuentaDecuenta
Echinopepon coulteri	61
Echinopepon milleflorus	72
Echinopepon minimus	32
Echinopepon pubescens	75
Echinopepon racemosus	101
Echinopepon wrightii	33
Gurania makoyana	14
Hanburia mexicana	39
Hanburia parviflora	6
Ibervillea hypoleuca	35
Ibervillea lindheimeri	47
Ibervillea millspaughii	26
Ibervillea sonorae	44
Lagenaria siceraria	28
Luffa cylindrica	64
Luffa operculata	36
Marah macrocarpa	11
Melothria pendula	461
Melothria trilobata	16
Microsechium palmatum	148
Momordica charantia	203
Parasicyos dieterleae	8
Peponopsis adhaerens	19
Polyclathra cucumerina	114
Psiguria triphylla	21
Rytidostylis gracilis	23
Rytidostylis longisepala	34
Schizocarpum filiforme	16
Schizocarpum liebmannii	23
Schizocarpum longisepalum	20
Schizocarpum palmeri	155
Schizocarpum parviflorum	91
Schizocarpum reflexum	46
Sechiopsis distincta	7
Sechiopsis tetraptera	45
Sechiopsis triquetra	101
Sechium compositum	23
Sechium chinantlense	18
Sechium edule	34
Sechium hintonii	12

acceptedNameUsage	CuentaDecuenta
Sechium mexicanum	17
Sicydium schiedeanum	51
Sicydium tamnifolium	152
Sicydium tuerckheimii	17
Sicyos barbatus	19
Sicyos chiriquensis	8
Sicyos galeottii	30
Sicyos guatemalensis	16
Sicyos laciniatus	96
Sicyos lirae	13
Sicyos longisepalus	65
Sicyos microphyllus	281
Sicyos parviflorus	46
Sicyos peninsularis	29
Sicyos sertulifer	10
Sicyosperma gracile	6
Tumamoca macdougalii	8
Vaseyanthus brandegeei	6
Vaseyanthus insularis	51

Anexo II. Valores de la prueba ROC (AUC) y umbral de corte

Especie	TRAINING	TEST	UMBRAL
<i>Apodanthera aspera</i>	0.979	0.970	0.271
<i>Apodanthera undulata</i>	0.900	0.914	0.406
<i>Brandegea bigelovii</i>	0.867	0.939	0.476
<i>Cayaponia attenuata</i>	0.913	0.845	0.512
<i>Cayaponia racemosa</i>	0.938	0.924	0.494
<i>Chalema synanthera</i>	0.967	0.945	0.332
<i>Cionosicycs excisus</i>	0.878	0.798	0.550
<i>Cionosicycs macranthus</i>	0.955	0.688	0.210
<i>Citrullus lanatus</i>	0.941	0.694	0.567
<i>Cucumis anguria</i>	0.953	0.863	0.380
<i>Cucumis dipsaceus</i>	0.967	0.927	0.277
<i>Cucumis melo</i>	0.921	0.841	0.408
<i>Cucumis sativus</i>	0.979	0.952	0.787
<i>Cucurbita argyrosperma</i>	0.893	0.873	0.481
<i>Cucurbita cordata</i>	0.949	0.844	0.412
<i>Cucurbita digitata</i>	0.809	0.724	0.705
<i>Cucurbita foetidissima</i>	0.893	0.838	0.518
<i>Cucurbita lundelliana</i>	0.929	0.845	0.614
<i>Cucurbita moschata</i>	0.938	0.963	0.406
<i>Cucurbita okeechobeensis</i>	0.965	0.896	0.380
<i>Cucurbita palmata</i>	0.943	0.826	0.335
<i>Cucurbita palmata</i>	0.953	0.930	0.570
<i>Cucurbita pedatifolia</i>	0.963	0.926	0.359
<i>Cucurbita pepo</i>	0.888	0.805	0.521
<i>Cucurbita radicans</i>	0.947	0.918	0.454
<i>Cucurbita scabridifolia</i>	0.954	0.940	0.522
<i>Cyclanthera dissecta</i>	0.883	0.837	0.536
<i>Cyclanthera integrifoliola</i>	0.938	0.917	0.460
<i>Cyclanthera langaei</i>	0.945	0.894	0.247
<i>Cyclanthera minima</i>	0.974	0.978	0.193

<i>Cyclanthera multifoliolata</i>	0.906	0.861	0.465
<i>Cyclanthera ribiflora</i>	0.941	0.869	0.409
<i>Cyclanthera rostrata</i>	0.965	0.976	0.657
<i>Cyclanthera steyermarkii</i>	0.976	0.721	0.311
<i>Cyclanthera tamnoides</i>	0.953	0.942	0.380
<i>Dieterlea fusiformis</i>	0.969	0.951	0.250
<i>Dieterlea maxima</i>	0.949	0.877	0.358
<i>Doyerea emetocathartica</i>	0.928	0.879	0.505
<i>Echinopepon cirrhopedunculatus</i>	0.923	0.946	0.413
<i>Echinopepon coulteri</i>	0.934	0.917	0.359
<i>Echinopepon milleflorus</i>	0.938	0.928	0.396
<i>Echinopepon minimus</i>	0.948	0.897	0.493
<i>Echinopepon pubescens</i>	0.931	0.900	0.384
<i>Echinopepon racemosus</i>	0.914	0.883	0.521
<i>Echinopepon wrightii</i>	0.965	0.704	0.272
<i>Gurania makoyana</i>	0.938	0.985	0.529
<i>Hanburia mexicana</i>	0.977	0.935	0.414
<i>Hanburia parviflora</i>	0.643	0.286	0.663
<i>Ibervillea hypoleuca</i>	0.978	0.978	0.188
<i>Ibervillea lindheimeri</i>	0.958	0.932	0.351
<i>Ibervillea millspaughii</i>	0.961	0.854	0.516
<i>Ibervillea sonora</i>	0.963	0.929	0.386
<i>Lagenaria siceraria</i>	0.880	0.695	0.575
<i>Luffa cylindrica</i>	0.908	0.897	0.513
<i>Luffa operculata</i>	0.969	0.956	0.425
<i>Marah macrocarpa</i>	0.978	0.998	0.370
<i>Melothria pendula</i>	0.826	0.797	0.597
<i>Melothria trilobata</i>	0.849	0.921	0.446
<i>Microsechium palmatum</i>	0.919	0.930	0.445
<i>Momordica charantia</i>	0.903	0.885	0.497
<i>Parasicyos dieterleae</i>	0.987	0.984	0.747
<i>Peponopsis adhaerens</i>	0.952	0.890	0.318

<i>Polyclathra cucumerina</i>	0.923	0.883	0.499
<i>Psiguria triphylla</i>	0.935	0.849	0.473
<i>Rytidostylis gracilis</i>	0.958	0.904	0.402
<i>Rytidostylis longisepala</i>	0.947	0.895	0.507
<i>Schizocarpum filiforme</i>	0.907	0.983	0.527
<i>Schizocarpum liebmannii</i>	0.978	0.931	0.343
<i>Schizocarpum longisepalum</i>	0.986	0.853	0.409
<i>Schizocarpum palmeri</i>	0.882	0.988	0.481
<i>Schizocarpum parviflorum</i>	0.941	0.942	0.365
<i>Schizocarpum reflexum</i>	0.939	0.932	0.446
<i>Sechiopsis distincta</i>	0.957	0.801	0.533
<i>Sechiopsis tetraptera</i>	0.962	0.901	0.171
<i>Sechiopsis triquetra</i>	0.906	0.888	0.454
<i>Sechium chinantense</i>	0.990	0.967	0.265
<i>Sechium compositum</i>	0.987	0.995	0.280
<i>Sechium edule</i>	0.966	0.981	0.385
<i>Sechium hintonii</i>	0.932	0.911	0.627
<i>Sechium mexicanum</i>	0.620	0.842	0.658
<i>Sicydium schiedeanum</i>	0.945	0.910	0.399
<i>Sicydium tamnifolium</i>	0.909	0.831	0.516
<i>Sicydium tuerckheimii</i>	0.933	0.944	0.459
<i>Sicyos barbatus</i>	0.971	0.875	0.417
<i>Sicyos chiriquensis</i>	0.993	0.995	0.493
<i>Sicyos galeottii</i>	0.987	0.980	0.127
<i>Sicyos guatemalensis</i>	0.932	0.964	0.595
<i>Sicyos laciniatus</i>	0.887	0.781	0.479
<i>Sicyos lirae</i>	0.987	0.963	0.096
<i>Sicyos longisepalus</i>	0.944	0.849	0.387
<i>Sicyos microphyllus</i>	0.864	0.832	0.565
<i>Sicyos parviflorus</i>	0.956	0.955	0.479
<i>Sicyos peninsularis</i>	0.953	0.966	0.865
<i>Sicyos sertulifer</i>	0.954	0.938	0.417
<i>Tumamoca macedougali</i>	0.960	0.935	0.322

Vaseyanthus brandegeei	0.991	0.987	0.663
Vaseyanthus insularis	0.981	0.957	0.486

Anexo III. Especies con valores inferiores de AUC a 0.800

Especie	TRAINING	TEST	UMBRAL
Cionosicycs excisus	0.878	0.798	0.550
Cionosicycs macranthus	0.955	0.688	0.210
Citrullus lanatus	0.941	0.694	0.567
Cucurbita digitata	0.809	0.724	0.705
Cyclanthera steyermarkii	0.976	0.721	0.311
Echinopepon wrightii	0.965	0.704	0.272
Hanburia parviflora	0.643	0.286	0.663
Lagenaria siceraria	0.880	0.695	0.575
Melothria pendula	0.826	0.797	0.597
Sechium mexicanum	0.620	0.842	0.658
Sicyos laciniatus	0.887	0.781	0.479