

Informe final* del Proyecto PE001

Diversidad genética de las especies de Cucurbita en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación*

Responsable:	Dr. Rafael Lira Saade
Institución:	Universidad Nacional Autónoma de México Facultad de Estudios Superiores Iztacala
Correo electrónico:	rlira@servidor.unam.mx
Fecha de inicio:	15 de marzo de 2018
Fecha de término:	31 de enero de 2022
Principales resultados:	Base de datos, Informe final
Forma de citar** el informe final y otros resultados:	Lira, S. R., Eguiarte, F. L., Montes, H. S., Aguirre, P. E., Gasca P. J., Sánchez, de la V. G., Castellanos, M. G., Barrera, R. J., Hernández, R. H. S., Villafán de la T. E., Tapia, A. F. y Monterrubio, L. B., 2021. "Diversidad genética de las especies de <i>Cucurbita</i> en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación". Universidad Nacional Autónoma de México. Facultad de Estudios Superiores Iztacala. Informe Final SNIB-CONABIO. Proyecto No. PE001. Ciudad de México.

Resumen:

México es centro de origen y diversificación del género *Cucurbita*. Dentro de sus límites geográficos se encuentran 15 taxa del género, de los cuales 5 son endémicos de México y otros 5 son taxa domesticados. Sin embargo, estos taxa habían sido poco estudiados desde el punto de vista genético hasta antes del desarrollo de la primera fase del presente proyecto. Durante la primera fase del proyecto se realizaron avances considerables en el análisis filogenético de todos los taxa (excepto *C. radicans*), así como en estudios poblacionales de tres taxa domesticados (*C. argyrosperma* ssp. *argyrosperma*, *C. moschata* y *C. pepo* ssp. *pepo*) y sus parientes silvestres más cercanos (*C. argyrosperma* ssp. *sororia* y *C. pepo* ssp. *fraterna*). De acuerdo con estos análisis, la diversificación del grupo ocurrió durante el Pleistoceno, siendo los grupos xerofíticos perennes los más basales y observándose una progresión de norte a sur en el área de distribución de los grupos. Adicionalmente, se detectaron niveles de variación genética considerable en los taxa domesticados y señales de flujo génico entre *C. argyrosperma* ssp. *argyrosperma* y el taxa silvestre *C. argyrosperma* ssp. *sororia*. Por otra parte, se ha avanzado considerablemente en el ensamble y anotación de los genomas de *C. argyrosperma* ssp. *argyrosperma*, *C. argyrosperma* ssp. *sororia* y *C. moschata*; datos que permitirán identificar regiones genómicas asociadas con el síndrome de domesticación y con la respuesta a distintas condiciones ambientales. En la presente fase del proyecto se obtendrán datos moleculares de regiones de mitocondria, cloroplasto y microsatélites nucleares para complementar los datos obtenidos durante la primera fase del proyecto en las especies cultivadas y sus parientes cercanos. Además, se obtendrán SNPs con el método tGBS para todas las especies del género distribuidas en México y se secuenciarán y analizarán genomas tanto de taxa domesticados como silvestres con el objetivo de: 1) Estudiar detalladamente la diversidad genética de las especies de *Cucurbita* que crecen en México y su relación con variables climáticas, incluyendo aspectos genómicos a nivel población; y 2) Avanzar en el estudio genómico de diferentes especies de *Cucurbita* y sus comparaciones en términos de evolución molecular. Se realizarán análisis de genética y genómica de poblaciones, así como filogenia y filogeografía de taxa silvestres y domesticados y de adaptación local para profundizar en la comprensión del proceso de domesticación de las especies del género *Cucurbita*, su biogeografía y demografía histórica y relaciones de ancestría-descendencia entre especies silvestres y domesticadas. También se van a llevar a cabo análisis de genómica comparativa, que permitan avanzar en la comprensión de los cambios genómicos asociados al síndrome de domesticación y adaptación a distintas condiciones ambientales.

-
- * 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 “Diversidad genética de las especies de *Cucurbita* en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación”

Proyecto FB1837/PE001/18

Responsables

Dr. Rafael Lira Saade

Profesor Titular C TC, Facultad de Estudios Superiores Iztacala, UNAM. Avenida de Los Barrios, #1 Colonia Los Reyes Iztacala, 54090, Tlalnepantla, Edo. de México.

Correo electrónico: rlira@servidor.unam.mx

Dr. Luis Enrique Eguiarte Fruns

Investigador Titular C TC, Departamento de Ecología Evolutiva, Instituto de Ecología, UNAM. Cd. Universitaria, Anexo al Jardín Botánico, 04510. Ciudad de México.

Correo electrónico: fruns@unam.mx

Dr. Salvador Montes Hernández

Investigador del Campo Experimental Bajío, INIFAP. Km. 6.5 Carr. Celaya-San Miguel de Allende. 38110, Celaya, Gto.

Participantes

Dra. Erika Aguirre Planter

Técnico Académico, Instituto de Ecología, UNAM

Correo electrónico: eaguirre@ecologia.unam.mx

Dr. Jaime Gasca Pineda

Investigador Postdoctoral, UBIPRO, FES-Iztacala, UNAM

Correo electrónico: jaimegasca@yahoo.com

M. en C. Guillermo Sánchez de la Vega

Estudiante de doctorado, Instituto de Ecología, UNAM

Correo electrónico: gsanchezdv@gmail.com

Dra. Gabriela Castellanos Morales

Investigadora asociada al proyecto,

Investigadora Asociada "B", Departamento de Conservación de la Biodiversidad, El Colegio de la Frontera Sur -Villahermosa.

Dr. Josué Barrera Redondo

Estudiante de doctorado, Instituto de Ecología, UNAM

Biól. Helena Socorro Hernández Rosales

Estudiante de doctorado, Instituto de Ecología, UNAM

M. en C. Carmina Martínez González

Estudiante de doctorado, FES-Iztacala, UNAM

M. en C. Emanuel Villafán de la Torre

Estudiante de doctorado, Instituto de Ecología, UNAM

Fernando Tapia Aguirre

Estudiante de licenciatura, FES-Iztacala, UNAM

Brenda Monterrubio Lezama

Estudiante de licenciatura, FES-Iztacala, UNAM

RESUMEN DEL PROYECTO

México es centro de origen y diversificación del género *Cucurbita*. En el país se encuentran 15 taxa del género, de los cuales cinco son endémicos de México y otros cinco son taxa domesticados. Los taxa domesticados habían sido poco estudiados desde el punto de vista genético y evolutivo antes del desarrollo de la primera fase del proyecto. En la primera fase se realizaron avances en el análisis filogenético, genético y evolutivo de todos los taxa mexicanos, así como en estudios poblacionales de tres taxa domesticados (*C. argyrosperma* ssp. *argyrosperma*, *C. moschata* y *C. pepo* ssp. *pepo*) y en dos de sus parientes silvestres cercanos (*C. argyrosperma* ssp. *sororia* y *C. pepo* ssp. *fraterna*). En esta segunda fase se han obtenido datos moleculares de regiones de mitocondria, cloroplasto y microsatélites nucleares para completar la información obtenida durante la primera fase en las especies cultivadas y sus parientes silvestres. Adicionalmente, con los objetivos de *i*) estudiar detalladamente la diversidad genética de las especies de *Cucurbita* que crecen en México y su relación con variables climáticas, incluyendo aspectos genómicos a nivel población, y *ii*) avanzar en el estudio genómico de diferentes especies de *Cucurbita* y sus comparaciones en términos de evolución molecular, obtuvimos marcadores genómicos (SNPs) a nivel poblacional usando métodos de secuenciación masiva (GBS) para las especies silvestres y domesticadas del género en México y estamos secuenciando y analizando genomas tanto de taxa domesticados como silvestres. Analizando la genética y genómica de poblaciones para diferentes especies domesticadas y silvestres, generamos por primera vez una filogenia incluyendo a todos los taxa del género. Iniciamos la descripción filogeográfica de taxa silvestres y domesticados y análisis de su adaptación local, para así profundizar en la comprensión del proceso de domesticación y evolutivo del género *Cucurbita*. Describimos aspectos de la biogeografía y demografía histórica, así como relaciones de ancestría-descendencia entre especies silvestres y domesticadas. También avanzamos en análisis de genómica comparada que están permitiendo comprender los cambios genómicos asociados a la adaptación a distintas condiciones ambientales en las especies silvestres y en las variedades locales de las especies cultivadas y aspectos de su domesticación. Además, hemos avanzado en la escritura y publicación de siete artículos científicos de primer nivel y tres de divulgación nacional, así como en tesis a diferentes niveles.

INTRODUCCIÓN

México es centro de origen y diversidad del género *Cucurbita*, ya que en sus límites geográficos se encuentran 15 de los 20 taxa que lo conforman. De forma tradicional, estos 20 taxa han sido divididos en dos grandes grupos: el de las especies xerofíticas, con raíces tuberiformes perennes de almacenamiento, y el de las especies mesofíticas, anuales o perennes de vida corta, con raíces fibrosas. Dentro de los 15 taxa mexicanos podemos encontrar a cuatro de los cinco taxa domesticados del género (*C. argyrosperma* ssp. *argyrosperma*, *C. pepo* ssp. *pepo*, *C. ficifolia* y *C. moschata*) y 11 taxa silvestres. Antes del inicio de la primera fase de este proyecto, era evidente la existencia de vacíos de información y la necesidad de estudios específicos acerca de los taxa silvestres nativos de México, así como la falta de actualización en los análisis genéticos realizados hasta ese momento.

Dentro del género, México es importante respecto a la distribución de los taxa silvestres, ya que la mayoría de estos (11 de 14) crece de manera natural en nuestro país: *C. cordata*, *C. palmata* y *C. digitata* (grupo Digitata) en la península de Baja California, Sonora y Chihuahua; *C. radicans*, *C. foetidissima*, *C. pedatifolia* y *C. scabridifolia* (grupo Foetidissima) en el Eje Neovolcánico, la Sierra Madre del Sur y el desierto Chihuahuense, *C. argyrosperma* ssp. *sororia* (grupo Argyrosperma), en la costa del Pacífico y la costa del Golfo de México; *C. okeechobeensis* ssp. *martinezii* y *C. lundelliana* (grupo Lundelliana) en la Sierra Madre Oriental, la península de Yucatán y en la región del Golfo de México y *C. pepo* ssp. *fraterna* (grupo Pepo) en Tamaulipas.

Durante la primera fase de este proyecto se trabajó principalmente con cuatro taxa domesticados (*C. pepo* ssp. *pepo*, *C. argyrosperma* ssp. *argyrosperma*, *C. moschata*, y *C. ficifolia*) y dos de sus parientes silvestres más cercanos (*C. argyrosperma* ssp. *sororia* y *C. pepo* ssp. *fraterna*).

En esta segunda fase del proyecto se concluyeron los análisis de cloroplasto, mitocondria y microsatélites que se iniciaron con taxa cultivados en la fase previa, se finalizó con la colecta y se ha avanzado considerablemente con la obtención de datos de todos los parientes silvestres del género distribuidos en México.

Nuestros resultados filogenéticos y de evolución molecular (Castellanos-Morales et al., 2018; Aguirre-Dugua et al., 2019) indican que los distintos taxa se ubican en grupos bien definidos: las especies xerofíticas perennes aparecen en una posición basal en las reconstrucciones filogenéticas, con tiempos de divergencia aparentemente asociados a cambios climáticos en el pasado. Los análisis muestran que en *C. pepo* han ocurrido al menos dos eventos de domesticación independiente: uno en México y otro en el territorio de E.U.A. y se detecta la filiación (previamente propuesta en otros estudios) de *C. moschata* con el grupo *Argyrosperma* (Castellanos-Morales et al., 2018, 2019).

Los análisis a nivel poblacional con cloroplasto, mitocondria y núcleo muestran que las variedades de *C. argyrosperma* ssp. *argyrosperma*, *C. moschata* y *C. pepo* ssp. *pepo* cultivadas en México se encuentran bien diferenciadas, con poco o nulo flujo génico entre ellas, mientras que se detectan señales de flujo génico entre la especie cultivada *C. argyrosperma* ssp. *argyrosperma* y su contraparte silvestre *C. argyrosperma* ssp. *sororia*. Tanto en *C. argyrosperma* como en *C. moschata* se observa una clara diferencia de las variedades cultivadas en la península de Yucatán. *Cucurbita pepo* ssp. *pepo* se encuentra bien diferenciada de su contraparte silvestre, *C. pepo* ssp. *fraterna* (Sánchez de la Vega et al., 2018; Castellanos-Morales et al., 2018, 2019; Hernández-Rosales et al., 2020).

Asimismo, se generaron datos genéticos a nivel poblacional con marcadores moleculares producto de secuenciación masiva (*Genotyping by sequencing*, GBS), se ensamblaron y analizaron los genomas de referencia de algunas especies silvestres y se realizaron análisis detallados de genómica comparada entre especies silvestres y cultivadas de México (Barrera-Redondo et al., 2019, 2021; Martínez-González et al., 2021).

En resumen, en este proyecto hemos logrado avances en la caracterización y la determinación de la distribución de la diversidad genética de las diferentes especies del género *Cucurbita* que crecen en México mediante el uso de marcadores moleculares y de datos genómicos. Una vez que se concluya el análisis de los datos poblacionales de los taxa silvestres será posible: *i)* analizar aspectos de la adaptación en la distribución de cada una de las especies y *ii)* proponer estrategias para la conservación de la diversidad

genética de las especies silvestres y domesticadas del género *Cucurbita* que crecen en México.

OBJETIVOS GENERALES DEL PROYECTO (SEGUNDA ETAPA)

- 1) Estudiar detalladamente la diversidad genética de las especies de *Cucurbita* que crecen en México y su relación con variables climáticas, incluyendo aspectos genómicos a nivel población.
- 2) Avanzar en el estudio genómico de diferentes especies de *Cucurbita* y sus comparaciones en términos de evolución molecular.

OBJETIVOS ESPECÍFICOS DEL PROYECTO Y ACTIVIDADES A REALIZAR

Estudios de diversidad genética en *Cucurbita*

- 1) Realizar colectas poblacionales de plantas del género *Cucurbita*, especialmente de especies silvestres.
- 2) Concluir la filogenia de *Cucurbita*, incluyendo a *C. radicans*, y realizar análisis tipo reloj molecular, incluyendo caracteres de cloroplasto y nucleares derivados de los estudios genómicos.
- 3) Determinar la distribución de la diversidad genética de las especies del género *Cucurbita* que crecen en México mediante el uso de marcadores moleculares y datos genómicos.
- 4) Proponer estrategias para apoyar en la conservación de la diversidad genética de las especies silvestres y domesticadas del género *Cucurbita* que crecen en México.

Estudios genómicos en *Cucurbita*

- 5) Concluir el ensamble de los genomas completos de *C. argyrosperma* ssp. *sororia* y *C. moschata* para realizar análisis de genómica comparada.
- 6) Obtener genomas de tres especies silvestres: *C. okechobeensis* ssp. *martinezii*, *C. digitata* y *C. radicans* para realizar análisis de genómica comparada.
- 7) Con estos genomas anotar SNPs obtenidos por tGBS para análisis poblacionales y buscar correlaciones entre estos SNPs y patrones ambientales para detectar señales de selección.

- 8) Avanzar en entender la evolución del genoma de las especies que conforman el grupo *Argyrosperma* dentro del género *Cucurbita*, los cambios genéticos en la domesticación y en diferentes características como forma y tamaño del fruto, sabor amargo, y tamaño de semilla.
- 9) Avanzar en entender la evolución del genoma de especies silvestres del género *Cucurbita*, los cambios asociados a hábitos perennes y anuales, así como los cambios genéticos asociados a distintas condiciones ambientales (xerófilos y mesófilos).

METAS ESPECÍFICAS DE LA SEGUNDA ETAPA DEL PROYECTO:

Estudios de diversidad genética

- 1) Realizar colectas poblacionales de plantas del género *Cucurbita*, especialmente en Baja California (*C. cordata*, *C. digitata* y *C. palmata*), así como en la región de centro, oeste y norte de México (*C. pedatifolia*, *C. foetidissima* y *C. radicans*) y en la región del sur de México y península de Yucatán (*C. okechobeensis* ssp. *martinezii* y *C. lundelliana*), para así tener más poblaciones de los taxa silvestres que ayuden a documentar y comparar los niveles de variación y estructura genética en los taxa silvestres y en los domesticados de este género.
- 2) Concluir análisis filogenéticos de *Cucurbita* incluyendo a *C. radicans* y realizar análisis tipo reloj molecular, buscando incluir caracteres de cloroplasto y nucleares (derivados de los estudios genómicos) para avanzar en la comprensión de la dinámica evolutiva (tiempos, patrones geográficos y tasas de cambio) del género.
- 3) Terminar de amplificar microsatélites nucleares y regiones del cloroplasto y de la mitocondria en las distintas especies silvestres y domesticadas del género en México que se trabajaron durante la primera fase del proyecto. Con estos datos vamos a estimar diferentes parámetros genéticos y evolutivos, como los niveles de endogamia de cada población, los niveles de flujo génico y estimar sus tamaños efectivos actuales y en el pasado, e inferir aspectos de su demografía evolutiva (si

las poblaciones han sido estables o han cambiado de tamaño, y cuando) con simulaciones y análisis de coalescencia.

- 4) Realizar mapas detallados de la distribución de la variación genética, y propuestas de conservación del género *Cucurbita*.

Estudios genómicos en Cucurbita

- 5) Obtener datos de PacBio para complementar el ensamble de los genomas de *C. argyrosperma* ssp. *sororia* y *C. moschata*, y así realizar análisis detallados de genómica comparada.
- 6) Obtener datos HiSeq y PacBio para ensamblar los genomas de *C. okeechobeensis* ssp. *martinezii*, *C. digitata* y *C. radicans* que serán utilizados como grupos externos en análisis de genómica comparada y servirán para analizar aspectos relacionados con la adaptación en los taxa silvestres.
- 7) Obtener datos de tGBS de las especies domesticadas *C. argyrosperma* ssp. *argyrosperma*, *C. moschata* y *C. pepo* ssp. *pepo* y de especies silvestres para las que contamos con muestras y para las nuevas colectas, con énfasis en los ancestros de los taxa domesticados y las especies endémicas de México. De esta manera tendremos una buena representación de la variación genética y podremos hacer estudios detallados de demografía histórica y de selección en las especies cultivadas y silvestres. Asimismo, se podrán comparar los niveles de variación genética e identificar SNPs de interés para el mejoramiento de los taxa cultivados y realizar estudios de selección y adaptación local, como puede ser resistencia a enfermedades y/o la respuesta a variables ambientales en los taxa de este género que se distribuyen en México.
- 8) Realizar análisis de genómica comparativa de los datos de secuenciación masiva *next-gen* de los genomas (HiSeq y PacBio) de taxa seleccionados del género (5 genomas) para explorar las similitudes y diferencias, así como identificar paralelismos a nivel genómico entre taxa silvestres y cultivados. Esto nos permitirá

profundizar en la comprensión de los cambios evolutivos relacionados con el proceso de domesticación y adaptación.

Generales

- 9) Construir una base de datos de la diversidad genética presente en las poblaciones de *Cucurbita* en México.
- 10) Concluir la elaboración de los proyectos entregables indicadores de éxito del proyecto.
- 11) Hacer informe final para la CONABIO.

RESULTADOS, SIGUIENDO LOS OBJETIVOS, METAS E INDICADORES DE ÉXITO DEL PROYECTO

1. Colectas poblacionales, base de datos de los registros obtenidos, germinación y extracción de ADN.

Durante esta segunda etapa, se realizaron 16 salidas de campo a 28 estados del país, donde se colectaron un total de 309 poblaciones/localidades de 14 taxa, principalmente silvestres. En la Tabla 1 se muestran los taxa colectados en las fases I y II de este proyecto, resultando en total en la colecta de 1,724 frutos. Cabe mencionar que se realizaron colectas para especies que hasta la fecha eran poco frecuentes en bancos de germoplasma como *C. radicans*, *C. pedatifolia*, *C. scabridifolia* y *C. pepo* ssp. *fraterna*.

Tabla 1. Número de registros de poblaciones/localidades obtenidas en las salidas a campo e individuos con muestras de ADN de los taxa del género *Cucurbita* en México durante las dos etapas del proyecto.

Especie /Taxa	Número de localidades	Frutos Colectados	Muestras de ADN (localidades)
<i>C. foetidissima</i>	65	344	26 (13)
<i>C. pedatifolia</i>	14	67	103 (14)
<i>C. radicans</i>	22	144	96 (14)
<i>C. scabridifolia</i>	2	6	0
<i>C. cordata</i>	18	45	68 (17)
<i>C. palmata</i>	0	0	0
<i>C. digitata</i>	2	4	4 (2)
<i>C. lundelliana</i>	25	101	78 (25)
<i>C. okechobeensis</i> ssp. <i>martinezii</i>	21	109	33 (10)
<i>C. argyrosperma</i> ssp. <i>sororia</i>	30	387	134 (7)
<i>C. argyrosperma</i> ssp. <i>argyrosperma</i>	32	259	352 (25)
<i>C. ficifolia</i>	15	18	32 (14)
<i>C. pepo</i> ssp. <i>fraterna</i>	2	10	29 (1)
<i>C. pepo</i> ssp. <i>pepo</i>	14	32	291 (13)
<i>C. moschata</i>	47	198	256 (24)

TOTAL	309	1,724	1,567 (208)
--------------	------------	--------------	--------------------

Los frutos colectados pasaron por un proceso de secado, extracción de semillas y etiquetado para su conservación temporal en el laboratorio de Evolución Molecular y Experimental del Instituto de Ecología (UNAM). Se depositarán de forma definitiva en el banco de germoplasma de la FES Iztacala y una vez que se cuente con la carta de depósito, será enviada a la CONABIO. A partir de las colectas realizadas, ahora se cuenta con una de las colecciones de semillas de *Cucurbita* silvestres más completas, con suficientes poblaciones de cada taxa, lo que permite tener una buena representación de su distribución y diversidad. Las únicas especies que no fue posible coleccionar de forma extensiva y del que se tiene menor representación fueron *C. palmata* y *C. digitata*, que se distribuyen en los estados de Baja California y Sonora. No fue posible realizar la salida programada en el mes de mayo de 2020 debido a la situación sanitaria del país, por el COVID-19.

Los datos referentes a las colectas poblacionales durante las dos etapas del proyecto han sido recopilados en una base de datos en formato Excel, la cual se entregará junto con este informe final. La base de datos cuenta con 145 registros de la primera etapa (2013-2017) y 187 de la segunda, para sumar 332 registros de 15 taxa distribuidos en México. Los datos han sido recopilados tal y como se especifica en el punto 3 de los términos de referencia del Apéndice 3 del Convenio específico FB1837/PE001/18 como se anexan (ver Base de datos colectas, Apéndice 1).

Adicionalmente, durante estas salidas se colectaron y procesaron 70 ejemplares de herbario, que serán incorporados al Herbario Nacional MEXU en el Instituto de Biología de la UNAM y una vez que se cuente con la carta de recepción de los ejemplares, esta será enviada a la CONABIO. Las localidades de origen de todos estos ejemplares de herbario están incluidas en la base de datos de registros de colecta entregada.

La germinación de semillas y la posterior extracción de ADN se ha realizado con diferentes poblaciones de cada taxa. El detalle del número de las muestras usadas en los análisis de diversidad genética se puede observar en la Tabla 1. En total se extrajo ADN y

se secuenciaron con diferentes marcadores (cloroplasto, mitocondria, microsátélites y SNPs), un total de 1,567 Individuos de 208 localidades de 14 taxa.

2. Filogenia de *Cucurbita*, incluyendo a *C. radicans*, y análisis de reloj molecular, buscando incluir caracteres de cloroplasto y nucleares (derivados de los estudios genómicos).

Este objetivo se cumplió exitosamente con el estudio filogenético del género *Cucurbita*, que incluye a *C. radicans*. Como parte de esta sección se realizó la tesis de Licenciatura de la alumna Leslie Mariel Paredes, quien obtuvo el grado en 2017. Se realizaron análisis detallados de reloj molecular y tiempos de divergencia que incluyeron caracteres de cloroplasto y nucleares. El resultado final de este objetivo fue el artículo científico “Historical biogeography and phylogeny of *Cucurbita*: insights from ancestral area reconstruction and niche evolution” que ya fue publicado en la revista *Molecular Phylogenetics and Evolution* (Castellanos Morales et al., 2018).

En este artículo analizamos la filogenia con datos moleculares (cloroplasto) mediante métodos de máxima verosimilitud (con PhyML) y Bayesianos (con MrBayes) y se calibraron y calcularon los tiempos de divergencia de los diferentes clados. También se examinan las tasas de diversificación de cada uno de los taxa. En este mismo artículo utilizamos a la filogenia como una herramienta para realizar la reconstrucción de áreas ancestrales, para hacer pruebas de diferenciación ambiental en cada uno de los grupos y para obtener los modelos de distribución al presente y al pasado de cada uno de los taxa. Anexamos el artículo con los agradecimientos al proyecto (Castellanos-Morales et al., 2018, Apéndice 2).

3. Determinar la distribución de la diversidad genética de las especies del género *Cucurbita* que crecen en México mediante el uso de marcadores moleculares y datos genómicos.

En la primera fase del proyecto quedó pendiente concluir la obtención de secuencias de microsátélites, mitocondria, cloroplasto y SNPs de *C. moschata*, *C. argyrosperma* ssp. *argyrosperma*, *C. argyrosperma* ssp. *sororia*, y *C. pepo* ssp. *pepo*. Además, en esta segunda

fase se buscaba obtener y analizar SNPs de especies silvestres, y la realización de mapas de distribución de la diversidad genética de las especies.

Ya concluimos los análisis de la diversidad genética con diferentes marcadores como microsatélites, cloroplasto y mitocondria de las especies cultivadas y silvestres asociadas. Los artículos que utilizan estos marcadores han sido publicados y/o están en etapa de revisión para envío a publicación.

Así, en *C. argyrosperma* se realizó el análisis de 19 poblaciones de *C. argyrosperma* ssp. *argyrosperma* y seis de su pariente silvestre *C. argyrosperma* ssp. *sororia* usando nueve loci de microsatélites nucleares. Este estudio forma parte del artículo "Genetic Resources in the "Calabaza Pipiana" Squash (*Cucurbita argyrosperma*) in Mexico: Genetic Diversity, Genetic Differentiation and Distribution Models", publicado en la revista *Frontiers in Plants Science* (Sánchez-de la Vega et al. 2018, Apéndice 2); estos resultados son parte de la tesis de doctorado de Guillermo Sánchez de la Vega (Apéndice 2), en proceso de titulación.

Los datos del análisis de 12 loci de microsatélites nucleares de 13 localidades de México y de cinco variedades comerciales de *C. pepo* son parte del artículo "Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico", el cual fue publicado los *Proceedings of the Royal Society B* (Castellanos-Morales, et al. 2019) que se anexa en el Apéndice 2.

También se analizaron 11 loci de microsatélites nucleares de 24 variedades locales de *C. moschata* distribuidas a lo largo de México. Su análisis se describe en el artículo "Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient" publicado en el *American Journal of Botany* 107(3) 1-16 (Hernández-Rosales, et al. 2020 en el Apéndice 2); estos resultados son parte de la tesis de doctorado de Helena Socorro Hernández Rosales (Apéndice 2).

3.1. Cloroplasto y mitocondria.

Se concluyó la obtención y análisis de marcadores moleculares de cloroplasto y mitocondria en las especies pendientes (*C. argyrosperma* ssp. *argyrosperma*, *C. argyrosperma* ssp. *sororia*, *C. moschata* y *C. pepo* ssp. *pepo*).

En *C. argyrosperma* ssp. *sororia* se estudiaron 5,314 pares de bases (pb) de cuatro regiones del cloroplasto (*atpI-atpH*, *psbJ-petA*, *petL-psbE* y *psbD-trnT*) y 920 pb de la región mitocondrial *trnL-trnF* en 139 individuos de 18 localidades a lo largo de su distribución completa. Los resultados forman parte de la tesis de Licenciatura de Fernando Tapia Aguirre (FES Iztacala, UNAM), quien se encuentra en proceso de titulación (la tesis completa se incluye en el Apéndice 2).

En *C. moschata* se analizó un total de 2,801 pb correspondientes a 3 regiones del cloroplasto (*rpl20-rps12*, *psbJ-petA*, *atpI-atpH*) y 1,027 pb de la mitocondria (región *trnL-trnF*) en 256 individuos de 24 poblaciones a lo largo de su distribución. Estos datos se analizan en la tesis doctoral de Helena S. Hernández Rosales, quien se encuentra en proceso de titulación, así como en su artículo “Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient” publicado en el *American Journal of Botany* (Hernández-Rosales, et al., 2020, Apéndice 2).

En el artículo “Evolutionary dynamics of transferred sequences between organellar genomes in *Cucurbita*” publicado en el *Journal of Molecular Evolution* de Aguirre-Dugua et al. 2019 (ver Apéndice 2), se realizaron análisis de evolución molecular y filogenéticos de 29 regiones de ADN. Este trabajo permitió identificar eventos de transferencia entre organelos (mitocondria y cloroplasto) en la filogenia de *Cucurbita* y cambios en sus tasas de sustitución de nucleótidos.

3.2. GBS y obtención de SNPs.

Se concluyeron los estudios de genómica poblacional usando GBS para las especies cultivadas *C. argyrosperma* ssp. *argyrosperma*, *C. argyrosperma* ssp. *sororia*, *C. moschata* y *C. pepo*.

Para obtener estos SNPs, inicialmente durante esta etapa del proyecto, continuamos con la germinación de semillas, y con la extracción y cuantificación de ADN. A partir del ADN de alta calidad, se obtuvieron datos tipo tGBS de diversas especies.

En particular, los SNPs de *C. argyrosperma* se utilizaron para determinar la historia demográfica durante su domesticación. Se usaron muestras recolectadas previamente en todo México correspondientes a 117 individuos de *argyrosperma*, 50 individuos de *sororia*, 19 individuos asilvestrados de la *ssp. argyrosperma* previamente reportados con un fenotipo feral y un genotipo cultivado. Se obtuvieron 2861 SNPs neutrales que se usaron para los análisis demográficos; se obtuvieron parámetros de diversidad genética, se evaluaron las relaciones genealógicas y la estructura genética entre poblaciones silvestres y cultivadas. Asimismo, se generaron modelos de coalescencia para indicar las regiones probables donde fue domesticada la especie. Este estudio forma parte del artículo publicado “The domestication of *Cucurbita argyrosperma* as revealed by the genome of its wild relative.” En la revista *Horticulture Research* (Barrera-Redondo et al. 2021, Apéndice 2). Secciones de este artículo forman parte de las tesis doctorales de Josué Barrera Redondo y de Guillermo Sánchez de la Vega (Apéndice 2).

Los SNPs de *Cucurbita pepo* fueron analizados como parte de la tesis de maestría de la alumna Carmina Martínez González “Diversidad y estructura genética en loci neutrales de *Cucurbita pepo* subsp. *pepo* en México”, quien se tituló en diciembre de 2019 (ver Martínez González, Tesis M. C. en Apéndice 2). Actualmente, este trabajo se encuentra en proceso de publicación en *Frontiers in Ecology and Evolution* con el título: “Recent and historical gene flow in cultivated zucchini (*C. pepo* ssp. *pepo*) and its wild relative, *C. pepo* ssp. *fraterna* in Mexico” (ver Martínez-González et al. 2021, Apéndice 2).

Los SNPs de *Cucurbita Moschata* se han analizado parcialmente y se tienen patrones preliminares de 96 individuos de 24 localidades de todo el país, que se genotiparon utilizando la tecnología de secuenciación tGBS. En total se obtuvieron 1,797 SNPs neutrales y con estos marcadores se obtuvieron los estadísticos básicos de diversidad. Además, se obtuvo la estructura genética y se realizaron análisis de asignación y PCA para determinar los grupos genéticos en las muestras. Una vez que se lleven a cabo

los análisis finales y se tengan conclusiones puntuales se escribirá un artículo científico. estos análisis forman parte de la tesis de Doctorado de Helena Hernández Rosales (Apéndice 2).

Adicionalmente se están realizando los análisis demográficos de diferentes especies silvestres, como se explica más adelante en el presente informe.

3.3 Congreso de Ecología.

La presentación de los primeros resultados de los análisis de SNPs para diferentes especies del género *Cucurbita* se hizo en el VII Congreso Mexicano de Ecología, realizado en Querétaro, del 29 de septiembre al 4 de octubre de 2019. Se organizó un simposio específico titulado “*Cucurbita* (las calabazas), un género modelo para el estudio de la ecología, la evolución y la domesticación” en colaboración con otros grupos de investigación dedicados a este género. En total se presentaron ocho ponencias orales de diferentes especies en dicho simposio.

4. Estrategias para apoyar en la conservación de la diversidad genética de las especies silvestres y domesticadas del género *Cucurbita* que crecen en México.

En los artículos científicos que hemos publicado y en los diferentes productos, como son las tesis donde derivadas de este proyecto se discuten la conservación de los diferentes taxa estudiados. Adicionalmente, para difundir y generar conciencia de la importancia de las calabazas, hemos publicado tres artículos de divulgación sobre las calabazas.

El primer artículo de divulgación fue publicado al final de la primera fase del proyecto en la revista de difusión científica del Instituto de Ecología de la UNAM, *Oikos* = (ver Sánchez de la Vega, 2017, Apéndice 2) donde se hace un resumen del uso del género desde su aparición en registros arqueo-botánicos hasta los análisis genómicos que

estamos realizando actualmente. Uno de los objetivos principales es sensibilizar al lector sobre la importancia de los recursos fitogenéticos y su conservación.

El segundo artículo de divulgación fue publicado en la Revista Especializada en Ciencias Químico-Biológicas “TIP” de la ENEP Zaragoza titulado “Domesticación y recursos genéticos y genómicos de México: el caso de las Calabazas” (ver Eguiarte et al., 2018, Apéndice 2). Cabe destacar que una fotografía del proyecto fue la portada de este número de la revista.

Un tercer artículo de divulgación trata sobre la diversidad e importancia en México de los parientes silvestres de *Cucurbita* titulado “Los parientes pobres de la agricultura: las calabazas silvestres, riqueza para nuestro futuro.” El artículo ha sido publicado en Biodiversitas (ver Xitlali Aguirre-Dugua et al. 2020, Apéndice 2).

Diez fichas de diferentes taxa del género *Cucurbita* de la lista Roja (Red list UICN), publicadas y puestas a disposición desde agosto de 2019, fueron generadas con la participación de integrantes del proyecto en febrero de 2017, como parte del grupo especialista en el «Taller regional para la evaluación de riesgo de especies de parientes silvestres de cultivos para la Lista Roja de la UICN». Este evento fue organizado por la CONABIO y la UICN dentro del marco del proyecto “Salvaguardando parientes silvestres de cultivos mesoamericanos” financiado por la Iniciativa Darwin del Reino Unido. Las 10 fichas en las que se colaboró pueden consultarse en:

<https://www.iucnredlist.org/search?query=Cucurbita>).

Asimismo, producto también del taller, se colaboró en el artículo “Extinction risk of Mesoamerican crop wild relatives”, en el cual se analizaron diferentes escenarios y riesgos a los que se enfrentan los taxa silvestres de *Cucurbita*, entre otros cultivos. Este trabajo está en proceso de publicación en la revista *Plant, People, Planet* y es coordinado por la Dra. Bárbara Goettsch de la UICN.

En este rubro, el trabajo desarrollado hasta la fecha y los resultados obtenidos en este proyecto han permitido colaborar activamente en la CONABIO con: i) la elaboración y revisión de las fichas correspondientes a los taxa de *Cucurbita* para la plataforma digital

Enciclovida (2019), *ii*) en el cartel de difusión de los parientes silvestres de cultivos de Mesoamérica (2019) y *iii*) en el cartel de difusión de diversidad de calabazas en México (2020).

5. Análisis de genómica comparada de *C. argyrosperma* ssp. *sororia*, *C. argyrosperma* ssp. *argyrosperma* y *C. moschata*.

El genoma de *Cucurbita argyrosperma* comenzó a ser ensamblado *de novo* en la primera fase del proyecto. Su ensamble se concluyó en la segunda fase del proyecto. Con este genoma como referencia se comparó el contenido de genes con otros cinco miembros domesticados de la familia Cucurbitaceae, y además se analizaron los genomas de otras tres especies domesticadas de *Cucurbita*. Se evaluó la dinámica evolutiva de codificación de proteínas, de secuencias largas intergénicas de ARN no codificante (lincRNA). Se encontró que las familias de genes de *C. argyrosperma* asociadas con la polinización y el transporte de transmembrana tuvieron tasas evolutivas significativamente más rápidas. Los resultados sugieren que la duplicación del genoma completo en *Cucurbita* resultó en tasas más rápidas de evolución de la familia de genes a través de la neofuncionalización de genes duplicados. Este trabajo se publicó en *Molecular plant* en 2019 (The Genome of *Cucurbita argyrosperma* (silver-seed gourd) reveals faster rates of protein-coding gene and long noncoding RNA turnover and neofunctionalization within *Cucurbita*; Barrera-Redondo et al., 2019 Apéndice 2).

Se realizó la secuenciación de los genomas de *C. argyrosperma* ssp. *sororia*, de un ejemplar de Puerto Escondido, Oaxaca, y de un ejemplar de *C. moschata* de Jiquipilas, Chiapas. De ambas plantas primero se obtuvo ADN de alta calidad que se envió a secuenciar en la plataforma Illumina en la Universidad de Berkeley. Adicionalmente se envió a secuenciar el genoma de *C. argyrosperma* ssp. *sororia* en tres celdas en la plataforma PacBio Sequel, obteniendo una profundidad de secuenciación de 75.4x. Para el genoma de *C. argyrosperma* ssp. *sororia*, se ensamblaron en *contigs* cortos y posteriormente se utilizaron las lecturas de PacBio para ensamblar *contigs* más largos. El

genoma nuclear de *C. argyrosperma* ssp. *sororia* se ensambló en 828 contigs, correspondientes a un tamaño acumulativo de 253 Mbp y una N50 de 1.3 Mbp.

Asimismo, se terminó de ensamblar el cloroplasto completo de *C. argyrosperma* ssp. *sororia* (silvestre); estos genomas se utilizaron para realizar un análisis de genómica comparada con *C. argyrosperma* ssp. *argyrosperma* (domesticada) para identificar cambios estructurales asociados al proceso de domesticación, utilizando el genoma de *C. moschata* como grupo externo. Estos análisis se realizaron en conjunto con los análisis de tGBS para generar una publicación enfocada a la domesticación de *C. argyrosperma* en Mesoamérica. El artículo producto de dicho trabajo ya ha sido publicado en la revista *Horticulture Research* (Barrera-Redondo et al. 2021, Apéndice 2). Secciones de este artículo forman parte de las tesis doctorales de Josué Barrera Redondo y de Guillermo Sánchez de la Vega (Apéndice 2). Adicionalmente se publicó la reseña “Gourds and Tendrils of Cucurbitaceae: How Their Shape Diversity, Molecular and Morphological Novelties Evolved via Whole-Genome Duplications” en *Molecular Plant* (Ver Barrera-Redondo, et al., 2020a en Apéndice 2).

6. Secuenciación de *C. okeechobeensis* ssp. *martinezii*, *C. digitata*, *C. ficifolia* y *C. radicans*.

Se extrajo ADN de alta calidad para su secuenciación genómica de las especies *C. ficifolia*, *C. okeechobeensis* ssp. *martinezii*, *C. digitata* y *C. radicans*. Estas muestras se enviaron secuenciar mediante métodos secuenciación paralela masiva, con el objetivo de ensamblar *de novo* los genomas de estos taxa. A partir del proceso de secuenciación a través de la plataforma de Illumina HiSeq4000 se obtuvieron alrededor de 30 millones de lecturas para las muestras secuenciadas de *C. digitata* y *C. o. ssp. martinezii*, de las cuales cerca del 97% presentaron parámetros de calidad adecuados para ser utilizados en el proceso de ensamblado de los genomas correspondientes (Tabla 2). En cuanto a las lecturas obtenidas por la plataforma de PacBio Sequel, para *C. digitata* se obtuvieron 5,112,097 y 8,265,255 para *C. o. ssp. martinezii*, las cuales fueron procesadas y utilizadas

junto con las lecturas provenientes de la plataforma de Illumina para ensamblar los genomas de ambas especies.

Tabla 2. Cantidad de lecturas de Illumina disponibles antes y después del proceso de filtrado de calidad para cada una de las especies ensambladas.

Espece	Número de lecturas crudas	Número de lecturas de alta calidad	% de lecturas de alta calidad
<i>C. digitata</i>	36,129,636	35,075,032	97.08
<i>C. okechobeensis ssp. martinezii</i>	32,794,285	31,814,727	97.01

Después de realizar un proceso de ensamblado jerárquico (de secuencias contiguas a scaffolds), se logró obtener un genoma compuesto por 769 fragmentos en el caso de *C. digitata* y de 223 fragmentos para el caso de *C. o. ssp. martinezii*. En ambos genomas se lograron identificar más del 95% de los genes conservados de copia única y representativos del grupo de las embriofitas, lo cual indica que ambos ensamblados contienen la mayoría de las secuencias de sus genomas.

Finalmente, las métricas de los dos ensamblados (Tabla 3) indican que el genoma de *C. digitata* es de aproximadamente 258 Mb, mientras que el de *C. okechobeensis ssp. martinezii* es de 284 Mb, valores similares al tamaño de los genomas reportados para las especies del género que ya han sido publicadas (*C. pepo*: 261 Mb; *C. argyrosperma*: 230 Mb; *C. moschata*: 269.9 Mb; *C. maxima*: 271 Mb).

Tabla 3. Métricas de los ensamblados de *Cucurbita digitata* (DIG) y *C. o. ssp. martinezii* (CP).

	DIG	CP
Número de scaffolds	769	223
Tamaño total de los scaffolds en pb	258,642,492	284,250,708
Scaffold más largo en pb	20,748,420	23,433,386
Scaffold más corto en pb	1,390	1,070
Número de scaffolds > 1K nt	769	223
Número de scaffolds > 10K nt	315	91
Número de scaffolds > 100K nt	36	22

Número de scaffolds > 1M nt	21	21
Número de scaffolds > 10M nt	17	19
Tamaño promedio de scaffold en pb	336,336	1,274,667
Media del tamaño de scaffold en pb	7,585	6, 580
N50 (longitud del scaffold) en pb	11,219,285	13,667,081
L50 (Número de scaffolds)	10	9
scaffold %A	31.94	31.83
scaffold %C	18	18.11
scaffold %G	18.04	18.09
scaffold %T	31.95	31.79
scaffold %N	0.07	0.17

Adicionalmente, ya se cuenta con una versión del genoma de *C. ficifolia* de 238.9 Mbp en 91 scaffolds con una N50 de 10.5 Mbp. Esperamos obtener el genoma de *C. radicans* secuenciado durante el presente año. Lo anterior permitirá contar con un genoma de referencia para cada clado dentro del género *Cucurbita*, con los cuales se podrán anotar SNPs obtenidos por tGBS para análisis poblacionales y buscar correlaciones con patrones ambientales para identificar señales de selección. Al mismo tiempo, se realizarán diversos análisis de genómica comparada.

Estos resultados forman parte del proyecto de investigación doctoral del M. en C. Emanuel Villafán de la Torre que se titula “Genómica evolutiva de las calabazas (género *Cucurbita*),” bajo la dirección del Dr. Luis E. Eguiarte Fruns y quien se encuentra cursando el segundo semestre del programa de Doctorado en Ciencias Biológicas de la UNAM. Por lo anterior los resultados finales de este punto se verán reflejados en los próximos semestres.

7. Análisis poblacionales con datos genómicos (SNPs) obtenidos por tGBS.

7.1 Especies cultivadas.

Como se ha mencionado anteriormente, contamos con datos de tGBS para cuatro taxa cultivados, que en conjunto representan una muestra de más de 450 individuos. Los datos de *C. argyrosperma* están en prensa en este momento (Barrera-Redondo et al. 2021, en apéndice 2). Los datos de *C. pepo* se presentan en un artículo en revisión (Martínez-

González et al., 2021, en apéndice 2). Los datos de *C. moschata* se encuentran en proceso de análisis. Para *C. ficifolia* se evaluó el flujo génico con otras taxa silvestres y se generaron datos de GBS para 142 individuos de 58 localidades a lo largo de México. Se generaron más de 6,000 SNPs para realizar análisis de estructura y establecer modelos de Coalescencia. Este trabajo será presentado en el congreso virtual de la Botanical Society of America en julio de este año.

7.2 Especies silvestres.

Realizamos los primeros análisis de los datos de tGBS de ocho taxa silvestres del género, utilizando 440 individuos provenientes de 114 localidades (ver Análisis preliminares de las taxa silvestres, en Apéndice 2). Una vez obtenidos los genotipos finales a partir de las lecturas iniciales (crudas), se obtuvieron de manera global 9,609 SNPs para los ocho taxa. Con estos datos, se hizo un Análisis de Componentes Principales (PCA), para representar gráficamente las diferencias de los genotipos. Para evaluar la diversidad genética se estimó la heterocigosis esperada (H_E), el número de polimorfismos nucleotídicos simples bialélicos (SNPs) y el coeficiente de endogamia (F_{IS}). Asimismo, para evaluar el nivel de estructura genética por taxa, se realizó un análisis de asignación bayesiana de genotipos. Adicionalmente, se implementó un Análisis Espacial de Componentes Principales (SPCA) para asociar a los grupos genéticos con su distribución espacial. Una vez con los grupos genéticos definidos, para cada especie fueron estimados los parámetros de diversidad mencionados junto con el parámetro de estructuración F_{ST} . Finalmente, se obtuvieron mapas de los grupos genéticos y su asociación geográfica.

Estos análisis poblacionales con datos genómicos en taxa silvestres nos servirán para buscar correlaciones con patrones ambientales y para detectar señales de selección. También nos van a ayudar a entender la evolución de los genomas, los cambios genéticos en la domesticación y diferentes características como forma y tamaño del fruto, amargura, tamaño semilla y los cambios asociados a hábitos perennes y anuales. Adicionalmente, se podrán abordar preguntas relacionadas a los cambios genéticos asociados a distintas condiciones ambientales (xerófilos y mesófilos).

Específicamente, los datos de tGBS de 14 poblaciones de la especie silvestre *C. radicans* forman parte de la tesis de Licenciatura que de la alumna Brenda Monterrubio Lezama (proceso de revisión de tesis). Este trabajo destaca porque se está utilizando un método de ensamble de alelos *de novo*. Con esta información se busca identificar SNPs neutrales y SNPs bajo posible selección. Con estos datos se estimó la diversidad y estructura genética y se están buscando correlaciones ambientales.

ENTREGABLES E INDICADORES DE ÉXITO DEL PROYECTO

1. Base de datos de los registros de colectas obtenidos durante las nuevas salidas al campo en la segunda etapa.

La base de datos de las colectas realizadas durante las dos etapas del proyecto está conformada por 145 registros de la primera etapa, principalmente con localidades de especies y subespecies cultivadas y con 187 registros de la segunda etapa, en su mayoría de especies silvestres. Esta base ya ha sido revisada anteriormente por la subcoordinación de inventarios Bióticos (SIB) y se han subsanado errores u omisiones. Conforme a los acuerdos tomados durante la reunión con personal de CONABIO, la nueva base de datos de esta segunda fase del proyecto se integrará con la base previa (ver Base de datos colectas en Apéndice 1) y en suma se tienen 332 registros. A lo largo del proyecto se compilaron registros fotográficos, que ya han sido entregados en informes anteriores.

Como resultado de este proyecto se cuenta a la fecha con un total de 70 ejemplares de herbario que serán incorporados al Herbario Nacional MEXU en el Instituto de Biología de la UNAM y una vez que se cuente con la carta de recepción de los ejemplares, esta será enviada a la CONABIO. Las localidades de origen de todos estos ejemplares están incluidas en la base de datos de registros que será entregada a la CONABIO.

2. Base de datos genéticos de las especies domesticadas y silvestres del género *Cucurbita* en México.

Se cuenta con las secuencias de cloroplasto, mitocondria, microsatélites nucleares y SNPs, (estos últimos en proceso de análisis en algunas especies) de *C. argyrosperma* ssp. *argyrosperma*, *C. argyrosperma* ssp. *sororia*, *C. moschata*, *C. pepo* ssp. *pepo*, *C. pepo* ssp. *fraterna*, *C. pepo* ssp. *ovifera* y ya se están incorporando las nuevas secuencias de SNPs de diferentes especies. Los archivos de estas bases fueron entregados en el informe parcial de septiembre de 2019. La dirección de los repositorios de datos de las diferentes especies que ya han sido analizados y publicados se adjunta en la tabla de entregables (Apéndice 1), donde se incluye esta información.

Con este informe final, los datos de variación genética no publicados o en proceso de análisis fueron puestos a disposición de la CONABIO por medio de la plataforma OSF (Open Science Foundation <https://osf.io/dashboard>) con el proyecto osf.io/9q8dc/. La Dra. Alicia Mastretta fue incluida en el proyecto como colaboradora con derechos de lectura y escritura (Contributor Read+Write), de tal manera que cuenta con acceso completo a todos los datos generados. Para todos los juegos de datos se incluyen las lecturas originales, así como los metadatos que describen la especie y procedencia de la muestra. En el caso de los datos que ya han sido analizados, se incluyen los scripts y líneas de comando para asegurar la reproducibilidad de los análisis. La información se encuentra organizada en carpetas de acuerdo con el tipo de datos: GBS de especies silvestres, GBS de especies cultivadas y genomas completos. Los análisis e información que se generen posteriormente serán entregados puntualmente a la CONABIO, incluyendo ligas a los repositorios para el acceso de datos crudos junto con los métodos y scripts usados para los análisis.

3. Una tesis de maestría. Estudio de la variación morfológica del género *Cucurbita* en México. Esta tesis se inició durante la primera fase del proyecto y continuará durante la segunda fase del proyecto.

Los caracteres utilizados presentaron muchos problemas en un primer análisis, por lo que vamos a requerir re-plantear su análisis. Aún no se tiene el alumno para esta tesis, por lo que posiblemente este trabajo lo realizaran en un futuro los diferentes asociados al

proyecto. Actualmente nos encontramos en fase de obtención de datos morfológicos de las especies silvestres. Adicionalmente se concluyó la tesis de maestría “Diversidad y estructura genética en loci neutrales de *Cucurbita pepo* subsp. *pepo* en México” la cual defendió exitosamente Carmina Martínez González (Apéndice 2) en diciembre de 2019. Esta tesis se inició durante la segunda fase del proyecto.

4. Tesis de doctorado en versión para ser enviada a sinodales. Análisis de la diversidad genética de *Cucurbita argyrosperma* Huber (Cucurbitaceae) en México. Esta tesis se inició durante la primera fase del proyecto, y concluirá durante la segunda fase del proyecto. Alumno: Guillermo Sánchez de la Vega

La tesis se encuentra escrita y se encuentra en trámites de titulación. Se anexa en el Apéndice 2 del presente informe.

5. Tesis de doctorado. Genética de poblaciones y filogeografía de *Cucurbita moschata* Duchesne ex Poir. (Cucurbitaceae) en México. Esta tesis se inició durante la primera fase del proyecto y concluirá durante el segundo año de la segunda fase del proyecto. Alumna: Helena S. Hernández Rosales.

La tesis se encuentra escrita y en trámites de titulación. Se anexa en el Apéndice 2 del presente informe.

6. Tesis de doctorado. Genómica comparada de la domesticación en *Cucurbita argyrosperma*. Esta tesis se inició durante la primera fase del proyecto, y el borrador del artículo científico se presentará durante el segundo año de la segunda fase del proyecto. Alumno: Josué Barrera Redondo.

El alumno presentó y defendió la tesis en octubre de 2020. Se anexa en el Apéndice 2 del presente informe.

7. Artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Análisis filogenético del género *Cucurbita*. Incluyendo a *C. radicans* que será amplificada durante la segunda fase del proyecto. Alumna: Leslie M. Paredes Torres.

El análisis filogenético, incluyendo a *C. radicans* fue publicado en *Molecular Phylogenetics and Evolution* en 2018 (Castellanos Morales et al. 2018) y fue entregado en el informe anual en abril de 2019. También se anexa en el Apéndice 2 del presente informe.

8. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Enviado a una revista internacional. Recursos genéticos y filogeografía de *C. argyrosperma* ssp. *argyrosperma* y *C. argyrosperma* ssp. *sororia*. Alumno: Guillermo Sánchez de la Vega.

Ya fue publicado en 2018 en la revista internacional e indexada *Frontiers in Plant Science* y fue entregado en el informe anual en abril de 2019. Se anexa en el Apéndice 2 del presente informe (Sánchez-de la Vega et al. 2018).

9. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Enviado a una revista internacional. Recursos genéticos y filogeográficos de *C. pepo* ssp. *pepo* y *C. pepo* ssp. *fraterna*. Gabriela Castellanos Morales y alumna Karen Y. Ruiz Mondragón.

Este manuscrito ya se publicó bajo el título "Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico" en la revista internacional indizada *Proceedings of the Royal Society B* (Castellanos-Morales, et al. 2019). Se anexa en el Apéndice 2 del presente informe.

10. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. "Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient". Alumna: Helena S. Hernández Rosales.

Este manuscrito ya se encuentra publicado con el título: "Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient" publicado en el *American Journal of Botany* (ver Hernández-Rosales, et al. 2020 en el Apéndice 2).

11. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Transcriptoma y genoma de *C. argyrosperma* ssp. *argyrosperma*.

Alumno: Josué Barrera Redondo.

Ya fue publicado en la revista *Molecular Plant*. Dicho artículo fue entregado en el informe anual en abril de 2019 (Barrera Redondo et al. 2019). Se anexa en el Apéndice 2 del presente informe. Adicionalmente se publicó la reseña "Gourds and Tendrils of Cucurbitaceae: How Their Shape Diversity, Molecular and Morphological Novelty Evolved via Whole-Genome Duplications" en *Molecular Plant*, (Ver Barrera-Redondo, et al., 2020a en Apéndice 2).

12. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Genómica comparada de especies domesticadas y silvestres de *Cucurbita*. Alumno: Josué Barrera Redondo.

El manuscrito ha sido publicado recientemente con el título "The domestication of *Cucurbita argyrosperma* as revealed by the genome of its wild relative" en la revista *Horticulture Research*. Se anexa en el Apéndice 2 del presente informe (Barrera-Redondo et al. 2021).

13. Manuscrito artículo científico producto de los datos obtenidos durante la primera fase del proyecto. Análisis comparativo de la variación genética presente en las cucurbitas silvestres en México.

Se llevaron a cabo los primeros análisis de los datos de tGBS de ocho taxa silvestres del género, utilizando 440 individuos provenientes de 114 localidades. Se anexa el documento "Análisis preliminares taxa silvestres" en Apéndice 2. Estos datos

posteriormente se publicarán en un artículo con reconocimiento explícito en los agradecimientos al proyecto.

14. Un artículo de difusión relatando los logros más relevantes del proyecto. (responsable RLS y LEE)

Se han publicado tres artículos de difusión: Sánchez de la Vega, 2017 y Eguiarte et al. 2018, que fueron publicados en las revistas de difusión *Oikos* del Instituto de Ecología, UNAM y TIP, de la FES Zaragoza UNAM respectivamente, y en un tercer artículo, enfocado en los taxa silvestres de Aguirre-Dugua et al. 2020, fue publicado en *Biodiversitas* (CONABIO) (se anexan en Apéndice 2).

Se tiene contemplado realizar un artículo de difusión con los resultados finales de todo el proyecto, cuyo contenido se definirá una vez se hayan concluido los análisis y tocará los temas de la distribución de la diversidad genética de calabazas silvestres y domesticadas de México y su importancia para la conservación de recursos fitogenéticos.

Asimismo, se colaboró con CONABIO para producir material de difusión e incluir una sección sobre calabazas en su página web (Enciclovida), así como para la elaboración de carteles de diversidad y un cartel sobre el proceso de domesticación de calabazas, y la inclusión de fotografías colectadas durante el proyecto, en el banco de imágenes, guías ilustradas de identificación de especies, etc.

Adicionalmente, se publicó en 2020 el artículo “Variedades locales y criterios de selección de especies domesticadas del género *Cucurbita* (Cucurbitaceae) en los Andes Centrales del Perú: Tomayquichua, Huánuco”, en *Botanical Sciences*, 98. Este trabajo de investigación es el resultado de la colaboración de dos alumnos del proyecto: Josué Barrera-Redondo y Helena Hernández-Rosales. (ver Barrera-Redondo et al. 2020b Variedades locales de *Cucurbita* en Perú, Apéndice 2).

16. Informe de avances (a los 6 y 18 meses).

Ya se entregó y aprobó el primer informe semestral y el tercero en el mes de septiembre de 2019.

17. Informe anual y entrega de productos comprometidos para el primer año del proyecto.

El documento del informe anual fue entregado en abril de 2019.

18. Informe final del proyecto y entrega de productos comprometidos al finalizar el proyecto.

Se realiza con el presente informe final.

19. Documento con “Propuesta de estrategias para apoyar la conservación de la diversidad genética de las especies silvestres y domesticadas del género *Cucurbita* que crecen en México”.

Como se describió antes, tenemos muchos avances y discusiones al respecto, especialmente para las especies cultivadas. El documento sobre la conservación de la diversidad genética está pendiente y se elaborará una vez que tengamos todos los datos analizados.

CONSIDERACIONES FINALES Y PERSPECTIVAS

El proyecto nos permitió explorar la distribución de la diversidad genética de las especies del género *Cucurbita* que crecen en México, mediante el uso de herramientas moleculares y datos genómicos. Analizamos detalladamente su diversidad genética, particularmente de las especies silvestres, vinculándolas con variables climáticas e incluyendo aspectos genómicos a nivel población. También logramos un avance notable en el estudio genómico de diferentes especies de *Cucurbita* de las cuales se tenían vacíos de información y se han generado comparaciones en términos de evolución molecular.

Realizamos un gran número de colectas poblacionales de plantas del género, especialmente de especies silvestres, donde el registro era escaso y poco actualizado. A partir de ello se logró obtener una filogenia del género que incluye a *C. radicans*, la cual no había sido incluida en trabajos similares anteriores. Los avances obtenidos y el

germoplasma colectado nos van a permitir en el futuro estudios muy detallados sobre la evolución y potencial de estas especies.

El trabajo en conjunto permitió el ensamble y publicación de dos genomas, además de estar en proceso los genomas de otros tres taxa. Con estos genomas fue posible avanzar en entender la evolución del genoma de las especies que conforman el grupo *Argyrosperma* dentro del género *Cucurbita*, identificando los cambios genéticos durante la domesticación. Utilizando las bases de SNPs con las que se cuenta, hemos realizado análisis poblacionales y de genómica comparada, buscando patrones ambientales para detectar señales de adaptación local y señales de cómo ha actuado la selección durante la domesticación, como en el caso de *C. argyrosperma* y *C. pepo*.

Aún queda terminar de entender la evolución del genoma de especies silvestres del género *Cucurbita*, así como los cambios asociados a hábitos perennes y anuales, o los cambios genéticos asociados a distintas condiciones ambientales. Con todo lo anterior, se pretende a corto, mediano y largo plazo seguir proponiendo estrategias para apoyar en la conservación de la diversidad genética de las especies silvestres y domesticadas del género *Cucurbita* que crecen en México. Asimismo, se pretende avanzar en un estudio detallado del Pangenoma de todas las especies cultivadas y silvestres del género. Para esto deberemos en el futuro hacer estudios del genoma completo para varios individuos de cada especie.

Los avances obtenidos y el germoplasma colectado permitirán en el futuro estudios detallados sobre la evolución y potencial de estas especies. Consideramos que ningún otro grupo de plantas en México tiene en este momento el detalle de conocimiento alcanzado para el género *Cucurbita*, y lo más notable es que casi toda esta investigación de frontera se realizó con el apoyo para este proyecto (dividido en dos partes) de la CONABIO.

Indudablemente, la información y datos generados en este proyecto son recursos muy importantes que previamente no se tenían, con los cuales podrán contestarse preguntas y tomar decisiones a futuro, basadas en datos sólidos. Con estos datos se

podrán continuar distintos estudios de interés para el conocimiento de especies domesticadas y sus parientes silvestres en México, particularmente el género *Cucurbita*.

Literatura citada:

- Aguirre-Dugua, X., Castellanos-Morales, G., Paredes-Torres, L.M., Hernández-Rosales, H.S., Barrera-Redondo, J., Sánchez-de la Vega, G., Tapia-Aguirre, F., Ruiz-Mondragón, K.Y., Scheinvar, E., Hernández, P. and Aguirre-Planter, E., 2019. Evolutionary dynamics of transferred sequences between organellar genomes in *Cucurbita*. *Journal of molecular evolution*, 87(9), 327-342.
- Barrera-Redondo, J., Ibarra-Laclette, E., Vázquez-Lobo, A., Gutiérrez-Guerrero, Y. T., Sánchez-de la Vega, G., Piñero, D., Montes-Hernández, S., Lira-Saade, R. and Eguiarte, L. E. 2019. The genome of *Cucurbita argyrosperma* (Silver-seed gourd) reveals faster rates of protein-coding gene and long noncoding RNA turnover and neofunctionalization within *Cucurbita*. *Molecular Plant*, 12(4), 506-520.
- Barrera-Redondo, J., Sánchez-de la Vega, G., Aguirre-Liguori, J., Castellanos-Morales, G., Gutiérrez-Guerrero, Y., Aguirre-Dugua, X., Aguirre-Planter, E., Tenailon, M., Lira-Saade, R., and Eguiarte, L. 2021. (en prensa). The domestication of *Cucurbita argyrosperma* as revealed by the genome of its wild relative. *Horticulture Research*. <https://doi.org/10.1038/s41438-021-00544-9>
- Castellanos-Morales G., Paredes, L., Gámez, N., Hernández-Rosales, H.S., Sánchez de la Vega, G., Barrera-Redondo, J., Aguirre-Planter, E., Vázquez-Lobo, A., Montes-Hernández, S., Lira-Saade, R. and Eguiarte, L.E. 2018. Historical biogeography and phylogeny of

Cucurbita: insights from ancestral area reconstruction and niche evolution. *Molecular Phylogenetics and Evolution*, 128, 38-54.

- Castellanos-Morales, G., Ruiz-Mondragón, K. Y., Hernández-Rosales, H. S., Sánchez-de la Vega, G., Gámez, N., Aguirre-Planter, E., Montes-Hernández, S., Lira-Saade R., Eguiarte, L. E. 2019. Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proceedings of the Royal Society B*, 286(1908), 20191440.
- Hernández-Rosales, H. S., Castellanos-Morales, G., Sánchez-de la Vega, G., Aguirre-Planter, E., Montes-Hernández, S., Lira-Saade, R., and Eguiarte, L. E. 2020. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *American Journal of Botany*, 107(3), 1-16.
- Martínez-González, C., Castellanos-Morales, G., Barrera-Redondo, J., Sánchez-de la Vega, G., Hernández-Rosales, H. S., Gasca-Pineda, J., Aguirre-Planter, E., Moreno-Letelier, A., Escalante, A.E., Montes-Hernández, S., Lira-Saade R. and Eguiarte L.E. 2021. Recent and historical gene flow in cultivated zucchini (*C. pepo* ssp. *pepo*) and its wild relative, *C. pepo* ssp. *fraterna* in Mexico. *Frontiers in Ecology and Evolution* 9, 307.
- Sánchez de la Vega, G., Castellanos-Morales G., Gámez, N., Hernández-Rosales H.S., Vázquez-Lobo, A., Aguirre-Planter, E., Montes-Hernández S., Lira-Saade, R., and Eguiarte, L.E. 2018. Genetic resources in the “calabaza pipiana” pumpkin (*Cucurbita argyrosperma*) in Mexico: Genetic diversity, genetic differentiation and distribution models. *Frontiers in Plant Science*, 9, 400, 18 págs.