

Diversidad y estructura genética en poblaciones de *Moringa oleifera* Lam. cultivadas en México a través de marcadores SNP

Diversity and genetic structure in populations of *Moringa oleifera* Lam. cultivated in Mexico through SNP markers

Ruiz Hernández, R. ¹ , Pérez Vázquez, A. ¹ , Hernández Rodríguez, M. ^{2*} ,
Rayas Amor A. A. ³ .

¹Campus Veracruz. Colegio de Postgraduados. Carretera Xalapa Veracruz km 88.5, Tepetates. 91690, Manlio Fabio Altamirano, Veracruz, México.

²Campus Montecillo. Colegio de Postgraduados. Calle, km 36.5, Montecillo. 56230, Texcoco de Mora, Estado de México, México.

³Unidad Lerma. Universidad Autónoma Metropolitana. Av. De las Garzas No. 10, Colonia El Panteón, C.P. 52005. Lerma de Villada, Estado de México, México.



Please cite this article as/Como citar este artículo: Ruiz-Hernández, R., Pérez-Vázquez, A., Hernández- Rodríguez, M., Rayas-Amor A. A. (2025). Diversity and genetic structure in populations of *Moringa oleifera* cultivated in Mexico through SNP markers. *Revista Bio Ciencias*, 12, e1747. <https://doi.org/10.15741/revbio.12.e1747>

Article Info/Información del artículo

Received/Recibido: Septiembre 7th 2024.

Accepted/Aceptado: June 19th 2025.

Available on line/Publicado: July 08th 2025.

RESUMEN

Moringa oleifera es una especie con amplia distribución mundial debido a su capacidad adaptativa y a sus propiedades nutricionales. En México, han ocurrido varias reintroducciones de esta especie y aunque se ha observado diversidad morfológica, se desconoce su diversidad a nivel de DNA. El objetivo fue determinar la diversidad y estructura genética de 14 poblaciones de *M. oleifera* cultivadas en México a través de SNP. Las semillas se germinaron en invernadero y el DNA se extrajo de hojas y tallos jóvenes mediante un protocolo basado en CTAB. Este DNA se utilizó para realizar el genotipado de cada población con la tecnología DArTseq™. Se obtuvieron 9,862 SNP, de ellos el 0.64 % fueron polimórficos. Dentro de las poblaciones se observaron niveles moderados de diversidad ($A_e=1.52$, $H_e=0.33$, $H_o=0.44$, índice de Shannon = 0.44). El trasfondo genético de las poblaciones permitió visualizar dos grupos a través del análisis de conglomerados y PCoA. La estructura genética de las poblaciones mediante STRUCTURE indicó que la ascendencia proviene de dos poblaciones originales ($K=2$). Las 14 poblaciones de moringa presentan diversidad genética que puede ser empleada para un manejo sostenible y su conservación.

PALABRAS CLAVE: Polimorfismo, genotipificación, recursos genéticos.

*Corresponding Author:

Martha Hernández-Rodríguez. Posgrado en Recursos Genéticos y Productividad – Genética. Campus Montecillo. Colegio de Postgraduados. km 36.5 carretera Federal México-Texcoco, 56264, Montecillo, Texcoco, Estado de México, México. Teléfono (595) 952 0265. E-mail: hernandez.martha@colpos.mx

ABSTRACT

Moringa oleifera is a species with a wide global distribution due to its adaptive capacity and nutritional properties. In Mexico, there have been several reintroductions of this species and although morphological diversity has been observed, its diversity at the DNA level is unknown. The objective was to determine the diversity and genetic structure of 14 populations of *M. oleifera* cultivated in Mexico through SNP. The seeds were germinated in a greenhouse and DNA was extracted from young leaves and stems using a CTAB-based protocol. This DNA was used to genotype each population with DArTseq™ technology. A total of 9,862 SNPs were obtained, of which 0.64 % were polymorphic. Moderate levels of diversity were observed within the populations ($A_e=1.52$, $H_e=0.33$, $H_o=0.44$, Shannon index = 0.44). The genetic background of the populations allowed two groups to be visualized through cluster analysis and PCoA. The genetic structure of the populations using STRUCTURE indicated that the ancestry comes from two original populations ($K = 2$). The 14 moringa populations present genetic diversity that can be used for sustainable management and conservation.

KEY WORDS: Polymorphism, genotyping, genetic resources.

Introducción

Para afrontar problemas como el hambre, la desnutrición, la baja productividad de alimentos y el cambio climático, el sistema agrícola se ve forzado en la identificación y utilización de nuevas especies. Aunado a esto, la poca variedad de alimentos consumidos puede ocasionar un desequilibrio en cuanto a la diversidad y cantidad de micronutrientes que requiere el organismo para su correcto funcionamiento (Calicioglu *et al.*, 2019). En consecuencia, los nuevos cultivos, además de satisfacer las tendencias y preferencias alimentarias, deben presentar alto rendimiento, mejor contenido nutricional, mayor tolerancia a los diversos tipos de estrés, un mínimo impacto ambiental y una mayor eficiencia de uso de los recursos naturales (Hendre *et al.*, 2019).

Moringa oleifera Lam. es un árbol originario de la India que puede crecer hasta altitudes cercanas a los 1,800 msnm, presenta tolerancia a altas temperaturas y puede crecer en suelos con bajos niveles de humedad y nutrientes (Muhl *et al.*, 2011). Por sus características adaptativas, moringa puede prosperar en diferentes climas y actualmente puede cultivarse en el 37 % de la superficie terrestre. Las hojas de moringa contienen cantidades importantes de proteínas, aminoácidos esenciales, lípidos, grasas, carbohidratos, fibra y metabolitos secundarios (Ruiz-Hernández *et al.*, 2022a). Con base en las propiedades nutrimentales y alimenticias, moringa se

encuentra dentro del listado del consorcio de cultivos huérfanos africanos y es considerada una alternativa para la alimentación del futuro (Hendre & Van Deynze, 2015). Esta planta presenta una baja producción debido a que la inversión agrícola, tecnológica y científica ha sido marginal y existen escasos informes de manejo agronómico (Hendre *et al.*, 2019), por lo que la oportunidad de aportar información acerca de este cultivo es valiosa.

Moringa es una especie diploide ($2n = 2X = 28$), de naturaleza alógama, cuyo genoma ha sido secuenciado al 91.78 %. El tamaño del genoma es de 281.92 Mb. La estimación del número de genes de moringa oscila alrededor de 19,465 genes (Tian *et al.*, 2015), aunque una resecuenciación posterior del genoma de esta especie determinó 32,062 genes putativos codificadores de proteínas (Shyamli *et al.*, 2021).

En México, se han observado diferencias morfológicas en las poblaciones cultivadas (Ruiz-Hernández *et al.*, 2021). La estructura genética de las poblaciones de moringa cultivadas en México presentó tres poblaciones originales a través de ocho microsatélites (Ruiz-Hernández *et al.*, 2024), sin embargo, un mayor número de marcadores moleculares podrían detallar la diversidad y estructura genética de las poblaciones de moringa establecidas en México. Este tipo de variación entre las poblaciones es consecuencia del genotipo de cada población y de las condiciones ambientales, por lo que la diversidad genética a nivel de DNA es importante para determinar el nivel de diversidad debido a que la secuencia de DNA no es afectada por el ambiente.

Para examinar esta diversidad se utilizan los marcadores moleculares. Cada marcador molecular es una secuencia de DNA con una ubicación conocida en el genoma que indica la presencia o proximidad de un gen o de una región regulatoria implicada en la expresión de un carácter. Los polimorfismos de un solo nucleótido o SNP, son un tipo de marcadores moleculares que permiten determinar el grado de variación a una escala fina. Aunque la mayoría de los SNP no son decisivos, algunos de ellos son favorables para la expresión del genotipo, ya sea se encuentren o no en regiones codificantes (Ousmael *et al.*, 2023).

Los SNP son muy abundantes en el genoma, tienen una tasa baja de mutación y se pueden obtener mediante diferentes plataformas de alto rendimiento (Teklemariam *et al.*, 2022). Actualmente, la genotipificación por secuenciación permite la obtención de secuencias polimórficas a nivel de un solo nucleótido de todo el genoma que facilitan el escrutinio de la diversidad genética en miles de loci a la vez (Nadeem *et al.*, 2018). Además del análisis de diversidad, este tipo de polimorfismos permiten la selección genómica, la elaboración de mapas de asociación del genoma, establecer relaciones filogenéticas y brindar información precisa acerca de la historia evolutiva de las poblaciones de interés (Yirgu *et al.*, 2023).

Con su aplicación ha sido posible asociar variantes favorables que más tarde se pueden emplear para precisar la selección de plantas con rasgos de interés alimentario como la concentración de nutrientes, la identificación de individuos con mayor capacidad de defensa contra patógenos o con mejor respuesta a condiciones restrictivas de humedad, altas temperaturas, fertilización limitada, salinización o incluso inundación (Huster *et al.*, 2021, Elbasyoni *et al.*, 2022).

Debido a que conocer la diversidad genética de los cultivos de moringa en México puede ser la base para el mejoramiento genético de la especie, el objetivo de esta investigación fue evaluar la diversidad y estructura genética de poblaciones de *M. oleifera* cultivadas en México a través de SNP.

Material y Métodos

Material biológico

Se seleccionaron 14 poblaciones de moringa en 11 estados de México (Veracruz, Chiapas, Guerrero, Sinaloa, Oaxaca, Guanajuato, San Luis Potosí, Hidalgo, Yucatán, Michoacán y Estado de México) (Tabla 1). Se seleccionaron poblaciones destinadas a la venta de productos y semillas de moringa y que estuvieran distanciados geográficamente. De cada población se recolectaron 500 g de semillas. Para su germinación, se seleccionaron 30 semillas, las cuales se desinfectaron con cloro al 10 % durante 5 min y fueron sembradas en bolsas de vivero de tamaño 22 cm x 22 cm con sustrato de tierra y composta 4:1. Treinta días después de la germinación se tomaron 30 mg de hoja y tallo de 10 individuos por población. Al final se formó una muestra compuesta de 300 mg por población, la cual se mantuvo a -80 °C hasta el momento de la extracción del DNA. Se realizó un mapa utilizando el sistema de referencia de coordenadas y la capa estados del INEGI con el programa QGIS.

Tabla 1. Ubicación geográfica de las poblaciones evaluadas de *M. oleifera* en México.

No.	Población	Estado	Municipio	Localidad	Longitud	Latitud
1	CHI	Chiapas	Tuxtla Gutiérrez	Colonia La Salle	-93.0868889	16.7429444
2	EMEX	Estado de México	Texcoco	San Pedro Santa Úrsula	-98.898385	19.497082
3	GTO	Guanajuato	Soledad de Gasca	Celaya	-100.8146904	20.502528
4	HGO	Hidalgo	San Felipe Orizatlán	Ahuatitla	-98.6660845	21.1630165
5	MICH1	Michoacán	Benito Juárez	El Rodeo	-100.4708226	19.3055772
6	MICHMP	Michoacán	Múgica	Múgica	-102.180997	18.928047
7	NL	Nuevo León	Escobedo	Francisco I. Madero	-100.2847444	25.78544732
8	OAX	Oaxaca	Santa Cruz Xoxocotlán	San Juan Bautista La Raya	-96.7280556	16.9791667
9	SLP	San Luis Potosí	Tanlajas	Guayajox	-98.73666667	21.71444444
10	SIN	Sinaloa	Culiacán	Valle	-107.37889	24.985
11	YUC1	Yucatán	Peto	Teshan	-88.62125	20.1486389
12	YUC2	Yucatán	Baca	Felipe Carrillo Puerto	-89.60700993	20.9954688
13	OAXH	Oaxaca	Santa María Huatulco	La Herradura	-96.3658333	15.7772222
14	GRO	Guerrero	Tecpan de Galeana	Mitla	-99.89343517	16.87894246

Fuente: Elaboración propia.

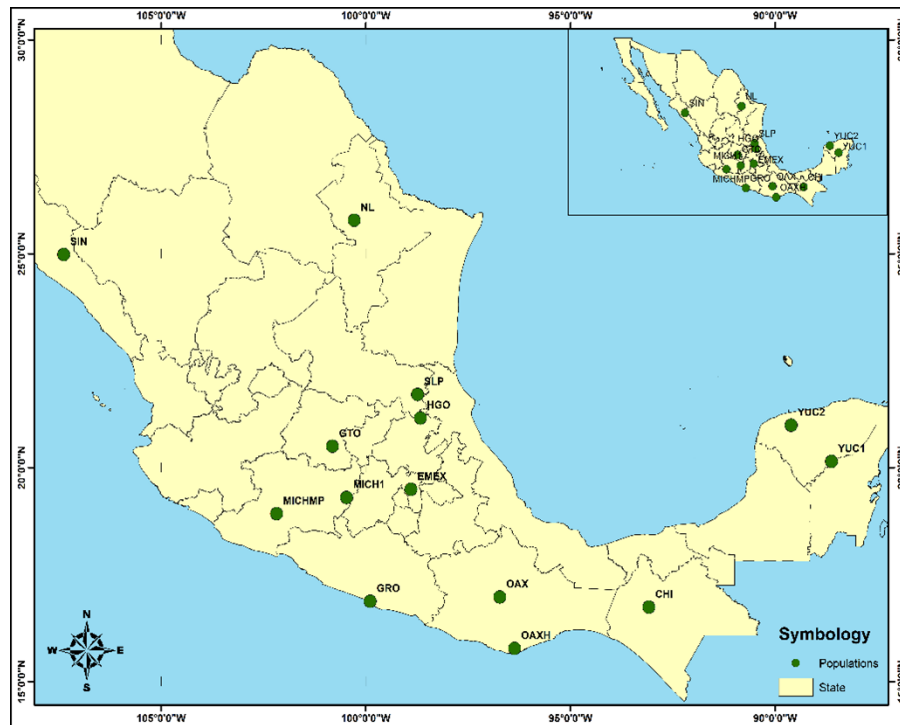


Figura 1. Ubicación geográfica de las poblaciones de *M. oleifera* cultivadas en México examinadas en este estudio.

Fuente: Elaboración propia.

Extracción de DNA

Las muestras compuestas fueron pulverizadas con un pistilo de propileno y nitrógeno líquido en tubos de polipropileno de 2 mL. El producto obtenido se utilizó para extraer el DNA siguiendo un protocolo de extracción a pequeña escala basado en CTAB (CIMMYT, 2006). Brevemente, se agregó 1 mL de solución lisis (100 mM de Tris-base, 700 mM de NaCl, 50 mM de EDTA, 2 % de CTAB y 1 % de β -mercaptoetanol) y se mezcló hasta resuspender completamente el tejido, inmediatamente después se incubó a 65 °C en baño María durante 60 min mezclando por inversión cada 10 min. Posteriormente, las muestras fueron centrifugadas a 11,500 rpm durante 10 min a temperatura ambiente (TA) para formar las tres fases y luego se transfirieron 800 μ L de la fase superior a un tubo limpio y esterilizado de 2.0 mL. Enseguida se purificó el DNA con 500 μ L de cloroformo: alcohol isoamílico 24:1 mezclando por inversión constante durante 10 min a TA. La mezcla resultante se centrifugó a 11,500 rpm durante 10 min a TA y 600 μ L del sobrenadante se colocaron en un tubo esterilizado de 1.5 mL. El DNA se precipitó de este sobrenadante con 300 μ L de isopropanol incubando durante 1 h a -20 °C. El DNA se centrifugó a 11,500 rpm durante 10 min a TA y se lavó con 1 mL de etanol 70 % frío. La pastilla de DNA limpio se recuperó

nuevamente mediante centrifugación a 8,500 rpm durante 5 min a TA. Finalmente, se retiró el sobrenadante y la pastilla se secó durante 2 h a TA y se disolvió en 70 μ L de agua HPLC. La concentración de DNA se cuantificó mediante absorbancia en un espectrofotómetro de ultra bajo volumen (NanoDrop™ 2000, Thermo Scientific, Waltham, MA, USA). La integridad del DNA se verificó en un gel de agarosa (Sigma, MO, EUA) al 1 % utilizando 2 μ L de DNA migrando a 80 volts durante 40 min. El gel se visualizó en un transiluminador MiniBisPro 16 mm® (DNR Bio Imaging Systems®, Israel) con luz UV.

Genotipificación

El genotipado fue realizado por el Servicio de Genotipificación para la Agricultura del CIMMYT utilizando el protocolo DArTseq™ con base en 60 μ L a 50 ng μ L⁻¹ de DNA enviado para análisis. El método de genotipado se hizo siguiendo el protocolo patentado por DArT (<https://www.diversityarrays.com/services/dartseq/>) utilizando las enzimas *Pst*I (CTGCA/G) y *As*eI (AT/TAAT) para fragmentar el DNA y secuenciación de segunda generación con el equipo NovaSeq™ 6000 (Illumina). El llamado de los SNP se realizó con el programa DArTsoft (<http://www.diversityarrays.com/software.html#dartsoft>). Las secuencias con SNP se recibieron por parte del servicio de secuenciación en una matriz de datos en formato de doble hilera con puntuaciones “-”, “0”, y “1”. Esta matriz se recodificó con una nueva puntuación para designar con 0 el genotipo homocigote dominante, 1 para el genotipo homocigote recesivo y 2 para el heterocigote.

Análisis de resultados

La matriz de SNP se limpió para obtener secuencias de alta calidad; para ello, se eliminaron aquellas secuencias que presentaron 25 % de datos faltantes y una frecuencia de 5 % del alelo menor (MAF < 0.05). De esta matriz se calculó la heterocigosidad esperada (H_e), heterocigosidad observada (H_o), número de alelos efectivos (A_e), número de alelos raros (A_r), índice de diversidad de Shannon y porcentaje de loci polimórficos con el programa BIO-R (Biodiversity Analysis with R for Windows) versión 3.2 (Pacheco-Gil *et al.*, 2016). Para visualizar las relaciones genéticas se hizo un análisis de coordenadas principales (PCoA) con el programa PAST versión 3.0 y un análisis de conglomerado utilizando una matriz de distancias Euclidianas con el programa JMP® versión 17.2.0. El análisis de estructura genética mediante ascendencias se realizó con el programa STRUCTURE versión 2.3.4 (Pritchard *et al.*, 2000) con los siguientes parámetros: longitud de burn-in period, 10,000; MCMC (Monte Carlo Markov Chains) réplicas después del burn-in period 250,000; modelo Admixture; número de K de 1 a 10 con 100,000 iteraciones. La visualización del valor Delta K se obtuvo con Structure Harvester versión 0.6.94 (Earl & vonHoldt, 2012).

Resultados y Discusión

Diversidad genética

El genotipado de las 14 poblaciones mexicanas de moringa permitió identificar 9,862 SNP. A partir de estos SNP se obtuvieron los estadísticos de diversidad genética. Se encontró que la H_e fluctuó de 0.18 (OAXH) a 0.43 (EMEX) con un promedio de 0.33, en tanto que la H_o osciló de 0.21 (OAXH) a 0.62 (EMEX) con un promedio de 0.44. En cuanto a los A_e , éste fue de 1.23 a 1.74 con un promedio de 1.51 y el Índice de diversidad de Shannon osciló de 0.21 a 0.62 con un promedio de 0.43. El A_r fue de 67.41 a 167.79 con promedio de 119.90, que indica una cantidad importante de alelos raros (Tabla 2). Entre las poblaciones, se encontró que OAXH tuvo los valores de diversidad más bajos ($H_e = 0.18$, $H_o = 0.21$) en tanto que EMEX presentó los valores más altos ($H_e = 0.43$, $H_o = 0.62$). En general, se encontró que los valores de heterocigosidad fueron de moderados a altos con OAXH como la población más homogénea y EMEX como la más diversa.

Tabla 2. Parámetros de diversidad en poblaciones de *M. oleifera* cultivadas en México utilizando SNP.

No.	Población	H_e	H_o	A_e	I. Shannon	A_r	% NA
1	CHI	0.42	0.60	1.72	0.60	163.33	0.00
2	EMEX	0.43	0.62	1.74	0.62	167.79	0.00
3	GTO	0.38	0.50	1.60	0.50	121.51	0.00
4	HGO	0.41	0.57	1.69	0.57	157.83	0.00
5	MICH1	0.37	0.49	1.58	0.49	119.01	0.00
6	MICHMP	0.38	0.51	1.61	0.51	120.89	0.00
7	NL	0.35	0.46	1.54	0.46	124.69	0.00
8	OAX	0.33	0.42	1.50	0.42	99.84	0.00
9	SLP	0.23	0.26	1.30	0.26	75.45	0.00
10	SIN	0.40	0.54	1.65	0.54	153.84	0.00
11	YUC1	0.26	0.31	1.36	0.31	88.39	0.00
12	YUC2	0.29	0.36	1.41	0.36	95.74	0.00
13	OAXH	0.18	0.21	1.23	0.21	67.41	0.00
14	GRO	0.24	0.28	1.32	0.28	80.92	0.00
Media		0.33	0.44	1.52	0.44	116.90	0.00

H_e : Heterocigosis esperada; H_o : Heterocigosis observada; A_e : Número de alelos efectivos; I. Shannon: Índice de diversidad de Shannon; A_r : número de alelos raros; % NA: proporción de datos faltantes.

Fuente: Elaboración propia con base en la salida del software BIO-R versión 3.2.

Esta investigación se enfocó en determinar la extensión de la diversidad y las relaciones genéticas que mantienen 14 poblaciones de moringa cultivadas en México utilizando SNP, por tanto, este sería el primer reporte utilizando este tipo de marcadores polimórficos en este cultivo en este país. Se identificaron 15,191 SNP de los cuales 9,862 SNP permitieron evaluar la diversidad y las relaciones genéticas de las 14 poblaciones cultivadas de moringa. Este número de polimorfismos se encontraron en todo el genoma por las ventajas de utilizar DArTseq (Gawroński *et al.*, 2016). Esta característica permite la evaluación de colecciones de germoplasma inexploradas o de cultivares de interés agroalimentario.

A nivel de poblaciones, la H_e máxima fue de 0.42, es decir, una diversidad moderada, en tanto que la H_e mínima fue de 0.18, es decir, una diversidad baja con base en la escala propuesta por Botstein *et al.* (1980). Uno de los factores que influye en la heterocigosidad dentro de las poblaciones de moringa es su polinización cruzada. Wu *et al.* (2018) mencionan que el porcentaje de autofecundación en esta especie es del 27 % y el rango de dispersión del polen en moringa puede alcanzar hasta 24.7 m, estas características garantizan el flujo de genes entre los individuos en una población. Los valores de heterocigosidad obtenidos en esta investigación indican que entre las 14 poblaciones existe diversidad que podría ser utilizada para diferentes propósitos.

En relación con los estadísticos de diversidad de los loci SNP, se encontró que después del filtrado de datos, el 65 % de ellos fueron polimórficos, lo que derivó en obtener una H_e de 0.313, H_o de 0.438, A_e de 1.507 e índice de Shannon de 0.695 (Tabla 3).

Tabla 3. Estadísticos de diversidad genética de *M. oleífera* de 9,862 marcadores SNP-GBS detectados en 14 poblaciones de cultivadas de *M. oleífera*.

Parámetro	Valor	Desviación estándar
% de loci polimórficos	0.645	
Heterocigosidad esperada (H_e)	0.313	0.002
Heterocigosidad observada (H_o)	0.438	0.003
Número de alelos efectivos (A_e)	1.507	0.004
Índice de diversidad de Shannon	0.695	0.003

Fuente: Elaboración propia con base en la salida del software BIO-R versión 3.2.

Se encontró que el contenido de información polimórfica (PIC) en los SNP fue de 0.64 después del filtrado de los datos. Gouda *et al.* (2021) reportaron 0.23 de PIC al evaluar especies de *Oryza* con 158 SNP y Tomar *et al.* (2021) obtuvieron un PIC promedio de 0.201 en 14,563 SNP al evaluar *Triticum aestivum*. Dube *et al.* (2023) mencionan que un valor alto de PIC permite

una evaluación eficiente de la diversidad genética en materiales vegetales. Muli *et al.* (2022) mencionan la importancia del filtrado de los marcadores para una determinación más objetiva sobre aspectos de diversidad genética en poblaciones vegetales. Por otra parte, Teklemariam *et al.* (2022) reportan que el total de SNP generados proporciona una relación entre las poblaciones evaluadas y que el filtrado de los datos es importante para disminuir el volumen de datos y conservar los SNP más informativos. También el índice de diversidad de Shannon de los SNP fue de 0.44, un valor similar a los 0.47 en *Triticum aestivum* al ser evaluado con SNP (Tomar *et al.*, 2021).

Al respecto, se ha reportado que la diversidad genética permite a las poblaciones naturales tener mayores probabilidades de tolerar los efectos desfavorables de plagas y enfermedades, o del cambio climático (Zhang *et al.*, 2017). A nivel del investigador, el conocimiento sobre el grado de variación dentro y entre poblaciones de la misma especie permite un manejo eficiente de los recursos fitogenéticos que favorezcan a los agricultores (Nadeem *et al.*, 2018) y una mejor planeación de la conservación y aprovechamiento de aquellos ecotipos con mayor polimorfismo o de aquellos que contengan alelos raros en baja frecuencia. En este sentido, Das & Mishra (2021) mencionan que la diversidad podría ser utilizada en cruzamientos dirigidos, experimentos híbridos, asignación de individuos en grupos heteróticos y en la conservación de germoplasma. La diversidad genética es una fuente de alelos que pueden tener una relación directa con la respuesta de las plantas al ambiente, con su tolerancia a estreses bióticos y abióticos, y con la producción de los alimentos (Begna, 2022). A través de la evaluación y discernimiento de las mejores variantes alélicas en moringa se pueden crear sistemas agroalimentarios con mayor grado de sostenibilidad (Peter *et al.*, 2023). El uso de la diversidad genética en los cultivos podría ser utilizada para incrementar la calidad nutricional de los alimentos mejorando la calidad de vida de las personas.

La diversidad genética en moringa podría explicar la variación en el contenido nutricional de las poblaciones evaluadas en México. Investigaciones previas han reportado diferencias en el contenido de minerales (Ndhlala & Tshabalala, 2023) y valores de proteína de 13.67 a 23.50 % para algunas poblaciones de moringa cultivadas bajo las mismas condiciones edafoclimáticas y de manejo agronómico (Ruiz-Hernández *et al.*, 2022b). Pasha *et al.* (2020) identificaron 36 genes que tienen relación con la producción de metabolitos secundarios. Sería de interés examinar la asociación que existe entre los SNP y los genes responsables de la síntesis de los metabolitos secundarios para generar cultivares con alelos favorables y responsables de mayores concentraciones de estos metabolitos. El número de SNP identificados en este estudio permitiría profundizar en el estudio de caracteres de importancia nutricional dado que ya se encuentran reportadas 18,541 secuencias codificadoras de genes en moringa (Chang *et al.*, 2018). Las poblaciones con mayor diversidad genética pueden servir de base para la selección de individuos con mayores atributos creando variedades y conservando el acervo genético de la población en un banco de germoplasma.

En adición, la diversidad identificada en estas poblaciones mediante los SNP, podría apoyar la evaluación de materiales cultivados en condiciones de clima templado frío, con la finalidad de seleccionar los materiales con mayor capacidad de aclimatación. Pasha *et al.*

(2020) identificaron 2326 factores de transcripción y reportaron que el 32 % están relacionados a respuestas del estrés biótico y abiótico. Zhang *et al.* (2019) reportaron las proteínas WRKY en moringa que participan en el crecimiento, desarrollo, generan respuestas ante diversos tipos de estrés (bióticos y abióticos) y regulan la producción de metabolitos secundarios.

Relaciones genéticas

La visualización de la relación genética de las 14 poblaciones caracterizadas con los SNP se realizó mediante diferentes tipos de análisis. El primero de ellos fue el agrupamiento jerárquico con base en la matriz de distancias siguiendo el método Ward. Los perfiles moleculares permitieron detectar que la población más distante fue GRO y las poblaciones más cercanas fueron GTO y MICHMP (Figura 2). Mediante este agrupamiento se identificaron dos grupos. El grupo I detectó que la distancia que une a las poblaciones GTO y MICHMP es la menor dentro de este agrupamiento, seguida muy cerca por MICH1; OAX se integra a este grupo también, pero queda muy distante indicando que parte de su perfil genético es diferente. El grupo I también agrupó como cercanas a EMEX y CHI, seguidas por SIN, NL y HGO, ésta última como la más distante del grupo y del agrupamiento en general. El grupo II detectó como muy parecidas a YUC1 y YUC2, seguidas de SLP, OAXH, y GRO, las cuales se integraron al grupo en ese orden de distancia.

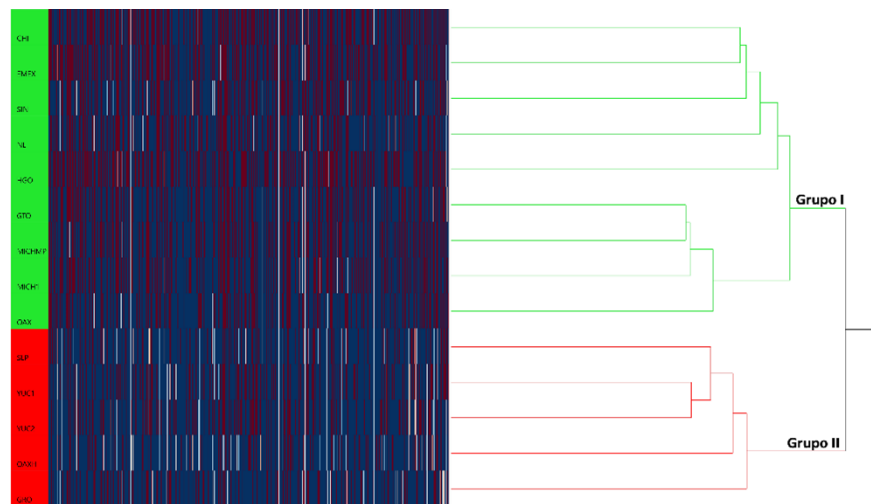


Figura 2. Dendrograma de 14 poblaciones de *M. oleifera* cultivadas en México caracterizadas molecularmente con 9,862 SNP.

En el dendrograma se observa un mapa de calor compuesto por 9,862 SNP donde la similitud del color corresponde al mismo tipo de alelo. Color azul: 0 = AA; color blanco: 1 = aa; y color rojo: 2 = Aa.

Fuente: Elaboración propia con base en la salida del software JMP.

La visualización de las relaciones mediante el análisis de coordenadas principales (PCoA), también detectó la formación de dos grupos separados en el plano de los dos ejes principales. En este caso, el eje 1 explicó el 18 % de la variación genética encontrada, en tanto que el eje 2 explicó el 11.33 %. Los grupos que se observaron quedaron conformados por las mismas poblaciones como en el agrupamiento jerárquico. La conformación de los grupos fue de la siguiente manera: el grupo I integró a las poblaciones MICH1, MICHMP y GTO con una gran cercanía, mientras que la población OAX se integró separada a las tres primeras poblaciones. Las poblaciones EMEX, HGO, NL, CHI y SIN se integraron también en el grupo I de manera dispersa. Estas poblaciones se distribuyeron de manera separada en el plano de los dos ejes. El grupo II incluyó a las poblaciones OAXH, SLP, YUC1, YUC2 y GRO, las cuales se integraron de manera cercana (Figura 3).

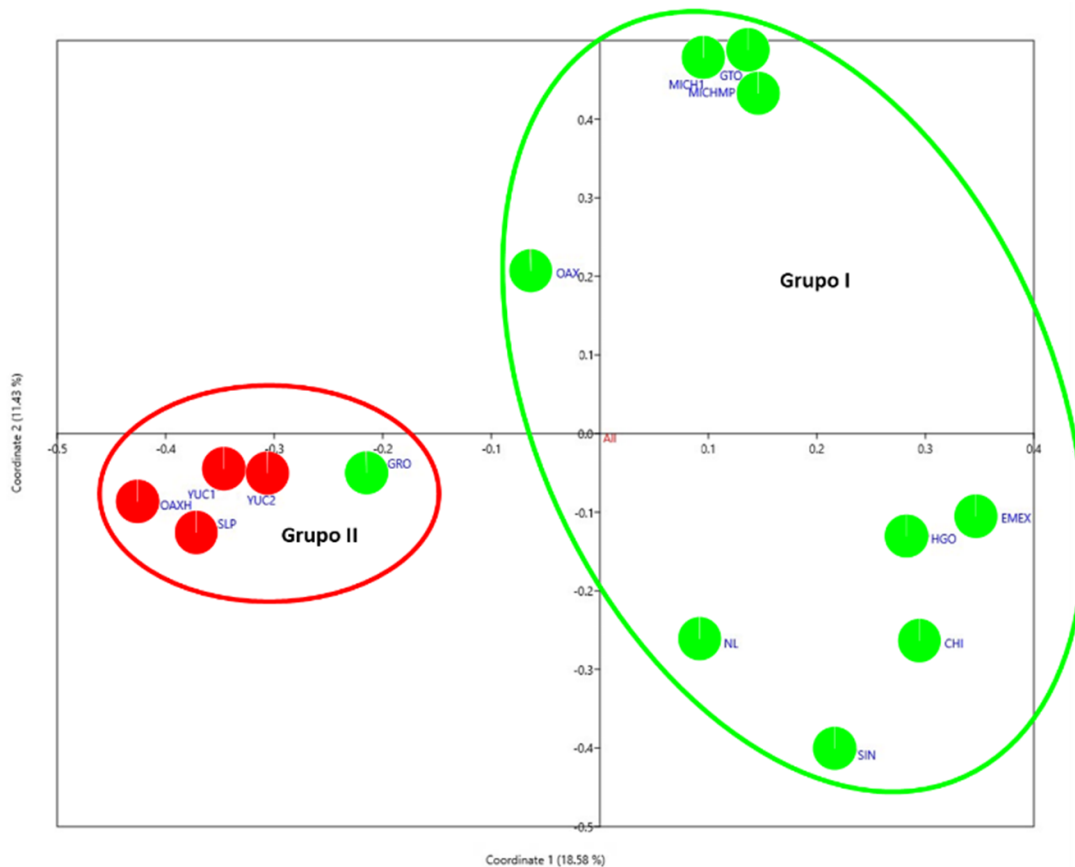


Figura 3. Diagrama de dispersión de las dos primeras coordenadas del PCoA de 14 poblaciones de *M. oleifera* cultivadas en México caracterizadas con 9862 SNP.

Los colores indican la proporción que comparten las poblaciones de los grupos genéticos originales (I: verde y II: rojo) obtenidos con el programa STRUCTURE.

Fuente: Elaboración propia con base en la salida del software PAST.

La clasificación de las 14 poblaciones de moringa vía análisis de conglomerados y PCoA determinó consistentemente dos grupos: MICH1, MICHMP, GTO, OAX, NL, SIN, CHI, HGO y EMEX formaron el grupo I, y OAXH, SLP, YUC1, YUC2 y GRO se agruparon en el II. El PCoA y el análisis de agrupamiento demostraron que la distancia geográfica entre las poblaciones evaluadas no es indicadora sobre los niveles de similitud a nivel genético. Moringa presentó un resurgimiento comercial en los últimos años ocasionando una distribución antropogénica y una reintroducción de nuevos materiales en México. En este trabajo se observó que la reintroducción reciente de materiales a México incrementó el acervo genético de moringa. El estudio de la diversidad genética en las poblaciones vegetales permite la identificación de las relaciones genéticas entre las poblaciones, la mezcla de germoplasma y la identificación de progenitores a través de una distancia genética (Delfini *et al.*, 2021).

Estructura genética

El análisis bayesiano identificó que el número más probable de grupos genéticos originales es $K = 2$, ya que se obtuvo el valor más alto de la tasa de cambio de la función de verosimilitud con respecto a K (delta K), por lo que cada población fue asignada a uno de los 2 de los grupos genéticos. En la Figura 4 se muestra los valores de Delta K calculados y computados con Structure Harvester (Evanno *et al.* 2005). Debido a que el agrupamiento jerárquico y el análisis PCoA agruparon a las poblaciones en tres grupos, se continuó el análisis bayesiano y en total fueron identificadas solo 2 poblaciones genéticas originales ($\Delta K = 2$) en las 14 poblaciones de moringa evaluadas. La Figura 5 muestra el porcentaje de contribución genética que comparten las 14 poblaciones de moringa con base en el valor de $K = 2$.

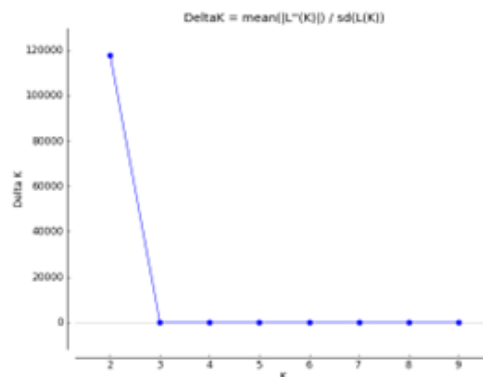


Figura 4. Valores de Delta K para el análisis de STRUCTURE de las poblaciones de *M. oleifera* cultivadas en México.

Fuente: Elaboración propia con base a los resultados Structure Harvester versión 0.6.94.

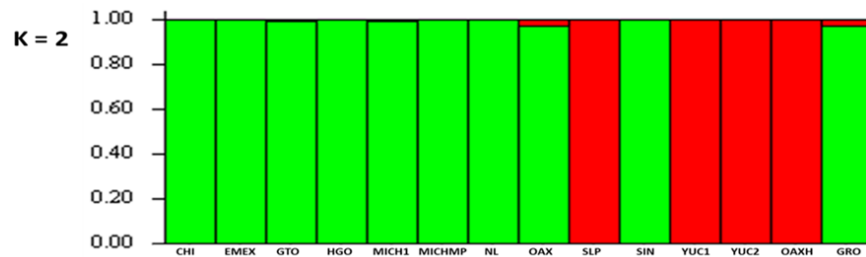


Figura 5. Agrupamiento tipo barplot de 14 poblaciones de *M. oleifera* cultivadas en México caracterizadas con 9,862 SNP según STRUCTURE.

Se observa la distribución de dos poblaciones originales ($K = 2$). Los números en el eje indican el coeficiente de ascendencia. Fuente: Elaboración propia con base a los resultados obtenidos de STRUCTURE versión 2.3.4.

El análisis de la distribución de la ascendencia genética mediante STRUCTURE permitió identificar la proporción compartida entre las poblaciones de moringa. Las poblaciones con puntajes iguales o superiores a 0.800 se consideran con alto grado de pureza y aquellas que presentan valores inferiores a 0.800 son identificadas como heterocigóticas. En el Tabla 4 se muestra la proporción del grupo genético original que comparten las poblaciones de moringa recolectadas en cultivos de México.

Tabla 4. Proporción estimada de ascendencia de las 14 poblaciones de *M. oleifera* Lam.

No.	Población	Grupo genético	
		I	II
1	CHI	1.000	0.000
2	EMEX	1.000	0.000
3	GTO	1.000	0.000
4	HGO	1.000	0.000
5	MICH1	1.000	0.000
6	MICHMP	1.000	0.000
7	NL	1.000	0.000
8	OAX	0.994	0.006
9	SLP	0.000	1.000
10	SIN	1.000	0.000
11	YUC1	0.000	1.000
12	YUC2	0.003	0.997
13	OAXH	0.000	1.000
14	GRO	0.994	0.006

Fuente: Elaboración propia.

El análisis de estructura genética reveló 2 poblaciones genéticas originales ($K = 2$) y que las poblaciones del grupo I del PCoA y conglomerado comparten la misma ascendencia genética original. Se sabe que, en México, los materiales de moringa llegaron desde los viajes de Nao de la China en el año 1500 (Olson & Fahey, 2011) y al menos en estas 14 poblaciones examinadas, aparentemente su origen se encuentra en los mismos ecotipos, o bien en la propagación de ellos. Debido a que las poblaciones de moringa se han distribuido debido a la acción humana se esperaba que el nivel de diversidad genética fuera bajo. Liu *et al.* (2023) mencionan que la distancia geografía entre las poblaciones incrementa la diferenciación genética entre las muestras recolectadas. Sin embargo, esto no fue así para las poblaciones de moringa examinadas.

No obstante, aunque las poblaciones del grupo I comparten la misma contribución genética original, es revelador saber que existe una contribución mínima de la población original II que otorga variación genética a la población OAX. En el grupo II se observó una contribución genética de la población original I a la población de GRO. La diversidad de las poblaciones originales I y II deben ser el punto de partida para planes de conservación y manejos dirigidos. Zhang *et al.* (2017) mencionan que la estructura genética de las poblaciones está influenciada por procesos de domesticación y selección. Otro de los factores que influyen en el flujo genético es la migración (Muli *et al.*, 2022). Sin embargo, no ocurre así en las poblaciones mexicanas de moringa. Las poblaciones de moringa presentes en el país aún no se encuentran en procesos de selección y mejoramiento y el conocimiento de su diversidad y estructura genética brinda información sobre qué poblaciones pueden ser candidatas a un mejoramiento. El intercambio de germoplasma o la generación de poblaciones con dos materiales genéticos contrastantes incrementa la diversidad genética y el número de alelos polimórficos. Sin embargo, limita la formación de grupos definidos en la estructura genética (Raatz *et al.*, 2019). La diversidad genética en las poblaciones enfocadas a la producción de alimentos está influenciada por el manejo humano y la selección de individuos ocasiona la pérdida de genes. La estructura genética permite comprender e identificar un patrón de distribución de las poblaciones vegetales (Cheng *et al.*, 2020).

Conclusiones

Las 14 poblaciones de moringa cultivadas en diversos estados de México son diversas genéticamente. Los SNP utilizados para este discernimiento revelaron que entre ellas hay una diversidad génica que va de baja a moderada con predominancia de esta última. También se identificó que esta diversidad se encuentra formando diferentes genotipos debido a que el valor de la H_o fue mayor que el de la H_e . Estas poblaciones, formaron dos grupos de relación genética, y dentro de ellos, algunas poblaciones son más distantes genéticamente. Estos agrupamientos, estuvieron en concordancia con el número de poblaciones de ascendencia original. Las poblaciones de moringa evaluadas presentan un potencial para el desarrollo de nuevas combinaciones genéticas que incrementen rasgos de interés productivo y nutricional, así como mejor tolerancia a factores ambientales y biológicos.

Contribución de los autores

“Conceptualización del trabajo, Ruiz Hernández, R. y Hernández Rodríguez, M.; desarrollo de la metodología, Ruiz-Hernández, R., Hernández Rodríguez, M. y Pérez Vázquez, A.; manejo de software, Ruiz Hernández, R. y Hernández Rodríguez; validación experimental, Ruiz Hernández, R. y Hernández Rodríguez; análisis de resultados, Ruiz Hernández, R. y Hernández Rodríguez; Manejo de datos, Ruiz Hernández, R. y Hernández Rodríguez; escritura y preparación del manuscrito, Ruiz Hernández, R. y Hernández Rodríguez; redacción, revisión y edición, Ruiz-Hernández, R., Hernández Rodríguez, M. y Pérez Vázquez, A.; administrador de proyectos, Rayas Amor, A.A.; adquisición de fondos, Rayas Amor, A.A.

Todos los autores de este manuscrito han leído y aceptado la versión publicada del mismo

Financiamiento

Esta investigación fue financiada con fondos propios.

Agradecimientos

Al Colegio de Postgraduados Campus Veracruz y Campus Montecillo, por el apoyo otorgado para la realización de este estudio.

Conflicto de interés

Los autores declaran no tener conflicto de interés.

Referencias

- Begna, T. (2022). Role of plant genetic diversity conservation and utilization. *International Journal of Novel Research in Life Sciences*, 9(1), 1-11. <https://www.noveltyjournals.com/upload/paper/ROLE%20OF%20PLANT%20GENETIC-01012022-7.pdf>
- Botstein, D., White., R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *The American Journal of Human Genetics*, 32(3), 314-331.
- Calicioglu, O., Flammini, A., Bracco, S., Bellù, L., & Sims, R. (2019). The future challenges of food and agriculture: an integrated analysis of trends and solutions. *Sustainability*, 11(1), 222. <https://doi.org/10.3390/su11010222>
- Chang, Y., Liu, H., Liu, M., Liao, X., Sahu, S. K., Fu, Y., Song, B., Cheng, S., Kariba, R., Muthemba, S., Hendre, P. S., Mayes, S., Ho, W. K., Yssel, A. E. J., Kendabie, P., Wang, S., Li, L., Muchugi, A.,

- Jamnadass, R., Lu, H., Peng, S., Van Deynze, A., Simons, A., Yana-Shapiro, H., Van de Peer, Y., Xu, X., Yang, H., Wang, J., & Liu, X. (2018). The draft genomes of five agriculturally important African orphan crops. *GigaScience*, 8(3), 1-16. <https://doi.org/10.1093/gigascience/giy152>
- Cheng, J., Kao, H., & Dong, S. (2020). Population genetic structure and gene flow of rare and endangered *Tetraena mongolica* Maxim. revealed by reduced representation sequencing. *BMC Plant Biology*, 20(1), Article 391. <https://doi.org/10.1186/s12870-020-02594-y>
- Centro Internacional de Mejoramiento de Maíz y Trigo [CIMMYT]. (2006). Protocolos de laboratorio: Laboratorio de Genética Molecular Aplicada del CIMMYT. (3a ed.). México, D.F. <https://repository.cimmyt.org/bitstream/handle/10883/593/91224.PDF?sequence=4&isAllowed=y>
- Das, D., & Mishra, R. (2021). SNP markers and its impact on crop improvement. *Asian Journal of Biological and Life Sciences*, 10(3), 539-548. <https://doi.org/10.5530/ajbls.2021.10.72>
- Delfini, J., Moda-Cirino, V., dos Santos Neto, J., Ruas, P.M., Sant'Ana, G.C., Gepts, P., & Gonçalves, L. S. A. (2021). Population structure, genetic diversity and genomic selection signatures among a Brazilian common bean germplasm. *Scientific Reports*, 11(1), 2964. <https://doi.org/10.1038/s41598-021-82437-4>
- Dube, S. P., Sibiya, J., & Kutu, F. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13(1), 17851. <https://doi.org/10.1038/s41598-023-44961-3>
- Earl, D. A., & vonHoldt, B. M. (2012). Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359-361. <https://doi.org/10.1007/s12686-011-9548-7>
- Elbasyoni, I. S., Eltaher, S., Morsy, S., Mashaheet, A. M., Abdallah, A. M., Ali, H. G., Mariey, S. A., Baenziger, P. S., & Frels, K. (2022). Novel single-nucleotide variants for morpho-physiological traits involved in enhancing drought stress tolerance in barley. *Plants*, 11(22), Article 3072. <https://doi.org/10.3390/plants11223072>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14(8), 2611-2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Gawroński, P., Pawełkowicz, M., Tofil, K., Uszyński, G., Sharifova, S., Ahluwalia, S., Tyrka, M., Wędzony, M., Kilian, A., & Bolibok-Brągoszewska, H. (2016). DArT markers effectively target gene space in the rye genome. *Frontiers in Plant Science*, 7(6), 1600. <https://doi.org/10.3389/fpls.2016.01600>
- Gouda, A. C., Warburton, M. L., Djedatin, G. L., Kpeki, S. B., Wambugu, P. W., Gnikoua, K., & Ndjioudjop, M. N. (2021). Development and validation of diagnostic SNP markers for quality control genotyping in a collection of four rice (*Oryza*) species. *Scientific Reports*, 11(1), 18617. <https://doi.org/10.1038/s41598-021-97689-3>
- Hendre, P. S. & Van Deynze, A. (2015). How a consortium is changing Africa's food systems? <http://www.africanorphancrops.org/category/resources/blog/page/3/>
- Hendre, P. S., Muthemba, S., Kariba, R., Muchugi, A., Fu, Y., Chang, Y., Song, B., Liu, H., Liu, M., Liao, X., Sahu, S. K., Wang, S., Li, L., Lu, H., Peng, S., Cheng, S., Xu, X., Yang, H., Wang, J., Liu, X., Simons, A., Shapiro, H. Y., Mumm, R. H., Van Deynze, A., & Jamnadass, R. (2019). African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. *Planta*, 250(3), 989-1003. <https://doi.org/10.1007/s00425-019-03156-9>
- Huster, A. R., Wallace, L. T., & Myers, J. R. (2021). Associated SNPs, Heritabilities, trait correlations, and genomic breeding values for resistance in snap beans (*Phaseolus vulgaris* L.) to root rot caused by

- Fusarium solani* (Mart.) f. sp. *phaseoli* (Burkholder). *Frontiers in Plant Science*, 12(2021), 697615. <https://doi.org/10.3389/fpls.2021.697615>
- Liu, H., Wang, Z., Zhang, Y., Li, M., Wang, T., & Su, Y. (2023). Geographic isolation and environmental heterogeneity contribute to genetic differentiation in *Cephalotaxus oliveri*. *Ecology and Evolution*, 13(3), e9869. <https://doi.org/10.1002/ece3.9869>
- Muhl, Q. E., Toit, E. S. D., & Robbertse, P. J. (2011). *Moringa oleifera* (Horseradish Tree) leaf adaptation to temperature regimes. *International Journal of Agriculture and Biology*, 13(6), 10-21. <https://www.up.ac.za/media/shared/Legacy/sitefiles/file/48/2056/31fijab1112510211024.pdf>
- Muli, J. K., Neondo, J. O., Kamau, P. K., Michuki, G. N., Odari, E., & Budambula, N. L. M. (2022). Genetic diversity and population structure of wild and cultivated *Crotalaria* species based on genotyping-by-sequencing. *PLoS ONE*, 17(9), e0272955. <https://doi.org/10.1371/journal.pone.0272955>
- Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doğan, Y., Comertpay, G., Yıldız, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N., Özkan, H., Chung, G., & Baloch, F. S. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment*, 32(2), 261–285. <https://doi.org/10.1080/13102818.2017.1400401>
- Ndhlala, A. R., & Tshabalala, T. (2023). Diversity in the nutritional values of some *Moringa oleifera* Lam. cultivars. *Diversity*, 15, 834. <https://doi.org/10.3390/d15070834>
- Olson, M. E., & Fahey, J. W. (2011). *Moringa oleifera*: un árbol multiusos para las zonas tropicales seca. *Revista Mexicana de Biodiversidad*, 82(4), 1071-1082. <https://doi.org/10.22201/ib.20078706e.2011.4.678>
- Ousmael, K., Whetten, R. W., Xu, Nielsen, U. B., Lamour, K., & Hansen O. K. (2023). Identification and high-throughput genotyping of single nucleotide polymorphism markers in a non-model conifer (*Abies nordmanniana* (Steven) Spach). *Scientific Reports*, 13(2023), 22488. <https://doi.org/10.1038/s41598-023-49462-x>
- Pacheco-Gil, R. Á., Alvarado, G., Rodríguez, F., & Burgueño, J. (2016). BIO-R (Biodiversity analysis with R for Windows) Version 3.3. CIMMYT Research Data & Software Repository Network, 14. <https://hdl.handle.net/11529/10820>
- Pasha, S. N., Shafi, K. M., Joshi, A. G., Meenakshi, I., Harini, K., Mahita, J., Sajeevan, R. S., Karpe, S. D., Ghosh, P., Nitish, S., Gandhimathi, A., Mathew, O. K., Prasanna, S. H., Malini, M., Mutt, E., Naika, M., Ravooru, N., Rao, R. M., Shingate, P. N., Sukhwal, A., Sunitha, M. S., Upadhyay, A. K., Vinekar, R. S., & Sowdhamini, R. (2020). The transcriptome enables the identification of candidate genes behind medicinal value of drumstick tree (*Moringa oleifera*). *Genomics*, 112(1), 621–628. <https://doi.org/10.1016/j.ygeno.2019.04.014>
- Peter A. F., Wagiran, A., Rahmat Z., Yusop M. R., & Ridzuan, R. (2023). Exploring genetic variation and therapeutic properties of *Moringa oleifera*: progress and future potential for crop improvements. *Pharmacognosy Reviews*, 17(34), 426-438. <https://doi.org/10.5530/phrev.2023.17.18>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Raatz, B., Mukankusi, C., Lobaton, J. D., Male, A., Chisale, V., Amsalu, B., Fourie, D., Mukamuhirwa, F., Muimui, K., Mutari, B., Nchimbi-Msolla, S., Nkalubo, S., Tumsa, K., Chirwa, R., Maredia, M. K., & He, C. (2019). Analyses of African common bean (*Phaseolus vulgaris* L.) germplasm using a SNP fingerprinting platform: diversity, quality control and molecular breeding. *Genetic Resources and Crop Evolution*, 66(3), 707–722. <https://doi.org/10.1007/s10722-019-00746-0>

- Ruiz-Hernández, R., Hernandez-Rodriguez, M., Cruz-Monterrosa, R. G., Diaz-Ramirez, M., Martinez-García, C. G., Garcia-Martinez, A., & Amor, A. A. R. (2022a). *Moringa oleifera* Lam.: una revisión de los factores ambientales y de manejo que influyen en el contenido nutricional de las hojas. *Tropical and Subtropical Agroecosystems*, 25(1), 1-15. <https://doi.org/10.56369/tsaes.4053>
- Ruiz-Hernández, R., Hernández-Rodríguez, M., Pérez-Vázquez, A., Cruz-Monterrosa, R. G., & Rayas-Amor, A. A., (2022b). Chemical composition of twelve accessions of *Moringa oleifera* Lam. grown in Mexico. *Agro Productividad*. 15(11), 11-20. <https://doi.org/10.32854/agrop.v15i10.2405>
- Ruiz-Hernández, R., Pérez-Vázquez, A., García-Pérez, E., Morales-Trejo, F., & Soto-Hernández, R. M. (2021). Caracterización morfológica de accesiones de *Moringa oleifera* provenientes del sur-sureste de México. *Revista Mexicana De Ciencias Agrícolas*, 12(7), 1210–2022. <https://doi.org/10.29312/remexca.v12i7.2632>
- Ruiz-Hernández, R., Pérez-Vázquez, A., Hernández-Rodríguez, M., García-Pérez, E., Morales-Trejo, F., & Soto-Hernández, R. M. (2024). Genetic diversity of *Moringa oleifera* Lam. in southeastern Mexico. *Agrociencia*, 58(6), 1-17. <https://doi.org/10.47163/agrociencia.v58i6.3101>
- Shyamli, P. S., Pradhan, S., Panda, M., & Parida, A. (2021). *De novo* whole-genome assembly of *Moringa oleifera* helps identify genes regulating drought stress tolerance. *Frontiers in Plant Science*, 12(1), 766999. <https://doi.org/10.3389/fpls.2021.766999>
- Teklemariam, S. S., Bayissa, K. N., Matros, A., Pillen, K., Ordon, F., & Wehner, G. (2022). The genetic diversity of Ethiopian barley genotypes in relation to their geographical origin. *PLoS ONE*, 17(5), e0260422. <https://doi.org/10.1371/journal.pone.0260422>
- Tian, Y., Zeng, Y., Zhang, J., Yang, C., Yan, L., Wang, X., Shi, C., Xie, J., Dai, T., Peng, L., Zeng Huan, Y., Xu, A., Huang, Y., Zhang, J., Ma, X., Dong, Y., Hao, S., & Sheng, J. (2015). High quality reference genome of drumstick tree (*Moringa oleifera* Lam.), a potential perennial crop. *Science China Life Sciences*, 58(7), 627–638. <https://doi.org/10.1007/s11427-015-4872-x>
- Tomar, V., Dhillon, G. S., Singh, D., Singh, R. P., Poland, J., Joshi, A. K., Tiwari B. S., & Kumar, U. (2021). Elucidating SNP-based genetic diversity and population structure of advanced breeding lines of bread wheat (*Triticum aestivum* L.). *PeerJ*, 9, e11593. <http://doi.org/10.7717/peerj.11593>
- Wu, J. C., Zhang, Y. P., Zheng, Y. X., & Peng, X. M. (2018). Pollen mediated gene flow in a small experimental population of *Moringa oleifera* Lam. (Moringaceae). *Industrial Crops and Products*, 117(1), 28-33. <https://doi.org/10.1016/j.indcrop.2018.02.077>
- Yirgu, M., Kebede, M., Feyissa, T., Lakew, B., Woldeyohannes, A. B. & Fikere, M. (2023). Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm. *BMC Genomic Data*, 24(1), 7. <https://doi.org/10.1186/s12863-023-01109-6>
- Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B. H. (2017). Back into the wild-apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, 10(1), 5-24. <https://doi.org/10.1111/eva.12434>
- Zhang, J., Yang, E., He, Q., Lin, M., Zhou, W., Pian, R., & Chen, X. (2019). Genome-wide analysis of the WRKY gene family in drumstick (*Moringa oleifera* Lam.). *PeerJ*, 7(1), e7063. <https://doi.org/10.7717/peerj.7063>