

ANÁLISIS DEL GENOMA DEL CLOROPLASTO DEL MAÍZ MORADO INIA 601 PARA RECONSTRUIR LA HISTORIA EVOLUTIVA DEL MAÍZ MORADO PERUANO

Lila M. Rodríguez Pérez, Juan D. Montenegro, Reinhard Simon, Lenin Chumbe Nolasco, Fernando Serna, Gabriel Delgado, Aura Liz García Serquén, Dina L. Gutiérrez

macielerodriguez2@gmail.com, jdmontenegroc@gmail.com, ldchumbe@gmail.com, agarcias@inia.gob.pe, dgutierrez@inia.gob.pe, drgb@inia.gob.pe

Instituto Nacional de Innovación Agraria, Dirección de Recursos Genéticos y Biotecnología, Subdirección de Biotecnología

RESUMEN

En el presente estudio, se ha reconstruido la secuencia completa del genoma plastidial del maíz morado peruano y se ha comparado con otros genomas plastidiales de maíces. El genoma plastidial tiene una longitud de 140,458 pb y muestra una estructura típica del genoma del cloroplasto: un par de regiones repetidas invertidas (IRa e IRb) de 22,594 pb, una región Larga de Copia Única (LSC) de 82,472 pb, y una región Corta de Copia Única (SSC) de 12,798 pb. Las relaciones filogenéticas fueron obtenidas a partir de alineamientos genómicos completos con los genomas plastidiales de otros miembros del género *Zea*. Los resultados indicaron que el maíz morado peruano está más relacionado a *Z. mays* subsp. *huehuetenangensis*. Es posible que eventos de hibridación introgresiva a lo largo de la evolución del maíz morado hayan jugado un papel en la adquisición del fenotipo morado en el clado *Z. mays*.

ABSTRACT

In this study, the complete sequence of the plastidial genome of Peruvian purple maize has been reconstructed and compared to other complete maize plastidial genomes. The plastid genome is 140,458 bp in length and shows a typical architecture of the chloroplast genome, with a pair of inverted repeated regions (IRa and IRb) of 22,594 bp, a Long Single Copy region (LSC) of 82,472 bp, and a Short Single Copy region (SSC) of 12,798 bp. The phylogenetic relationships were obtained from whole genome alignments with plastidial genomes of other members of the genus *Zea*. The results indicated that Peruvian purple maize is more related to *Z. mays* subsp. *huehuetenangensis*. It is possible that introgressive hybridization events throughout the evolution of purple maize have played a role in the acquisition of the purple phenotype in the *Z. mays* clade.

INTRODUCCIÓN

El maíz morado (*Zea mays* L. var. *subnigrovioleaceo*) es una variedad nativa de América cuyo color particular se debe a la alta concentración de antocianina, un pigmento que le otorga su color característico y propiedades antioxidantes (Bonilla et al. 2015). El maíz morado INIA 601 es una de las variedades mejoradas generadas por el Instituto Nacional de Innovación Agraria del Perú - INIA, la cual se destaca por su alto rendimiento (hasta 5.2 toneladas por hectárea), amplia capacidad de adaptación a diferentes climas y alta intensidad en el color morado de la tusa y del grano (Abanto et al. 2003). Recientemente, el INIA 601 fue identificado como la variedad que presenta el mayor contenido de antocianinas, con 6.34 % de antocianinas en coronta y 3.03 % en panca (INIA, 2019). Por lo expuesto, es importante tener un mejor conocimiento sobre el origen de esta variedad para elucidar el origen evolutivo del fenotipo morado y del control transcripcional de genes de antocianinas para poder utilizarlos eficientemente en programas de mejoramiento genético. Por ello, el presente trabajo tiene como objetivo emplear el genoma del cloroplasto del maíz morado INIA 601 para reconstruir su historia evolutiva y su relación filogenética con otros maíces de América.

MATERIALES Y MÉTODOS

Se realizó el secuenciamiento del genoma completo (WGS) de un pool de 8 individuos de maíz morado INIA 601, empleando la tecnología Illumina y el kit TruSeq DNA PCR Free. Se generó una biblioteca de lecturas emparejadas (Paired-end) con un tamaño de inserto de 350 pb. El ensamblaje *denovo* guiado por referencia del genoma del cloroplasto se realizó con el programa NOVOPlasty v3.3 (Dierckxsens et al. 2017) a partir de las bibliotecas limpias. La anotación del genoma del cloroplasto se realizó con GeSeq v1.74 (Tillich et al. 2017). El diagrama circular del genoma se generó con OGDRAW v1.3.1 (Greiner et al. 2019). En el análisis filogenético, se comparó el genoma plastidial INIA 601 recién ensamblado, con otros genomas plastidiales del género *Zea*. Estos fueron alineados empleando MAFFT v7 (Kato y Standley, 2013), y el análisis de Maximum Likelihood (ML) se realizó empleando los programas IQ-TREE v1.6.12 (Nguyen et al. 2015), ModelFinder (Kalyaanamoorthy et al. 2017) y RAxML-NG v0.6.0 (Kozlov et al. 2019). El árbol filogenético fue editado en el programa FigTree v1.4.4 (Rambaut 2018).

RESULTADOS Y DISCUSIÓN

El genoma plastidial ensamblado presenta la arquitectura cuatripartita característica de los genomas plastidiales de plantas terrestres, con dos regiones invertidas repetidas IRa e IRb (22,594 pb) separadas por la región Larga de Copia Simple LSC (82,472 pb) y la región Corta de Copia Simple SSC (12,798 pb). El contenido de GC del genoma fue de 38.43%. El genoma ensamblado contiene 127 genes de los cuales 84 son genes codificantes de proteínas, 35 son genes de tRNA y 8 son genes de rRNA. El análisis filogenético mostró que el maíz morado forma

un grupo monofilético con *Z. mays* subsp. *mays* y *Zea mays* subsp. *huehuetenangensis*. Sin embargo, se observó que el maíz morado presenta una mayor cercanía a esta última subespecie de *Zea mays*. Esto indicaría que el maíz morado es resultado de eventos de introgresión de especies silvestres. Procesos similares de introgresión ya han sido reportados en el género *Zea* (Hufford et al. 2012b). Debido a que el número de genomas plastidiales de miembros de este género disponibles en las bases de datos es limitado, estos resultados pueden ser ampliados y mejor definidos en futuros trabajos. No obstante, cabe resaltar que la topología generada fue congruente con estudios similares empleando el genoma plastidial completo de diferentes maíces americanos (Orton et al. 2016; Bosacchi et al. 2015), y con estudios realizados en base a SNP's en el genoma nuclear (Hufford et al. 2012a).

CONCLUSIÓN

Dada la cercanía filogenética del maíz morado INIA 601 a *Z. mays* subsp. *huehuetenangensis*, es probable que el maíz morado peruano pueda haber experimentado hibridación introgresiva en algún punto de su historia evolutiva.

El genoma completo del cloroplasto del maíz morado INIA 601 es un recurso genético que, además de ser empleado para inferir relaciones filogenéticas, también puede ser empleado como fuente de marcadores moleculares (SSR y repeticiones en Tándem) en futuros programas de mejoramiento o en estudios filogeográficos y de especiación.

REFERENCIAS BIBLIOGRÁFICAS

- Abanto, W; Medina, A; Injante, P. 2003. Maíz INIA 601. Variedad mejorada de maíz morado para la sierra norte del Perú. Estación Experimental Agraria Baños del Inca. Dirección general de proyección y servicios agrarios.
- Bonilla, PE; Quispe, F; Negrón, L; Zavaleta, AI. 2015. Compuestos bioactivos y análisis sensorial de una bebida funcional de maíz morado (*Zea mays* L.) y estevia (*Stevia* SP.). *Ciencia e Investigación* 2015; 18(1): 37-42.
- Bosacchi, M; Gurdon, C; Maliga, P. 2015. Plastid Genotyping Reveals the Uniformity of Cytoplasmic Male Sterile-T Maize Cytoplasms. *Plant Physiology* 169: 2129–2137.
- Dierckxsens, N; Mardulyn, P; Smits, G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* 45(4): e18.
- Greiner, S; Lehwark, P; Bock, R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47: W59-W64.
- Hufford, MB; Bilinski, P; Pyhäjärvi, T; Ross-Ibarra, J. 2012a. Teosinte as a model system for population and ecological genomics. *Trends in Genetics* 28(12): 606–615.
- Hufford, MB; Xu, X; Van Heerwaarden, J; Pyhäjärvi, T; Chia, JM; Cartwright, RA; Ross-Ibarra, J. 2012b. Comparative population genomics of maize domestication and improvement. *Nature Genetics* 44(7): 808–811.
- INIA. 2019. INIA identifica variedad de maíz morado con mayor contenido de antocianinas y potencial de rendimiento (en línea). INIA, Lima, Perú; 13 de jun. 2019. Disponible en <http://www.inia.gob.pe/2019-nota-075/>
- Kalyaanamoorthy, S; Minh, BQ; Wong, TKF; Haeseler, A; Jermin, LS. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587-589.
- Katoh, K; Standley, DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*. *Molecular Biology and Evolution* 30(4): 772-780.
- Kozlov, AM; Darriba D; Flouri, T; Morel, B; Stamatakis, A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, btz305, 1-3.
- Nguyen, LT; Schmidt, HA; Haeseler, A; Minh, BQ. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution* 32: 268-274.
- Orton, LM; Burke, SV; Wysocki, WP; Duvall, MR. 2016. Plastid phylogenomic study of species within the genus *Zea*: rates and patterns of three classes of microstructural changes. *Current Genetics* 63(2): 311-323.
- Rambaut, A. 2018. FigTree as a graphical viewer of phylogenetic trees and a program for producing publication-ready figures. Disponible en: <https://github.com/rambaut/figtree>
- Tillich, M; Lehwark, P; Pellizzer, T; Ulbricht-Jones, ES; Fischer, A; Bock, R; Greiner, S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* 45: W6-W11.