

BV52. Optimización de un protocolo de secuenciación de genoma completo de baja redundancia para girasol cultivado

Fass, M.I.*; Aguirre, N.; Gonzalez, S.; Martinez, M.C.; Vera, P.; Puebla, A.; Lia, V.V.; Marcucci Poltri, S.; Paniego, N.

Instituto de Agrobiotecnología y Biología Molecular (IABIMO), Instituto Nacional de Tecnología Agropecuaria (INTA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. *fass.monica@inta.gob.ar

El girasol cultivado, *Helianthus annuus* L., es un importante cultivo oleaginoso. Es valorado por su alto contenido de ácido linoleico, oleico, palmítico, esteárico y otros ácidos grasos insaturados. Además, mantiene un rendimiento estable a lo largo de distintos ambientes, favoreciendo su producción globalmente. Distintos programas de mejoramiento en el mundo se dedican a generar y mantener colecciones de líneas en pos de incorporar características agrónomicamente significativas a los cultivares comerciales, haciendo uso de las variantes alélicas presentes en los materiales. El conocimiento genómico de las líneas de mejoramiento y pre-mejoramiento existentes resulta de gran utilidad para dilucidar los mecanismos moleculares de los caracteres agronómicos y seleccionar las líneas más promisorias. El genoma de girasol, de aproximadamente 3.6 Gb, fue secuenciado completamente a partir de la implementación de las tecnologías de secuenciación de tercera generación. La posibilidad de generar secuencias más largas que las secuencias repetitivas de girasol, los métodos mejorados de ensamblado y los recursos computacionales avanzados se convirtieron en herramientas valiosas para la generación de una referencia genómica completa para el cultivo. Actualmente, se cuenta con información genómica de distintas líneas mayoritariamente generadas con baja cobertura. Con el objetivo de caracterizar a nivel de genoma completo a la línea HA89, una de las líneas públicas más utilizadas tanto en mejoramiento como en investigación básica de girasol, se ha propuesto ensayar un protocolo de secuenciación a baja escala y baja redundancia basado en la generación de librerías libres de amplificación. Para la obtención de moléculas de alto peso molecular se realizó una extracción de núcleos seguido de la purificación del ADN. Posteriormente se generaron bibliotecas mediante el protocolo Nextera Mate Pair Gel-Free Sample Preparation (NEB, EEUU) y se procedió a una secuenciación de prueba con un equipo MiSeq (Illumina, EEUU). Se obtuvieron 177149 lecturas, de las cuales 95% pudo ser alineado al genoma de referencia XRQ v1. Estos resultados indican una buena calidad de la biblioteca y permitirán aumentar la cobertura de la secuenciación. A su vez, el protocolo de extracción de ADN implementado permite su combinación con tecnologías de secuenciación de tercera generación, asegurando un ensamblado del genoma de mejor calidad, así como también la caracterización de variantes estructurales.

La colección de datos genómicos obtenidos de diversos materiales posibilitará la construcción de un pangenoma y de expandir el uso de la diversidad genética en el mejoramiento.