

Construcción de un nuevo mapa genético de ligamiento en albaricoquero basado en SNPs

P. Irisarri¹, P. Errea¹, T. Zhebentyayeva² y A. Pina¹

¹Unidad de Hortofruticultura, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Instituto Agroalimentario de Aragón-IA2 (CITA-Universidad de Zaragoza), Zaragoza, España. ²Clemson University Genomics & Computational Biology Laboratory, Clemson, SC 29634, USA

Palabras clave: genómica, genotipado por secuenciación (GBS), segregación, *Prunus armeniaca* L

Resumen

En este estudio se describe la creación del primer mapa de ligamiento genético en albaricoquero (*Prunus armeniaca* L.) mediante genotipado por secuenciación (GBS) a partir de una población F1 de 138 individuos obtenidos del cruce intra-específico de los cultivares de albaricoquero 'Moniquí' y 'Paviot' (MoxPa). Debido al contraste fenotípico de los parentales, este cruce segrega para varios caracteres agronómicos importantes como la autocompatibilidad floral, la tendencia a la compatibilidad del injerto y el color del fruto. Los resultados de genotipado en los parentales y en la descendencia proporcionaron un conjunto de 26.593 SNPs. Para los cálculos asociados a la construcción de los mapas genéticos se utilizó el programa JoinMap 4.1. Los mapas desarrollados incluyen ocho grupos de ligamientos y tienen una cobertura de 780.2 cM y 690.4 cM para 'Moniquí' y para 'Paviot', respectivamente. Los mapas obtenidos proporcionan un valioso conjunto de SNPs útiles para la identificación de QTLs (Quantitative Trait Loci) que podrán servir como herramienta en los programas de mejora de albaricoquero que actualmente se llevan a cabo.

INTRODUCCIÓN

El albaricoquero (*Prunus armeniaca* L.) es una de las especies frutales de clima templado más importantes, con una producción mundial total alrededor de 3.88 millones de toneladas en 2016 (FAOSTAT). Dentro de los problemas que presenta este cultivo se encuentra la poca disponibilidad de variedades comerciales compatibles con un amplio número de patrones. Sin embargo, existe poca información sobre la base genética de este carácter. Hasta el momento, se han llevado a cabo numerosos estudios para determinar la heredabilidad de gran variedad de caracteres de interés agronómico (fecha de floración, necesidades de frío, resistencia al virus de la sharka) y se han construido varios mapas de ligamiento en albaricoquero usando AFLPs y microsatélites (SSRs) (Vilanova et al., 2003; Salazar et al., 2013). En los últimos años, el desarrollo de la tecnología de genotipado por secuenciación (GBS) basada en la reducción previa de la complejidad genómica, permiten la identificación simultánea de un gran número de SNPs mediante plataformas de alto rendimiento, siendo una alternativa rentable para desarrollar miles de marcadores SNPs útiles para la construcción de mapas de ligamiento (He et al., 2014). El objetivo principal de este trabajo ha sido la construcción de un nuevo mapa genético de ligamiento a partir de una población de albaricoquero derivada del cruce intra-específico 'Moniquí' y 'Paviot' genotipada mediante GBS. Estos nuevos mapas saturados proporcionan una herramienta

muy valiosa para la identificación de loci de caracteres cuantitativos (QTL) y otras aplicaciones para programas de mejora en albaricoquero.

MATERIAL Y MÉTODOS

Material vegetal y construcción de librerías

Las muestras seleccionadas para el análisis fueron 138 individuos de una población F1 de albaricoque procedentes de un cruce intra-específico de los cultivares 'Moniquí' y 'Pavot' (MoxPa). El ADN genómico fue extraído de hojas utilizando un protocolo CTAB estandarizado y se purificó usando el kit DNeasy plant (Quiagen, Germantown, MD USA). La cuantificación se realizó en geles de agarosa al 1% empleando el marcador lambda λ (life technologies), con el fin de obtener un mínimo de 1 μ g de ADN a 100 ng/ μ L y se determinó su calidad realizando digestión enzimática con las enzimas ApeK1, Mse1 y Pst1 (New England Biolabs) de muestras seleccionadas al azar. Una vez se obtuvo ADN con concentración y calidad necesaria se enviaron las muestras de cada individuo a la Universidad de Clemson ('Computational Biology Laboratory') para ser procesadas por el protocolo GBS. Las muestras de ADN fueron digeridas con la enzima de restricción ApeK1 que reconoce sitios de 5 pb (CGWGC) y la preparación de las librerías se llevó a cabo siguiendo el protocolo de Elshire et al. (2011). Los primers y adaptadores se eliminaron y las lecturas se demultiplexaron utilizando el comando 'process_rad-tags' implementado por el programa Stacks (Catchen et al., 2013). En total, el 97% de las lecturas se conservaron después del control de calidad (QC > 30) y la presencia del sitio de restricción ApeK1. Para el procesamiento de datos se renombraron como "ind_1 ... 138". EL ADN de los padres fue secuenciado tres veces para incrementar el número y precisión de los SNPs identificados.

Alineamiento y búsqueda de SNPs

Las secuencias filtradas se alinearon utilizando el genoma de melocotonero (*Prunus persica*) v2.0 con el programa genómico de alineación de nucleótidos de lectura breve GSNAP versión 2015-07-23 en parámetros predeterminados (Wu and Nacu, 2010). Además, se utilizó Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009) para alinear las lecturas de GBS y comparar los resultados con GSNAP. También verificamos el enfoque de Stacks 'de novo' (*SNP calling* sin genoma de referencia). Para el filtrado y validación de los SNPs '*SNP calling*' se utilizaron los parámetros de Stacks "m5, n4" (número mínimo de lecturas en stacks en 5 y 4 errores en comparación con la referencia conocida).

Análisis de ligamiento y construcción de mapas genéticos

El análisis de ligamiento genético se realizó empleando la estrategia de "two-way pseudo testcross" (Grattapaglia and Sederoff, 1994). Se construyeron mapas de ligamiento para cada parental (cultivar femenino 'Moniquí' y cultivar masculino 'Pavot') usando CP (polinización cruzada) como tipo de población. Solo los marcadores polimórficos en un padre (configuraciones <lm x ll> y <nn x np>) se utilizaron para la construcción del mapa. Individuos con más del 10 % de datos perdidos fueron excluidos del análisis y para convertir las unidades de recombinación a distancias genéticas se utilizó la función de Kosambi. La representación gráfica del mapa se realizó con el programa Mapchart v2.5 a partir de los datos obtenidos con el programa JoinMap4 (Van Ooijen, 2006).

RESULTADOS Y DISCUSIÓN

En el presente trabajo se ha desarrollado un nuevo mapa genético en albaricoquero usando la tecnología de genotipado por secuenciación (GBS), la cual permite identificar miles de marcadores tipo polimorfismos de un solo nucleótido (SNPs) (Elshire et al., 2011). Los genotipos de los padres y de la descendencia se cortaron en los sitios de restricción Apek1, que produjo el mayor número de fragmentos y se observó que es una de las enzimas más adecuadas para construir librerías de ADN para GBS (Guajardo et al., 2015). Las lecturas unidireccionales se procesaron usando Stacks v1.35, las lecturas filtradas se alinearon con el genoma de melocotonero v2.0 y se realizó un análisis de la distribución a lo largo de los grupos de ligamiento de melocotonero. El porcentaje de lecturas únicas alineadas correctamente con el genoma de referencia fue mayor en los archivos bam generados con GSNAP en comparación con BWA (51.27% vs. 40.58% respectivamente). Por lo tanto, usamos la alineación GSNAP para el genotipado de SNPs. Además, el análisis Stacks 'de novo' produjo un conjunto de datos de SNP menos robusto y fue descartado. Al utilizar los parámetros por defecto del 'pipeline' de GBS cuando se usó como referencia el genoma de melocotonero, se identificaron 26.593 SNPs, con un porcentaje de datos perdidos de 10%. Se observó una media de lecturas GBS de 1.934.186 por individuo. Después de aplicar los filtros de cobertura mínima, el número de marcadores tipo SNPs obtenidos fue 7.618 de alta calidad distribuidos en los ocho grupos de ligamiento, que contienen 1.551 (23%) SNPs con segregación materna <lm x ll>, 1.349 (18%) con segregación paterna <nn x np y 4.718 (62%) con segregaciones tipo <ab x cd>, <ef x eg> y <hk x hk>, heterocigotos en ambos padres (Figura 1). El número de SNPs osciló entre 622 en el grupo de ligamiento 7 a 1.835 en el grupo de ligamiento 1. Los 2.900 marcadores con segregación materna y paterna se utilizaron para la construcción de los mapas genéticos de ligamiento, dos mapas parentales y un mapa consenso para aplicaciones posteriores. Los mapas del parental femenino 'Moniquí' (Mo) y masculino 'Paviot' (Pa) estaban compuestos de 557 y 501 SNPs y cubrían 780.2 y 690.4 cM, respectivamente. Los mapas de 'MoxPa' proporcionan un valioso conjunto de SNPs basados en secuencias útiles para la identificación de loci de rasgos cuantitativos (QTL) y análisis genéticos adicionales relevantes en programas de mejora de albaricoquero. En este sentido, la población genotipada en este trabajo segrega para varios caracteres agronómicos relacionados con la compatibilidad de injerto (línea necrótica, discontinuidad en la madera y corteza) y se está determinando las regiones del genoma implicadas en la expresión de este carácter.

AGRADECIMIENTOS

Esta investigación ha sido financiada por los proyectos de investigación RTA2012-00097 y RTA2015-00046 del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) y el Grupo de Investigación A12 de Aragón, España.

Referencias

- Catchen, J.M., Hohenlone, P.A., Bassham, S., Amores, A. and Cresko, W. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22:3124–3140.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and Mitchell, S.E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:1–10.
- FAO, n.d. Food and Agriculture Organization of the United Nations. FAO Stat. database World Wide Web. URL <http://faostat.fao.org>
- Grattapaglia, D. and Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and

- Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*. 137:1121–37.
- Guajardo, V., Solís, S., Sagredo, B., Gainza, F., Muñoz, C. and Gasic, K. and Hinrichsen, P. 2015. Construction of High Density Sweet Cherry (*Prunus avium* L.) Linkage Maps Using Microsatellite Markers and SNPs Detected by Genotyping-by-Sequencing (GBS). *PLoS One* 10: e0127750.
- He, J., Zhao, X., Laroche, A., Lu, Z., Liu, H. and Li, Z. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant. Sci.* 5:1–8.
- Li, H. and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754–1760.
- Salazar, J.A., Ruiz, D., Egea, J. and Martínez-Gómez, P. 2013. Transmission of Fruit Quality Traits in Apricot (*Prunus armeniaca* L.) and Analysis of Linked Quantitative Trait Loci (QTLs) Using Simple Sequence Repeat (SSR) Markers. *Plant Mol. Biol. Report.* 31:1506–1517.
- Van Ooijen, J. 2006. JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations.
- Vilanova, S., Romero, C., Abbott, AG., Llácer, G. and Badenes, M.L. 2003. An apricot (*Prunus armeniaca* L.) F2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. *Theor. Appl. Genet.* 107:239–47.
- Wu, T.D. and Nacu, S. 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*. 26:873–881.

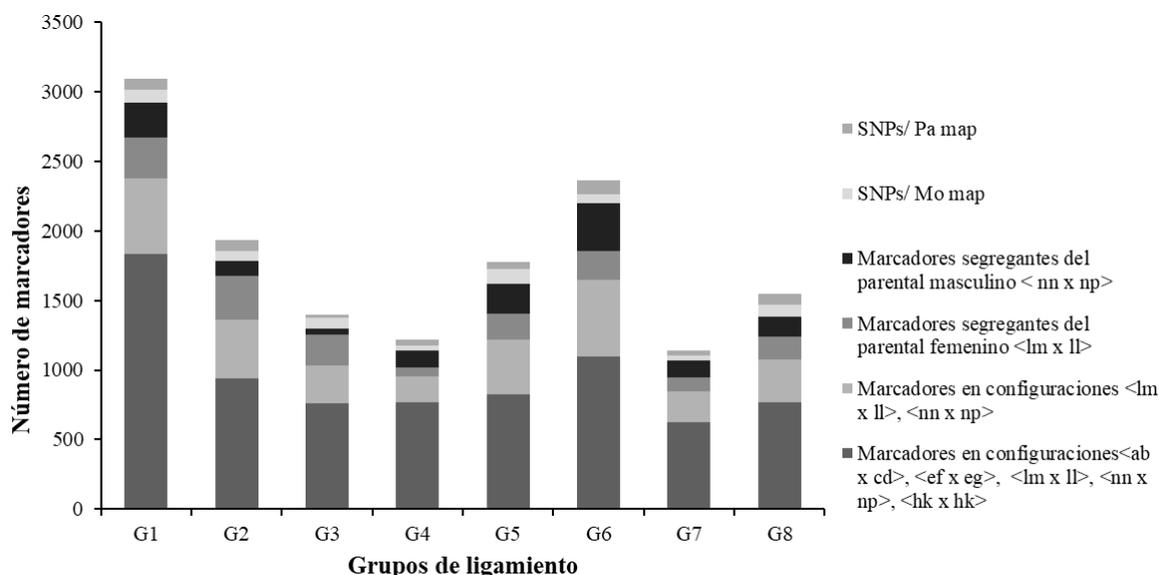


Fig. 1. Distribución de marcadores SNPs en función de la segregación en los 8 grupos de ligamiento dentro de la población F1 del cruce ‘Moniquí’ x ‘Pavlot’ (Mo x Pa).