

EXPLORANDO GERMOPLASMA DE PESSEGUEIRO VIA GENOTIPAGEM POR SEQUENCIAMENTO

Liane Bahr Thurow¹; Ksenija Gasic²; Maria do Carmo Bassols Raseira³; Sandro Bonow³; Caroline Marques Castro³

¹ Doutoranda em Agronomia-Fitomelhoramento na UFPel e bolsista CAPES; Pelotas, RS, Brasil; lianepel@yahoo.com.br

² Associate Professor-Molecular Peach Breeding; School of Agricultural, Forest, and Environmental Sciences-Clemson University; Clemson, SC, United States of America

³ Pesquisadores na Embrapa Clima Temperado; Pelotas, RS, Brasil.

O pessegueiro (*Prunus persica* L. Batsch) é uma das espécies frutíferas de maior relevância em regiões de clima temperado e subtropical. Atualmente, é considerada a espécie arbórea frutífera geneticamente melhor caracterizada e é utilizada como modelo de referência em estudos genômicos (*THE INTERNATIONAL PEACH GENOME INITIATIVE*, 2013).

Com o sequenciamento completo do genoma do pessegueiro e o crescente avanço do conhecimento sobre genômica, métodos de redução da complexidade genômica, utilizando enzimas de restrição, vem sendo utilizadas com resultados promissores para estudos de GWAS (*Genome Wide Association Studies*). Dentre estas metodologias, a Genotipagem por Sequenciamento (GBS) apresenta grande aplicabilidade no melhoramento genético, permitindo a descoberta simultânea de marcadores SNPs e genotipagem, além de apresentar baixo custo por amostra e alta qualidade de SNPs gerados (ELSHIRE et al., 2011). No Brasil, o germoplasma de pessegueiro ainda é pouco explorado em nível de DNA.

Trabalhos anteriores de caracterização, principalmente morfológica e fenológica, evidenciaram grande variabilidade entre os acessos disponíveis, os quais possuem origens diversas. Diante do exposto, visando explorar este germoplasma de pessegueiro conservado no Banco Ativo de Prunóideas da Embrapa, um painel associativo com 220 genótipos foi selecionado baseado em fenótipos contrastantes para diferentes caracteres, como resistência a bacteriose e a podridão parda, assim como tolerância a fatores abióticos, a exemplo, baixa exigência em frio hibernal e tolerância ao calor na floração com o intuito de obter marcadores SNPs (Single Nucleotide Polymorphism), através de redução da complexidade genômica via GBS.

O objetivo deste trabalho consistiu na análise e filtragem dos dados genômicos obtidos por GBS para detectar e selecionar marcadores SNPs, quando comparado o alinhamento de sequências contra as duas versões do genoma disponíveis: versão inicial (Peachv1.0) e a versão mais recente (Peachv2.0a1), utilizando este painel associativo de 220 genótipos de pessegueiro. O DNA genômico foi isolado a partir de folhas liofilizadas de acordo com protocolo descrito por Dellaporta et al. (1983), com modificações. Posteriormente, as amostras de DNA foram normalizadas para 10ng/ul e digeridas com a enzima de restrição *ApeKI* para construção das bibliotecas genômicas, de acordo com o protocolo de GBS descrito por Elshire et al. (2011).

As bibliotecas obtidas foram sequenciadas utilizando a plataforma IlluminaHiSeq2000. Dados brutos foram processados utilizando GBS pipeline parâmetros implementados no software TASSEL 4.0 (BRADBURY et al., 2007) e, posteriormente, *reads* sequenciados foram alinhados às duas versões do genoma de referência disponíveis Peachv1.0 e Peach v2.0 a1 (www.rosaceae.org). Quando a versão inicial (Peachv1.0) foi utilizada como genoma de referência para alinhamento dos *reads* sequenciados, foram identificados 94.358 locos SNPs.

Marcadores SNPs que apresentavam uma porcentagem maior que 20% de dados faltantes foram removidos da análise (*call rate* > 80%), contabilizando 33.568 SNPs (33K) polimórficos com distribuição uniforme ao longo dos oito principais *scaffolds* do genoma, com variação entre 3.212 no *scaffold5* e 6.926 SNPs no *scaffold1*. Deste total de 33K SNPs, 44% apresentaram frequência alélica inferior a 5%.

A proporção de alelos heterozigotos por genótipo variou de 7,65% a 26,58%. Com a versão mais recente (Peachv2.0a1) disponibilizada com o intuito de melhorar vários problemas anteriormente detectados, como a escala na montagem dos cromossomos e anotação dos genes e sequências repetidas, foram detectados 85.023 locos SNPs. Quando filtrados considerando-se um *call rate*>80%, foram selecionados 28.598 SNPs (28K) polimórficos, com uma amplitude de cobertura variando de 2.759 no *scaffold*8 até 6.027 SNPs no *scaffold*1. Deste total de 28KSNPs, 48,4% apresentaram frequência alélica inferior a 5%.

A proporção de alelos heterozigotos encontrada variou entre 8,08% e 25,29%. A versão inicial (Peachv1.0) foi melhorada usando grande quantidade de dados de mapeamento molecular permitindo a integração de regiões previamente não-mapeadas dentro dos oito grupos de ligação, assim como dispor corretamente a orientação de outros *scaffolds* e fixar regiões posicionadas incorretamente ao longo destes oito grupos de ligação.

Como resultado destes esforços de mapeamento, a versão Peachv2.0a1 tem agora uma abrangência de 99,2% de sequências mapeadas com 97,9% da ordenação dos nucleotídeos confirmada, proporcionando um excelente recurso para o entendimento dos aspectos genéticos, estruturais e funcionais, além de facilitar a conexão entre genótipo e fenótipo para associação de caracteres de grande relevância para o melhoramento genético desta espécie.

Os marcadores SNPs de alta qualidade obtidos via GBS permitirão uma melhor compreensão e conhecimento do germoplasma disponível. As análises de dados em desenvolvimento visam determinar a estrutura genética deste painel associativo assim como, em combinação com dados fenotípicos, servirão de base para estudos de mapeamento associativo (GWAS).

Agradecimentos

O suporte para esta pesquisa fornecido pelo CNPq, Embrapa Clima Temperado e Clemson University.

Referências

BRADBURY, P. J.; ZHANG, Z.; KROON, D. E.; CASSTEVENS, T. M.; RAMDOSS, Y.; BUCKLER, E. S. TASSEL - Software for association mapping of complex traits in diverse samples. **Bioinformatics**, v.23, p.2633-2635, 2007.

DELLAPORTA, S. L.; WOOD, J.; HICKS, J.B. A plant DNA mini preparation: Version II. **Plant Molecular Biology Reporter**, v.1, p.19-21, 1983.

ELSHIRE, R. J.; GLAUBITZ, J. C.; SUN, Q.; POLAND, J. A.; KAWAMOTO, K.; BUCKLER, E. S.; MITCHELL, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS One**, v.6, n.5, e 19379, 2011.

THE INTERNATIONAL PEACH GENOME INITIATIVE; VERDE, I.; ABBOTT, A.G.; SCALABRIN, S. et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. **Nature Genetics**, v.45, p.487-494, 2013.