

## SELEÇÃO GENÔMICA AMPLA PARA O MELHORAMENTO DE SOJA [*Glycine max* (L.) MERRILL]

PASSIANOTTO, A. L. DE L.<sup>1</sup>; MARCELINO-GUIMARÃES, F.C<sup>1</sup>; OLIVEIRA, M.F.<sup>1</sup>;  
ARIAS, C. A. A.<sup>1</sup>; BELZILE, F.<sup>2</sup>; ABDELNOOR, R.V.<sup>1</sup>

<sup>1</sup> Empresa Brasileira de Pesquisa Agropecuária – Centro Nacional de Pesquisa da Soja, Embrapa Soja, Caixa postal 231, Rodovia Carlos João Strass, Distrito de Warta, Londrina, PR, Brasil. <sup>2</sup>Université Laval. andrepassianoto@hotmail.com

A soja [*Glycine max* (L.) Merrill] é uma das principais culturas brasileiras, gerando empregos diversos e contribuindo massivamente para o desenvolvimento do país. O entendimento de suas características agronômicas e morfológicas, bem como a resistência e/ou tolerância a estresses bióticos e abióticos, é de fundamental importância para que o seu potencial produtivo seja almejado. Assim, com o objetivo de explorar a amplitude do genoma da soja e entender melhor suas características, estudos moleculares vem sendo feitos ao longo das ultimas décadas, tendo como base diferentes tipos de marcadores moleculares.

Marcadores moleculares têm por objetivo auxiliar programas de melhoramento através da seleção assistida, facilitando assim rápidos ganhos genéticos (JANNINCK et al., 2010). A detecção e a exploração da variação genética sempre foi uma parte fundamental do melhoramento de plantas. Variações presentes no DNA associadas a caracteres de interesse agronômico foram abordadas durante as duas últimas décadas com diferentes tipos de marcadores moleculares (VARSHNEY et al., 2009).

Mais recentemente, o aumento do conhecimento do genoma via análises “in silico” permitiu a identificação de variações que representam o tipo mais abundante de variação genética existente, os polimorfismos de base única (Single Nucleotide Polymorphism - SNP). Estes possuem uma ampla aplicabilidade em pesquisas genômicas modernas. Acessando as informações genômicas da soja foi possível, por exemplo, associar variações alélicas ao rendimento e aos ganhos de produção visando médias maiores de produtividade (HAO, et al., 2012). Vários estudos envolvendo a descoberta e uso de SNPs já foram relatados em diferentes espécies de plantas e animais. Em soja, estudos já foram realizados com enfoque na identificação de caracteres agronômicos de interesse, como produtividade e resistência a patógenos (HYTEN et al., 2010, HAO, et al.,

2012). Técnicas que envolvem o sequenciamento parcial de genomas através da redução de complexidade têm sido usadas com resultados promissores para estudos de GWAS. Dentre as técnicas de redução de complexidade mais utilizadas, a genotipagem por sequenciamento (GBS – *Genotyping by Sequencing*) (ELSHIRE et al., 2011) se destaca pela sua multiplexidade, baixo custo por amostra e alta qualidade dos SNPs gerados. O presente trabalho está inserido no contexto de inovação tecnológica aplicada ao melhoramento genético da soja, tendo como objetivo inicial a descoberta e validação de marcadores SNPs via metodologia GBS, seguida do mapeamento por associação. Por meio da análise genotípica e fenotípica dos materiais disponibilizados, busca-se a detecção de SNPs associados às características de interesse.

A coleção de genótipos utilizada neste trabalho foi obtida do banco ativo de germoplasma da Embrapa – Soja e consistiu em 240 cultivares comerciais de soja e 170 plantas introduzidas. O material vegetal de cada um dos genótipos foi macerado e 100mg de tecido foliar foram utilizados para a extração de DNA. O tecido vegetal utilizado para a construção da biblioteca foi obtido com o kit de extração DNeasy Plant Mini Kit (Qiagen cat. No. 69106). O DNA foi posteriormente quantificados através do Thermo Scientific Nanodrop 8000 spectrophotometer (Wilmington, DE). As amostras foram normalizadas a 10ng/ul e posteriormente digeridas com a enzima de restrição *ApeKI*, de acordo com o protocolo descrito por Elshire et al., 2011, para construção da biblioteca genômica. As cinco bibliotecas genômicas obtidas foram sequenciadas em cinco canais, em três flowcells distintas utilizando o equipamento Illumina HiSeq2000 (McGill University- Genome Quebec Innovation Center, Montreal, QC, Canada).

O processo de filtragem dos SNPs foi de acordo com o descrito por SONAH et al.,

2013. O pipeline GBS identificou uma lista com 127.734 SNPs, no entanto a frequência alélica ajustada em 0,05 proveu 15.947 SNPs validados na população, o que permitiu a formulação e criação do banco de dados genotípicos via marcadores moleculares de base única. Utilizando os SNPs validados, estudos analisaram a estrutura populacional dos genótipos. A análise de componentes principais forneceu meios de identificar o background genético existente nos materiais brasileiros e plantas introduzidas. A Figura 1 utiliza como eixos os componentes principais 1 e 2, identificando cada um dos materiais e diferenciando-os em cultivares comerciais e plantas introduzidas. Um cluster contendo materiais comerciais é claramente formado, no entanto algumas plantas introduzidas compartilham a mesma região, indicando uma constituição genética semelhante aos materiais brasileiros. Contudo uma ampla quantidade de materiais exóticos apresentam características genéticas diferentes dos materiais comerciais utilizados no Brasil, indicando fontes de variabilidade genética fundamentais para a condução de estudos com intuito de identificar novas fontes de resistência a patógenos. A genotipagem de cada um dos indivíduos bem como sua fenotipagem serão carregados junto ao banco de dados do laboratório de seleção assistida da Embrapa Soja para estudos futuros. Desse modo, concluímos que a metodologia GBS pode prover dados genotípicos fundamentais para auxilio aos Programas de Melhoramento

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, e19379, 2011.

HAO, D.; CHENG, H.; YIN, Z.; CUI, S.; ZHANG, D.; WANG, H.; YU, D. Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor. Appl. Genet.*, v.124, p.447–458, 2012.

HYTEN, D.L.; CANNON, S.B.; SONG, Q.; WEEKS, N.; FICKUS, E.W.; SHOEMAKER, R.C.; SPECHT, J.E.; FARMER, A.D.; MAY, G.D.; CREGAN, P.B. High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics*, v.11, p.38, 2010.

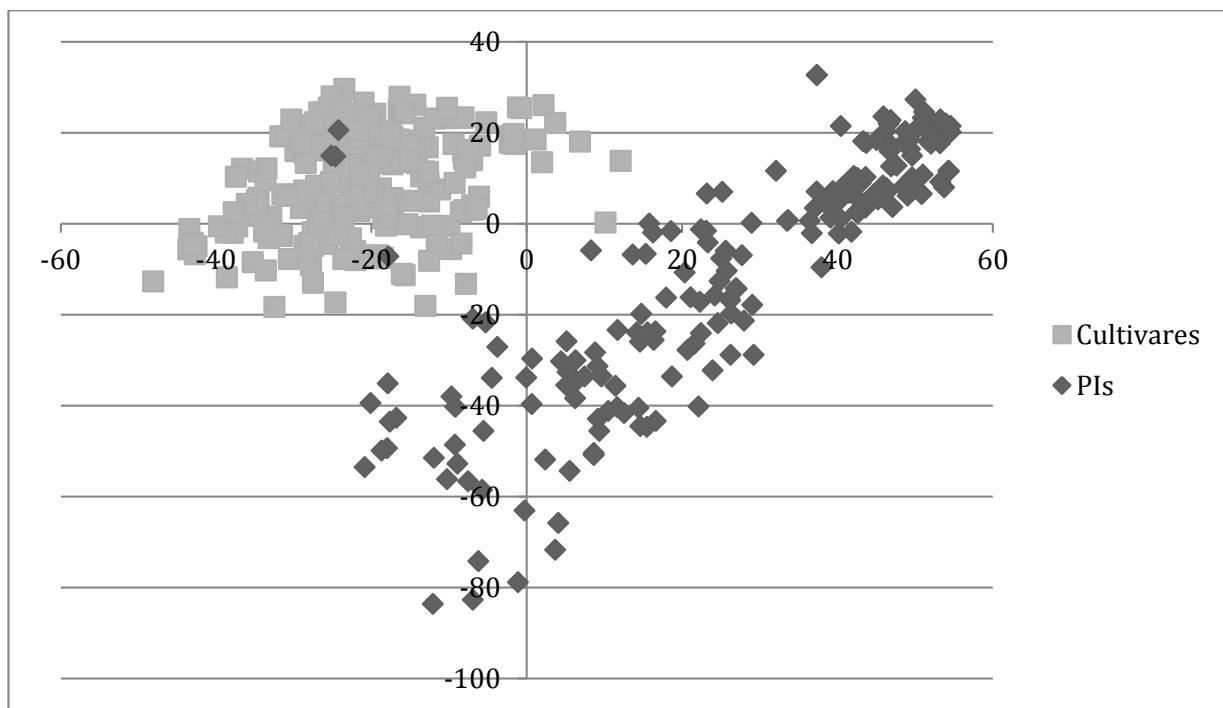
JANNINK, J-L.; LORENZ, A.J.; IWATA, H. Genomic selection in plant breeding: from theory to practice. *Brief Funct Genomics*, v.9, p.166–177, 2010.

SONAH, H.; BASTIEN, M.; IQUIRA, E.; TARDIVEL, A.; LÉGARÉ, G.; BOYLE, B.; NORMANDEAU, E.; LAROCHE, J.; LAROSE, S.; JEAN, M.; BELZILE, F. An improved genotyping by sequencing (GBS) approach offering increased efficiency of SNP discovery and genotyping. *PLoS One*, v.8, n.1, e54603, 2013.

VARSHNEY, R.K.; NAYAK, S.N.; MAY, G.D.; JACKSON, S.A. Next generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol.*, v.27, p. 522–530, 2009.

## Referências

ELSHIRE, R.J.; GLAUBITZ, J.C.; SUN, Q.;



**Figura 1.** Estrutura populacional da população de *Glycine max* (L.) Merrill estudada de acordo com os componentes principais 1 e 2, cultivares comerciais brasileiras e plantas introduzidas (PI).foram plotadas de acordo com sua constituição genotípica.