

Identificação de loci relacionados à produtividade sob deficit hídrico utilizando uma abordagem GBS - GWAS

Gabriel Feresin Pantalão¹, Tereza Cristina de Oliveira Borba², Cleber Morais Guimarães³, Marcelo Gonçalves Narciso⁴, Rosana Pereira Vianello⁵, Claudio Brondani⁶

A seca é um fator ambiental que limita a produção das culturas, como a do arroz de terras altas (*Oryza sativa* L.). O conhecimento de fatores envolvidos na tolerância à deficiência hídrica e das respostas das plantas a esse estresse podem fornecer subsídios aos programas de melhoramento para o desenvolvimento de cultivares tolerantes, e, conseqüentemente, com uma maior produtividade sob essas condições. Estudos de associação genômica ampla ou *Genome Wide Association Studies* (GWAS) analisam variações na sequência do DNA em todo o genoma, em um esforço para identificar associações a caracteres fenotípicos de interesse. Tecnologias de sequenciamento de nova geração (NGS), como a genotipagem por sequenciamento (GBS), têm sido recentemente utilizadas para identificar, validar e avaliar um grande número de SNPs, os quais podem ser utilizados em GWAS. Esse trabalho objetivou detectar, via GBS, o polimorfismo de marcadores SNPs em 175 acessos de arroz de terras altas componentes da CNAE (Coleção Nuclear de Arroz da Embrapa) e associá-los à produtividade sob deficit hídrico, identificando regiões genômicas que podem estar relacionadas a esse caráter. A caracterização molecular realizada através do sequenciamento das bibliotecas GBS dos 175 acessos de arroz de terras altas e também após a imputação de dados faltantes através do *software* FastPHASE 1.3 forneceu um total de 150.325 SNPs polimórficos distribuídos nos 12 cromossomos de arroz. A distribuição física dos SNPs teve uma média de 402 SNPs/Mpb, variando entre um mínimo de 327 SNPs/Mpb no cromossomo 5 e um máximo de 463 SNPs/Mpb no cromossomo 11. A partir dos dados de 150.325 SNPs, foi estimada a estruturação dos indivíduos pelo modelo bayesiano do STRUCTURE e também foi calculada a matriz de parentesco (Kinship) pelo algoritmo EMMA, como etapas preliminares da análise GWAS para reduzir o número de falsos positivos na análise. GWAS foi realizada pelo *software* TASSEL 4.0 pelo Modelo Linear Misto (MLM) utilizando os dados fenotípicos derivados do experimento de avaliação de produtividade, conduzido no ano de 2012 em Porangatu (GO). Foram identificados 13 SNPs significativamente associados à produtividade sob deficit hídrico e 13 SNPs significativamente associados à produtividade em ambiente sem deficiência hídrica. Após a obtenção dos SNPs significativos, a posição de cada SNP no cromossomo foi utilizada para identificar quais estariam em sequências transcritas já identificadas em arroz. Em relação ao ambiente sem deficit hídrico, 10 genes foram identificados como tendo SNPs significativos, enquanto que no ambiente com deficit hídrico, seis genes tiveram SNPs significativos. Os SNPs desses genes foram então utilizados para se observar o padrão de frequência alélica entre os indivíduos mais produtivos e menos produtivos do experimento de deficiência hídrica realizado. Considerando apenas os SNPs ancorados a genes, no experimento sem deficiência hídrica, foram encontrados 8 SNPs (80%) com frequência alélica diferencial entre os 10 acessos mais produtivos e os 10 acessos menos produtivos, enquanto que no experimento com deficiência hídrica foram encontrados 4 SNPs (66,7%) com frequência alélica diferencial entre os 10 acessos mais produtivos e os 10 acessos menos produtivos. Adicionalmente, a partir das sequências de aminoácidos, foi realizada uma análise BLAST para identificar em banco de dados de arroz e de *Arabidopsis*, informações mais detalhadas dos 16 genes. De acordo com essa análise, os genes foram relacionados à produtividade, estresses abióticos e resistência a doenças, podendo ser classificados como gene candidatos para uso efetivo na seleção assistida por marcadores. Adicionalmente, esses genes podem ser superexpressos para avaliar sua capacidade de aumentar a tolerância à seca, e em caso positivo, gerar cultivares comerciais de arroz geneticamente modificadas mais produtivas sob condições de deficit hídrico.

¹ Estudante de Pós-Graduação em Genética e Melhoramento de Plantas na Universidade Federal de Goiás. Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, gabrielferesin@hotmail.com

² Engenheira de alimentos, Doutora em Agronomia, pesquisadora da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, tereza.borba@embrapa.br

³ Engenheiro agrônomo, Doutor em Biologia Vegetal, pesquisador da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, cleber.guimaraes@embrapa.br

⁴ Engenheiro eletrônico, Doutor em Computação Aplicada, pesquisador da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, marcelo.narciso@embrapa.br

⁵ Bióloga, Doutora em Ciências Biológicas, pesquisadora da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, rosana.vianello@embrapa.br

⁶ Engenheiro agrônomo, Doutor em Ciências Biológicas, pesquisador da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, claudio.brondani@embrapa.br