

MARCADORES SNPs DERIVADOS DE CAPTUREseq ASSOCIADOS A TOLERÂNCIA À SECA EM ARROZ

Ariadna Faria Vieira¹, Paula Arielle Mendes Ribeiro Valdisser², Tereza Cristina de Oliveira Borba², Anna Cristina Lanna², Rosana Pereira Vianello², Leandro Gomide Neves³, Claudio Brondani^{2*}

¹Universidade Federal de Goiás. ²Embrapa Arroz e Feijão. ³Rapid Genomics. *E-mail do autor para correspondência: claudio.brondani@embrapa.br.

O arroz é um dos cereais de maior importância socioeconômica a nível mundial, sendo o Brasil o nono maior produtor. O arroz de terras altas possui variabilidade genética para a tolerância à seca, mas esse caráter não está presente em todas as cultivares comercializadas. A incorporação desse caráter, portanto, é um desafio para os programas de melhoramento. O método de sequenciamento de DNA por captura (ou *Capture-seq*) permite a identificação de polimorfismos em regiões candidatas ou regiões genômicas de interesse, utilizando plataformas de sequenciamento de nova geração (NGS). Como parte do esforço de incorporar a tolerância à seca em genótipos elite do programa de melhoramento da Embrapa, esse trabalho objetivou identificar genes e alelos relacionados a esse caráter. O estudo foi realizado em 300 acessos de arroz de terras altas, avaliados quanto a tolerância à seca em experimentos de campo por 3 anos (2010 a 2012), e genotipados por sequenciamento de DNA por captura de 2.500 unigenes previamente identificados como relacionados à tolerância a deficiência hídrica, permitindo realizar um estudo de mapeamento associativo. O sequenciamento por captura, realizado na empresa Rapid Genomics (EUA), utilizou a plataforma Illumina *Hi-seq* 2500, e como resultado, foram identificados 16.055 SNPs e InDels distribuídos ao longo dos 12 cromossomos do arroz. Para o mapeamento associativo (GWAS), houve inicialmente a filtragem dos dados utilizando a frequência mínima alélica ($MAF \geq 0.05$), e por meio do *software* Tassel, foram identificados quatro SNPs significativos (genes localizados nos cromossomos 5, 7, 8 e 10) pelo critério de *False Discovery Rate* ($FDR \leq 0.05$). Estes genes foram anotados (BLAST2GO), e são putativamente traduzidos nas proteínas: Domínio proteico BURP (LOC_Os05g12640), Cis-zeatina O-glucosiltransferase (LOC_Os07g46610), hidrolase e família proteica alfa/beta (LOC_Os10g37720) e proteína de ligação ao DNA (LOC_Os08g44910). Como perspectiva, esses SNPs identificados serão convertidos em ensaio TaqMan para validar sua aplicação na rotina de seleção assistida para tolerância à seca no melhoramento genético do arroz da Embrapa.

Palavras-chave: Sequenciamento por captura; GWAS; SNPs

Agradecimentos: À Embrapa Arroz e Feijão, Capes, CNPq e a Universidade Federal de Goiás.