

IDENTIFICAÇÃO DE REGIÕES GENÔMICAS POR GWAS EM TRIGO ASSOCIADAS À TOLERÂNCIA À SECA - PRIMEIROS RESULTADOS

Luciano Consoli¹, Jorge Gonzalez Aguilera^{1,2}, Pedro Arnaldo Duarte Barcellos^{1,3}, Ana Lúcia Variani Bonato¹, Antonio Nhani Jr¹, Gisele Abigail Montan Torres¹, Jorge Fernando Pereira¹, Márcio Nicolau¹ e Márcio Só e Silva¹

¹Embrapa Trigo, BR 285, km 294, CEP 99050-970, Passo Fundo - RS. Email: luciano.consoli@embrapa.br

².Bolsista Pós-Doutorado PNDP CNPq/CAPES– 159358/2011-6

³Escola Estadual Técnica Celeste Gobbato, BR 158, km 6, CEP 98300-000, Palmeira das Missões - RS.

Estudos de mudanças climáticas globais indicam um agravamento dos problemas relacionados com irregularidades na distribuição das chuvas, e conseqüentemente, aumentam o risco de ocorrência de estresse hídrico durante o ciclo das culturas. Além disso, um dos maiores desafios para a expansão da triticultura na região dos Cerrados está relacionado com tolerância à restrição hídrica (RH). Por ser uma característica complexa, envolvendo interações solo-água-planta-ambiente, a seleção de plantas tolerantes à RH é um constante desafio para os programas de melhoramento. Marcadores moleculares e populações biparentais para estudos de QTLs vêm sendo amplamente usados em trigo, desde a publicação do primeiro mapa genético em 1989. Estudos de genética de associação, usando coleções de genótipos não aparentados, surgiram como opção aos cruzamentos biparentais, possibilitando tanto a exploração de base genética ampla quanto o acúmulo de recombinações entre os genótipos, contribuindo na redução da distância entre as marcas e os genes de interesse. Com o aumento da disponibilidade de métodos de genotipagem em larga escala baseados em marcadores moleculares do tipo SNPs, que proporcionam uma maior saturação do genoma, e o uso de modelos lineares mistos para estimar o efeito das marcas observa-se, na literatura, um crescente número de estudos

de associação genômica ampla (GWAS - Genome-Wide Association Studies). Essa tendência também é observada em trigo, principalmente após o lançamento no mercado entre 2013 e 2014 de plataformas de genotipagem de alta densidade pelas empresas Illumina e Affymetrix.

O objetivo deste trabalho foi identificar regiões genômicas relacionadas com a tolerância à restrição hídrica via estudos de associação numa coleção de 98 acessos de trigo, composta por cultivares brasileiras, variedades tradicionais, trigos sintéticos, variedades estrangeiras e linhagens avançadas do programa de melhoramento da Embrapa Trigo.

O experimento foi conduzido em condição controle e de RH sob estrutura denominada “shelter”, ou abrigo móvel, na Embrapa Trigo. A semeadura foi realizada em 25 de julho de 2014, com delineamento em blocos casualizados com duas repetições de uma linha de um metro e espaçamento de 20 cm entre linhas. A RH foi aplicada 45 dias após a semeadura, através da movimentação do “shelter” durante a ocorrência de precipitação por um sistema automático de detecção de chuvas. Para evitar a umidade noturna, o “shelter” era deslocado para as parcelas do tratamento RH durante o período noturno. A umidade volumétrica do solo foi monitorada com sensores de haste de 20 cm HidroFarm da Falker tanto para o a condição de RH ($Uv=21,4\%$) quanto para o tratamento controle ($Uv=32,1\%$), que recebeu irrigação quando necessário. Os demais tratamentos culturais foram realizados de acordo com as recomendações para a cultura do trigo. Foram avaliadas oito características: Número de dias para espigamento (Ndesp), Número de dias para florescimento (Ndflor), Altura (cm) de planta na maturidade (Alt), Número de plantas por parcela (Nplantas), Número de espigas por parcela (NEsp), Número de grãos por parcela (Ngrãos), peso(g) de grãos por parcela (Pgrãos) e Peso (g) de Mil Grãos (PMG).

As médias dos experimentos e as análises de variância dos dados fenotípicos foram realizadas com o programa Genstat 18th Edition (Tabela 1). Os resultados indicam uma redução significativa na média para todas as características avaliadas, em relação ao tratamento controle, com exceção para Nplantas e PMG, que apresentaram valores superiores no tratamento

com RH em relação ao controle. Também verificou-se a presença de variabilidade genética significativa entre os acessos para todas as características avaliadas, mesmo para aquelas com elevado coeficiente de variação.

Do total de 98 acessos fenotipados, foram obtidos dados genotípicos para 78 materiais com o Axiom Wheat Breeder's Genotyping Array da empresa Affymetrix, contendo 35.143 SNPs. Informações sobre a localização dos SNPs no genoma de trigo foram obtidas no site da base de dados CerealsDB (Wilkinson et al., 2012). Os SNPs foram pré-analisados com o software Analysis Suite da Affymetrix, usando os valores recomendados para os parâmetros de qualidade avaliados pelo software. Posteriormente, os dados genotípicos foram filtrados para frequência mínima de alelo de 0,05, removendo dessa forma todos SNPs monomórficos e alelos raros. Um subconjunto de dados com 1.706 SNPs, apresentando desequilíbrio de ligação com $r < 0,90$ entre as marcas, foi usado para cálculo da matriz de kinship e análise de componentes principais. Um total de 7.813 marcas, apresentando zero dados perdidos, foi usado nas análises de associação usando o software TASSEL5 (Zhang et al., 2010) juntamente com os valores médios dos acessos para as características avaliadas, para os dois tratamentos, usando modelo linear misto com os coeficientes de kinship e com valores da análise de componentes principais (número de componentes igual a 3). Foram detectados em média 8,75 SNPs associados com as características avaliadas na condição controle, distribuídas entre 3 e 7 cromossomos. Já para os valores obtidos para as características avaliadas na condição de RH, foram detectados em média 12,9 SNPs distribuídos entre 1 e 14 cromossomos. No total foram observadas 173 SNPs associados com probabilidade inferior a $p < 0,001$ alocados nos cromossomos, correspondendo a valores de $LOD > 3$ (Tabela 2). Na Figura 1 são representados gráficos Manhattan plot obtidos após as análises de associação para a característica NEsp.

Os resultados desse estudo evidenciam que: i) o experimento conduzido em campo sob os abrigos móveis foram capazes de reproduzir o efeito de

ocorrência de seca; ii) com exceção para Nplantas e PMG, todas as características avaliadas apresentaram redução significativa na média sob RH; iii) os estudos preliminares de GWAS evidenciaram um conjunto de SNPs e regiões genômicas associadas à RH nas condições experimentais adotadas. A validação desses resultados será realizada com a condução de experimentos na região dos Cerrados, onde além da restrição da água, há uma forte influência da temperatura do ambiente na expressão da tolerância à seca. O uso destes marcadores poderá auxiliar na seleção de plantas mais tolerantes à seca.

Referências bibliográficas

- WILKINSON, P.A.; WINFIELD, M.O.; BARKER, G.L.A.; ALLEN, A.M.; BURRIDGE, A.; COGHILL, J.A.; BURRIDGE, A.; EDWARDS, K.J. CerealsDB 2.0: an integrated resource for plant breeders and scientists. BMC Bioinformatics, 13:219, 2012.
- ZHANG, Z.; ERSOZ, E.; LAI, C-Q.; TODHUNTER, R.J.; TIWARI, H.K.; GORE, M.A.; BRADBURY, P.J.; YU, J.; ARNETT, D.D.; ORDOVAS, J.M.; BUCKLER, E.S. Mixed linear model approach adapted for genome-wide association studies. Nature Genetics, v42:355-360, 2010.

TABELA 1. Valores médios para as características Altura de planta (Alt, cm), Número de Dias para Espigamento (NDEsp), Número de Dias para Florescimento (NDFlor), Número de espigas (Nespigas), Número de grãos (Ngrãos), Número de plantas (Nplantas), Peso de grãos (Pgrãos, g) e Peso de Mil Grãos (PMG, g) para os tratamentos controle e restrição hídrica. Os valores do teste F para o contraste entre as médias dos tratamentos e o coeficiente de variação são apresentados.

Característica	Médias		F	CV%
	Controle	Estresse hídrico		
Alt**	87,06	83,21	25,91**	9,06
NDEsp**	79,94	73,19	38,96**	3,74
NDFlor**	79,55	77,62	40,39**	3,81
Nespigas**	77,16	53,17	182,27**	27,17
Ngrãos**	1289,0	904,0	86,62**	37,95
Nplantas**	22,73	27,02	40,38**	26,26
Pgrãos(g)**	37,18	29,48	34,07**	40,43
PMG(g)**	29,11	32,48	102,00**	10,66

** , p<0,001 Asteriscos após o nome das características indicam a presença significativa de variabilidade genética entre os acessos. Asteriscos após os valores de F indicam que as médias entre os tratamentos diferem estatisticamente.

TABELA 2. Número de SNPs associados por característica apresentando $p < 0,001$ (LOD score > 3) e número de cromossomos representados.

Característica	Número de SNPs	Número de cromossomos
Alt_C	15	6
Alt_S	5	4
Ndesp_C	7	6
Ndesp_S	8	5
Ndflor_C	5	5
Ndflor_S	7	5
Nesp_C	6	5
Nesp_S	13	6
Ngraos_C	6	3
Ngraos_S	1	1
Nplantas_C	5	6
Nplantas_S	22	11
Pgraos_C	15	7
Pgraos_S	2	2
PMG_C	10	5
PMG_S	45	14

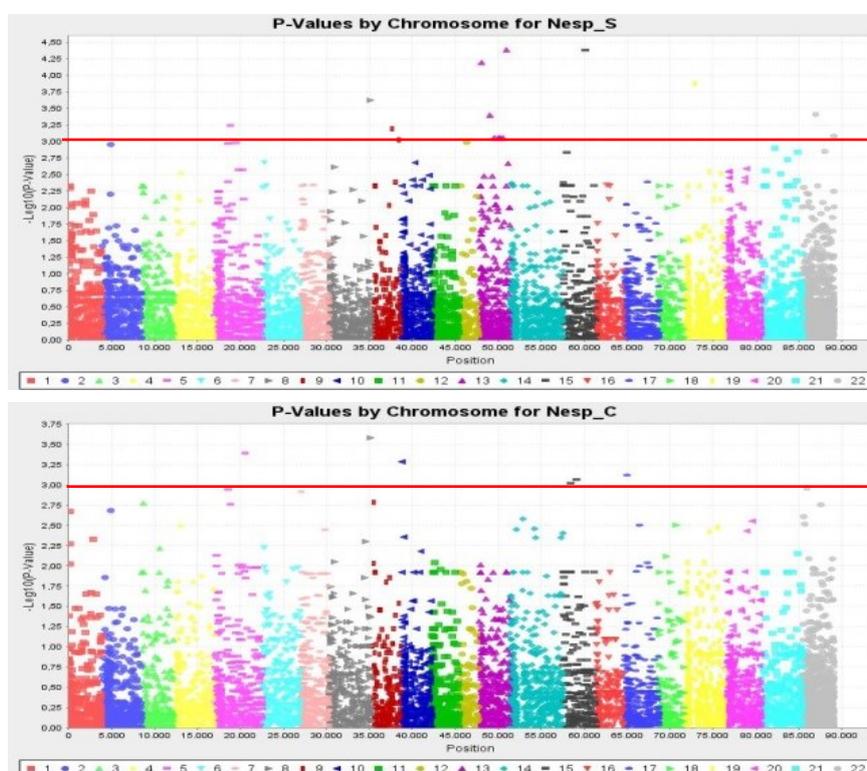


FIGURA 1. Gráfico Manhattan plot obtido para a característica Número de espigas avaliada na condição controle (C) e de restrição hídrica (S). A linha vermelha indica valor de LOD score = 3, correspondendo ao limite de p -value $< 0,001$.