

ANÁLISE GENÔMICA PARA COMPOSTOS BIOATIVOS EM SORGO UTILIZANDO DADOS DE MAPEAMENTO ASSOCIATIVO

RESUMO

O sorgo é o quinto cereal mais produzido no mundo, não possui glúten e é uma rica fonte de compostos bioativos. A compreensão da regulação genética das vias de biossíntese de compostos bioativos é importante para o desenvolvimento de cereais mais nutritivos e funcionais. Os compostos bioativos têm sido estudados e avaliados quanto aos seus benefícios para a saúde e a. Desse modo, este trabalho teve por objetivo analisar dados resultantes do mapeamento associativo para as características: taninos, fenólicos totais, cloreto de apigenenidina e cloreto de luteolinidina em sorgo através da triagem de polimorfismos de nucleotídeos únicos (SNPs) associados a essas características e a prospecção de genes candidatos. Foram identificados SNPs associados a taninos e fenólicos totais no cromossomo 4, na região de aproximadamente 61Mb, próximo ao *Tannin1*, gene responsável por regular a biossíntese de taninos em sorgo. No cromossomo 6, próximo da região de 57Kb, foram identificados SNPs associados ao teor de 3-deoxiantocianidinas. Nessa região genômica, foram identificados genes candidatos que podem estar relacionados à biossíntese desse composto bioativo em sorgo. Os resultados obtidos nesse trabalho contribuirão para a identificação de variedades de sorgo com potencial para melhorar a nutrição e promover a saúde humana.

INTRODUÇÃO

Os cereais são alimentos básicos para grande parte da população mundial (1). Dentre os cereais mais cultivados, o sorgo [*Sorghum bicolor* (L.) Moench] é uma rica fonte de compostos bioativos, com diversos genótipos apresentando diferentes níveis de ácidos fenólicos, flavonoides (antocianinas e tanino) e amido resistente (2, 3), podendo ser classificado como um alimento funcional (4). Os compostos bioativos são constituintes extranutricionais que geralmente ocorrem em pequenas quantidades nos alimentos (5, 6) e têm sido estudados e avaliados quanto aos seus benefícios para a promoção da saúde e prevenção de doenças (6), visto que muitos possuem efeitos antioxidantes (7,8).

A maior parte do sorgo cultivado no mundo é destinada basicamente à alimentação animal e, somente em poucos países da África e da Ásia, ele é utilizado na alimentação humana (9). Entretanto, tem-se observado um aumento crescente no interesse em utilizar o sorgo para o consumo humano, principalmente devido à demanda por alimentos mais saudáveis, sem glúten, com maior valor nutricional e fonte de compostos bioativos (10). A acessibilidade, a disponibilidade e as características químicas dos vários compostos bioativos apresentam ampla variabilidade entre os diferentes genótipos de sorgo, o que resulta em comportamentos distintos na matriz alimentar e na interação com outros ingredientes alimentares (11). Tal variabilidade genética deve ser explorada para o desenvolvimento de cultivares de sorgo com características mais adequadas para a alimentação humana. Dessa forma, a pesquisa genética do sorgo e a compreensão dos mecanismos moleculares responsáveis pelas variações nos níveis desses compostos, tornam-se importantes ferramentas para o melhoramento assistido por marcadores moleculares e o desenvolvimento de variedades com maior conteúdo de compostos bioativos.

O mapeamento associativo é um método de mapeamento genético que tem potencial de identificar polimorfismos em regiões genômicas próximas a genes responsáveis por variações fenotípicas, inferindo o fenótipo a partir do genótipo. Assim, é possível identificar variações funcionais específicas, ligadas a diferenças fenotípicas relacionadas a caracteres de interesse (12, 13). Estudos têm sido direcionados para a identificação e caracterização de genes envolvidos com a biossíntese desses compostos. A genética relacionada com a biossíntese dessas substâncias tem sido estudada em espécies economicamente importantes como os cereais: milho, cevada e trigo. Estudos de mapeamento associativo têm resultado na identificação de regiões genômicas associadas à composição de grãos em diferentes espécies de cereais, como o milho (14), o sorgo (15), entre outros (16). Em sorgo, o gene *Tannin1* que regula a via de biossíntese de taninos foi clonado. Ele localiza-se na região 61Mb do cromossomo 4 do sorgo (3). Outros estudos também foram conduzidos em sorgo para identificar regiões genômicas associadas com polifenóis, antocianinas, taninos e atividade antioxidante (12, 17, 13). Contudo, muitos SNPs (*Single Nucleotide Polymorphism*), especialmente para capacidade antioxidante, fenólicos e antocianinas, apresentaram pequenos efeitos, demonstrando a necessidade de estudos adicionais (12). Estudos genéticos podem contribuir para a seleção e o desenvolvimento de cultivares de maneira rápida e com menores custos comparando-se com algumas análises químicas tradicionalmente utilizadas para a identificação e a quantificação de compostos funcionais nos alimentos que são onerosas, demoradas, além de gerarem um grande volume de resíduos.

Com isso, torna-se importante a busca por informações a respeito dos teores desses compostos nos diferentes genótipos de sorgo do painel constituído por 240 linhagens com ampla variabilidade genética e de teores dos compostos bioativos. Ademais, os estudos na área podem promover um maior interesse pelo cultivo do sorgo destinado à alimentação humana, devido ao potencial para gerar cultivares mais nutritivos e com propriedades funcionais, aumentando o seu valor comercial e o interesse da indústria de alimentos.

OBJETIVO

Objetivou-se analisar dados resultantes de análises de mapeamento associativo (GWAS) para: taninos, fenólicos totais, cloreto de apigenenidina e cloreto de luteolinidina em linhagens de sorgo através da triagem dos SNPs que se mostraram mais significativos estatisticamente nas análises de GWAS, identificando suas localizações no genoma. Além disso, foram realizadas pesquisas por genes candidatos relacionados à biossíntese dos compostos bioativos.

RESULTADO E DISCUSSÃO

O mapeamento associativo foi conduzido utilizando um conjunto de aproximadamente 300.000 SNPs gerados por genotipagem via GBS (*Genotyping-By-Sequencing*, 18) do genoma de 240 linhagens geneticamente contrastantes para os compostos bioativos taninos, fenólicos totais, cloreto de apigenenidina e cloreto de luteolinidina. A quantificação dos teores de taninos (mg CE/g) e fenólicos totais (mg GAE/g) foi realizada através do método de espectrofotometria, e os compostos cloreto de apigenenidina ($\mu\text{g/g}$) e cloreto de luteolinidina ($\mu\text{g/g}$) foram quantificados em sistema HPLC (*High Performance Liquid Chromatography*; 19). Para as análises de mapeamento associativo foi usado o *software* TASSEL (20).

O mapeamento associativo resultou na identificação de SNPs estatisticamente associados para cada característica. Nos intervalos de 10Mb (Mega pares de bases)

próximos a esses SNPs, foram identificados genes candidatos para taninos (Tabela 1), fenólicos totais (Tabela 2), cloreto de apigenenidina ($\mu\text{g/g}$) (Tabela 3) e cloreto de luteolinidina ($\mu\text{g/g}$) (Tabela 4).

Tabela 1 - Genes candidatos identificados nas análises de associação para taninos (mg CE/g) do painel de linhagens de sorgo granífero.

Cromossomo	Gene ID ¹	Posição ²	Anotação	$-\log_{10}$ (p-value) ³
4	Sb.04g02	61Mb	Fator de transcrição da família MYB.	19.46
4	Sb.04g03	61Mb	Ciclo de divisão celular 20, cofator do complexo APC. Proteína contendo repetição WD.	19.46
4	Sb.04g04	61Mb	Relacionado à proteína de ligação de Hélice Básica (BHLH)	19.46

¹Identificação do gene. ²Localização do gene no genoma. ³Valor de probabilidade p ($-\log_{10}$) obtido nos resultados do mapeamento associativo.

Tabela 2- Genes candidatos identificados nas análises de associação para fenólicos totais (mg GAE/g) do painel de linhagens de sorgo granífero.

Cromossomo	Gene	Posição	Anotação	$-\log_{10}$ (p-value)
4	Sb.04g34	61Mb	Fator de transcrição da família MYB.	15.52
4	Sb.04g35	61Mb	Proteína de ligação de hélice básica (BHLH)	15.52
4	Sb.04g36	85Mb	Proteína fosfatase PP2A	12.37

¹Identificação do gene. ²Localização do gene no genoma. ³Valor de probabilidade p ($-\log_{10}$) obtido nos resultados do mapeamento associativo.

As análises de mapeamento associativo para taninos permitiram a identificação de SNPs com associação significativa na região genômica de 61Mb no cromossomo 4 do sorgo. Esses resultados corroboram trabalhos prévios em que o gene *Tannin1*, que regula a via de biossíntese de taninos, foi clonado (3). A distância entre os SNPs identificados e o gene *Tan1* foi de 0.8Mb, ou seja, geneticamente muito próximos. Os resultados do mapeamento associativo para os compostos fenólicos totais foram semelhantes aos obtidos para taninos, em que os SNPs mais significativos localizam-se no cromossomo 4, também na região de 60Mb. Isso faz sentido devido ao fato de o tanino ser um tipo de ácidos fenólicos e apresentar a parte inicial da via de biossíntese em comum com esses compostos (12), portanto, eles podem apresentar elementos regulatórios da via de biossíntese em comum.

Tabela 3 - Genes candidatos identificados nas análises de associação para cloreto de apigenenidina ($\mu\text{g/g}$) do painel de linhagens de sorgo granífero.

Cromossomo	Gene ID ¹	Posição ²	Anotação	$-\log_{10}$ (p-value) ³
6	Sb.06g53	57Mb	Fator de iniciação de tradução eucariótica SUI1	6.54
6	Sb.06g55	57Mb	Proteína semelhante à ubiquitina Nedd8.	6.54
6	Sb.06g58	57Mb	Proteína C44H4.4	6.54

¹Identificação do gene. ²Localização do gene no genoma. ³Valor de probabilidade p ($-\log_{10}$) obtido nos resultados do mapeamento associativo.

O estudo de GWAS para cloreto de apigenenidina possibilitou a identificação de SNPs associados no cromossomo 6 do sorgo na região de 57Mb. As pesquisas por genes localizados nessa região do genoma permitiram a descoberta de genes candidatos possivelmente relacionados com a biossíntese de 3-deoxiantocianidinas. Em região genômica próxima a essa, de 54Mb, um gene, denominado Sb06g029550, foi clonado em sorgo. Esse gene altera a cor das folhas do sorgo sob ataque de patógenos. Ele codifica a enzima flavona-4-redutase que converte flavanona em flavan-4-ol na via de biossíntese de 3-deoxiantocianidina (21).

CONCLUSÃO

Este trabalho possibilitou a identificação de SNPs resultantes das análises de GWAS e genes candidatos que podem estar relacionados com a biossíntese dos compostos bioativos fenólicos totais, taninos, cloreto de apigenenidina e cloreto de luteolinidina. Esses resultados serão importantes para subsidiar pesquisas futuras que visem o desenvolvimento de marcadores moleculares que podem auxiliar de maneira mais rápida e mais econômica o melhoramento genético de sorgo. Ademais, o trabalho contribuirá para a identificação de genótipos superiores de sorgo tendo como principal foco a qualidade nutricional e funcional dos grãos voltados para a alimentação humana.

AGRADECIMENTOS: O presente trabalho foi realizado com apoio financeiro das seguintes agências de fomento: FAPEMIG, CAPES e CNPq.

REFERÊNCIA BIBLIOGRÁFICA

1. Cereals & Grains. **Improving Nutrition through Biofortification:** Preharvest and Postharvest Technologies. Disponível em: <https://www.cerealsgrains.org/publications/cfw/2019/May-June/Pages/CFW-64-3-0025.aspx>. Acesso em 31 out., 2019.
2. DE MORAIS CARDOSO, Leandro et al. Sorghum (*Sorghum bicolor* L.): Nutrients, bioactive compounds, and potential impact on human health. **Critical reviews in food science and nutrition**, v. 57, n. 2, p. 372-390, 2017.
3. WU, Yuye et al. Presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. **Proceedings of the National Academy of Sciences**, v. 109, n. 26, p. 10281-10286, 2012.
4. PRZYBYLSKA-BALCEREK, Anna; FRANKOWSKI, Jakub; STUPERSZABLEWSKA, Kinga. The influence of weather conditions on bioactive compound content in sorghum grain. **European Food Research and Technology**, v. 246, n. 1, p. 13-22, 2020.

5. DYKES, L; ROONEY, L. Phenolic Compounds in Cereal Grains and Their Health Benefits. **Cereal Foods World**, v. 52, n. 3, p. 105–111, 2007.
6. KRIS-ETHERTON, P. M. *et al.* Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. **The American Journal of Medicine**, v. 113, n. 1, p. 71S–88S, 2002.
7. HARBORNE, J.B.; WILLIAMS, C.A. Advances in flavonoid research since 1992. **Phytochemistry**, v. 55, n. 6, p. 481–504, 2000.
8. SLAVIN, J., *et al.* Whole-grain consumption and chronic disease: Protective mechanisms. **Journal Nutrition and Cancer**, v. 27, n. 1, p. 14-21, 1997.
9. DICKO, M. H. *et al.* Sorghum grain as human food in Africa: relevance of content of starch and amylase activities. **African Journal of Biotechnology**, v. 5, n. 5. p. 384-395, 2006.
10. ORTIZ-CRUZ, Raquel Alejandra *et al.* Effect of Extrusion Processing Conditions on the Phenolic Compound Content and Antioxidant Capacity of Sorghum (*Sorghum bicolor* (L.) Moench) Bran. **Plant Foods for Human Nutrition**, p. 1-6, 2020.
11. GIACONIA, Michele Amendoeira *et al.* Overcoming restrictions of bioactive compounds biological effects in food using nanometer-sized structures. **Food Hydrocolloids**, p. 105939, 2020.
12. HABYARIMANA, Ephrem *et al.* Genome-wide association mapping of total antioxidant capacity, phenols, tannins, and flavonoids in a panel of *Sorghum bicolor* and *S. bicolor* × *S. halepense* populations using multi-locus models. **PLoS One**, v. 14, n. 12, p. e0225979, 2019.
13. RHODES, Davina H. *et al.* Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] germplasm. **Journal of agricultural and food chemistry**, v. 62, n. 45, p. 10916-10927, 2014.
14. COOK, Jason P. *et al.* Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. **Plant physiology**, v. 158, n. 2, p. 824-834, 2012.
15. TAO, Yongfu *et al.* Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. **Plant biotechnology journal**, v. 18, n. 4, p. 1093-1105, 2020.
16. GYAWALI, Sanjaya *et al.* Genome wide association studies (GWAS) of element contents in grain with a special focus on zinc and iron in a world collection of barley (*Hordeum vulgare* L.). **Journal of Cereal Science**, v. 77, p. 266-274, 2017.
17. MORRIS, Geoffrey P. *et al.* Dissecting genome-wide association signals for loss-of-function phenotypes in sorghum flavonoid pigmentation traits. **G3: Genes, Genomes, Genetics**, v. 3, n. 11, p. 2085-2094, 2013.
18. ELSHIRE, Robert J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS one**, v. 6, n. 5, p. e19379, 2011.
19. YANG, Liyi *et al.* Sorghum phenolics demonstrate estrogenic action and induce apoptosis in nonmalignant colonocytes. **Nutrition and cancer**, v. 64, n. 3, p. 419-427, 2012.
20. BRADBURY, Peter J. *et al.* TASSEL: software for association mapping of complex traits in diverse samples. **Bioinformatics**, v. 23, n. 19, p. 2633-2635, 2007.
21. KAWAHIGASHI, Hiroyuki *et al.* The sorghum gene for leaf color changes upon wounding (P) encodes a flavanone 4-reductase in the 3- deoxyanthocyanidin biosynthesis pathway. **G3: Genes, Genomes, Genetics**, v. 6, n. 5, p. 1439-1447, 2016.