

# ANÁLISE GENÔMICA COMPARATIVA DE GENES *bZIP* ENTRE ARROZ E ARABIDOPSIS

Ariadne Ribeiro Henriques<sup>1</sup>, Fabiane Igansi de Castro dos Santos<sup>2</sup>, Vívia Ebeling Viana<sup>3</sup>, Luciano Carlos da Maia<sup>4</sup>, Antonio Costa de Oliveira<sup>5</sup>

Palavras-chave: zinco, fatores de transcrição, análise *in silico*

## INTRODUÇÃO

O arroz (*Oryza sativa* L.), um dos cereais mais cultivados no mundo, é o alimento básico e principal fonte de calorias dos países em desenvolvimento (FAO, 2015). O Brasil produz cerca de 12 milhões de toneladas/ano de arroz, sendo a região Sul responsável por 67% desta produção, movimentando milhões de dólares ao ano e gerando grande contribuição para o desempenho econômico brasileiro (CONAB, 2015). Entretanto, a produtividade desta cultura é limitada por diversos fatores bióticos e abióticos tais como a deficiência nutricional dos solos.

Atualmente, a deficiência de zinco é reconhecida como um dos principais desafios globais para a saúde pública e para agricultura do século 21 (PFEIFFER e MCCLAFFERTY, 2007). Segundo Fungetto et al (2010) este é um dos micronutrientes cuja deficiência mais tem limitado a produção das culturas no Brasil. Na cultura do arroz o cultivo sucessivo, devido à sistematização e o alagamento, a acidez do solo, o uso de fertilizantes altamente concentrados e de cultivares nutricionalmente exigentes, vem contribuindo para diminuição da biodisponibilidade de zinco nos solos (MARCHESAN et al., 2001; SILVA et al., 2003). A baixa biodisponibilidade desse nutriente limita a absorção de zinco pelas plantas levando a um inadequado funcionamento de muitas funções fisiológicas essenciais nas células que resulta em significativas reduções no crescimento, na produtividade e no teor de zinco dos grãos (HENRIQUES et al., 2012). Para enfrentar a baixa biodisponibilidade de zinco, as plantas aumentam a expressão de vários genes envolvidos na homeostase deste metal visando à manutenção da homeostase celular.

Muitos ajustamentos fisiológicos e moleculares associados com respostas a estresses em plantas são controladas a nível transcricional. Desta maneira, a aplicação de técnicas biotecnológicas para a utilização de recursos genéticos de plantas e exploração dos genes com intuito de conseguir cultivares com melhor conteúdo e/ou tolerância à deficiência de zinco é uma abordagem eficaz que pode ajudar a desvendar a complexa rede de homeostase de zinco em plantas, além de contribuir para minimizar severas perdas a nível mundial e aliviar o problema de desnutrição em seres humanos. Assunção et al (2010) descobriram os primeiros fatores de transcrição envolvidos na homeostase de zinco em plantas, o *bZIP19* e *bZIP23*. Ensaios realizados com *Arabidopsis thaliana* mostraram que o aumento da expressão controlada desses fatores de transcrição tem sido muito eficaz no aumento da tolerância à deficiência de zinco em *Arabidopsis* (ASSUNÇÃO et al., 2013). Assim, conhecendo que os fatores de transcrição *bZIP19* e *bZIP23* e seus genes alvos são conservados no reino vegetal (ASSUNÇÃO et al., 2010) uma abordagem abrangente *in silico* para identificar as relações existentes entre *Oryza sativa* e *Arabidopsis thaliana* é potencialmente promissora. Portanto, este trabalho teve como objetivo identificar e realizar uma análise comparativa e funcional dos genes *bZIP19* e *bZIP23* no genoma de arroz utilizando como referência o genoma de *A. thaliana*.

1 Dra. em Fisiologia Vegetal, Universidade Federal de Pelotas, caixa postal 354, e-mail: ariadnehenriques@yahoo.com.br

2 Msc. em Agronomia, Universidade Federal de Pelotas.

3 Msc. em Agronomia, Universidade Federal de Pelotas.

4 Prof. Dr. em Agronomia. Universidade Federal de Pelotas.

5 Prof. PhD. em Agronomia. Universidade Federal de Pelotas.

## MATERIAL E MÉTODOS

Com o objetivo de identificar sequências homólogas, os genes *bZIP19* e *bZIP23* de *Arabidopsis thaliana* selecionados no banco de dados TAIR - The Arabidopsis Information Resource – (<http://www.arabidopsis.org/>) foram usados para buscar homólogos em *Oryza sativa* no banco de dados Phytozome (<http://www.phytozome.net/>) através da ferramenta Blastp e parâmetros default, sendo as sequências redundantes manualmente descartadas. As sequências de proteínas encontradas foram analisadas no banco de dados de famílias proteicas PFAM (<http://pfam.xfam.org/>) que inclui as anotações e os alinhamentos de sequências proteicas geradas a partir do modelo Hidden Markov (HMM). Com o intuito de aumentar a confiabilidade nas sequências selecionadas foi realizado um alinhamento domínio versus proteína no programa Blast2seq ([www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi](http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi)) sendo que como “query” foi utilizado a sequência consenso do domínio do *AtbZIP19* identificada através do banco de dados Pfam e com “subject” foram inseridos, uma a uma, as sequências de proteínas previamente encontradas. Por meio do Blast2seq foi analisada a identidade, a cobertura e o e-value das sequências.

As sequências de proteínas putativas bZIPs de *Arabidopsis* e arroz serviram como entrada para a análise de motivos conservados realizada com MEME (<http://meme.sdsc.edu/meme/mee.html>) versão 6.0. A largura máxima de um motivo foi ajustada para 50 e o número máximo de motivos foi ajustado para três, os outros parâmetros foram utilizados como padrão. Para corroborar os resultados, uma árvore filogenética foi construída a partir das proteínas ortólogas *bZIP19/bZIP23* preditas. O alinhamento de sequências múltiplas ocorreu através da ferramenta CLUSTALW do MEGA ver. 6.0. e foi realizado através da matriz PAM e o método Neighbor Joining. Os pontos de ramificação foram testados para significância por bootstrapping com 1000 repetições.

## RESULTADOS E DISCUSSÃO

Foram encontradas dez sequências potencialmente homólogas, entretanto a maioria das sequências apresentou similaridade menor que 70%, não sendo confiáveis para inferir relações funcionais e evolutivas entre as sequências. Para aumentar a confiabilidade um novo alinhamento, domínio versus proteína, foi realizado no programa Blast2seq.

Os domínios proteicos são considerados as unidades funcionais das proteínas, pois em geral estão associados a uma função (LESK, 2008). Os domínios, que são a parte conservada de uma determinada sequência de proteína, podem funcionar e existir independentemente do resto da cadeia proteica servindo de base para geração de novas sequências pela natureza. Os fatores de transcrição *AtbZIP19/ AtbZIP23* possuem um domínio conservado bZIP, região básica zipper de leucina, em todas Magnoliophytas (ASSUNÇÃO et al., 2010). Desta maneira, através do PFAM foi possível confirmar a presença da família *bZIP2* (Basic Leucine Zipper), domínio identificado por Pfam 7716, em seis das dez sequências selecionadas previamente no banco de dados do Phytozome. Através do alinhamento domínio versus proteína no programa Blast2seq foi possível notar que três das seis sequências apresentaram alta similaridade e identidade (acima > 70%), boa cobertura (100-83%) e e-value, indicando que as sequências não foram emparelhadas ao acaso podendo portanto serem sequências ortólogas exercendo entre si funções similares (Tab.1).

Tabela 1. Similaridade, cobertura e e-value das sequências de aminoácidos deduzidos de aminoácido bZIP23/bZIP19 de *Arabidopsis thaliana*.

Espécie	Sequence ID	Cobertura	Similaridade	e-value
<i>Arabidopsis thaliana</i>	AT2G16770.1 (bZIP23)	100	85	9e-29
<i>Arabidopsis thaliana</i>	AT4G35040.1 (bZIP19)	100	89	8e-40
<i>Arabidopsis thaliana</i>	AT3G51960.1 (bZIP24)	96	62	1e-18
<i>Oryza sativa</i>	LOC_Os06g50310.1 E2F bZIP48	96	77	2e-25
<i>Oryza sativa</i>	LOC_Os05g41540.1	94	78	5e-25
<i>Oryza sativa</i>	LOC_Os01g58760.1	94	71	3 e-23

Os fatores de transcrição *bZIP19* e *bZIP23* de *Arabidopsis thaliana* pertencem ao grupo F dos bZIPs que possuem 2 motivos característicos ricos em histidina (CysHis-motifs) na região básica N-terminal localizados entre a posição 55-68 e 71-81. Assunção et al (2013), sugerem que estes motivos participam da homeostase de zinco funcionando como sensores celulares de zinco sendo portanto, uma característica de grande importância. Das três sequências apenas duas, LOC\_Os06g50310.1 e LOC\_Os05g41540.1, apresentaramos motivos ricos em histina, conforme desejado. Estes estão localizados na sequência logo do motivo 3 (Fig. 1) e apresentaram um nível alto de conservação. A Sequência proteica LOC\_Os01g58760.1, que não apresentou motivos conservados Cys-His, foi descartada.



Figura 1 – Sequência logo de domínios conservados das proteínas de arroz preditas de aminoácido bZIP23/bZIP19 de *Arabidopsis thaliana*.

Para corroborar os resultados, uma árvore filogenética foi construída a partir das proteínas ortólogas *bZIP19/ bZIP23* preditas. Através da filogenia é possível inferir grupos ortólogos e explorar a relação evolucionária entre as proteínas bZIP. A árvore filogenética ficou agrupada em dois grupos estáveis (Fig. 2). No primeiro grupo ficou agrupada a sequência LOC\_Os06g50310, mais próximas de *AtbZIP19* e *AtbZIP23* e no segundo grupo ficou agrupada a proteína LOC\_Os05g41540 com maior proximidade com a sequência *AtbZIP24* cujo é um fator de transcrição também pertencente ao grupo F dos bZIP porém o mesmo não parece estar envolvido na homeostase de zinco mas com estresse salino (Assunção et al, 2010).

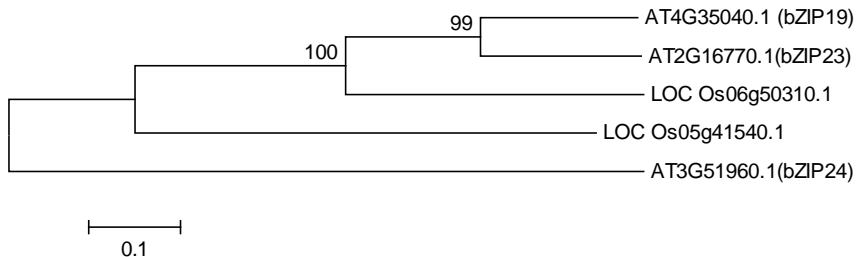


Figure 2. Relação evolucionária de Taxa

## CONCLUSÃO

Baseado na premissa que sequências com alta identidade tendem a apresentar estruturas parecidas e, por conseguinte funções biológicas similares, existe grande possibilidade do gene LOC\_Os06g50310.1 estar envolvido na regulação da homeostase de zinco em plantas assim como AtbZIP19 e AtbZIP23, possibilitando através de técnicas biotecnológicas auxiliar o desenvolvimento de cultivares tolerantes a deficiência de zinco e que possam crescer em ambientes com baixa disponibilidade de zinco sem grandes penalidades.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ASSUNÇÃO, A. G. L. et al. Model of how plants sense zinc deficiency. **Metallomics**, v.5, n.(9), p. 1110 -1116, 2013.
- ASSUNÇÃO, A. G. L. et al. Arabidopsis thaliana transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency. **Proceedings of the National Academy of Sciences**, v.107, p.10296-10301, 2010.
- CONAB. **Levantamentos de safra**: 2º Levantamento grãos safra 2014/15. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1253&ordem=criterioSafra1>>. Acesso em: 21 jun. 2015.
- FAO (Food and Agriculture Organization of the United Nations). **Rice in the World**. Disponível em <<http://www.fao.org>>. Acesso em: 09 fev. 2015.
- FUNGUETTO, C. et al. Desempenho de sementes de arroz irrigado recobertas com zinco. **Revista Brasileira de Sementes**, Londrina, v. 32, n. 2, p. 117-123, 2010.
- HENRIQUES, A. R. et al. Strategies to increase zinc deficiency tolerance and homeostasis in plants. **Brazilian Journal of Plant Physiology**, v.24, n.1, p. 3-8, 2012.
- LESK, A. (Ed.). **Introdução a Bioinformática**. Porto Alegre: Artmed, 2008.
- MARCHESAN, E. et al. Adubação foliar com micronutrientes em arroz irrigado, em área sistematizada. **Ciência Rural**, v.31, n.6, p.969-972, 2001.
- PFEIFFER, W. H.; MCCLAFFERTY, B. In: KANG, M.S.; PRIYADARSHAN, P.M. (Eds.). **Breeding major food staples for the 21<sup>st</sup> century**. Oxford, Blackwell Scientific, 2007. p. 61-91.
- SILVA, L.S. et al. Alterações nos teores de nutrientes em dois solos alagados, com e sem plantas de arroz. **Ciência Rural**, v.33, n.3, p.487-490, 2003.