

## ANÁLISE EVOLUTIVA DOS GENES AGPase NO GÊNERO ORYZA

KARINE ELISE JANNER DE FREITAS<sup>1</sup>; JENNIFER LUZ LOPES<sup>2</sup>, CAMILA PEGORARO<sup>2</sup>, VIVIAN EBELING VIANA<sup>2</sup>; CARLOS BUSANELLO<sup>2</sup>; ANTONIO COSTA DE OLIVEIRA<sup>3</sup>

<sup>1</sup>Universidade Federal de Pelotas – [karinejanner@gmail.com](mailto:karinejanner@gmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [jenniferlopesagronomia@gmail.com](mailto:jenniferlopesagronomia@gmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [pegorarocamilanp@gmail.com](mailto:pegorarocamilanp@gmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [vih.viana@gmail.com](mailto:vih.viana@gmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [carlosbuzzza@gmail.com](mailto:carlosbuzzza@gmail.com)

<sup>3</sup>Universidade Federal de Pelotas – [acostol@gmail.com](mailto:acostol@gmail.com)

### 1. INTRODUÇÃO

As preferências do consumidor pela qualidade dos grãos de arroz (*Oryza sativa* L.) variam em cada região do mundo com base em seus hábitos alimentares. Por isso, a melhoria na qualidade do grão de arroz tornou-se um dos componentes de pesquisa em quase todos os programas de melhoramento do arroz no mundo (PANDEY et al., 2012). Características como conteúdo de amilose (AC) e temperatura de gelatinização (GT) que exercem grande efeito sobre a qualidade de cozimento e consumo (ECQ) são controladas pelas propriedades físico-químicas do amido no endosperma dos grãos de arroz (PANDEY et al., 2012). A qualidade do amido nas culturas varia com base em diferentes proporções do conteúdo de amilose e amilopectina e/ou a estrutura da amilopectina (Yu et al., 2011). A biossíntese do amido é um processo complexo e altamente regulado, que requer atividades coordenadas entre enzimas múltiplas, incluindo a ADP-glicose pirofosforilase (AGPase). A AGPase é codificada por uma família gênica, e consiste em três subunidades grandes, codificadas pelos genes *AGPL1*, *AGPL3* e *AGPL4*, e uma subunidade pequena codificada pelo gene *AGPS2a*. A proteína AGPS desempenha uma função catalítica, enquanto que as enzimas AGPLs são responsáveis principalmente por modular as propriedades regulatórias alostéricas da AGPase.

Um dos requisitos básicos para o melhoramento de plantas é a presença de variabilidade genética. As cultivares comerciais de arroz apresentam base genética estreita devido a forte pressão de seleção para caracteres de interesse agronômico. Assim, a ampliação da variabilidade nessa espécie pode ser obtida através da introgressão de segmentos genômicos a partir de espécies selvagens do gênero *Oryza*. Essa proposta se aplica inclusive para caracteres associados a qualidade de grão. No entanto, pouco se sabe sobre a estrutura dos genes envolvidos na síntese do amido, bem como a estrutura das proteínas codificadas por esses genes nas espécies selvagens do gênero *Oryza*. Determinar essas características poderia auxiliar na definição da base genética que determina a ECQ, o que facilita o trabalho de melhoristas e biotecnologistas no sentido de manipular a qualidade de grão de forma previsível conforme as exigências do mercado. Por isso, o objetivo deste trabalho foi traçar o perfil evolutivo dos genes AGPase no gênero *Oryza* através de análise estrutural e filogenética.

### 2. METODOLOGIA

Inicialmente, os quatro genes AGPase em *O. sativa* ssp. *japonica* foram obtidos através do banco de dados RAP-DB (*The Rice Annotation Project*

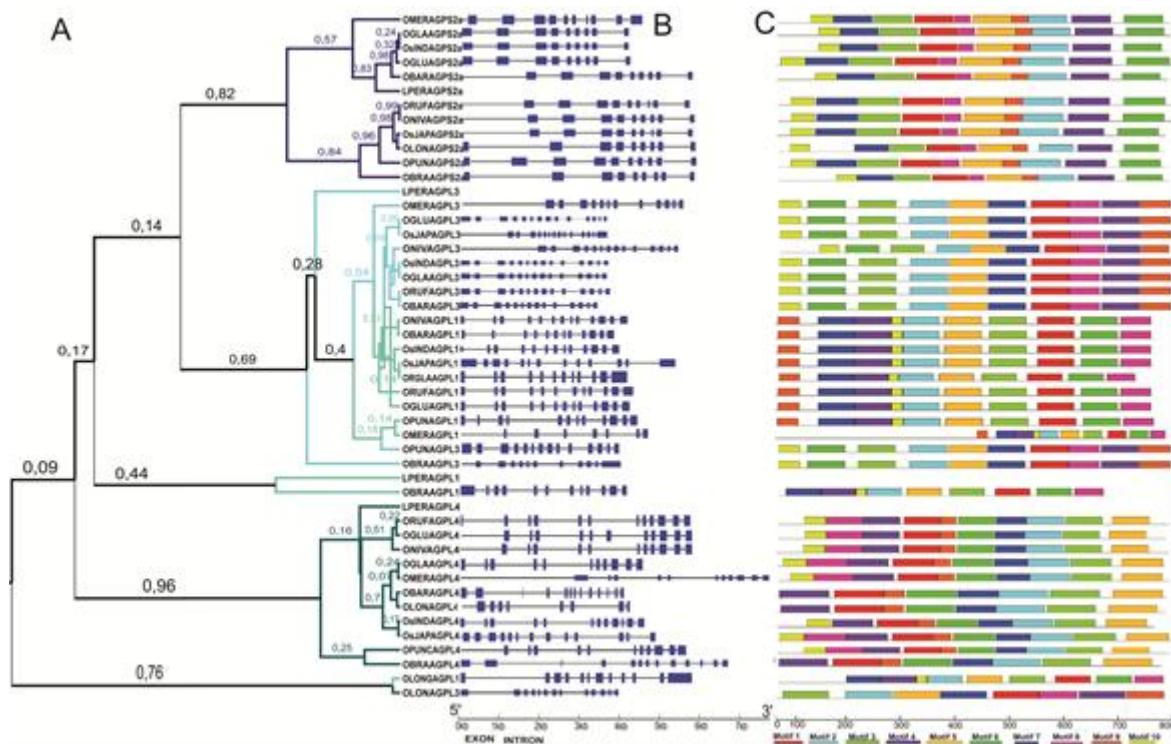
Database - <https://rapdb.dna.affrc.go.jp/index.html>). Para cada gene foi feita uma análise de BLAST no banco de dados ENSEMBL PLANTS (<http://plants.ensembl.org/index.html>) contra o genoma de onze espécies do gênero *Oryza* (*O. meridionalis*, *O. brachyantha*, *O. longistaminata*, *O. glumaepatula*, *O. glaberrima*, *O. sativa indica*, *O. rufipogon*, *O. barthii*, *O. punctata* e *O. nivara*). Os quatro genes AGPase selecionados em cada uma das espécies de *Oryza* e do grupo externo formado pela espécie *Leersia perrieri* foram submetidos a alinhamento global com ClustalW (LARKIN et al., 2007) com auxílio do software MEGA 7 (KUMAR et al., 2016). O melhor modelo de substituição foi obtido através de análise no software MEGA7 (KUMAR et al., 2016). O modelo apropriado foi selecionado para uso na análise Bayesiana utilizando o pacote BEAST (DRUMMOND & RAMBAUT, 2007) com 1.000.000 de réplicas de bootstrap. A estrutura do gene foi determinada através do software de anotação Augustus (STANKE & MORGESTERN, 2015) e visualizada através do software GSDS 2.0 (HU et al., 2014). As sequências das proteínas AGPase de cada espécie foram obtidas no banco ENSEMBL PLANTS. A identificação dos motivos das proteínas foi realizada utilizando Multiple Motif In Elicitation versão 4.11.1 (MEME; <http://meme-suite.org/tools/meme>) (BAILEY et al., 2009), considerando o número máximo de motivos igual a 10.

### 3. RESULTADOS E DISCUSSÃO

Nesse estudo, um total de 45 genes codificares das subunidades AGPase foram identificados nas espécies de *Oryza*. A árvore filogenética construída com base na sequência proteica indica que as sequências podem ser agrupadas em 4 clados (Figura 1A) conforme o posicionamento do grupo externo, *Leersia perrieri*. O primeiro deles compreende a proteína AGPS2a, que apresenta dois ramos (Figura 1A), sendo essa divisão decorrente do tamanho do gene (Figura 1B). Em um dos ramos incluem-se as espécies que possuem o gene menor que 5 Kb e apresentam 9 exons conservados, sendo que a espécie *O. meridionalis* apresenta um exon a mais na porção terminal do gene (Figura 1B). No outro ramo, estão as espécies que possuem o gene maior que 5 Kb, e também apresentam 9 exons conservados, sendo que a espécie *O. punctata* possui 10 exons (Figura 1B). A variação estrutural entre os dois grupos e dentro do segundo ramo ocorre devido à presença de um longo ítron na porção inicial do gene (Figura 1B). Apesar disso, os genes AGPS2a em *Oryza* são bastante conservados.

O segundo clado compreende o ramo que une as proteínas AGPL1 e AGPL3, sendo que o ramo ancestral dá origem ao terceiro clado, formando pela proteína AGPL1 de *O. brachyantha* com AGPL1 do grupo externo (Figura 1A). A estrutura do gene AGPL1 apresenta diferença no tamanho dos ítrons e posição dos exons, que variam de 12 a 13 exons, exceto em *O. meridionalis* que apresenta apenas 7 exons (Figura 1B). O número de exons do gene AGPL3 varia entre 11 a 14, sendo que a maioria das espécies apresentam 14 exons, e a maior variação entre estas ocorre devido à presença de um longo ítron na porção 5' (Figura 1B). No entanto, algumas espécies como *O. meridionalis*, *O. brachyantha* e *O. longistaminata* parecem ter perdido alguns exons durante a evolução (Figura 1B).

O quarto clado é composto pela proteína AGPL4 (Figura 1A). O gene desta proteína apresenta variações na posição e no número de exons (8 a 13 entre as espécies) (Figura 1B). Além das proteínas AGPL4, também foram agrupados OLONAGPL1 e OLONAGPL3.



**Figura 1.** Relação filogenética, estrutura dos genes e análise dos motivos conservados nas proteínas AGPase em *Oryza*. **A.** Árvore filogenética das proteínas foi construída com o pacote BEAST através de análise bayesiana utilizando réplicas de *bootstrap* de 1.000.000. Os ramos das proteínas AGPS2a, GPL3, GPL1 e GPL4 estão marcados em azul escuro, azul claro, verde claro e verde escuro, respectivamente. **B.** Estrutura éxons-ítrons dos genes AGPase no gênero *Oryza*. Os blocos azuis representam os éxons e as linhas escuras representam os ítrons. **C.** Motivos proteicos indicados em diferentes cores.

Uma análise foi realizada a fim de determinar o nível de diversificação dos motivos nas proteínas do grupo AGPase. Foram revelados 10 motivos (1-10), bastante conservados entre as espécies para a proteína AGPS2a (Figura 1C). Em relação as proteínas GPL1 e GPL3, apenas *O. longistaminata* e *O. brachyantha* apresentam 9 motivos, enquanto que para a proteína GPL4, *O. barthi*, *O. longistaminata* e *O. brachyantha* apresentam apenas 8 motivos (Figura 1C).

A divergência entre os genes que codificam proteínas da pequena e da grande subunidade da AGPase já havia sido identificada entre uma ampla gama de espécies de plantas (Qu et al., 2018). Aqui, foi demonstrada existência de divergência evolutiva entre as proteínas GPLs (GPL1, GPL3 e GPL4) e AGPS (AGPS2a) dentro do mesmo gênero (*Oryza*), indicando que as espécies selvagens de arroz podem servir de fonte de variabilidade genética para qualidade de grão e arroz.

#### 4. CONCLUSÕES

Os genes e proteínas AGPase apresentam variações estruturais entre as diferentes espécies do gênero *Oryza*. Essas variações precisam ser melhor avaliadas, pois podem ser fonte de variabilidade genética para uso no melhoramento para qualidade do grão de arroz.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- Drummond, A.J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC evolutionary biology**, v.7, n. 214 p 1-8, 2007.
- Pandey, M. K., Rani, N. S., Madhav, M. S., Sundaram, R. M., Varaprasad, G. S., Sivarajanji, A. K. P., Bohra A., Kumar G.R. & Kumar, A. Different isoforms of starch-synthesizing enzymes controlling amylose and amylopectin content in rice (*Oryza sativa* L.). **Biotechnology advances**, v.30, n.6, p. 1697-1706, 2012.
- Yu, G; Olsen, K.M. & Schaal, B.A. Molecular evolution of the endosperm starch synthesis pathway genes in rice (*Oryza sativa* L.) and its wild ancestor, *O. rufipogon* L. **Molecular Biology Evolution**, v.28, p. 659–671, 2011.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. Clustal W and Clustal X version 2.0. **Bioinformatics**, v. 23, n. 21, p. 2947-2948, 2007.
- Kumar, S., Stecher, G., & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular biology and evolution**, v. 33, n. 7, p. 1870-1874, 2016.
- Stanke, M., Steinkamp, R., Waack, S., & Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. **Nucleic acids research**, v. 32, n. 2, p. 309-312, 2004.
- Hu, B; Jin, J; Guo, A. Y; Zhang, H; Luo, J; & Gao, G. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics**, v. 31, n. 8, p. 1296-1297, 2014.
- Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. & Noble, W.S. MEME SUITE: tools for motif discovery and searching. **Nucleic acids research**, v. 37, n. 2, p. 202-208, 2009.
- Qu, J., Xu, S., Zhang, Z., Chen, G., Zhong, Y., Liu, L., Zhang, R., Xue, J. & Guo, D. Evolutionary, structural and expression analysis of core genes involved in starch synthesis. **Scientific Reports**, v. 8, p. 12736.