



UFG

**UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E
MELHORAMENTO DE PLANTAS**

**EFEITO DA SUPEREXPRESSÃO DO GENE *OVP1* EM
PARÂMETROS FISIOLÓGICOS E CARACTERES
RELACIONADOS A PRODUTIVIDADE DE GRÃOS EM
ARROZ DE TERRAS ALTAS**

DHIÔVANNA CORRÊIA ROCHA

Orientador:
Dr. Claudio Brondani



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE AGRONOMIA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese

2. Nome completo do autor

DHIÔVANNA CORRÊIA ROCHA

3. Título do trabalho

EFEITO DA SUPEREXPRESSION DO GENE OVP1 EM PARÂMETROS FISIOLÓGICOS E CARACTERES RELACIONADOS A PRODUTIVIDADE DE GRÃOS EM ARROZ DE TERRAS ALTAS

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.

Termo de Ciência e de Autorização (TECA) EA 1911748 SEI 23070.001780/2021-10 / pg. 1



Documento assinado eletronicamente por **DHIOVANNA CORREIA ROCHA, Discente**, em 02/03/2021, às 16:29, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Claudio Brondani, Usuário Externo**, em 02/03/2021, às 17:01, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1911748** e o código CRC **A4E371F8**.



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE AGRONOMIA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese

2. Nome completo do autor

DHIÔVANNA CORRÊIA ROCHA

3. Título do trabalho

EFEITO DA SUPEREXPRESSÃO DO GENE OVP1 EM PARÂMETROS FISIOLÓGICOS E CARACTERES RELACIONADOS A PRODUTIVIDADE DE GRÃOS EM ARROZ DE TERRAS ALTAS

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Claudio Brondani, Usuário Externo**, em 04/04/2023, às 07:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **DHIOVANNA CORREIA ROCHA, Discente**, em 04/04/2023, às 16:58, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3643335** e o código CRC **0017E8A0**.

Referência: Processo nº 23070.001780/2021-10

SEI nº 3643335

DHIÔVANNA CORRÊIA ROCHA

**EFEITO DA SUPEREXPRESSÃO DO GENE *OVP1*
EM PARÂMETROS FISIOLÓGICOS E
CARACTERES RELACIONADOS A
PRODUTIVIDADE DE GRÃOS EM ARROZ DE
TERRAS ALTAS**

Dissertação apresentada à Coordenação do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, como exigência para obtenção do título de Mestre em Genética e Melhoramento de Plantas com área de concentração em Genética e Melhoramento de Plantas.

Orientador:

Dr. Claudio Brondani

Goiânia, GO – Brasil

2021

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Rocha, Dhióvanna Corrêia

Efeito da superexpressão do gene OVP1 em parâmetros fisiológicos e caracteres relacionados a produtividade de grãos em arroz de terras altas [manuscrito] / Dhióvanna Corrêia Rocha. - 2021.

73 f.

Orientador: Prof. Claudio Brondani.

Dissertação (Mestrado) - Universidade Federal de Goiás, Escola de Agronomia (EA), Programa de Pós-graduação em Genética e Melhoramento de Plantas, Goiânia, 2021.

Bibliografia. Apêndice.

Inclui siglas, fotografias, abreviaturas, símbolos, gráfico, tabelas, lista de figuras, lista de tabelas.

1. Engenharia genética. 2. Produtividade de grãos. 3. RT-qPCR. 4. RNAseq. 5. Oryza sativa. I. Brondani, Claudio, orient. II. Título.

CDU 575



UNIVERSIDADE FEDERAL DE GOIÁS

ESCOLA DE AGRONOMIA

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº 07/2021 da sessão de Defesa de Dissertação de DHIÔVANNA CORRÊLA ROCHA, que confere o título de Mestre em Genética e Melhoramento de Plantas, na área de concentração em Genética e Melhoramento de Plantas.

Aos 05/02/2021 cinco dias do mês de fevereiro de dois mil e vinte e um, a partir da 08:15, por videoconferência, realizou-se a sessão pública de Defesa de Dissertação intitulada “EFEITO DA SUPEREXPRESSION DO GENE OVPI EM PARÂMETROS FISIOLÓGICOS E CARACTERES RELACIONADOS A PRODUTIVIDADE DE GRÃOS EM ARROZ DE TERRAS ALTAS”. Os trabalhos foram instalados pelo Orientador, Dr. Claudio Brondani (Embrapa Arroz e Feijão) com a participação dos demais membros da Banca Examinadora: Dra. Rosana Pereira Vianello (Embrapa Arroz e Feijão), membro titular externo; Dr. Guilherme Souza Prado (Embrapa Arroz e Feijão), membro titular externo. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido a candidata APROVADA pelos seus membros. Proclamados os resultados pelo Dr. Claudio Brondani, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos 05/02/2021 cinco dias do mês de fevereiro de dois mil e vinte e um.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por Claudio Brondani, Usuário Externo, em 05/02/2021, às 11:56, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por Rosana Pereira Vianello, Usuário Externo, em 05/02/2021, às 12:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por Guilherme Souza Prado, Usuário Externo, em 05/02/2021, às 13:24, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

A autenticidade deste documento pode ser conferida no site

https://sei.ufg.br/sei/controlador.php?acao=documento_imprimir_web&acao_origem=arvore_visualizar&id_documento=1950827&infra_sistema=1000... 1/2

9/03/2021

SEI/UFG - 1806418 - Ata de Defesa de Dissertação



https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 1806418 e o código CRC C4B17F89.

Referência: Processo nº 23070.001780/2021-10

SEI nº 1806418

AGRADECIMENTOS

Meus pais foram as pessoas que me incentivaram a seguir pelo caminho dos estudos. Foram eles que me deram a base para que eu conseguisse chegar até aqui. Agradeço imensamente toda dedicação, carinho e cuidado que tiveram comigo. Obrigada por estarem sempre ao meu lado.

Agradeço ao meu orientador Dr. Cláudio Brondani por ter aceitado me orientar e por ter me ensinado tanto. Obrigada, principalmente, por ter me dado a oportunidade de participar de projetos paralelos que me proporcionaram expandir minha experiência com técnicas moleculares. Gostaria de agradecer também à Dr.^a Rosana Vianello pelos ensinamentos no Laboratório de Biotecnologia e por todo o apoio que me ofereceu durante esses dois anos de pesquisa; ao Dr. Alexandre Coelho que me ajudou com as análises dos dados do RNAseq, pois sem ele não teria sido possível fazer tal análise, muito obrigada; e ao Dr. Odilon Peixoto pela grande ajuda com as análises estatísticas.

Agradeço aos meus companheiros de laboratório Thaís Ignez e João Augusto, por todo aprendizado e assistência. Agradeço também aos funcionários da Embrapa Arroz e Feijão, em especial às pessoas que estavam diariamente me ajudando e que foram essenciais para o desenvolvimento dessa dissertação: José Simião, Paula Arielle, Luana Alves, Gesimária Coelho, José Paixão, Sylvana de Paiva, Jadelson Sousa e João Antônio Mendonça. Saibam que não tenho palavras para agradecer por todo o conhecimento e experiência adquirida com vocês.

Ao programa de Pós-Graduação em Genética e Melhoramento de Plantas, à Embrapa Arroz e Feijão, à Universidade Federal de Goiás e à CAPES, muito obrigada por tornarem essa pesquisa possível.

SUMÁRIO

1	INTRODUÇÃO	16
2	REVISÃO BIBLIOGRÁFICA	19
2.1	A CULTURA DO ARROZ	19
2.2	O CARACTERE PRODUTIVIDADE DE GRÃO	20
2.3	O GENE <i>OVPI</i>	22
2.3.1	Proteínas H⁺-PPases: estrutura e função.....	22
2.3.2	Funções da proteína H⁺-PPase no metabolismo da planta	24
2.4	MÉTODOS DE TRANSFORMAÇÃO GENÉTICA DE PLANTAS	28
2.5	QUANTIFICAÇÃO DA EXPRESSÃO GÊNICA POR RT-qPCR	29
2.6	ANÁLISE DO TRANSCRITOMA VIA RNAseq.....	32
3	MATERIAL E MÉTODOS	34
3.1	TRANSFORMAÇÃO GENÉTICA DO ARROZ COM O GENE <i>OVPI</i>	34
3.2	CONDIÇÕES EXPERIMENTAIS.....	35
3.3	ANÁLISES FENOTÍPICAS	35
3.3.1	Análise dos dados morfológicos e dos componentes da produtividade	35
3.3.2	Análises dos dados fisiológicos	36
3.4	ANÁLISE ESTATÍSTICA.....	37
3.5	EXPRESSÃO GÊNICA	37
3.5.1	Extração, quantificação e análise da qualidade do RNA	37
3.5.2	Síntese do cDNA	38
3.5.3	Otimização da RT-qPCR	38
3.5.4	Quantificação da expressão do gene <i>OVPI</i> via RT-qPCR.....	40
3.5.5	Análise do transcrito via RNAseq.....	40
4	RESULTADOS	42
4.1	ANÁLISES FENOTÍPICAS	42
4.1.1	Dados morfológicos e produção de grãos	42
4.1.2	Análise dos dados fisiológicos	43

4.2	EXPRESSÃO GÊNICA	47
4.2.1	Quantificação da expressão do gene <i>OVPI</i> via RT-qPCR.....	47
4.2.2	Análise do transcrito via RNAseq.....	53
5	DISCUSSÃO	61
6	CONCLUSÕES.....	68
7	PERSPECTIVAS.....	69
	REFERÊNCIAS	70
	APÊNDICES	82

LISTA DE TABELAS

Tabela 1. Componentes da produtividade e dados morfológicos.	43
Tabela 2. Resultado dos dados fisiológicos analisados durante o estágio R6.	44
Tabela 3. Concentrações dos <i>primers</i> de referência e do gene <i>OVPI</i> escolhidas para serem utilizadas nas reações de RT-qPCR.	49
Tabela 4. <i>Primers</i> de referência usados nas reações de RT-qPCR.	51
Tabela 5. <i>Primers</i> desenhados para quantificação da expressão do gene <i>OVPI</i>	51
Tabela 6. Valores de Ct dos <i>primers</i> desenhados para o gene <i>OVPI</i>	52
Tabela 7. Genes diferencialmente expressos que foram positivamente regulados nas plantas geneticamente modificadas.	57
Tabela 8. Genes diferencialmente expressos que foram negativamente regulados nas plantas geneticamente modificadas.	58
Tabela 9. Categorias GO enriquecidas nas plantas geneticamente modificadas.	59
Tabela 10. Categorias GO reprimidas nas plantas geneticamente modificadas.	60

LISTA DE FIGURAS

Figura 1. Função da enzima H ⁺ -PPase no vacúolo celular.....	23
Figura 2. Modelo de funcionamento da enzima H ⁺ -PPase proposto para <i>V. radiata</i>	24
Figura 3. Função da enzima H ⁺ -PPase em células companheiras do floema.	27
Figura 4. Modelo do vetor utilizado na transformação genética das plantas OVP1-E4.....	34
Figura 5. Estádios nos quais foram coletados os dados fisiológicos e o tecido foliar para análise de expressão gênica.	36
Figura 6. Genótipos avaliados.....	42
Figura 7. Correlação entre os dados morfológicos e fisiológicos.	43
Figura 8. Resultados dos dados fisiológicos avaliados no ensaio 2018/2019.....	45
Figura 9. Resultados dos dados fisiológicos avaliados no ensaio 2019/2020.....	46
Figura 10. Eletroforese realizada para confirmar a ausência de gDNA nas amostras de cDNA.....	47
Figura 11. Curvas de dissociação dos <i>primers</i> utilizados nas reações de RT-qPCR..	48
Figura 12. Curva padrão realizada para determinação da eficiência dos <i>primers</i> utilizados nas reações de quantificação da expressão do gene <i>OVPI</i>	50
Figura 13. Expressão relativa do gene <i>OVPI</i> no ensaio 2018/2019.....	52
Figura 14. <i>Heatmap</i> representando o perfil de expressão das quatro bibliotecas para os genes diferencialmente expressos.....	54
Figura 15. Genes pertencentes à via fotossintética que foram <i>up-regulated</i> nas plantas geneticamente modificadas (OVP1-E4)..	55

LISTA DE APÊNDICES

Apêndice A. Resultado da quantificação e análise de qualidade das amostras de RNA provenientes do ensaio 2018/2019.....	83
Apêndice B. Resultado da quantificação e análise de qualidade das amostras de RNA provenientes do ensaio 2019/2020.....	84
Apêndice C. Resultado da análise de integridade do RNA via Bioanalyzer.....	85
Apêndice D. Otimização da temperatura de anelamento dos <i>primers</i> desenhados para o gene <i>OVPI</i>	86
Apêndice E. Expressão relativa do gene <i>OVPI</i> no ensaio 2019/2020.....	87
Apêndice F. Análise conjunta da expressão relativa do gene <i>OVPI</i> durante o estágio R6.....	88

LISTA DE SIGLAS E ABREVIATURAS

A	Taxa fotossintética/taxa de assimilação
ACT	Actina
ACi	Eficiência de carboxilação
ANOVA	Análise de variância
AVP1	<i>Arabidopsis Vacuolar Pyrophosphatase I</i>
cDNA	DNA complementar
Ci	Concentração intercelular de dióxido de carbono
CO₂	Dióxido de carbono
DNA	Ácido desoxirribonucleico
E	Taxa de transpiração
eEF1-α	<i>Eukaryotic Elongation Factor 1 alpha</i>
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
GAPDH	Gliceraldeído-3-fosfato-desidrogenase
gDNA	DNA genômico
Gs	Condutância estomática
GM	Geneticamente modificado
H⁺	Próton
H⁺-PPase	Próton pirofosfatase
IRGA	<i>Infrared Gas Analyzer</i>
Mg²⁺	Íon de magnésio
NGM	Não geneticamente modificado
OVP1	<i>Oryza Vacuolar Pyrophosphatase I</i>
PCR	<i>Polymerase Chain Reaction</i>
PSI II	Fotossistema II
qPCR	PCR Quantitativa
RNA	Ácido ribonucleico
SDS	<i>Sequence Detection Software</i>
iWUE	Eficiência instantânea no uso de água
WUEintr	Eficiência intrínseca no uso de água

RESUMO

ROCHA, D.C. **Efeito da superexpressão do gene *OVPI* em parâmetros fisiológicos e caracteres relacionados a produtividade de grão em arroz de terras altas.** 2021. 88f. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Escola de Agronomia, Universidade Federal de Goiás, Goiânia, 2021¹.

O aumento populacional, as mudanças climáticas e a redução da área para cultivo são alguns dos desafios que a agricultura vem enfrentando nos últimos anos, fazendo com que seja necessário desenvolver cultivares que sejam mais produtivas para atender a demanda global por alimentos. A cultivar de arroz de terras altas BRSMG Curinga foi transformada via *Agrobacterium tumefaciens* para superexpressar o gene *OVPI*. O gene *OVPI* codifica uma proteína capaz de bombear prótons através de membranas usando a energia da quebra do pirofosfato inorgânico (PP_i), e pode causar diversas alterações metabólicas e morfológicas nas plantas. Este trabalho objetivou avaliar alguns caracteres relacionados a produtividade de grãos, a fisiologia e a expressão gênica do arroz geneticamente modificado (GM) para superexpressar o gene *OVPI* (OVP1-E4) em relação ao BRSMG Curinga não geneticamente modificado (NGM). Foram realizados dois experimentos em condições controladas na Embrapa Arroz e Feijão, nas safras 2018/2019 e 2019/2020. O delineamento experimental foi baseado em blocos completos ao acaso com oito repetições para o primeiro ensaio e quatro repetições para o segundo. Foram obtidos os dados de alguns caracteres relacionados a produtividade de grãos em arroz, e dados fisiológicos foram coletados com o aparelho IRGA durante a fase reprodutiva (estádios R2, R3, R4, R6, R7, R8 e R9), possibilitando as estimativas da eficiência de carboxilação (A/Ci) e da eficiência no uso instantâneo (A/E) e intrínseco de água (A/Gs). O gene *OVPI* foi quantificado via RT-qPCR e o transcrito foi avaliado por RNAseq. Os dados coletados nos dois ensaios foram avaliados em uma análise de variância conjunta (p<0,05). O evento OVP1-E4 se destacou por produzir 21% mais grãos por panícula e 31% mais grãos cheios do que o NGM. Modificações fisiológicas significativas foram encontradas durante o enchimento de grãos (R6). Nesse estágio, o evento OVP1-E4 apresentou um incremento de 61% na taxa fotossintética e de 89% na eficiência de carboxilação, sem a ocorrência de um aumento na condutância estomática e na taxa transpiratória. As plantas GM também foram mais eficientes no uso de água e o gene *OVPI* foi mais expresso no evento OVP1-E4 durante o enchimento de grãos (R6). Genes envolvidos na fotossíntese tiveram a expressão aumentada no evento OVP1-E4. A superexpressão do gene *OVPI* resultou em plantas com alterações fisiológicas e fenotípicas que abrem caminho para o desenvolvimento de cultivares de arroz fisiologicamente mais eficientes e potencialmente mais produtivas.

Palavras-chave: engenharia genética, produtividade de grãos, RT-qPCR, RNAseq, *Oryza sativa*.

¹ Orientador: Prof. Dr. Claudio Brondani. Embrapa Arroz e Feijão.

ABSTRACT

ROCHA, D.C. **Effect of *OVP1* gene overexpression on physiological parameters and traits related to grain yield in upland rice.** 2021. 88f. Dissertation (Master in Genetics and Plant Breeding) – School of Agronomy, Federal University of Goiás, Goiania, 2021¹.

Population growth, climate change and reduction of area for cultivation are some of challenges that agriculture has been facing in recent years, making it necessary to develop cultivars to meet global demand for food. The highland rice cultivar BRSMG Curinga was transformed via *Agrobacterium tumefaciens* to overexpress the *OVP1* gene. *OVP1* gene encodes a protein that pumps protons across membranes using the energy from breaking down inorganic pyrophosphate (P_i) and trigger various metabolic and morphological changes in plants. This work aimed to evaluate the performance of some traits related to grain yield, physiology, and gene expression of genetically modified (GM) rice that overexpressed the *OVP1* gene (OVP1-E4) compared to the non-genetically modified (NGM). Two experiments were carried out under controlled conditions at Embrapa Rice and Beans in the 2018/2019 and 2019/2020 seasons. The experimental design was based on complete randomized blocks with eight repetitions for the first experiment and four for the second. Some traits related to grain yield in rice and physiological data were collected with the IRGA device during the reproductive phase (stages R2, R3, R4, R6, R7, R8 e R9), making it possible to estimate the carboxylation efficiency (A/C_i) and the instantaneous (A/E) and intrinsic water (A/G_s) use efficiency were estimated. The *OVP1* gene was quantified by RT-qPCR and the transcriptome was evaluated through RNAseq. A joint analysis of variance (p<0.05) was performed for the collected data in the two experiments. The OVP1-E4 event stood out for producing 21% more grains per panicle and 31% more filled grains than NGM. Significant physiological changes were found during grain filling (R6). At this stage, the OVP1-E4 event showed an increase of 61% in the photosynthetic rate and was 89% more efficient in carboxylation than NGM, without the occurrence of an increase in stomatal conductance and in the transpiratory rate. GM plants were also more efficient in using water. *OVP1* gene was more expressed in the OVP1-E4 event during the grain filling period, and genes involved in photosynthesis were up-regulated in GM plants. Overexpression of *OVP1* gene resulted in plants with physiological and phenotypic changes that pave the way for the development of rice cultivars more physiologically more efficient and more productive.

Keywords: genetic engineering, grain yield, RT-qPCR, RNAseq, *Oryza sativa*.

¹ Orientador: Prof. Dr. Claudio Brondani. Embrapa Arroz e Feijão.

1 INTRODUÇÃO

Aproximadamente 2 bilhões de pessoas não tiveram acesso regular a alimentos em 2019 (FAO, 2020). Além disso, estima-se que em 2050 a população mundial será composta por 9,7 bilhões de pessoas, um acréscimo de 26% sobre a quantidade atual (UN, 2019). Logo, o desafio é fazer com que todos tenham acesso a alimentação regular e de qualidade, ao mesmo tempo em que a cadeia produtiva, em particular a do arroz, deve se preparar para atender ao aumento populacional.

Considerando que o arroz faz parte da alimentação diária de milhões de pessoas ao redor do mundo, a produtividade de grãos se torna um caractere importante comum aos programas de melhoramento genético dessa cultura (ABIARROZ, 2019; CONAB, 2019; FAOSTAT, 2017; Fiaz et al., 2019). No Brasil, a produção total de arroz na safra 2017/2018 foi de 12 milhões de toneladas, porém teve uma redução nas duas últimas safras (MAPA, 2020a), e no contexto da pandemia da Covid-19, um conjunto de fatores levou o preço do arroz a subir consideravelmente (MAPA, 2020b), dificultando o acesso de muitos brasileiros a este cereal. Isso fez com que o Brasil retirasse a alíquota de importação até o final do ano de 2020 para que o produto pudesse ser importado (MAPA, 2020b) com objetivo de atender a demanda interna e evitar a majoração dos preços. Além disso, confirma a necessidade de se aumentar a produção de arroz, já que alimentos caros são um dos principais motivos que causam insegurança alimentar e desnutrição (FAO, 2020).

Nesse contexto, a agricultura enfrenta ainda outros desafios: as mudanças climáticas e a redução da área para cultivo devido à urbanização e a necessidade de preservação dos ecossistemas. Por isso, é importante desenvolver cultivares que sejam capazes de apresentar alta produtividade mesmo estando sob condições ambientais adversas (Carvalho et al., 2020; UN, 2019).

O arroz de terras altas, também conhecido como arroz sequeiro, é cultivado em um sistema alternativo ao de irrigação por inundação, e é caracterizado por demandar uma menor quantidade de recursos hídricos (Carvalho et al., 2020). Em um contexto de mudanças climáticas no qual a escassez de água pode inviabilizar a produção do arroz irrigado, é importante ter disponível plantas que combinem uma alta produtividade de grãos com tolerância ao déficit hídrico, uma característica típica das cultivares de arroz de terras altas

(Fageria et al., 2014; Ramirez-Villegas et al., 2018; Tuong & Bouman, 2003).

No Brasil, o arroz irrigado tem uma produtividade média de 7,7 t/ha com uma área cultivada de 1,4 milhão de hectares. Já o arroz de terras altas tem uma produtividade média de 2,3 t/ha, e apesar de ter apresentado aumento na produtividade nos últimos anos, perdeu 63% da área destinada à sua produção (Carvalho et al., 2020; CONAB, 2020). Esses fatores demonstram que é possível e necessário investir no aumento da produtividade de cultivares de arroz de terras altas, fazendo com que o Brasil tenha capacidade de incrementar sua produção de arroz ao mesmo tempo que reduz os custos de produção em relação ao arroz irrigado, por meio do desenvolvimento de cultivares que sejam mais tolerantes ao déficit hídrico.

O aumento da produtividade de grãos de cultivares de arroz de terras altas pode ser realizado por meio da engenharia genética. Esta por sua vez diminui o tempo de geração de novas variedades de plantas e permite a introdução de sequências de DNA entre indivíduos de espécies em que não é possível obter cruzamentos férteis naturalmente. Quando a sequência inserida pertence a outra espécie, obtém-se uma planta transgênica e o *pool* gênico da espécie receptora é ampliado. Entretanto, quando a sequência inserida é da própria espécie, obtém-se uma planta cisgênica. Ambos os tipos de transformação citados geram os organismos geneticamente modificados (OGMs) (Schouten & Jacobsen, 2006).

Quando os genes conhecidos em uma espécie ocorrem em outra espécie, são ditos homólogos, e quando esses genes são derivados de um gene pertencente a um ancestral comum que sofreu especiação, são conhecidos como ortólogos. Normalmente, os ortólogos apresentam função conservada (Dessimoz et al., 2012; Glover et al., 2016).

O gene *OVPI* (*Oryza Vacuolar Pyrophosphatase I*) é responsável por codificar uma proteína chamada de H⁺ pirofosfatase (H⁺-PPase) (Choura & Rebai, 2005; Gaxiola et al., 2016; Pan et al., 2011). A proteína H⁺-PPase foi bem caracterizada em *Arabidopsis thaliana*, sendo considerada uma enzima conservadora de energia por conseguir bombear prótons para o interior do vacúolo celular usando a energia resultante da quebra do pirofosfato inorgânico (PP_i) (Schilling et al., 2017). Sua função molecular primária interfere em outras atividades celulares, e por isso causa diversas alterações metabólicas e morfológicas em plantas que foram transformadas para superexpressar ou silenciar este gene.

Como consequência, as plantas geneticamente modificadas com o *OVPI* apresentaram maior biomassa, produtividade e se mostraram tolerantes ao estresse salino, ao

déficit hídrico, ao calor e ao frio (Esmaeli et al., 2019; Gaxiola et al., 2016; Pizzio et al., 2015; Raza et al., 2016; Schilling et al., 2017). Portanto, o gene *OVPI*, devido a sua importância em diversas rotas metabólicas, é um candidato para ser usado pela engenharia genética com o objetivo de aumentar a produtividade de grãos em arroz.

Uma das estratégias para se avaliar a expressão gênica após a modificação genética é a técnica de RT-qPCR (*Quantitative Reverse Transcription Polymerase Chain Reaction*), que permite fazer a quantificação da expressão de genes de maneira específica e sensível (Martin & Wang, 2011; Morgante & Blawid, 2016; Vinecky et al., 2011). Além disso, as diferenças metabólicas entre as plantas GM e NGM podem ser analisadas por meio do estudo dos seus transcritomas usando o sequenciamento do RNA (RNAseq). Essa técnica permite identificar o perfil de expressão de uma amostra, possibilitando a identificação dos genes diferencialmente expressos e às vias metabólicas às quais eles pertencem (Martin & Wang, 2011; Stark et al., 2019; Wang et al. 2009).

Com base nisso, o objetivo deste trabalho foi avaliar o efeito da superexpressão do gene *OVPI* nos componentes da produtividade de grãos, nos parâmetros fisiológicos e na expressão gênica da cultivar de arroz de terras altas BRSMG Curinga geneticamente modificada.

2 REVISÃO BIBLIOGRÁFICA

2.1 A CULTURA DO ARROZ

O arroz (*Oryza sativa* L.) é uma monocotiledônea autógama pertencente à família Poaceae. Seus grãos são constituídos principalmente de carboidratos, embora sejam fonte de vitaminas, minerais e proteínas (CONAB, 2015). Mais da metade da população mundial depende desse cereal para suprir suas necessidades alimentares diárias. Além disso, ele está diretamente relacionado com a segurança alimentar dos países em desenvolvimento (Muthayya et al., 2014). No Brasil, aproximadamente 95% da população consome arroz diariamente, tendo um consumo *per capita* de 34 kg/ano (CONAB, 2015; ABIARROZ, 2019).

Em média, a produção de arroz brasileira é de 12 milhões de toneladas por ano e a maior parte dessa produção é destinada ao consumo interno (ABIARROZ, 2019; CONAB, 2019; FAOSTAT, 2018). Em 2018, o Brasil ocupou a nona posição no *ranking* dos dez maiores produtores do mundo (FAOSTAT, 2018). Porém, a produção brasileira de arroz está diminuindo como consequência de fatores climáticos adversos e da redução da área plantada nos principais estados produtores do país (Rio Grande do Sul, Santa Catarina, Mato Grosso e Tocantins) (CONAB, 2019; IBGE 2019). Estima-se que a área cultivada de arroz diminuiu 38% no Brasil ao longo das últimas safras (CONAB, 2019), e um dos principais motivos apontados é a opção dos agricultores por culturas mais rentáveis, como a soja (Carvalho et al., 2020; CONAB, 2020).

O gênero *Oryza* possui 23 espécies silvestres e duas cultivadas, a *Oryza sativa* L., proveniente da Ásia; e *Oryza glaberrima* Steud, proveniente da África (Nadir et al., 2018; Sakai & Itoh, 2010). A espécie *O. sativa* é dividida em duas subespécies principais: a *Oryza sativa* ssp. *japonica*, que surgiu na China e é cultivada principalmente em regiões temperadas, sendo produzida no Brasil em condições de sequeiro; e a *Oryza sativa* spp. *indica*, que é originária da Índia e concentra sua produção em regiões tropicais e subtropicais, sendo produzida no Brasil em locais irrigados (Reifschneider et al., 2015; Sakai & Itoh, 2010; Muthayya et al., 2014).

Porém, foi sugerida outra divisão baseada em dados moleculares, na qual Garris et al. (2005) separaram *O. sativa* em cinco grupos (indica, aus, aromático, japonica temperada e japonica tropical), de acordo com a estrutura genética obtida por 169 marcadores microssatélites. De acordo com essa classificação, os genótipos tradicionais de sequeiro no Brasil pertencem ao grupo japônica tropical (Ferreira et al., 2005).

Quanto ao cultivo do arroz, existem dois principais sistemas utilizados: o irrigado e o sequeiro (ou de terras altas). O sistema de cultivo irrigado apresenta seus campos continuamente alagados durante o desenvolvimento da planta e, por isso, é característico de várzeas. Já o sistema de cultivo de terras altas é dependente da ocorrência de chuvas, e como consequência, está sujeito a perda de produtividade se ocorrer déficit hídrico durante a fase reprodutiva (Muthayya et al., 2014; Xia et al., 2016). As cultivares de arroz irrigado são adaptadas ao solo anaeróbico, enquanto as cultivares de arroz de terras altas são mais resistentes ao déficit hídrico (Fageria et al., 2014).

Segundo Carvalho et al. (2020), o arroz de terras altas apresenta como vantagens a utilização reduzida de água, menor custo de produção e o favorecimento de uma distribuição segura e estratégica de arroz por todo país. Além disso, por ser cultivado em condições aeróbicas, não emite gás metano como acontece com a produção de arroz irrigado.

Para 2020, a produção de arroz estimada foi de 11,2 milhões de toneladas (CONAB, 2020), o que representa uma redução de 7,14% em relação à média nacional de 12 milhões. Dessa produção, aproximadamente 10,3 milhões de toneladas foram provenientes do cultivo irrigado e 0,9 do cultivo sequeiro (CONAB, 2020). Isso demonstra a necessidade de desenvolver cultivares de arroz de terras altas que sejam mais produtivas, a fim de reduzir a dependência do grão oriundo do sistema de cultivo irrigado (Carvalho et al., 2020; Fageria et al., 2014).

2.2 O CARACTERE PRODUTIVIDADE DE GRÃO

A produtividade de grão é um caractere controlado por muitos genes, ou seja, é um caractere quantitativo. Normalmente, um caractere quantitativo é muito influenciado pelas alterações ambientais e possui baixa herdabilidade (Xing & Zhang, 2010). Por isso, para se ter um cultivo com alta produtividade deve-se combinar um bom genótipo com condições ambientais adequadas (Xing & Zhang, 2010).

Em arroz, são considerados os seguintes componentes da produtividade: o número de panículas, o número de grãos por panícula e o peso total dos grãos. O número de

panículas depende da habilidade da planta em formar perfilhos; o número de grãos por panícula é determinado principalmente pelo número de espiguetas (Xing & Zhang 2010); e o peso total dos grãos está diretamente relacionado ao enchimento dos grãos (Li et al., 2018). O enchimento dos grãos (estádio R6) é dependente da capacidade de divisão celular do endosperma, do equilíbrio entre produção e distribuição de fotoassimilados, e da capacidade da planta em produzir amido (Okamura et al., 2018; Zhang et al., 2016b).

A divisão celular do endosperma é regulada por fitormônios. Zhang et al. (2016b) demonstraram que existe uma correlação positiva entre a taxa de divisão celular, o enchimento de grãos e o teor de auxina e citocinina nos grãos. Esses hormônios se concentram primeiramente nas espiguetas superiores, promovendo divisão celular e aumentando as células do endosperma. Posteriormente, se concentram nas espiguetas inferiores. Plantas mutantes que apresentaram picos de concentração desses hormônios ao mesmo tempo em todas as espiguetas da panícula apresentaram menor percentual de grãos cheios e peso de grãos (Zhang et al., 2016b). Esse resultado demonstrou que é necessário haver um período de enchimento de grãos diferente entre as espiguetas de uma panícula, para que as plantas consigam transportar nutrientes para os grãos.

O percentual de grãos cheios também é limitado pelo suprimento de carbono e pela capacidade de dreno dos órgãos (Okamura et al., 2018). Como as espiguetas superiores enchem grão primeiro, se não houver suprimento suficiente para encher as espiguetas inferiores em seguida, pode-se ter uma redução no número de grãos cheios e no peso dos grãos. Dessa forma, é importante que haja um equilíbrio entre a produção de fotoassimilados e a capacidade de dreno das plantas, para que tanto as espiguetas superiores quanto as inferiores consigam encher os grãos (Murchie et al., 2002; Okamura et al., 2018).

Adicionalmente, é importante que as plantas tenham a capacidade de produzir amido a partir de sacarose (Ahmed et al., 2014; Okamura et al., 2018), já que o amido é a forma com que os carboidratos são armazenados nos grãos e a sacarose é a forma com que o açúcar produzido na fotossíntese é transportado (Ahmed et al., 2014). Okamura et al. (2018) demonstraram que a cultivar de arroz estudada que apresentou maior número de grãos cheios foi aquela que teve maior quantidade de amido nos entrenós, ou seja, a que conseguia transportar e transformar eficientemente o carbono fixado.

Durante a formação das panículas, as plantas conseguem armazenar carboidratos na bainha das folhas e nos colmos para o enchimento de grãos. Porém, a principal fonte de

carbono para o enchimento de grão é a fotossíntese, que ocorre durante este período do desenvolvimento (Murchie et al., 2002).

Genes relacionados ao número de panículas, ao número de grãos por panícula e ao peso total dos grãos já foram previamente usados na transformação genética do arroz com o objetivo de aumentar a produtividade (Ashikari et al., 2005; Fan et al., 2006; Heng et al., 2018; Huo et al., 2017; Miura et al., 2010). Dentre eles, pode-se citar: o *OsSPL14* (*Squamosa Promoter Binding Protein-like 14*), que provoca o aumento das ramificações da panícula; o *OsALMT7* (*Oryza Aluminum-activated malated transporter 7*), que está relacionado ao aborto das espiguetas; e o *GS3* (*Grain size 3*), que, quando superexpresso, induz à produção de grãos mais longos e por isso interfere no peso total dos grãos.

2.3 O GENE *OVP1*

O gene de arroz *OVP1* faz parte de uma família gênica que tem como função codificar proteínas H^+ -pirofosfatases (H^+ -PPases). As H^+ -PPases estão presentes principalmente na membrana dos vacúolos celulares, o tonoplasto (Choura & Rebai, 2005; Gaxiola et al., 2016; Pan et al., 2011). Entre as espécies de plantas para as quais o gene *OVP1* já foi descrito, pode-se citar *A. thaliana*, *Zea mays*, *Sorghum bicolor*, *Vitis vinifera* e *Vigna radiata* (Gaxiola et al., 2016; Pan et al., 2011). Em arroz, a identificação do gene *OVP1*, de acordo com a nomenclatura do *Rice Genome Annotation Project* (RGAP), é LOC_Os06g43660, enquanto seu ortólogo em *A. thaliana* possui a identificação AT1G15690, de acordo com a nomenclatura do *The Arabidopsis Information Resources* (TAIR).

Quanto à sua estrutura física, o gene *OVP1* apresenta 5770 resíduos de nucleotídeos localizados no cromossomo 6 do arroz, sendo constituído por 8 éxons e 7 íntrons, que proporcionam a ocorrência de três isoformas diferentes (1, 2 e 3) devido ao *splicing* alternativo. A região codante (CDS) das isoformas 1, 2 e 3 possuem 2349, 2220 e 2163 nucleotídeos, resultando em proteínas com 782, 739 e 720 resíduos de aminoácidos, respectivamente (*Rice Genome Annotation Project*, 2019).

2.3.1 Proteínas H^+ -PPases: estrutura e função

As proteínas H^+ -pirofosfatases (EC:7.1.3.1) têm como função bioquímica transportar H^+ do citoplasma para o interior do vacúolo a partir da clivagem do pirofosfato inorgânico (PP_i), e por isso são classificadas como translocases (Choura & Rebai, 2005;

Gaxiola et al., 2016; Pan et al., 2011). O Mg^{2+} é o cofator necessário para que exerçam sua função catalítica e, de maneira geral, elas são constituídas por um sítio de ligação ao substrato (complexo Mg^{2+}/PP_i) e por um domínio no qual os prótons são transportados pela membrana (Pan et al., 2011).

A reação catalisada pelas H^+ -PPases no vacúolo celular é:



Dessa forma, elas reduzem a concentração de pirofosfato inorgânico no citoplasma das células e contribuem para a acidificação dos vacúolos (Figura 1) (Lin et al., 2012; Schilling et al., 2017).

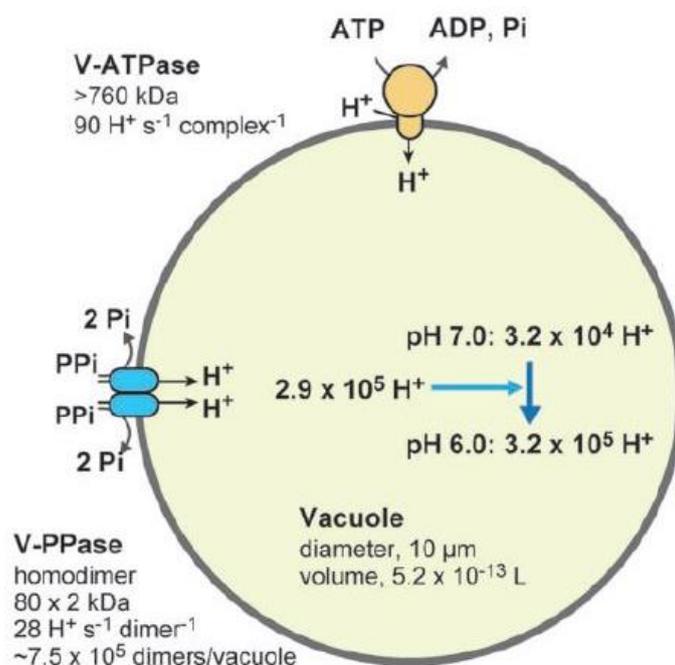


Figura 1. Função da enzima H^+ -PPase no vacúolo celular. Quando está presente no tonoplasto, a enzima H^+ -PPase hidrolisa PP_i para retirar a energia resultante e bombear prótons ao interior do vacúolo celular.
Fonte: Segami et al., 2018.

Quanto à sua estrutura, as H^+ -PPases são homodímeros com aproximadamente 75 Å de altura e 85 Å de largura (Lin et al., 2012), constituídos por 16 hélices transmembrana, uma região com muitos resíduos hidrofílicos voltados para o citoplasma e outra região mais hidrofóbica voltada para o lúmen do vacúolo. Essa região hidrofóbica é importante para não ocorrer a passagem contrária de prótons do vacúolo para o citoplasma (Lin et al., 2012). O canal onde ocorre a clivagem de PP_i e por onde atravessam os prótons

é composto por 6 hélices centrais que possuem resíduos de aminoácidos altamente conservados entre os genes ortólogos das H^+ -PPases (Lin et al., 2012).

Lin et al. (2012) propuseram um modelo de funcionamento para essas proteínas, em *V. radiata*, que explica o acoplamento entre a quebra do PP_i e a translocação de H^+ (Figura 2). Na primeira fase, chamada de estado de repouso, o canal está fechado para o lúmen do vacúolo e um próton encontra-se ligado a ele. Na segunda fase, denominada de estado I, o canal é semiaberto para o lúmen vacuolar e ocorre a quebra do pirofosfato em fosfato com a ajuda de moléculas de água. Por fim, na terceira fase, os prótons atravessam a membrana para o interior do vacúolo e então a proteína volta ao estado de repouso.

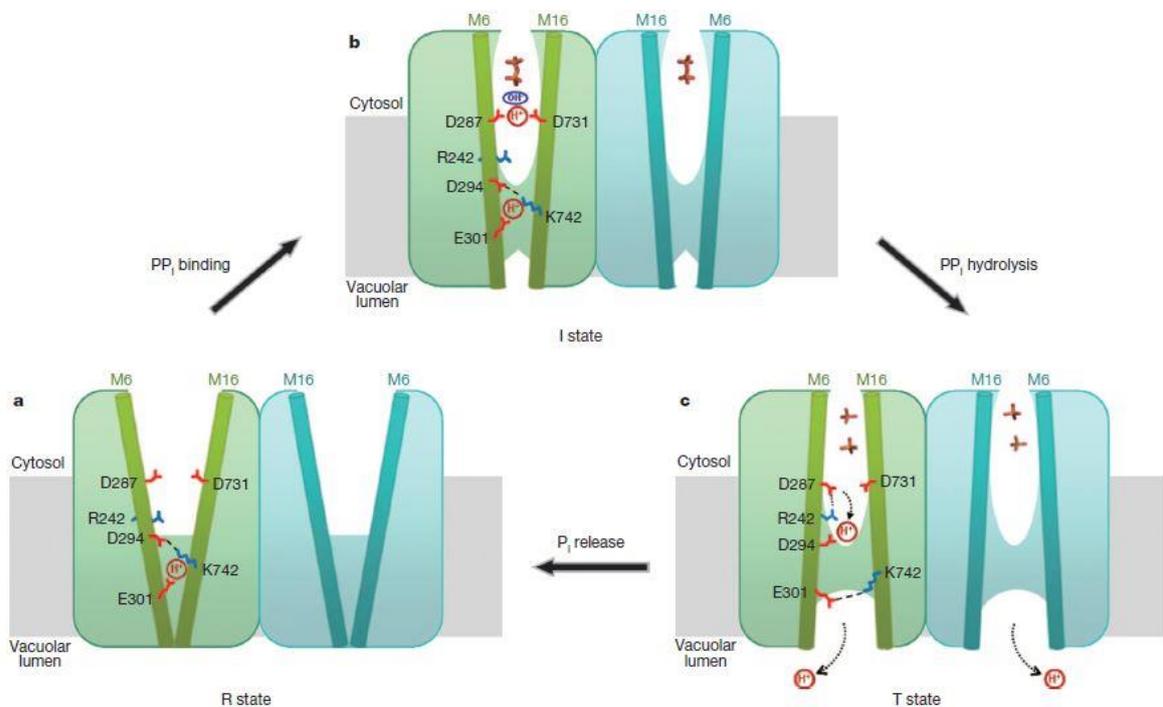


Figura 2. Modelo de funcionamento da enzima H^+ -PPase proposto para *V. radiata*. Fonte: Lin et al., 2012.

2.3.2 Funções da proteína H^+ -PPase no metabolismo da planta

As plantas possuem duas classes de enzimas responsáveis por acidificar o interior dos vacúolos: a H^+ -ATPase e a H^+ -PPase (Gaxiola et al., 2001; Li et al., 2016; Schilling et al., 2017). Como o próprio nome sugere, a H^+ -ATPase depende de ATP para exercer sua atividade catalítica, enquanto a H^+ -PPase usa a energia produzida pela quebra do pirofosfato para realizar a mesma função (Gaxiola et al., 2001), o que possibilita a

manutenção do potencial eletroquímico no tonoplasto mesmo em condições de estresse, quando a disponibilidade de ATP é baixa (Schilling et al., 2017).

De acordo com as funções moleculares da proteína codificada pelo gene *OVPI*, pode-se dividir a ação da enzima no metabolismo da planta em duas classes: a translocação de prótons e a quebra do PP_i em P_i (Schilling et al., 2017).

A translocação de prótons é importante por contribuir para a formação de um gradiente eletroquímico que é usado no transporte secundário de outras moléculas. Isso é essencial para o processo de compartimentalização que é realizado nos vegetais como forma de regular a distribuição de nutrientes e até mesmo de íons tóxicos (Gaxiola et al., 2001; Gaxiola et al., 2016; Schilling et al., 2017). Ao mesmo tempo, como o pirofosfato é um inibidor natural da gliconeogênese, a atividade de clivagem do PP_i em $2P_i$ pela H^+ -PPase diminui a inibição dessa via (Schilling et al., 2017). Nas plantas, a gliconeogênese se refere à formação de sacarose a partir dos lipídeos. Por isso, é uma via importante para a germinação das sementes (Asaoka et al., 2016).

Plantas de *A. thaliana* GM para superexpressar o gene *AVPI* apresentaram maior biomassa, sugerindo um aumento na fixação de carbono (Esmaeli et al., 2019; Gaxiola et al., 2016; Pizzio et al., 2015; Schilling et al., 2017). Quando submetidas a estresses abióticos, elas provocaram a acidificação do solo, exibiram mais raízes laterais e aumentaram a condutância estomática e a transpiração (Gaxiola et al., 2016; Pizzio et al., 2015; Raza et al., 2016; Schilling et al., 2017). Em contrapartida, plantas mutantes ou que tiveram esse gene silenciado apresentaram o fenótipo contrário (Gaxiola et al., 2016; Li et al., 2005; Pizzio et al., 2015).

Em condições de estresse salino, o gradiente de prótons permite a translocação de Na^+ para o vacúolo celular através de uma bomba (Na^+/H^+) (Gaxiola et al., 2016; Schilling et al., 2017), proporcionando a retenção de Na^+ dentro do vacúolo, diminuindo seus efeitos tóxicos e facilitando a captação celular de K^+ (Gaxiola et al., 2001).

Esmaeli et al. (2019) mediram a concentração de Na^+ e K^+ em folhas de *A. thaliana*, irrigadas com água salina, e que cossuperexpressavam os genes *OsSIZ1* e *AVPI*, descobrindo que elas retinham 20% e 30% mais íons Na^+ e K^+ , respectivamente, se comparadas aos indivíduos-controle não transformados. Esses autores obtiveram plantas com alta produtividade de grão, mesmo quando submetidas ao estresse salino, hídrico e ao calor. O mesmo foi observado com eventos de arroz transformados com o gene *AVPI* de *A. thaliana*. Em um trabalho realizado por Kim et al. (2014), as plantas foram tolerantes ao

estresse salino e apresentaram aumento no peso do colmo, no peso fresco da raiz, no número de panículas e no peso total dos grãos.

Além da interferência no transporte dos íons citados acima, Gaxiola et al. (2001) demonstraram que a diferença de potencial eletroquímico produzido pelas H^+ -PPases também contribui para a atividade da bomba Ca^{2+}/H^+ , aumentando em 36% a concentração de Ca^{2+} no interior vacuolar de plantas geneticamente modificadas com o gene *AVPI*. Como consequência do sequestro de cátions pelo vacúolo, ânions são captados pela célula e por esse motivo, a planta passa a absorver mais água para manter o potencial hídrico, fazendo com que o protoplasto e a parede celular sofram distensão para reter a água e gerando, assim, o chamado turgor celular. O turgor celular é responsável pelo crescimento e pela rigidez do vegetal, além de ser um dos fatores que controlam a abertura estomática (Gaxiola et al., 2001). O acúmulo de água intracelular favorece a adaptação das plantas ao déficit hídrico (Raza et al., 2016). Plantas de *A. thaliana* GM com o *AVPI* foram capazes de sobreviver e produzir sementes viáveis mesmo sendo submetidas ao déficit hídrico por 10 dias (Gaxiola et al., 2001).

Outro exemplo do efeito da superexpressão do *AVPI* foi observado em cana-de-açúcar, em que três eventos independentes mostraram um aumento médio da biomassa aérea e radicular de 78,3% e 82,6%, respectivamente, em relação às plantas não transformadas. Também houve um acréscimo significativo no parâmetro Brix de 13%. Esse parâmetro está relacionado ao conteúdo de sacarose da cana-de-açúcar (Raza et al., 2016).

A sacarose, por sua vez, é a forma com que os carboidratos, produzidos na fotossíntese, são transportados a longas distâncias (Khadilkar et al., 2016; Scholz-Starke et al., 2019). O carregamento de sacarose (SUC) nas células companheiras (CC) do floema é realizado por intermédio de bombas SUC/H^+ (Khadilkar et al., 2016; Pizzio et al., 2015; Scholz-Starke et al., 2019). Por isso, é necessário que haja uma diferença de potencial eletroquímico de prótons entre essas células e o apoplasto para que elas possam exercer sua função, e tal gradiente é gerado pelas H^+ -ATPases (Figura 3).

Em *A. thaliana*, foi comprovado que as proteínas H^+ -PPases também estão presentes na membrana plasmática das células companheiras do floema, onde atuam sintetizando PP_i (Khadilkar et al., 2016; Pizzio et al., 2015; Primo et al., 2019; Scholz-Starke et al., 2019). Esse PP_i é usado pelas CC para produzir o ATP através da oxidação da sacarose. O ATP, por sua vez, é usado pelas H^+ -ATPases para manter o gradiente de prótons necessário para o carregamento da sacarose no floema. Dessa forma, as H^+ -PPases atuam

contribuindo para distribuição da sacarose entre os órgãos-fonte e os órgãos-dreno (Figura 3) (Khadilkar et al., 2016; Scholz-Starke et al., 2019).

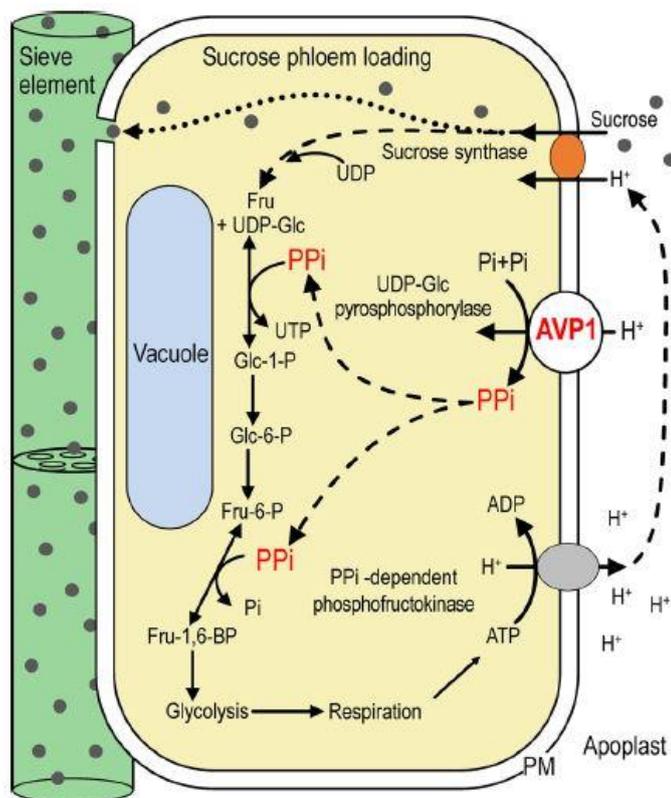


Figura 3. Função da enzima H⁺-PPase em células companheiras do floema. No floema, a enzima H⁺-PPase atua sintetizando o PP_i necessário para produção de ATP. Fonte: Primo et. al, 2019.

Adicionalmente, a superexpressão do *OVPI* conferiu tolerância ao frio em arroz, aumentando significativamente a taxa de sobrevivência de plantas GM cultivadas a 4°C devido a manutenção da integridade das membranas e à diminuição da acidez do citoplasma. A acidez citoplasmática diminuiu devido ao acúmulo de prótons no vacúolo e do aumento de prolina celular (Zhang et al., 2011). A prolina é tida como um aminoácido que regula a acidez do citoplasma, e esses autores relataram um aumento de três vezes na concentração de prolina no arroz GM superexpressando o *OVPI*.

Por fim, a expressão do gene *OVPI* está relacionada à ação da auxina, um hormônio que atua no desenvolvimento das plantas. Isso porque plantas de *A. thaliana* que tiveram o *AVPI* silenciado ou superexpresso apresentaram modificações morfológicas significativas, como a falta de desenvolvimento completo das flores e o aumento do número de folhas e da área foliar (Li et al., 2005), respectivamente. A explicação dada por esses autores é que as H⁺-PPases produzidas pelo *AVPI* são modificadoras da acidez intracelular

e extracelular, e como a auxina se move de acordo com o pH, a ação das H⁺-PPases acaba interferindo na atividade desse hormônio de forma sistêmica.

2.4 MÉTODOS DE TRANSFORMAÇÃO GENÉTICA DE PLANTAS

Nos últimos 20 anos, cerca de 2,1 bilhões de hectares foram cultivados com OGMs em 26 países. A maior área plantada com cultivares geneticamente modificadas (GM) é a de soja, seguida de milho, algodão e canola. De acordo com o International Service For The Acquisition of Agri-Biotech Applications (ISAAA, 2019), o Brasil está entre os países que mais plantam OGMs no mundo, ficando atrás apenas dos Estados Unidos. Além disso, é o país que apresenta maior crescimento anual de cultivo desses organismos (Chaudhary & Kumar, 2019). Ainda não há nenhum evento de arroz geneticamente modificado aprovado para cultivo no Brasil (ISAAA, 2021). Os países que têm autorização para cultivar arroz GM até o momento são China, Estados Unidos e Irã, sendo que as transformações genéticas aprovadas estão relacionadas à resistência a herbicidas e pragas (ISAAA, 2021).

Os principais métodos de transformação de plantas ocorrem via *Agrobacterium* spp. e por bombardeamento de partículas (biobalística) (Esmali et al., 2019; Gaxiola et al., 2001). O desenvolvimento da transformação via *Agrobacterium tumefaciens*, em laboratório, foi possível porque esses organismos são bactérias de solo que naturalmente infectam tecidos de plantas ao introduzirem parte de seu DNA plasmidial (T-DNA) no genoma vegetal, fazendo com que a planta produza os nutrientes essenciais para seu desenvolvimento (Altpeter et al., 2016; Safitri et al., 2016). Neste caso, para a transformação genética é necessária a produção de um vetor de expressão que contenha a sequência desejada (gene de interesse), além de um gene repórter e/ou um marcador de seleção que permitirá selecionar as plantas transformadas (Dedicova et al., 2015; Safitri et al., 2016).

A transformação via *A. tumefaciens* exige que o vetor seja inserido nas células bacterianas e que o tecido vegetal seja exposto a estas bactérias para então ser regenerado, com base em protocolos de cultura de tecido. Como *A. tumefaciens* infectam naturalmente dicotiledôneas, são propostos vários métodos para adaptar esta técnica para monocotiledôneas. Eles levam em conta os principais fatores limitantes que são: o tipo de explante, a cepa da bactéria, o vetor e a temperatura do cultivo (Safitri et al., 2016; Singh & Prasad, 2016).

A biobalística é baseada no bombardeamento de micropartículas de ouro ou tungstênio cobertas com o DNA, o RNA, ou a proteína que se deseja inserir no organismo a

ser transformado (Altpeter et al., 2016; Liu et al., 2019). Liu et al. (2019), ao transformarem arroz por biobalística, observaram que múltiplas formas de danos foram encontradas, incluindo quebra de cromossomo e variação extrema no número de cópias inseridas. Porém, apesar da inserção de múltiplas cópias não ser desejada, devido à possibilidade da planta silenciar a expressão do gene ou devido à possibilidade de ocorrerem inserções em genes essenciais para o desenvolvimento do vegetal, os danos graves no genoma causados pela biobalística favorece a inserção de genes exógenos inteiros no genoma por estimular o mecanismo de reparo do DNA por homologia (Altpeter et al., 2016; Liu et al., 2019).

A inserção randômica de transgenes feita tanto pela transformação via *A. tumefaciens* quanto pela biobalística pode gerar vários eventos de transformação que diferem entre si pelo local de inserção do transgene no genoma da planta e pela quantidade de cópias inseridas (Altpeter et al., 2016; Liu et al., 2019). Considerando que a regulação da expressão gênica depende, por exemplo, da acessibilidade da cromatina e da interação entre complexos proteicos e o promotor (Nelson & Cox, 2013), pode-se inferir que diferentes eventos de transformação geram níveis variados de expressão gênica.

2.5 QUANTIFICAÇÃO DA EXPRESSÃO GÊNICA POR RT-qPCR

Para avaliar a expressão dos genes que foram usados nas transformações genéticas, pode-se utilizar a técnica conhecida como RT-qPCR. Segundo Morgante & Blawid (2016), a RT-qPCR combina a amplificação de um fragmento alvo com métodos de quantificação que utilizam a emissão de fluorescência associada à síntese do alvo para quantificá-lo ao longo dos ciclos de PCR (*Polymerase Chain Reaction*). Essa associação faz com que o aumento exponencial da sequência seja registrado em tempo real, diferentemente da PCR convencional que permite visualizar o produto da amplificação (*amplicon*) apenas após a realização de uma eletroforese (Nolan et al., 2013).

Dois exemplos de sistemas de quantificação em tempo real são o SYBR Green e o TaqMan. O sistema SYBR Green se baseia na capacidade da molécula SYBR se intercalar na dupla fita de DNA. Após intercalado, esse reagente emite uma fluorescência quando excitado pela luz. O comprimento de onda de máxima excitação do SYBR é 500 nm e de máxima emissão é 520 nm (Kozera & Rapacz, 2013; Morgante & Blawid, 2016).

Já o sistema TaqMan utiliza uma sonda que se liga a uma região interna do *amplicon*. Essa sonda possui, na extremidade 5', um fluoróforo repórter e, na extremidade 3', uma molécula silenciadora chamada *quencher*. Quando a DNA polimerase cliva a

extremidade 5', o fluoróforo é liberado, fazendo com que ele fique longe do silenciador e, por isso emita um sinal. Esse sistema exige uma reação-controle, na qual é realizada uma qPCR sem a DNA polimerase para verificar se está ocorrendo emissão de fluorescência devido à degradação da sonda (Kozera & Rapacz, 2013; Morgante & Blawid, 2016; Nolan et al., 2013).

O sistema SYBR Green é mais barato, porém é necessário que, além de assegurar uma temperatura de anelamento específica, seja realizada uma curva de dissociação (curva de *melting*) para analisar a formação de dímeros de *primers* (Bustin et al., 2009; Nolan et al., 2013), uma vez que essas estruturas impedem que a eficiência da qPCR seja de 100%.

Além desses testes, é essencial que a concentração inicial das amostras seja estimada adequadamente e que elas apresentem boa qualidade, ou seja, não tenham contaminantes, como proteínas, e que o RNA/DNA inicial esteja íntegro, para não haver comprometimento da eficiência da qPCR (Kozera & Rapacz, 2013; Morgante & Blawid, 2016; Nolan et al., 2013).

Um exemplo de equipamento adequado para estimar a concentração de DNA ou RNA é o Qubit (Invitrogen), porque utiliza fluoróforos intercalantes ou ligantes aos ácidos nucleicos, permitindo detectá-los de maneira precisa (Nolan et al., 2013). Ele é diferente dos espectrofotômetros, que medem a concentração dos ácidos nucleicos a partir da absorbância da luz UV a 260 nm, provocando geralmente a superestimação da quantidade de DNA ou RNA que está presente na amostra devido à quantificação de outras substâncias (Bustin et al., 2009; Nolan et al., 2013).

Apesar disso, para verificar se existe contaminação nas amostras de ácidos nucleicos pode-se usar o espectrofotômetro, já que ele permite determinar a presença de proteínas por meio da relação 260/280 e a contaminação como sais e fenol a partir da relação 260/230 (Bustin et al., 2009; Nolan et al., 2013). Essas medidas são importantes, porque a contaminação com outros resíduos pode inibir a qPCR.

A qualidade das amostras ainda deve ser avaliada quanto ao nível de degradação ou integridade. Para isso, pode-se usar a eletroforese ou o equipamento microfluídico Bioanalyzer (Agilent) (Nolan et al., 2013). O Bioanalyzer fornece uma medida da integridade do RNA total com base na integridade do rRNA (subunidades 18S e 28S para eucariotos), emitindo valores de 1 a 10 para cada amostra. Quanto maior o valor, mais íntegro está o RNA (Mueller et al., 2004).

Após a verificação da qualidade, existem duas formas de analisar a expressão gênica por qPCR: por meio de um único passo, no qual a reação para a síntese do cDNA ocorre no mesmo tubo da qPCR e se utiliza um *primer* específico para o gene de interesse; ou através de dois passos, nos quais o cDNA é produzido e só depois usado para quantificação (Nolan et al., 2013). A segunda metodologia oferece maior flexibilidade para otimização da qPCR. Além disso, por permitir o uso de *primers* randômicos ou oligo d(T), ela possibilita a análise de expressão de vários genes a partir de uma única amostra (Morgante & Blawid, 2016; Ruiz-Villalba et al., 2017).

Para calcular os resultados dos ensaios de quantificação via qPCR, pode-se usar a quantificação absoluta ou relativa. A quantificação absoluta é realizada através da interpolação dos resultados em uma curva padrão que contém amostras com concentrações conhecidas, feita com base em um método independente (Bustin et al., 2009; Nolan et al., 2013). Já a quantificação relativa se baseia em genes de referência (*housekeeping*/controles internos) para fazer a estimativa dos resultados de cada amostra/gene em estudo, além de uma amostra de controle negativo da reação (Kozera & Rapacz, 2013; Morgante & Blawid, 2016; Nolan et al., 2013).

Os genes de referência são aqueles que apresentam expressão estável em diferentes condições, tecidos ou fase do desenvolvimento. São genes constitutivos, ou seja, envolvidos em atividades essenciais para a sobrevivência das células (Kozera & Rapacz, 2013; Nolan et al., 2013) e, por isso, são usados como normalizadores no método comparativo de Ct (*threshold cycle*) da RT-qPCR. É recomendado usar no mínimo dois genes de referência ao mesmo tempo para minimizar os erros provenientes de eventuais diferenças de expressão destes genes entre amostras ou experimentos (Kozera & Rapacz, 2013; Nolan et al., 2013).

O Ct é o ciclo no qual o sinal da fluorescência emitida passa o *threshold*, isto é, o desvio padrão médio do sinal da fluorescência repórter normalizada (*background*) (Kozera & Rapacz, 2013). Em outras palavras, é o ciclo no qual é possível detectar a fluorescência emitida, devido à amplificação exponencial do alvo, pela primeira vez. De modo que a quantidade inicial da sequência tem relação direta com a maior produção dos fragmentos amplificados. Por isso, amostras que apresentarem maior quantidade do alvo inicial, precisarão de menos ciclos para começar a emitir fluorescência, ou seja, terão um Ct menor.

No caso do método comparativo de Ct, é necessário que seja estimada a eficiência de reação dos *primers* usados para o gene de interesse e para os genes

normalizadores, já que os valores de eficiência devem ser próximos para que haja confiança na comparação (Morgante & Blawid, 2016). A eficiência ideal é de 100% por indicar que a cada ciclo o número de *amplicons* é dobrado (Kozera & Rapacz, 2013).

2.6 ANÁLISE DO TRANSCRITOMA VIA RNAseq

Entender o transcrito é essencial para conhecer e interpretar os elementos funcionais do genoma e sua atuação espaço-temporal (Martin & Wang, 2011). O RNAseq é uma técnica que permite ter uma visão global dos transcritos de uma determinada amostra em um momento específico, possibilitando conhecer os genes diferencialmente expressos e suas isoformas. Além disso, o RNAseq permite determinar quais genes foram positivamente regulados (*up-regulated*) ou negativamente regulados (*down-regulated*) sob determinada condição ambiental (Martin & Wang, 2011; Stark et al., 2019; Wang et al. 2009).

A forma mais comum de se obter um transcrito via NGS (*Next Generation Sequencing*) ocorre por meio do sequenciamento do cDNA. Os principais passos realizados para sequenciar cDNA são: extração do RNA total, síntese do cDNA, ligação de adaptadores às duas pontas (*paired end*) ou em apenas uma das pontas (*single end*), amplificação via PCR e sequenciamento (Martin & Wang, 2011; Wang et al. 2009).

Exemplos de plataformas de sequenciamento de cDNA incluem Illumina, Pacific Biosciences (PacBio) e Oxford Nanopore (ONT). A Illumina gera fragmentos pequenos (*short reads*), enquanto a PacBio e Oxford Nanopore geram fragmentos longos (*long reads*) (Stark et al., 2019). O surgimento da tecnologia da Oxford Nanopore possibilitou sequenciar moléculas de mRNA diretamente (dRNAseq), porém é uma técnica recente, e, portanto, a estratégia mais usada continua sendo o sequenciamento do cDNA com etapas de amplificação por PCR (Stark et al., 2019).

O sequenciamento de *short reads* possui menor taxa de erro e gera mais fragmentos por corrida. Entretanto, a montagem do transcrito é mais complicada se comparada aos *long reads*, que precisam juntar menos fragmentos e por isso acabam sendo mais confiáveis na montagem. Por outro lado, o sequenciamento de *long reads* apresenta maior taxa de erro (Stark et al., 2019). Assim, a escolha por uma ou outra estratégia deve considerar diversos aspectos, como objetivo da pesquisa e o custo.

Após a obtenção das sequências, o arquivo contendo o resultado do sequenciamento e as informações de qualidade de cada base possui formato FASTqc (Martin & Wang, 2011; Stark et al., 2019; Wang et al. 2009). A partir dele, pode-se montar o

transcritoma usando ou não um genoma de referência. No primeiro caso, as sequências obtidas são alinhadas em um genoma de referência e, a partir desse alinhamento, é feita a montagem por meio de *softwares*, como o STAR (Dobin et al., 2013). No segundo caso, chamado de sequenciamento *de novo*, usam-se os grafos de Bruijn para a montagem. Os grafos de Bruijn se baseiam na sobreposição de pequenas sequências nucleotídicas (k-mers) obtidas a partir dos *reads* para montar o transcritoma (Grabherr et al., 2013; Martin & Wang 2011).

Uma das vantagens da montagem do transcritoma usando um genoma de referência é uma menor preocupação com contaminantes ou artefatos na amostra, já que eles não deverão se alinhar na referência e, portanto, não deverão ser incluídos no transcritoma. Porém, o sucesso da montagem dependerá da qualidade do genoma de referência e da eficácia do alinhamento (Grabherr et al., 2013; Martin & Wang 2011).

Já a principal vantagem da montagem *de novo* é que novos transcritos podem ser descobertos. Contudo, essa estratégia exige alta cobertura de sequenciamento (quantidade de vezes que um nucleotídeo é sequenciado) para obter os transcritos completos, além de ser sensível a contaminantes e demandar máquinas com muitos *gigabytes* de memória RAM (Grabherr et al., 2013; Martin & Wang 2011; Stark et al., 2019). Por esse motivo, alguns *pipelines* de montagem de transcritoma combinam as duas estratégias ao usar *softwares* que fazem a montagem *de novo*, como o Trinity, e em seguida usar softwares, como o Bowtie2, para fazer alinhamentos que permitem verificar a qualidade do transcritoma (Martin & Wang, 2011).

Após a condução do sequenciamento *de novo*, é essencial eliminar as bases com baixa qualidade, a região dos adaptadores e os contaminantes (Martin & Wang, 2011), a fim de realizar a montagem e quantificar a expressão de cada gene. Em seguida, é necessário normalizar os dados para não se ter uma medida enviesada da expressão entre amostras distintas (Stark et al., 2019) e, então, determinar quais genes foram diferencialmente expressos (*up* e *down-regulated*). Ao fim da análise pode-se fazer a anotação de tais genes determinando as proteínas putativas, as vias metabólicas de atuação e as categorias GO (*Gene Ontology*).

3.2 CONDIÇÕES EXPERIMENTAIS

O experimento realizado na safra 2018/2019 foi conduzido na Plataforma de Fenotipagem Semi-Automatizada SITIS, na Embrapa Arroz e Feijão (49° 17'W, 16° 28', 779 m de altitude), localizada em Santo Antônio de Goiás (GO), entre outubro de 2018 e março de 2019, com base no delineamento de blocos completos ao acaso, com oito repetições para cada genótipo: OVP1-E4, geneticamente modificado (GM), e BRSMG Curinga, não geneticamente modificado (NGM - controle). Cada parcela foi constituída por uma coluna de PVC (40 cm altura x 30 cm de diâmetro) contendo duas plantas. O solo utilizado foi da classe latossolo vermelho distrófico. A irrigação ocorreu uma vez ao dia para repor 80% da capacidade de campo do solo, que foi calculada após 24 horas do encharcamento do solo presente na coluna, com o auxílio de uma balança digital. A capacidade de campo é definida como o máximo de água retida pelo solo após o excesso ter sido drenado (Mello et al., 2002).

As plantas GM pertencentes à geração T4 foram submetidas à aplicação com herbicida glufosinato de amônia (produto comercial Liberty®, BASF) com o objetivo de certificar que todas as plantas eram transformadas. Para isso, foi aplicada com pincel na parte adaxial das folhas de arroz, uma solução de glufosinato de amônio 2% com Tween-20 a 0,1%.

Entre outubro de 2019 e março de 2020 realizou-se outro experimento com o evento OVP1-E4 e a cultivar BRSMG Curinga como controle. A planta mais produtiva de cada genótipo do experimento anterior teve suas sementes utilizadas para a avaliação na geração T5. O delineamento experimental foi baseado em blocos completos ao acaso com quatro repetições para cada genótipo. O experimento foi conduzido novamente na Embrapa Arroz e Feijão, no telado 14 (T14), sob licença para manipulação de OGMs. O solo e as condições de irrigação foram os mesmos do experimento anterior.

3.3 ANÁLISES FENOTÍPICAS

3.3.1 Análise dos dados morfológicos e dos componentes da produtividade

O tamanho da folha bandeira foi estimado com base na média do comprimento e da largura de três folhas por parcela durante a maturação fisiológica dos grãos. Para medir o comprimento, foi utilizada uma régua (escala em cm) e para medir a largura foi utilizado um paquímetro digital (escala em mm). O número de perfilhos e panículas foram contados no final da fase reprodutiva.

Com o objetivo de estimar o número de grãos cheios e vazios, três panículas por planta foram separadas aleatoriamente para contagem dos grãos. Logo após, foi obtida a média do número de grãos cheios, do número de grãos vazios e do número total de grãos por panícula. Por fim, todos os grãos cheios foram separados manualmente e pesados em uma balança digital de precisão para se obter o peso total de grãos por planta (produção de grãos).

3.3.2 Análises dos dados fisiológicos

No experimento que ocorreu em 2018/2019, parâmetros fisiológicos foram medidos em uma planta por parcela durante os estádios (segundo Counce et al., 2000): R3 (formação das panículas), R6 (enchimento de grãos) e R8 (maturação fisiológica com ao menos um grão marrom desenvolvido). Já no experimento de 2019/2020, foram realizadas medições nas fases R2 (formação da folha bandeira), R4 (florescimento), R6 (enchimento dos grãos,) R7 (início da maturação fisiológica dos grãos com ao menos um grão amarelo) e R9 (final da maturação fisiológica dos grãos com panículas contendo a maioria dos grãos da cor marrom) (Figura 5). Desse modo, houve uma ampliação do número de fases avaliadas no segundo ensaio para que pudesse verificar se em outro estádio iria ocorrer alguma alteração fisiológica significativa entre os genótipos, ao mesmo tempo em que se repetiu as medições da fase R6, com o objetivo de verificar se os resultados encontrados no experimento 2018/2019 iriam se repetir.



Figura 5. Estádios nos quais foram coletados os dados fisiológicos e o tecido foliar para análise de expressão gênica. No ensaio 2018/2019 foram coletados dados nas fases R3, R6 e R8. No ensaio 2019/2020 foram coletados nas fases R2, R4, R6, R7 e R9.

Utilizando o aparelho *Infrared Gas Analyzer* (IRGA LCpro-SD, ADC BioScientific) foram obtidas a taxa fotossintética (A) ($\mu\text{mol} \times \text{m}^{-2}\text{s}^{-1}$), a taxa de transpiração (E) ($\text{mmol} \times \text{m}^{-2}\text{s}^{-1}$), a condutância estomática (Gs) ($\text{mol} \times \text{m}^{-2}\text{s}^{-1}$) e a concentração de carbono interno (Ci) (vpm). A densidade do fluxo de fótons de luz ativo utilizada foi de

1200 μmol [quanta] m^{-2} e a área foliar foi 3,8 cm^2 . Foram estimadas a eficiência de carboxilação (A/Ci), a eficiência no uso instantâneo de água (A/E) ($\mu\text{mol} \times \text{m}^{-1}$) e a eficiência no uso intrínseco de água (A/Gs) ($\mu\text{mol} \times \text{m}^{-1}$).

3.4 ANÁLISE ESTATÍSTICA

Para avaliar os dados obtidos nos dois experimentos, foi realizada uma análise de variância conjunta no *software* Rstudio (v.4.0.2), seguida do teste de Tukey ($p < 0,05$). Os pacotes estatísticos utilizados foram: Readxl (Wickham et al., 2019), Lattice (Sarkar et al., 2020) e Agricolae (Mendiburu, 2020). Já os dados dos parâmetros fisiológicos obtidos apenas no experimento de 2019/2020 foram avaliados por uma ANOVA simples. Os pacotes utilizados para esse teste foram: Expedes.pt (Ferreira et al., 2018) e Easyanova (Arnhold, 2019). O programa GraphPad Prism foi utilizado para construção de alguns gráficos. Além disso, as análises de correlação de Pearson foram feitas no RStudio usando os pacotes Corrplot (Wei, 2017) e Hmisc (Harrell, 2020).

3.5 EXPRESSÃO GÊNICA

Para análise de expressão gênica, folhas das plantas GM e NGM foram coletadas simultaneamente à obtenção dos dados fisiológicos. No experimento realizado em 2018/2019, a quantificação da expressão do gene *OVPI* ocorreu para as três fases do desenvolvimento em que houve coleta foliar (R3, R6 e R8). Já no experimento 2019/2020, as fases nas quais o gene *OVPI* foi quantificado via RT-qPCR foram R2, R6 e R7.

O tecido foliar de uma planta por parcela foi usado para extração do RNA total. Logo após a coleta, as folhas foram embaladas em papel alumínio, congeladas imediatamente em nitrogênio líquido e armazenadas em ultrafreezer (-80°C). Além das réplicas biológicas, três réplicas técnicas foram feitas para as reações de quantificação da expressão do gene *OVPI* e duas réplicas técnicas foram feitas para as reações anteriores à quantificação, que são as reações de otimização da RT-qPCR.

3.5.1 Extração, quantificação e análise da qualidade do RNA

Com o auxílio de um pistilo e um almofariz, as amostras foram maceradas em nitrogênio líquido até a obtenção de um pó fino. Em seguida, 100 mg do tecido macerado foram usados para extração do RNA total com o kit RNAeasy[®] (QIAGEN[®]). Durante o processo de extração, as amostras foram tratadas com a enzima DNase I (Invitrogen).

A qualidade foi estimada com base na relação 260/280 e 260/230 fornecida pelo espectrofotômetro UV-Vis NanoVue™ Plus (GE Healthcare) (Apêndices A e B). Para a verificação da integridade do RNA, algumas amostras foram analisadas no Bioanalyzer 2100 (Agilent Technologies Ltda) com o objetivo de certificar que apresentavam valores de RIN (*RNA Integrity Number*) acima de 6 (Apêndice C), indicando que estavam íntegras, ou seja, com baixo nível de degradação. Por fim, a quantificação do RNA total foi realizada no aparelho Qubit® 2.0 Fluorometer (Invitrogen) utilizando um kit de reagentes de alta sensibilidade (Qubit RNA HS Assay, Life Technologies) (Apêndices A e B).

3.5.2 Síntese do cDNA

A síntese do cDNA foi conduzida utilizando o kit GO Script™ Reverse Transcription System (Promega) e utilizou-se 0,5 µg x µl⁻¹ de RNA em conjunto com *primers* aleatórios (*random primers*). A reação de transcrição reversa ocorreu no termociclador T1000™ Bio-Rad, nas seguintes condições: 70°C a 5 min (desnaturação do RNA), 25°C por 5 min (anelamento dos *primers*), 42°C por 1 hora (extensão) e 70°C por 15 min (inativação da enzima). Logo após, a concentração de todas as amostras de cDNA foram ajustadas para 5 ng x µL⁻¹.

Para avaliar a ocorrência de amostras de cDNA contaminadas com DNA genômico, foram realizadas reações de PCR convencional utilizando o *primer* eEF-1α (GenBank™ ID AK061464), que amplifica o gene codificador do fator de alongamento 1α dos eucariotos. Essa amplificação permite diferenciar as sequências que possuem íntrons das que não possuem e que, portanto, correspondem as moléculas de DNA genômico e cDNA, respectivamente.

3.5.3 Otimização da RT-qPCR

Os *primers* foram desenhados na ferramenta *online* OligoPerfect (ThermoFisher Scientific (<https://apps.thermofisher.com/apps/oligoperfect/#!/design>)). Para o gene *OVPI*, foram desenhados quatro pares de *primers* (AVP_1, AVP_2, AVP_3 e AVP_4), em que os *primers forward* (F) anelam em uma junção éxon-éxon e os *reverse* (R) anelam no éxon seguinte.

Para testar a capacidade de amplificação dos pares de *primers* recém-sintetizados e a verificação da temperatura de anelamento ideal para cada par de oligonucleotídeos, foi realizada uma PCR utilizando as seguintes temperaturas de anelamento: 54°C, 56°C e 60°C.

O volume total de cada reação foi 10 μ l, contendo por 5 μ M dos *primers* F e R, 20 ng de cDNA, 5 μ l da Taq Master Mix 2x, 1 μ l do reagente Qsolution (Qiagen) e água milli-Q.

O termociclador Veriti 96 Well (Applied Biosystems®) foi usado para gerar o gradiente de temperatura. A desnaturação do DNA ocorreu a 95°C por 15 min, seguida de 40 ciclos de amplificação de 94°C por 30 segundos, 54°C/56°C/60°C por 1,5 min, 72°C por 1,5 min com extensão final de 60°C por 10 min. Para visualizar o resultado, as amostras provenientes da PCR foram analisadas por eletroforese em gel de agarose a 1%, corado com brometo de etídio (Apêndice D). A imagem do gel foi obtida em transluminador L.PIX (Loccus Biotecnologia).

O reagente SyberGreen PowerUp™ Master Mix 2x da Applied Biosystems® foi usado nas reações de RT-qPCR. As reações de RT-qPCR ocorreram em placas de 96 poços que foram seladas com filme óptico apropriado (MicroAmp® Optical 96-Well, Applied Biosystems®). O termociclador 7500 Real Time PCR System (Applied Biosystems®) foi usado para atingir as seguintes condições de termociclagem: 50°C por 2 min (incubação da UDG), 95°C por 2 min (desnaturação), 40 ciclos de 95°C por 15 segundos e 60°C por 1 minuto (anelamento); seguidos da curva de dissociação que se iniciou com 95°C por 15 segundos e prosseguiu com 60°C por 1 min, 95°C por 30 segundos e 60°C por 15 segundos. A análise dos dados foi feita no programa *Sequence Detection* (SDS) v2.05 (Applied Biosystems®).

Para avaliar a formação de dímeros de *primers* e obter a concentração que possibilitaria atingir o menor Ct, foram realizadas reações de RT-qPCR com três concentrações diferentes de *primers*: 300 nmol/ μ L, 600 nmol/ μ L e 900 nmol/ μ L. O volume final de cada reação teve 10 μ l, contendo 5 ng do cDNA, 5 μ l do PowerUp™ SYBR® Green Master Mix 2x e água *Nuclease Free*.

Em seguida, foi obtida a curva padrão para o cálculo da eficiência dos *primers* normalizadores e dos *primers* desenhados para o *OVPI*, utilizando-se as concentrações dos *primers* escolhidas no passo anterior. Uma diluição seriada do cDNA com cinco pontos (5 ng, 2,5 ng, 1,25 ng, 0,625 ng, 0,3215 ng) foi realizada para obter a curva padrão. O *software Sequence Detection* (SDS) v2.05 (Applied Biosystems®) forneceu o valor da eficiência e r^2 para cada *primer*.

3.5.4 Quantificação da expressão do gene *OVPI* via RT-qPCR

O volume final de cada reação de qPCR foi de 10 µl, contendo a concentração escolhida para cada *primer* (*forward* e *reverse*, Tabela 3), 5 ng do cDNA, 5 µl do PowerUp™ SYBR® Green Master Mix 2x e água *Nuclease Free*.

A expressão dos genes *actin* (ACT, GenBank™ ID X15865) (Imin et al., 2006; Zhang et.al, 2009), *eukaryotic elongation factor-1α* (eEF-1α, GenBank™ ID AK061464) (Jain et.al, 2006; Zhang et al., 2009) e *glyceraldehyde-3-phosphate dehydrogenase* (GAPDH, GenBank™ ID AK064960) (Kim et al., 2003) foi usada para normalizar a expressão do gene *OVPI*, sendo que a estabilidade de tais genes foi avaliada no *software* NormFinder.

A expressão relativa foi calculada no programa DataAssist™ V3.01 (Life Technologies) e as plantas BRSMG Curinga foram utilizadas como controle. Os dados de expressão de ambos os experimentos foram avaliados individualmente no *software* Rstudio (v 4.0.2) por uma ANOVA, seguida do teste de Tukey ($p < 0,05$).

3.5.5 Análise do transcrito via RNAseq

A planta que apresentou maior produção (peso total de grãos cheios) e maior número de grãos cheios no experimento 2018/2019 em cada genótipo foi analisada em duplicata por RNAseq, ou seja, fez-se uma comparação da expressão gênica das plantas GM e NGM que deram origem às sementes do experimento realizado no ano seguinte.

O sequenciamento foi realizado pela empresa Genone Soluções em Biotecnologia utilizando a plataforma Illumina Hi-Seq 2000™, pelo método *paired-end*. Para isso, 3 µg do RNA total, proveniente das folhas coletadas no estádio R6, foram armazenados em tubos RNA Stable® (Sigma Aldrich) e liofilizados a vácuo.

Após obtenção dos dados de sequenciamento, filtros de qualidade foram aplicados por meio do *software* Trimmomatic (Bolger et al., 2014) para retirar as sequências dos adaptadores, as bases com baixa qualidade e as *reads* menores que 150 pares de base (pb). Logo após, a montagem *de novo* do transcrito foi realizada no programa Trinity (Haas et al., 2013).

A qualidade da montagem, chamada de representatividade dos *reads*, foi avaliada a partir do alinhamento dos fragmentos filtrados no transcrito obtido, usando-se o programa Bowtie2 (Langmead et al., 2019). Para avaliar o grau de completude dos genes, também chamado de cobertura, o transcrito montado foi alinhado ao transcrito de

referência do arroz (*O. sativa ssp japonica*, Nipponbare - MSU Rice Genome version 7.0) utilizando a ferramenta BLAST.

Uma vez verificada a qualidade da montagem, os genes foram quantificados através da ação conjunta dos programas RSEM (Li & Dewey, 2011) e Bowtie2, que utilizaram o próprio transcrito como referência para fazer a quantificação. Em seguida, a análise de expressão diferencial foi feita com o programa EdgeR (Robinson et al., 2010), que considerou como significativo apenas os genes diferencialmente expressos (DEGs) com p-valor para FDR (*False Discovery Rate*) menor que 0,001 e *fold change* maior que 4.

As ORFs (*Open Reading Frame*) codificantes de no mínimo 100 resíduos de aminoácidos, foram obtidas com o *software* TransDecoder (GitHub) que, entre outros resultados, fornece um arquivo contendo as regiões CDS (*Coding DNA Sequence*) e outro contendo a sequência polipeptídica correspondente. Em seguida, as sequências polipeptídicas foram alinhadas no banco de dados Uniprot (Bateman, 2018) utilizando-se a ferramenta BLASTp e usadas para predição dos domínios conservados das proteínas putativas presentes no banco de dados Pfam (Finn et al., 2013). O resultado dessa etapa foi a versão proteica do transcrito, gerada com base no resultado do BLASTp e nos domínios conservados das proteínas putativas.

A anotação dos dados se prosseguiu com um BLASTx do transcrito, um BLASTp do transcrito em versão proteica e uma nova busca pelos domínios conservados das proteínas. O Trinotate (GitHub) foi usado para fazer um banco de dados, no qual todos os resultados obtidos pudessem ser armazenados e as categorias GO (*Gene Ontology*) pudessem ser obtidas (Carbon et al., 2009; Mi et al., 2019). Foi feita uma análise de expressão diferencial dos genes e das vias enriquecidas pelo Trinity usando o Gene Ontology e o banco de dados KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa & Goto, 2000; Kanehisa et al., 2019; Kanehisa, 2019). Adicionalmente, o programa MapMan (Usadel et al., 2012) foi utilizado para visualização dos mapas metabólicos e os pacotes “gplots” (Warnes et al., 2020) e “RColorBrewer” (Neuwirth, 2014) do RStudio foram usados para criar o *heatmap*.

4 RESULTADOS

4.1 ANÁLISES FENOTÍPICAS

4.1.1 Dados morfológicos e produção de grãos

Utilizando a análise conjunta dos experimentos que ocorreram em 2018/2019 e em 2019/2020 pôde-se avaliar quais foram os caracteres melhorados nas plantas geneticamente modificadas (OVP1-E4) em relação às plantas-controle não geneticamente modificadas (NGM) (Figura 6, Tabela 1). Dos oito caracteres avaliados, em três as plantas OVP1-E4 foram estatisticamente superiores às NGM: número de grãos cheios por panícula (31% maior; $p < 0,01$), número total de grãos por panícula (21% maior; $p < 0,01$), e comprimento da folha bandeira (9% maior, $p < 0,05$).



Figura 6. Genótipos avaliados. Esquerda: cultivar BRSMG Curinga geneticamente modificada para superexpressar o gene *OVP1* (evento OVP1-E4). Direita: cultivar BRSMG Curinga não geneticamente modificada (NGM). Estádio R2.

Tabela 1. Resultado da análise conjunta dos componentes da produtividade de grãos e dos dados morfológicos avaliados nas safras 2018/2019 e 2019/2020.

Dados	NGM	OVP1-E4	p-valor
Perfilhos	38,50	36,00	0,41
Panículas	38,50	36,00	0,41
Largura folha Bandeira (mm)	1,66	1,75	0,33
Comprimento folha Bandeira (cm)	25,32	27,86	0,01*
Número de grãos cheios por panícula	174,05	228,21	0,001*
Número de grãos vazios por panícula	39,39	30,29	0,39
Número total de grãos por panícula	213,44	258,50	0,004*
Produção (g/planta)	91,03	91,85	0,86

*resultados significativos para o genótipo com p-valor <0,05

A análise de correlação mostrou que a produção de grãos se correlacionou positivamente com a largura ($p < 0,05$) e com o comprimento da folha bandeira ($p < 0,05$) para ambos os genótipos (Figura 7). Adicionalmente, para o evento OVP1-E4 houve correlação positiva significativa entre o comprimento da folha bandeira e o número de grãos cheios; e o comprimento da folha bandeira e o número de grãos total (Figura 7).

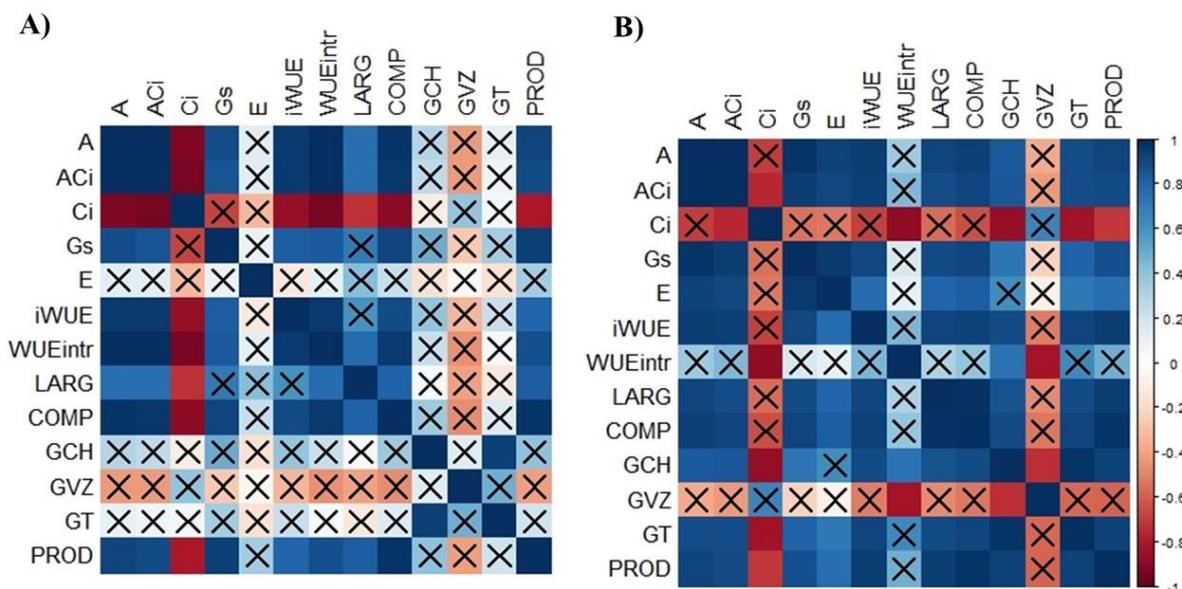


Figura 7. Correlação entre os dados morfológicos e fisiológicos. Azul: correlações positivas. Vermelho: correlações negativas. X: correlações não significativas ($p > 0,05$). A) Correlações do NGM. B) Correlações do evento OVP1-E4. LARG: largura da folha Bandeira. COMP: comprimento da folha Bandeira. GCH: número de grãos cheios. GVZ: número de grãos vazios. GT: número de grãos total. PROD: produção em peso de grãos cheios por planta. Dados fisiológicos obtidos no estágio R6.

4.1.2 Análise dos dados fisiológicos

Nos estádios R2, R3, R4, R7, R8 e R9 não houve diferença estatística entre os genótipos OVP1-E4 e NGM para os dados fisiológicos analisados (Figuras 8 e 9). Porém,

em R6 (enchimento de grãos), pôde-se observar mudanças significativas na maioria dos parâmetros fisiológicos avaliados (Tabela 2).

Tabela 2. Resultado da análise conjunta dos dados fisiológicos medidos durante o estágio R6 nos dois experimentos realizados (2018/2019 e 2019/2020).

Dados	NGM	OVP1-E4	p-valor
Taxa fotossintética (A)	8,15	13,15	0,002*
Carbono interno (Ci)	315,71	256,76	0,001*
Eficiência de carboxilação (ACi)	0,029	0,055	0,003*
Condutância estomática (Gs)	0,182	0,178	0,822
Taxa transpiratória (E)	3,18	3,11	0,802
iWUE	2,59	3,88	0,001*
WUEintr	41,53	73,15	0,001*

*resultados significativos para o genótipo com p-valor <0,05

Durante o enchimento dos grãos, o evento OVP1-E4 apresentou um aumento de 61,35% na taxa fotossintética em relação às plantas NGM ($p < 0,01$), reduziu sua concentração de carbono interno em 22,96% ($p < 0,01$), e aumentou em 89,6% a eficiência de carboxilação ($p < 0,01$). As plantas GM também apresentaram um aumento na eficiência no uso instantâneo (iWUE) e intrínseco (WUEintr) de água de 49,81% ($p < 0,01$) e 76,14% ($p < 0,01$) em relação às NGM, respectivamente (Tabela 2).

Houve uma correlação positiva entre a eficiência no uso instantâneo de água (iWUE) na fase R6 e a produção de grãos em ambos os genótipos; e uma correlação negativa entre a eficiência no uso intrínseco da água (WUEintr) e o número de grãos vazios nas plantas OVP1-E4 (Figura 7).

Além disso, as seguintes correlações positivas foram significativas para o evento OVP1-E4: taxa fotossintética e número de grãos cheios; taxa fotossintética e número de grãos total; taxa fotossintética e largura da folha bandeira; taxa fotossintética e comprimento da folha bandeira; eficiência de carboxilação e o número de grãos cheios; eficiência de carboxilação e número de grãos total; eficiência de carboxilação e largura da folha bandeira; e eficiência de carboxilação e comprimento da folha bandeira (Figura 7).

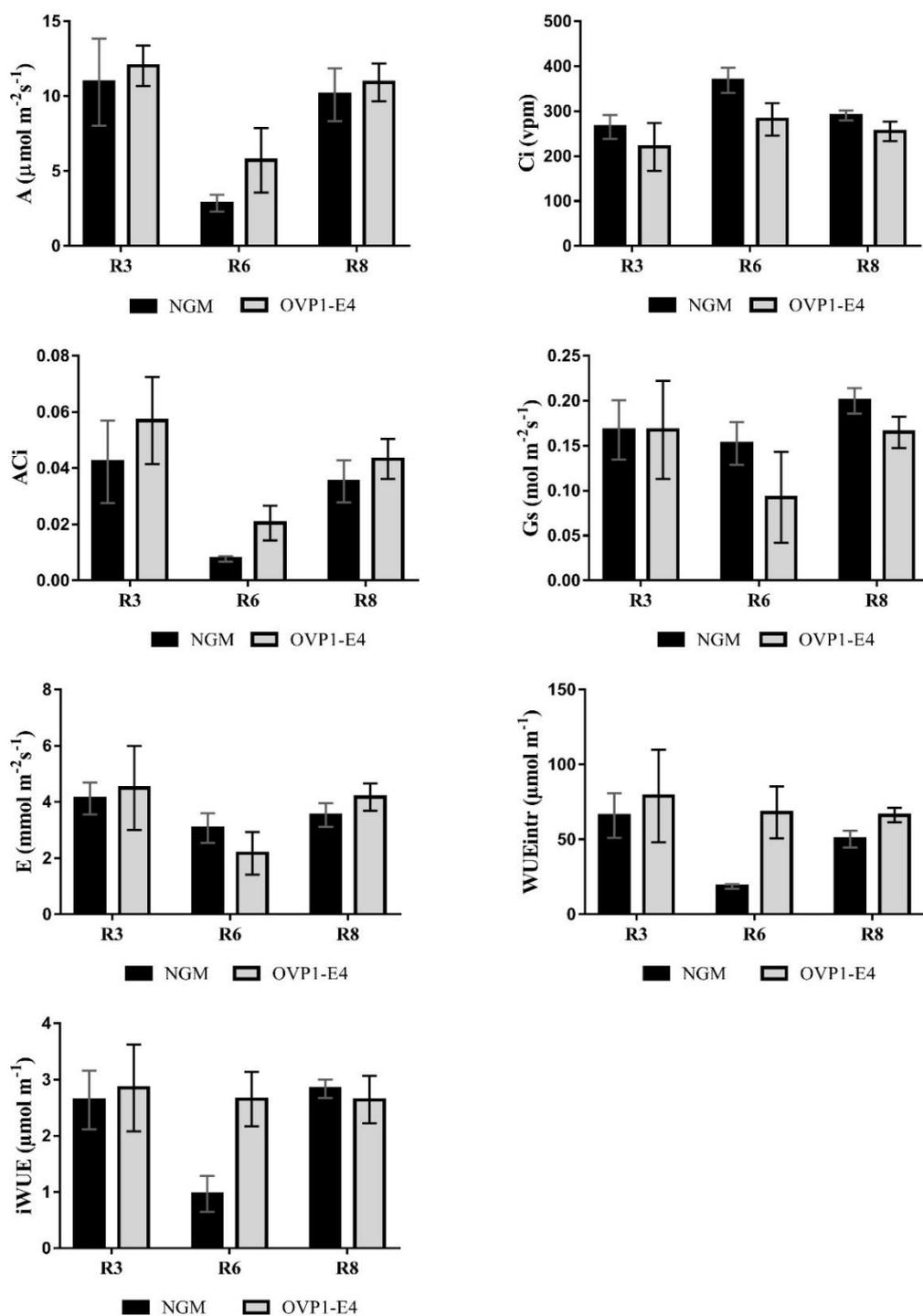


Figura 8. Resultados dos dados fisiológicos avaliados no ensaio 2018/2019. As barras indicam o desvio padrão. A: taxa fotossintética; Ci: concentração de carbono interno; ACi: eficiência de carboxilação; Gs: condutância estomática; E: taxa transpiratória; WUEintr: eficiência no uso intrínseco de água; iWUE: eficiência no uso instantâneo de água.

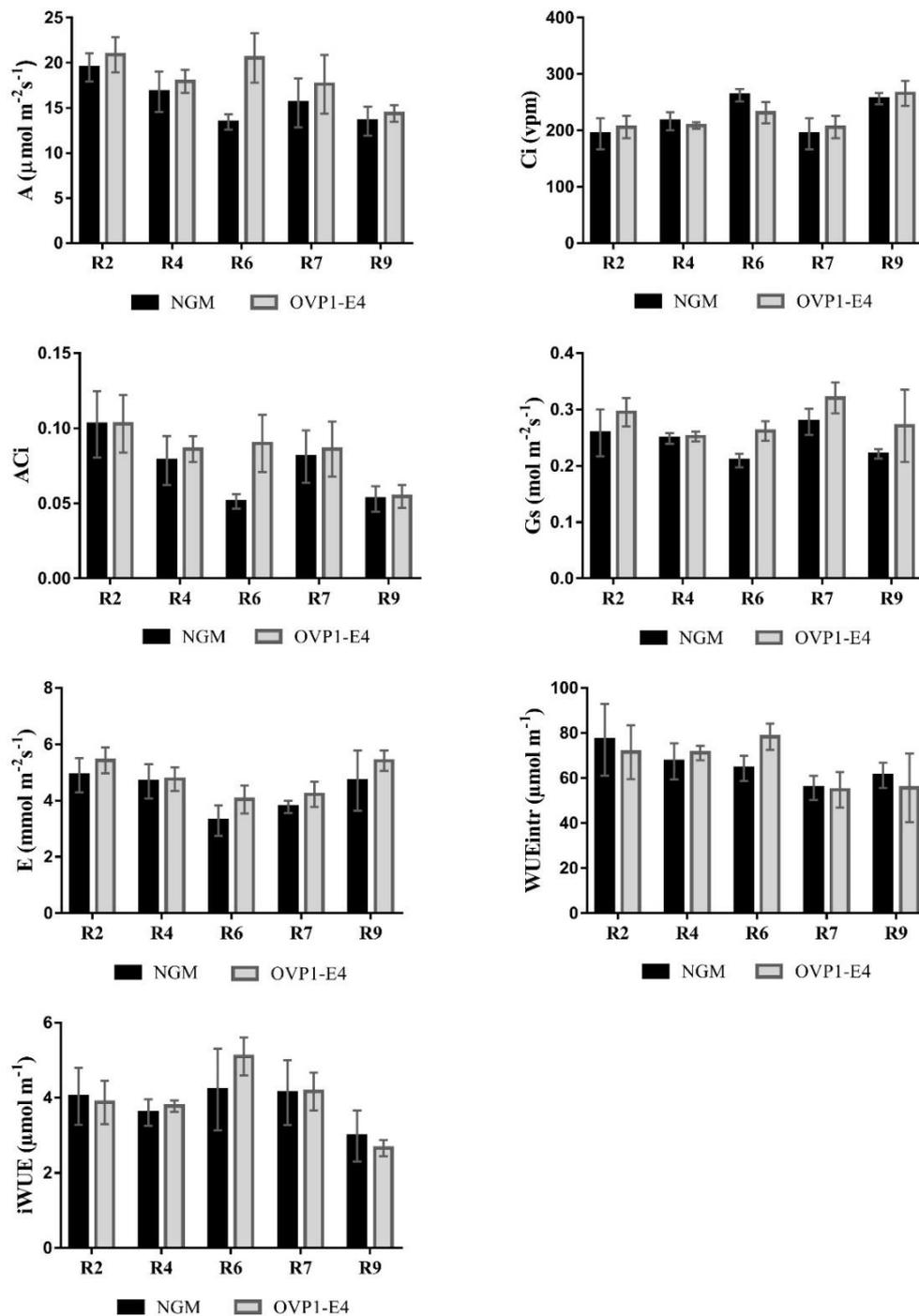


Figura 9. Resultados dos dados fisiológicos avaliados no ensaio 2019/2020. As barras indicam o desvio padrão. A: taxa fotossintética; Ci: concentração de carbono interno; ACi: eficiência de carboxilação; Gs: condutância estomática; E: taxa transpiratória; WUEintr: eficiência no uso intrínseco de água; iWUE: eficiência no uso instantâneo de água.

4.2 EXPRESSÃO GÊNICA

4.2.1 Quantificação da expressão do gene *OVPI* via RT-qPCR

Para quantificar a expressão do gene *OVPI* durante a fase reprodutiva das plantas, o RNA total foi extraído a partir do tecido foliar totalizando 48 amostras ao considerar quatro réplicas biológicas por genótipo, três épocas de coleta e dois experimentos. A maioria das amostras de RNA apresentaram valores adequados de pureza (Apêndice A e B) e integridade (Apêndice C). Adicionalmente, a presença de uma única banda de amplificação nas amostras de cDNA, resultante da PCR realizada com *primer eEF-1 α* , indicou a ausência de gDNA contaminante (Figura 10).

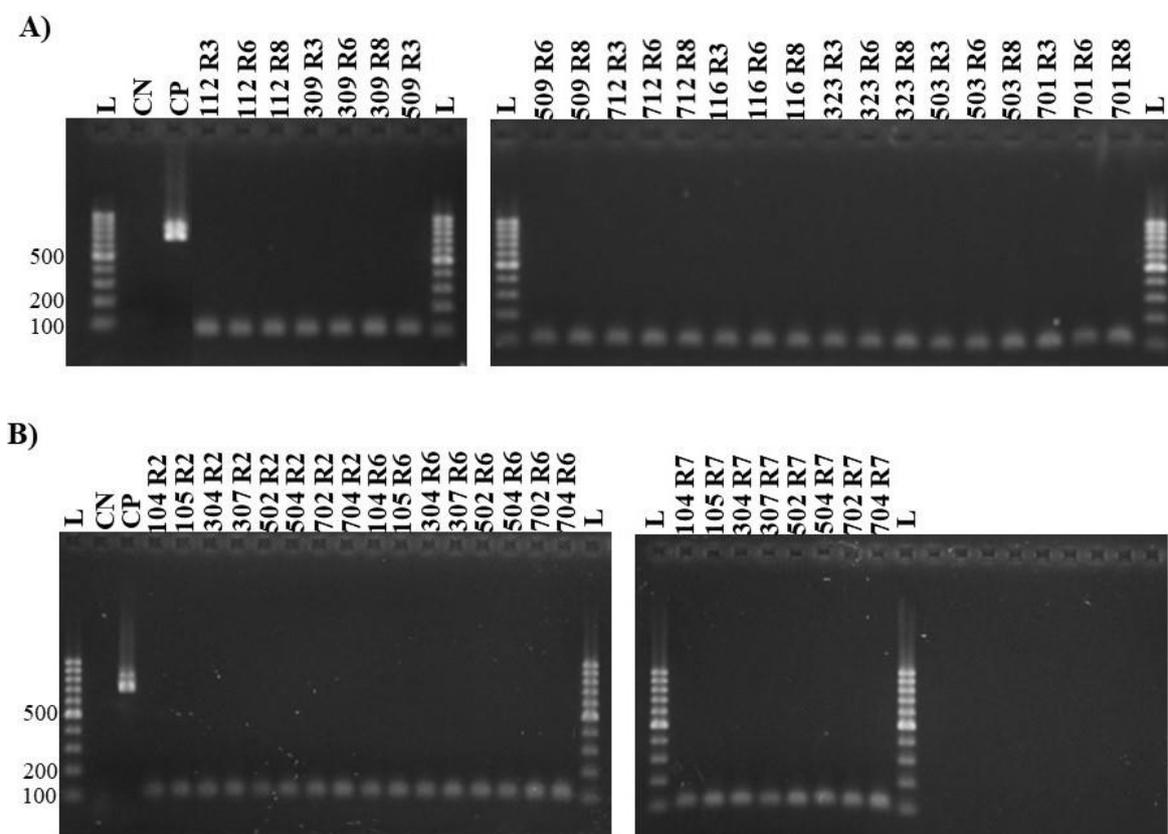


Figura 10. Eletroforese realizada para confirmar a ausência de gDNA nas amostras de cDNA. Gel de agarose a 2% corado com brometo de etídio. As bandas entre 100 pares de base (pb) e 200 pb correspondem à amplificação do primer *eEF-1 α* em amostras de cDNA. A) amostras da safra 2018/2019. B) amostras da safra 2019/2020. L) Marcador de peso molecular Ladder 100 Thermo Scientific. CN) Controle negativo com água substituindo o cDNA. CP) Controle positivo com DNA genômico de arroz.

Uma vez verificada a pureza das amostras, foram realizadas reações de otimização da qPCR para determinar a concentração ideal de cada *primer* a ser utilizada nas reações de quantificação. Essa análise foi feita a partir das curvas de dissociação, que foram realizadas após a ciclagem de temperatura. Portanto, foi possível escolher a concentração dos *primers* (Tabela 3) que, na curva de dissociação, não apresentou picos indicativos da presença de dímeros ou de amplificação inespecífica (Figura 11).

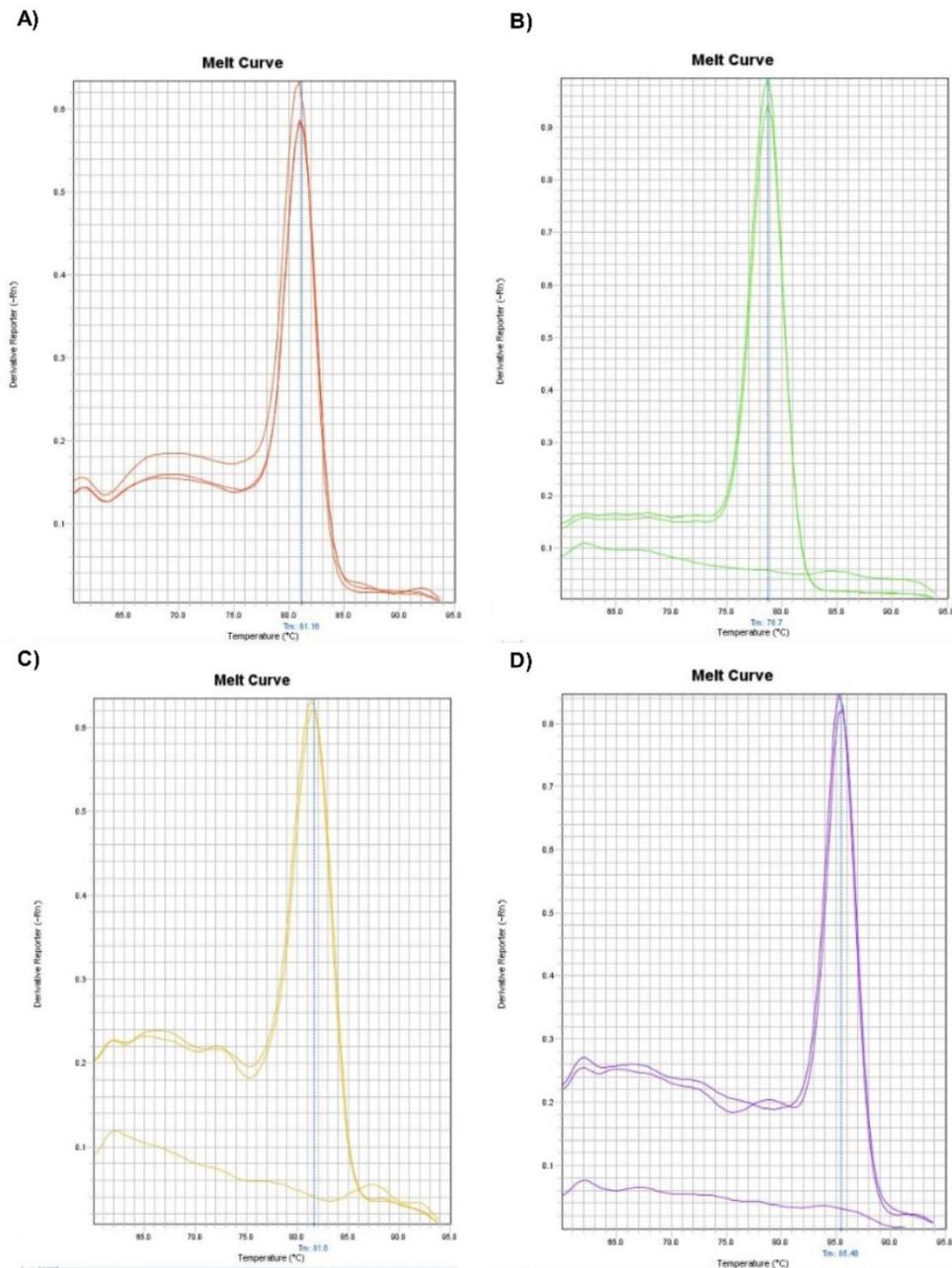


Figura 11. Curvas de dissociação dos *primers* utilizados nas reações de RT-qPCR: A) eEF1- α ; B) GAPDH; C) ACT; D) AVP_4. A presença de um único pico é um indicativo de que os pares de *primers* não anelaram entre si.

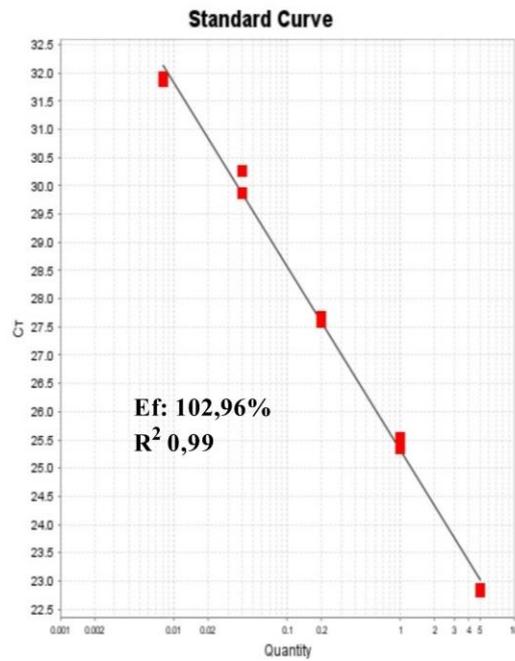
Tabela 3. Concentrações dos *primers* de referência e do gene *OVPI* escolhidas para serem utilizadas nas reações de RT-qPCR.

Identificação do <i>primer</i>	Concentração (nM)	
	<i>Forward</i>	<i>Reverse</i>
eEF1- α	600	600
GAPDH	300	300
ACT	600	600
AVP_4	300	300

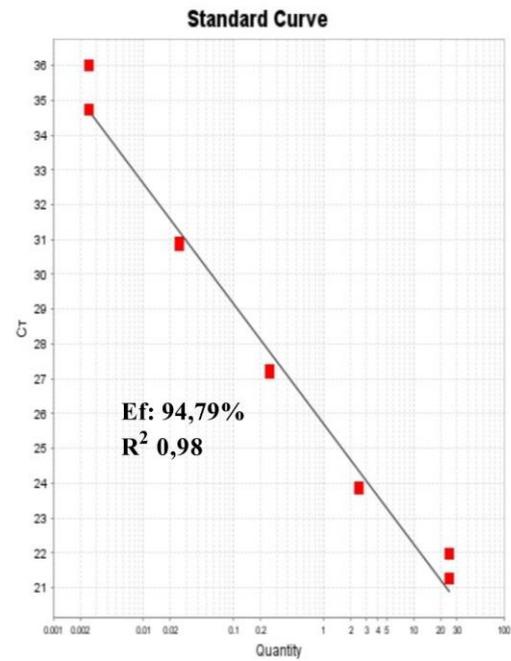
Em seguida, uma diluição seriada 1:2 com cinco pontos foi realizada para determinar a eficiência de amplificação dos *primers* usados neste trabalho. Os valores de eficiência dos *primers* de referência variaram de 94,79% a 102,96% (Tabela 4), ou seja, foram próximos ao valor esperado de 100%. Além disso, os coeficientes de determinação (r^2) foram acima de 0,98, indicando que a variação nos valores de Ct, na curva padrão (Figura 12), pode ser explicada pela variação na concentração do cDNA. Já os valores de eficiência dos *primers* desenhados para o gene *OVPI* variaram de 96,03% a 150,47%, com coeficientes de determinação entre 0,95 e 0,98 (Tabela 5). Dos quatro pares de *primers* desenhados para o *OVPI*, o AVP_4 foi escolhido para ser usado nas reações de quantificação da expressão gênica. Os critérios de escolha foram: não formação de dímeros nas concentrações testadas, coeficiente de determinação alto, eficiência próxima de 100% e menor Ct médio (Tabelas 5 e 6).

Após as reações de otimização e a escolha do *primer* do gene *OVPI*, foram realizadas as reações de quantificação utilizando os três genes de referência mencionados. Porém, para o cálculo da expressão relativa, apenas os dois genes que apresentaram maior estabilidade entre as amostras foram aplicados. Com base no programa NormFinder, os genes de referência recomendados para análise da expressão gênica do ensaio 2018/2019 foram o *eEF1- α* e o *ACT*, com valor de estabilidade combinada de 0,005; e para o experimento 2019/2020 foram o *eEF1- α* e o *GAPDH*, com valor de estabilidade combinada de 0,004. Por isso, eles foram usados como normalizadores das reações de quantificação da RT-qPCR, em seus respectivos ensaios.

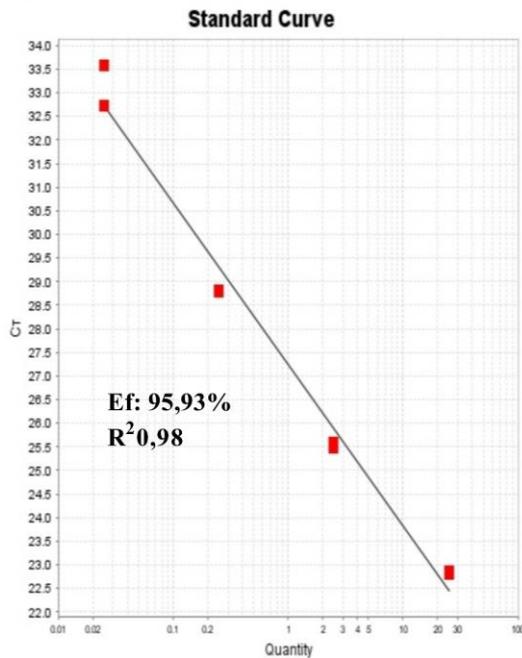
A)



B)



C)



D)

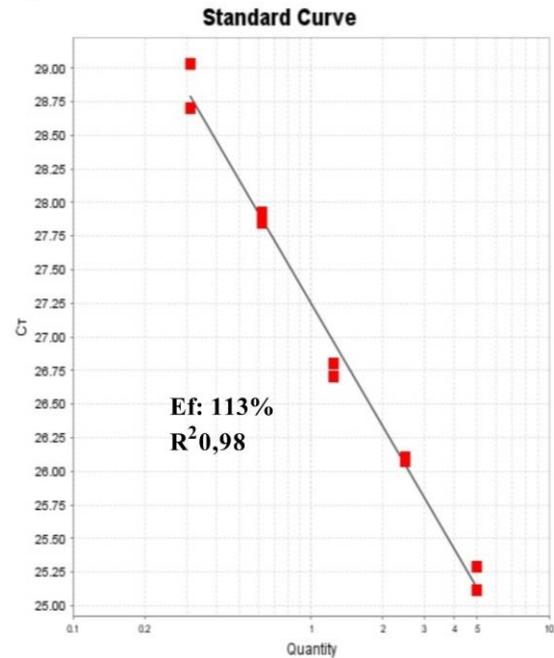


Figura 12. Curva padrão realizada para determinação da eficiência dos *primers* utilizados nas reações de quantificação da expressão do gene *OVPI*. A) eEF1- α ; B) GAPDH; C) ACT; D) AVP_4. Ef (eficiência do par de *primer*); R² (coeficiente de determinação).

Tabela 4. *Primers* de referência usados nas reações de RT-qPCR.

Identificação	<i>Primers</i>	Sequência F/R (5'→3')	<i>Amplicon</i> (pb) ^a	T _m (°C) ^b	r ²	Eficiência (%)
EEF1- α	F	TTTCACTCTTGGTGTGAAGCAGAT	103	81,35	0,996	102,96
	R	GACTTCCTTCACGATTTTCATCGTAA				
GAPDH	F	AAGCCAGCATCCTATGATCAGATT	79	78,43	0,982	94,79
	R	CGTAACCCAGAATACCCTTGAGTTT				
ACT	F	TGCGATAATGGAAGTGGTATGG	147	82,96	0,985	95,93
	R	ACAGCCCTGGGCGCAT				

^aTamanho esperado do *amplicon* em cDNA^bTemperatura de *melting* (dissociação) em °C.**Tabela 5.** *Primers* desenhados para quantificação da expressão do gene *OVPI*.

Identificação	<i>Primers</i>	Sequência F/R (5'→3')	<i>Amplicon</i> (pb) ^a	T _m (°C) ^b	r ²	Eficiência (%)	Dímeros ^c
AVP_1	F	CCGAAGGAGCAACTTCTTTC	103	78,55	0,954	150,47	+
	R	CCCTCAACAGATCCAAGGAA					
AVP_2	F	TGGCAACTTTTCTTGTGTGTG	115	82,07	0,986	117,42	+
	R	AGGAATCAGCAACGTCTTGG					
AVP_3	F	GGAAGGGTTTCGCCATTG	100	85,21	0,979	96,03	+
	R	GGTGTGAGAACATCGACAGTG					
AVP_4*	F	ACATTGAGGCTGGTGCATCT	101	85,88	0,989	113,52	-
	R	GAGAGGATCCCCAATGGTGT					

* *Primer* escolhido para quantificação do gene *OVPI*.^aTamanho esperado do *amplicon* em cDNA^bTemperatura de *melting* (dissociação) em °C.^cFormação de dímeros de *primers* nas reações de RT-qPCR (presente ou ausente).

Tabela 6. Valores de Ct dos *primers* desenhados para o gene *OVPI*.

<i>Primer ID</i>	Média (Ct)*	SD	Min (Ct)	Max (Ct)
AVP_1	25,08	1,16	23,9	27,09
AVP_2	26,43	0,96	25,44	27,43
AVP_3	26,36	0,38	25,95	26,74
AVP_4	24,53	0,67	23,85	25,29

*Média dos Cts obtidas a partir da realização de quatro qPCR por *primer*.

Negrito = média do Ct do par de *primer* escolhido para quantificação do gene *OVPI*.

No experimento realizado na safra 2018/2019, as plantas GM não apresentaram alterações significativas na expressão do gene *OVPI* em relação ao NGM durante a maioria das fases do desenvolvimento analisadas neste trabalho, tendo diferença significativa de expressão apenas no estágio R6. O evento OVP1-E4 expressou 3,65 vezes mais o gene *OVPI* do que as plantas NGM durante o enchimento dos grãos (Figura 13).

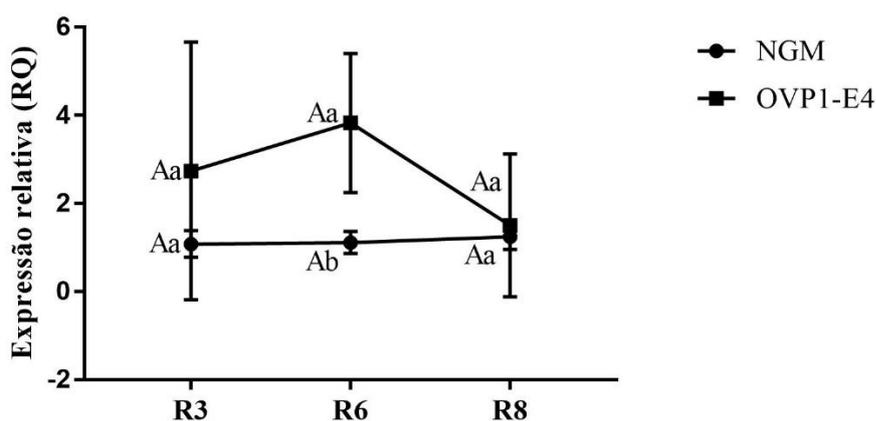


Figura 13. Expressão relativa do gene *OVPI* no ensaio 2018/2019. As letras minúsculas indicam diferença estatística entre os genótipos e as letras maiúsculas indicam diferença estatística entre as fases de desenvolvimento (n=4, p<0,05).

Já no ensaio realizado em 2019/2020, não houve diferença estatística entre os genótipos em nenhuma das fases analisadas (Apêndice E). Porém, na análise conjunta em R6, verificou-se uma diferença significativa (Apêndice F), tendo as plantas GM expressado o gene *OVPI* 1,35 vezes mais do que NGM de acordo com o teste de Tukey (p<0,05).

Além disso, a expressão do gene *OVPI* nas plantas GM do ensaio 2018/2019 apresentou uma correlação positiva com a taxa fotossintética ($r^2= 0,9638$; p<0,05) e com a eficiência de carboxilação ($r^2= 0,99$; p<0,05) durante o enchimento de grãos.

4.2.2 Análise do transcrito via RNAseq

As plantas GM e NGM mais produtivas do ensaio 2018/2019 foram analisadas em duplicata no estágio R6 via RNAseq, uma vez que foi nessa fase do desenvolvimento que houve maior expressão do gene *OVPI* e alterações fisiológicas significativas para o evento OVP1-E4.

O sequenciamento do transcrito das quatro bibliotecas gerou entre 41,94 e 82,62 milhões de *reads*. Em média, 93,80% das bases sequenciadas tiveram um valor de *Phred* maior que 30 (Q30), confirmando a qualidade do sequenciamento. Após a montagem do transcrito, estatísticas descritivas permitiram verificar que os *contigs* obtidos apresentaram uma média de 1199,89 bases e um N50 de 2370 bases. Quanto à representatividade, as bibliotecas tiveram mais de 99% das *reads* alinhadas no transcrito, indicando a qualidade da montagem usando o *software* Trinity. Além disso, a análise de completude mostrou que 74% dos genes têm mais de 50% de cobertura.

Uma vez verificada a qualidade do sequenciamento e da montagem do transcrito, o próximo passo foi a quantificação dos transcritos. Assim, a expressão dos genes foi quantificada utilizando o *software* RSEM em TPM (*Transcripts per Kilobase Million*). Em seguida, fez-se a análise de expressão diferencial usando o EdgeR. No total, foram identificados 126 genes diferencialmente expressos entre os dois genótipos. Desses, 51 genes foram positivamente regulados (*up-regulated*) no evento OVP1-E4 e 75 genes foram negativamente regulados (*down-regulated*). De acordo com o dendrograma (Figura 14), as réplicas de cada genótipo agruparam entre si e a maioria dos seus DEGs apresentaram um perfil de expressão semelhante.

Segundo a análise de expressão diferencial, o gene *OVPI* foi *up-regulated* nas plantas GM com um valor de *fold change* (\log_2) igual a 3,41 (Tabela 7). A classificação do *gene ontology* (GO:0009678) diz que este gene pertence à categoria FM (função molecular) e atua realizando o transporte de prótons por membranas por meio da hidrólise do pirofosfato inorgânico. Além disso, entre os genes que foram *up-regulated* nas plantas GM (Tabela 7), deve-se destacar aqueles envolvidos na fase fotoquímica (LOC_Os08g40160.1) e bioquímica (LOC_Os06g04270.1) da fotossíntese (Figura 15); no transporte de cátions (LOC_Os05g39540.1) e proteínas (LOC_Os09g28610.1); e no transporte de compostos tóxicos através da proteína de efluxo da família MATE (*Multidrug and toxic compound extrusion gene*, LOC_Os08g37432.2). Adicionalmente, dois genes envolvidos na biossíntese

de flavonol (LOC_Os09g07450.1 e LOC_Os10g41020.1), uma classe de compostos pertencentes aos flavonoides, também foram *up-regulated* nas plantas GM.

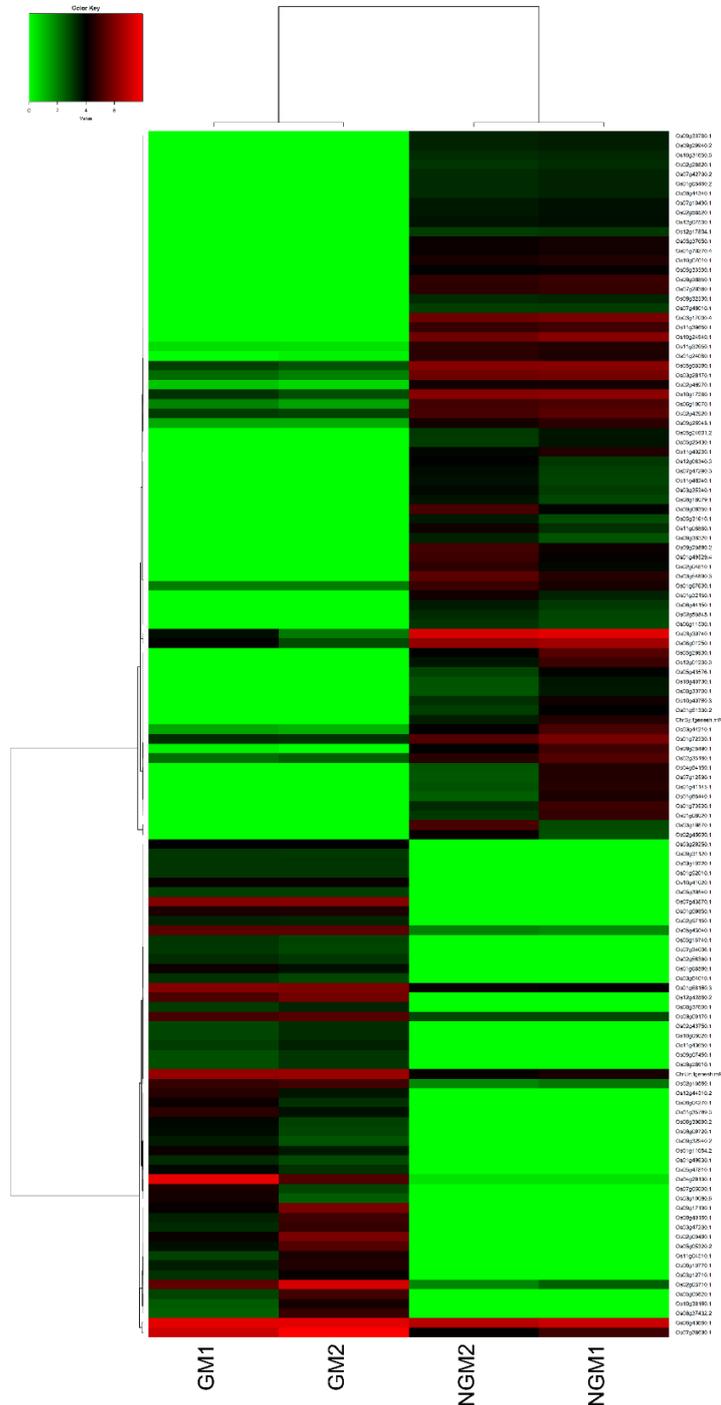


Figura 14. *Heatmap* representando o perfil de expressão das quatro bibliotecas para os genes diferencialmente expressos. GM1: réplica 1 do evento OVP1-E4. GM2: réplica 2 do evento OVP1-E4. NGM1: réplica 1 cultivar BRSMG Curinga não geneticamente modificada. NGM2: réplica 2 cultivar BRSMG Curinga não geneticamente modificada. Verde: genes *down-regulated*. Vermelho: genes *up-regulated*.

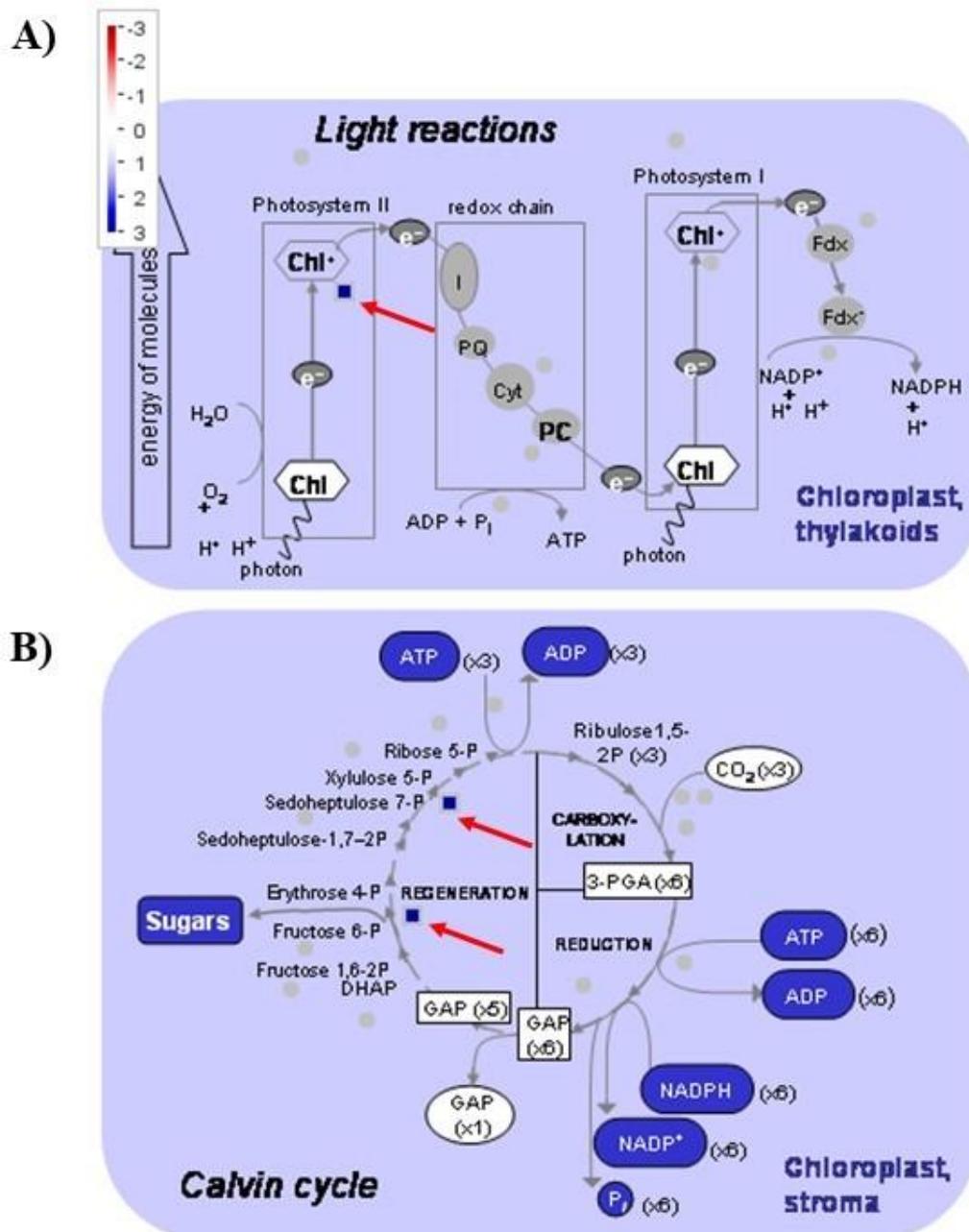


Figura 15. Genes pertencentes à via fotossintética que foram *up-regulated* nas plantas geneticamente modificadas (OVP1-E4). A: fase fotoquímica apresentando o gene LOC_Os08g40160.1, codificador de uma proteína estrutural do fotossistema II (PSII). B: fase bioquímica com o gene LOC_Os06g04270.1, codificador de uma transcetolase que atua na fase de regeneração da Ribulose 1,5-bisfosfato. As setas indicam os genes *up-regulated*.

Quanto aos fatores de transcrição que tiveram sua expressão aumentada no evento OVP1-E4, pode-se citar o fator de transcrição responsivo ao etileno (ERF, LOC_Os03g09170.1), o fator de transcrição da família bZIP (LOC_Os02g43750.1) e o fator de transcrição ZOS1-22 da família dos dedos de zinco C₂H₂ (LOC_Os01g68160.3). Já as

proteínas atuantes nas vias de transdução de sinais que foram regulados positivamente (Tabela 7) estiveram presentes dois F-BOX: OsFBX320 (LOC_Os09g17190.1) e OsFBX272 (LOC_Os08g09720.1); uma proteína com domínio SPX (LOC_Os03g29250.1); e uma proteína com domínio PPR (LOC_Os05g05320.2).

Diferentemente do que ocorreu com os genes que foram regulados positivamente, a maioria dos genes que foram *down-regulated* nas plantas GM, eram codificadores de fatores de transcrição, proteínas atuantes nas vias de transdução de sinais ou ainda não havia função definida (Tabela 8). Entre eles, pode-se citar os codificadores das proteínas quinases: CAMK_CAMK_like.16 (LOC_Os02g58520.1) e CAMK_KIN1/SNF1/Nim1_like_AMPKh.4-CAMK (LOC_Os08g37800.1); e os reguladores de resposta OsRR4 (LOC_Os01g72330.1) e OsRR2 (LOC_Os02g35180.1).

Adicionalmente, também foram negativamente regulados dois fatores de transcrição responsivos ao hormônio auxina (LOC_Os01g70270.4 e LOC_Os02g04810.1), o gene codificador da proteína promotora de crescimento independente de auxina (LOC_Os09g29940.2), e outros envolvidos no desenvolvimento da planta, como o LOC_Os09g25490.1 (CESA9 – celulose sintase 9).

Os genes diferencialmente expressos podem ser agrupados em categorias hierárquicas que permitem fazer sua anotação de acordo com a função molecular (FM), o processo biológico realizado (PB) e a localização celular (CC – componente celular) do produto codificado (Zhao et al., 2020). As categorias GO enriquecidas nas plantas GM foram aquelas envolvidas na homeostase e no transporte de íons através de membranas. Entre esses íons estão o cádmio, o zinco, o ferro e o manganês (Tabela 9). Também foram enriquecidas vias relacionadas com a desintoxicação celular e à resposta celular a estímulos externos. Já as categorias reprimidas (*depleted*) no evento OVP1-E4, foram em sua maioria aquelas relacionadas à regulação biológica (Tabela 10).

Tabela 7. Genes diferencialmente expressos que foram positivamente regulados (*up-regulated*) nas plantas geneticamente modificadas.

Função relacionada	Gene ID	Descrição*	log₂ (fold change)
Fotossíntese	LOC_Os06g04270.1	Transcetolase, precursor do cloroplasto	11,29
	LOC_Os08g40160.1	Proteína do lúmen do tilacoide (PsbP)	11,98
Metabolismo do carbono	LOC_Os01g11054.2	Fosfoenolpiruvato carboxilase	11,43
Transportadores	LOC_Os06g43660.1	Pirofosfatase H ⁺ inorgânica	3,41
	LOC_Os05g39540.1	Transportador de cátions de metal	10,43
	LOC_Os07g34006.1	Proteína da família de transportadores	10,44
	LOC_Os09g28610.1	Proteína de transporte de proteína	10,41
	LOC_Os08g37432.2	Proteína da família de efluxo MATE	11,62
Flavonoides	LOC_Os09g07450.1	Flavonol sintase	10,41
	LOC_Os10g41020.1	Flavonol sintase/flavonona 3-hidroxilase	11,66
Fatores de transcrição e vias de transdução de sinal	LOC_Os02g43750.1	Fator de iniciação de transcrição	10,54
	LOC_Os03g47200.1	Domínio de fator de transcrição bZIP	11,80
	LOC_Os01g59850.1	Proteína ativadora de GTPase	11,98
	LOC_Os03g29250.1	Proteína com domínio SPX	11,55
	LOC_Os09g17190.1	OsFBX320 – proteína com domínio F-box	12,85
	LOC_Os08g09720.1	OsFBX272 – proteína com domínio F-box	10,90
	LOC_Os05g05320.2	Domínio de repetição de PPR contendo proteína	12,30
	LOC_Os03g09170.1	Fator de transcrição responsivo ao etileno (ERF)	2,52
	LOC_Os01g68160.3	ZOS1-22 - proteína de dedo de zinco C ₂ H ₂	2,04
	LOC_Os01g49830.1	Domínio de ligação de DNA B3	10,53
	LOC_Os05g16740.1	SHR5 receptor de quinase	10,48

*Informação retirada do RGAP (Rice Genome Annotation Project, 2020)

Tabela 8. Genes diferencialmente expressos que foram negativamente regulados (*down-regulated*) nas plantas geneticamente modificadas.

Função relacionada	Gene ID	Descrição*	log₂ (fold change)
Fatores de transcrição e vias de transdução de sinal	LOC_Os01g49529.4	OsWAK10d - OsWAK cinase citoplasmática receptor semelhante ao OsWAK-RLCK	12,09
	LOC_Os02g58520.1	CAMK_CAMK_like.16 - CAMK inclui proteínas quinases dependentes de cálcio/calmodulina	11,14
	LOC_Os08g37800.1	CAMK_KIN1/SNF1/Nim1_like_AMPKh.4-CAMK quinases dependentes de cálcio/calmodulina	10,70
	LOC_Os01g73530.1	Transportador ABC, proteína de ligação de ATP	11,86
	LOC_Os02g46970.1	Domínio de ligação de AMP contendo proteína	4,71
	LOC_Os10g40780.3	Componente não complexo CCR4, proteína contendo domínio Not1	11,32
	LOC_Os12g06340.3	Fator de transcrição de homeodomínio semelhante a BEL1	11,07
	LOC_Os01g51300.2	Domínio WD, domínio de repetição G-beta contendo proteína	11,03
	LOC_Os05g25430.1	precursor de receptor de proteína quinase semelhante a At3g46290	10,85
	LOC_Os03g19870.1	Proteína de ligação de ATP	11,88
	LOC_Os01g72330.1	Regulador de resposta OsRR4 tipo A	2,45
	LOC_Os02g35180.1	Regulador de resposta OsRR2 tipo A	2,82
	LOC_Os02g45600.1	proteína contendo domínio rhoGAP	11,00
	LOC_Os07g12590.1	OsFBX225 - domínio F-box	11,33
Auxina	LOC_Os01g70270.4	Fator de resposta a auxina	11,74
	LOC_Os02g04810.1	Fator de resposta a auxina 5	11,83
	LOC_Os09g29940.2	Proteína promotora de crescimento independente de auxina	10,95
Desenvolvimento	LOC_Os05g29930.1	Proteína abundante de embriogênese tardia	12,32
	LOC_Os12g01200.3	Precursor de serina / treonina-proteína quinase semelhante ao receptor induzido por senescência	11,95
	LOC_Os09g25490.1	CESA9 - celulose sintase	8,39

*Informação retirada do RGAP (Rice Genome Annotation Project, 2020)

Tabela 9. Categorias GO enriquecidas nas plantas geneticamente modificadas.

GO	Descrição	termo GO*
GO:0006876	Homeostase celular de íon cádmio	PB
GO:0055073	Homeostase do íon cádmio	PB
GO:0015691	Transporte de íon cádmio	PB
GO:0070574	Transporte transmembrana de íon cádmio	PB
GO:0005385	Atividade do transportador transmembrana de íon zinco	FM
GO:0006882	Homeostase celular do íon zinco	PB
GO:0055069	Homeostase do íon zinco	PB
GO:0006879	Homeostase celular de íon de ferro	PB
GO:0005381	Atividade do transportador transmembrana de íons de ferro	FM
GO:0005384	Atividade do transportador transmembrana de íon manganês	FM
GO:1990748	Desintoxicação celular	PB
GO:0098869	Desintoxicação de oxidante celular	PB
GO:0004802	Atividade da transcetolase	FM
GO:0009535	Membrana de tilacoide de cloroplasto	CC
GO:0055035	Membrana plastídio tilacoide	CC
GO:0034357	Membrana fotossintética	CC
GO:0042651	Membrana tilacoide	CC
GO:0071496	Resposta celular a estímulos externos	PB
GO:0031669	Resposta celular aos níveis de nutrientes	PB
GO:0031667	Resposta aos níveis de nutrientes	PB

*PB (processo biológico), CC (componente celular), FM (função molecular).

Tabela 10. Categorias GO reprimidas (*depleted*) nas plantas geneticamente modificadas.

GO	Descrição	termo GO*
GO:0017076	Ligante de purina	FM
GO:0032555	Ligante de ribonucleotídeo de purina	FM
GO:0035639	Trifosfato ligante de ribonucleosídeo de purina	FM
GO:0019538	Processo metabólico de proteínas	PB
GO:0036094	Ligação de molécula pequena	FM
GO:0065007	Regulação biológica	PB
GO:1901564	Processo metabólico de compostos organonitrogênicos	PB
GO:0030554	Ligação de nucleotídeo de Adenil	FM
GO:0032559	Ligação adenil ribonucleotídeo	FM
GO:0005524	Ligante de ATP	FM
GO:0048518	Regulação positiva do processo biológico	PB
GO:0048856	Desenvolvimento da estrutura anatômica	PB
GO:0005515	Ligação proteica	FM
GO:0016787	Atividade de hidrolase	FM
GO:0032502	Processo de desenvolvimento	PB
GO:0050794	Regulação do processo biológico	PB
GO:0008152	Processo metabólico	PB
GO:0071704	Processo metabólico de substância orgânica	PB
GO:0044237	Processo metabólico celular	PB

*PB (processo biológico), CC (componente celular), FM (função molecular).

5 DISCUSSÃO

O aumento da produtividade de grão é o principal desafio dos programas de melhoramento genético do arroz do mundo todo. Apesar de ser um caráter quantitativo, a manipulação de genes específicos podem contribuir para obter cultivares com maior potencial produtivo. Neste trabalho, a superexpressão do gene *OVP1* na cultivar de arroz de terras altas BRSMG Curinga fez com que a quantidade de grãos totais (21%) e grãos cheios (31%) produzidos por panícula aumentasse significativamente, provavelmente devido à maior produção de fotoassimilados.

O aumento na taxa fotossintética também foi observado por outros autores, quando *AVP1* foi superexpresso em *A. thaliana* (Khadilkar et al., 2016), arroz irrigado (Kim et al., 2014), algodão (Pasapula et al., 2011) e cana-de-açúcar (Raza et al., 2016). Porém, diferentemente do que encontraram Pasapula et al. (2011) e Raza et al. (2016), em algodão e cana-de-açúcar, respectivamente, a superexpressão do gene *OVP1* no evento OVP1-E4 não apresentou modificações significativas na condutância estomática e na transpiração, indicando que ele foi capaz de melhor aproveitar o CO₂ assimilado mesmo sem ter ocorrido um aumento nas trocas gasosas, o que provavelmente foi devido ao aumento da eficiência de carboxilação. A eficiência da carboxilação é a razão entre a taxa fotossintética e a concentração interna de carbono. Sendo assim, altas taxas fotossintéticas combinadas com baixas concentrações internas de CO₂ indicam elevada eficiência de carboxilação.

A alta eficiência de carboxilação em conjunto com a maior taxa fotossintética durante o estágio R6 sugere que as plantas do evento OVP1-E4 tiveram maior capacidade de fixar carbono do que o NGM. Dessa forma, provavelmente proporcionaram maior quantidade de carbono fixado para o enchimento dos grãos e, como resultado, apresentaram uma maior quantidade de grãos cheios. Esse resultado não foi corroborado por Kim et. al (2014), que, ao verificarem o maior potencial produtivo de arroz superexpressando o *AVP1* no campo, encontraram plantas GM produzindo mais panículas, porém sem diferença significativa para o percentual de grãos cheios e o número de espiguetas.

Como observado para o evento OVP1-E4, Regmi et al. (2020), em pesquisa com trigo, obtiveram um aumento no número de grãos totais em plantas superexpressando o

AVPI, sendo que este aumento poderia estar relacionado à ação da auxina (Zhang et al., 2016b) durante a formação da panícula (estádio R3), ou seja, quando o número de espiguetas está sendo determinado. Isso porque as proteínas H^+ -PPases podem facilitar o deslocamento da auxina por meio da alteração do pH intracelular e extracelular, causando diversas modificações morfológicas como o aumento na área foliar em plantas de *A. thaliana* transformadas com o *AVPI* (Asaoka et al., 2016; Li et al., 2005) e com o comprimento da folha bandeira no evento OVP1-E4.

O aumento no comprimento da folha bandeira é interessante para plantas de arroz, pois esta folha é a principal fornecedora de fotoassimilados para a panícula e, por isso, exerce um importante papel na produção dos grãos (Al-Tahir, 2014; Rahman et al., 2013), o que ficou evidente devido a correlação positiva entre largura e comprimento da folha bandeira e a produção de grãos.

Em suma, o maior número de grãos cheios produzidos pelo evento OVP1-E4 provavelmente foi ocasionado pela combinação do maior comprimento da folha bandeira, da maior taxa fotossintética e da alta eficiência de carboxilação. Além disso, é um indicativo de que as plantas GM conseguiram fazer uma distribuição mais equilibrada dos fotoassimilados entre suas espiguetas, devido a ação das H^+ -PPases no floema em conjunto com as H^+ -ATPases e as bombas de SUC/H^+ , contribuindo para o aumento do transporte de açúcar para as panículas (Khadilkar et al., 2016; Pizzio et al., 2015; Primo et al., 2019; Scholz-Starke et al., 2019).

Adicionalmente, o evento OVP1-E4 foi mais eficiente no uso de água nos dois parâmetros analisados ($iWUE$ e WUE_{intr}), mesmo estando sob condições normais de irrigação. Esse resultado é um indicativo de que as plantas OVP1-E4 estarão mais aptas a suportarem um período com menor disponibilidade de água em épocas críticas, como na fase reprodutiva. A eficiência no uso de água é essencial para cultivares de arroz de terras altas, e como dependem principalmente da água das chuvas, é fundamental que possuam mecanismos que diminuam o impacto da ocorrência de estiagens na produtividade de grãos (Zhang et al., 2018).

Após as análises morfológicas e fisiológicas das plantas GM em relação às NGM, a expressão relativa do gene *OVP1* foi avaliada em alguns estádios da fase reprodutiva, com o objetivo de analisar a variação da expressão deste gene no período mais crítico para a produtividade de grãos. Durante o estágio R6, a maior expressão do gene *OVP1* exógeno (no OVP1-E4) em relação ao endógeno (no NGM) foi identificada tanto por RT-

qPCR quanto por RNAseq, confirmando o sucesso da superexpressão do gene *OVP1*. O promotor responsável pela superexpressão deste gene é constitutivo, ou seja, seu transcrito é continuamente expresso ao longo do ciclo da planta e em todos os tecidos. Contudo, provavelmente esse transcrito não é utilizado pelo arroz nos estádios onde ele não é necessário, o que explicaria por que sua expressão foi significativamente maior apenas em R6.

Para entender os mecanismos envolvidos no evento OVP1-E4 que o levaram a ter maior número de grãos cheios, fez-se o RNAseq dos dois genótipos avaliados (OVP1-E4 e NGM), e foram encontrados genes diferencialmente expressos e categorias GO enriquecidas relacionadas à fotossíntese, à homeostase de íons, à desintoxicação celular e à produção de flavonoides.

O gene LOC_Os08g40160.1, que foi *up-regulated* nas plantas GM, está presente no cromossomo 8 de *O. sativa* e é responsável por codificar uma proteína estrutural da família PsbP, componente do fotossistema II (PSII) (MSU Rice Genome Annotation Project, 2015; Mentewab et al., 2014; Quero et al., 2020; Zhang et al., 2016a). Essa proteína está envolvida na oxidação da molécula de água realizada pelo PSII que, por sua vez, tem como função retirar elétrons da água e doá-los para a clorofila A presente no centro de reação da luz, iniciando o transporte de elétrons (Mentewab et al., 2014). O ortólogo do LOC_Os08g40160.1 em *A. thaliana* (AT1G76450), foi *down-regulated* em plantas mutantes que apresentavam baixa taxa de crescimento (Zhang et al., 2016a) e, em milho, com taxa fotossintética reduzida passando por déficit hídrico (Zhang et al., 2018). Esses exemplos indicam que essa proteína tem um papel essencial para a realização da fotossíntese e, consequentemente, para o desenvolvimento normal das plantas.

Já o gene LOC_Os06g04270.1, que também foi *up-regulated* nas plantas OVP1-E4, codifica uma transcetolase que atua na fase de regeneração do Ciclo de Calvin-Benson em duas etapas: produzindo eritrose-4P a partir da frutose-6P; e xilulose-5P a partir de sedoetulose-7P (Rocha et al., 2014; UNIPROT, 2020). Essa rota pode estar relacionada com o fato das plantas OVP1-E4 terem apresentado maior eficiência de carboxilação, mesmo não tendo ocorrido maior condutância estomática, já que elas provavelmente tinham maior quantidade de ribulose-1,5 bifosfato regenerada disponível para fixação do CO₂. O ortólogo deste gene em sorgo (Sb06g004280) foi relatado, juntamente com outros genes do ciclo Calvin-Benson, como *up-regulated* em uma linhagem de sorgo granífero tolerante ao estresse salino, e que se caracterizou por acumular açúcares durante o estresse (Sui et al.,

2015), demonstrando a importância da transcetolase para a produção e acúmulo de carboidratos.

Como mencionado anteriormente, as plantas GM também apresentaram categorias GO enriquecidas relacionadas à desintoxicação celular. O transporte de moléculas para dentro de vacúolos e para fora da célula estão entre os mecanismos usados neste processo, que é essencial para a sobrevivência e desenvolvimento das plantas (Wang et al., 2016). No presente trabalho, um gene que codifica uma proteína da família MATE, que são enzimas responsáveis por diversos processos fisiológicos, incluindo o transporte de metabólitos secundários (Li et al., 2017; Wang et al., 2016), foi *up-regulated* no evento OVP1-E4, o que poderia ajudar no mecanismo de proteção dessas plantas. *Populus* spp., por exemplo, transformada para superexpressar um gene MATE, apresentou tolerância ao íon Al^{3+} , já que um dos mecanismos de eliminação do Al^{3+} é a extrusão deste íon em conjunto com ácidos orgânicos como o citrato, que, por sua vez, é realizado por uma proteína desta família (Li et al., 2017).

Em arroz, Huang et al. (2019) identificaram 55 genes da família proteica MATE, cujas proteínas codificadas atuam na desintoxicação celular de compostos tóxicos produzidos pelo próprio metabolismo ou por algum fator exógeno. Essas enzimas dependem de um gradiente eletroquímico de cátions (H^+ ou Na^+) para exercer sua atividade transportadora (Wang et al., 2016). De acordo com uma análise filogenética realizada por Wang et al. (2016), as proteínas MATE foram divididas em quatro subfamílias, sendo que o gene LOC_Os08g37432, que foi superexpresso no evento OVP1-E4, codifica uma proteína que faz parte do grupo MATE II. O grupo MATE II, por sua vez, se caracteriza por enzimas localizadas na membrana plasmática, e que usam os flavonoides como substratos, ou seja, atuam transportando flavonoides, que são metabólitos secundários.

De maneira geral, os metabólitos secundários atuam na defesa contra patógenos, na proteção contra radiação UV e na pigmentação dos órgãos reprodutores (Park et al., 2019; Upadhyay et al., 2019). Porém, em excesso, estes compostos podem se tornar tóxicos para a célula (Upadhyay et al., 2019) e como observado nos resultados do presente trabalho, dois genes de biossíntese de flavonoides foram positivamente regulados no evento OVP1-E4, o que é um indicativo de que as plantas GM procuraram uma forma de se proteger do excesso de flavonoides que provavelmente estavam produzindo a partir das proteínas transportadoras MATE.

Entretanto, um detalhe interessante é que os dois genes que foram superexpressos codificam enzimas que levam à produção de uma classe específica desses metabólitos: a classe do flavonol. Um deles (LOC_Os10g41020.1) codifica a enzima F3H (flavanona-3-hidroxilase) que produz diidroflavonol a partir de flavanonas, e o outro gene (LOC_Os09g07450.1) codifica a enzima FLS (flavonol sintase), que catalisa a produção do flavonol a partir do diidroflavonol. A proteína flavanona-3-hidroxilase é relatada como uma enzima-chave para a regulação do acúmulo de flavonoides nos vegetais, sendo parte também da via de biossíntese de antocianinas, por exemplo (Han et al., 2017; Park et al., 2019). Em Coníferas, esta enzima foi essencial para a defesa contra o fungo *Endoconidiophora polonica*, que normalmente infecta as plantas após o ataque do besouro da casca (Hammerbacher et al., 2019).

Quanto aos fatores de transcrição que tiveram expressão alterada, foi relatado por Zhu et al. (2003) que fatores de transcrição do tipo bZIP são altamente expressos no endosperma do arroz, podendo ser importantes para o desenvolvimento dos grãos. No presente trabalho, foi observado que um gene codificador de um fator de transcrição bZIP também foi *up-regulated* nas folhas das plantas geneticamente modificadas durante o estágio R6, dando suporte à hipótese de que estes fatores de transcrição podem ser importantes para o desenvolvimento dos grãos de arroz.

O gene LOC_Os03g09170.1, que codifica um dos fatores de transcrição de resposta ao etileno (ERF – *Ethylene Response Factor*), também foi positivamente regulado nas plantas GM. Entre outras funções, o etileno atua como um mediador para resposta à diversos estresses, porém o fator de transcrição ERF também pode ser regulado por outros hormônios vegetais como os ácidos abscísico e salicílico (Muller & Munné-Bosch, 2015; Pegoraro et al., 2013). Em arroz, o gene LOC_Os03g09170.1, foi *up-regulated* na presença do patógeno *Magnaporthe grisea* e *down-regulated* quando em condições de anoxia (Pegoraro et al., 2013).

Dentre os genes que tiveram expressão alterada nas plantas GM e estão envolvidos nas vias de transdução de sinais, está o codificador do domínio proteico SPX. O domínio proteico SPX foi *up-regulated* no evento OVPI-E4, o qual é essencial para manter a homeostase de fosfato inorgânico (P_i) a nível celular (Secco et al., 2012) e que, normalmente, é responsivo à falta do P_i (Shi et al., 2014). As proteínas SPX são divididas em quatro classes de acordo com a presença ou ausência de domínios proteicos adicionais (Secco et al., 2012; Shi et al., 2013). Neste trabalho, a SPX positivamente regulada foi a

OsSPX5, que faz parte da classe I por apresentar apenas o domínio SPX, e além disso possui um parálogo que é a OsSPX3 (Secco et al., 2012; Shi et al., 2014). Esses dois domínios proteicos podem formar homodímeros ou heterodímeros quando as plantas são expostas às mudanças externas na concentração de P_i , alterando o transporte desta molécula da raiz para a parte aérea ao reprimir o fator de transcrição PHR2 (*Protein Phosphate Starvation Response 2*), que tem papel central na regulação da homeostase do P_i (Secco et al., 2012; Shi et al., 2014). De acordo com Shi et al. (2014), a repressão duplicada do PHR2 pelas proteínas OsSPX3/OsSPX5 aumenta a sinalização que indica a falta do P_i . Assim, nas plantas OVP1-E4 esta sinalização pode estar ainda mais intensa devido à indução do gene codificador da OsSPX5.

A regulação positiva dos domínios SPX nas plantas OVP1-E4 é interessante, porque outros autores vêm relatando que a superexpressão de órtologos do gene *OVP1* proporciona tolerância às plantas cultivadas sob baixo teor de fosfato (Gaxiola et al., 2012; Pei et al., 2012; Yang et al., 2007; Yang et al., 2014). Portanto, uma hipótese para explicar essa tolerância é a alteração da atividade dos domínios proteicos SPX.

Outra classe de genes que tiveram expressão alterada no evento OVP1-E4 foram os codificadores das proteínas F-BOX, que são importantes porque atuam na regulação de várias vias metabólicas em plantas após o processo de tradução. Isso porque, para a degradação de proteínas, é utilizado um sistema conhecido como Ubiquitina/26S Proteassoma, do qual fazem parte as proteínas F-BOX (Jain et al., 2007), que atuam conferindo especificidade para marcação das proteínas-alvo que serão degradadas. Neste caso, as plantas GM regularam positivamente a expressão da OsFBX320 (LOC_Os09g17190.1) e OsFBX272 (LOC_Os08g09720.1), e ao mesmo tempo regularam negativamente a expressão da OsFBX225 (LOC_Os07g12590.1).

Quanto aos genes que foram negativamente regulados no evento OVP1-E4, durante o enchimento dos grãos, deve-se ressaltar aqueles relacionados ao desenvolvimento das plantas. Por exemplo, dois genes envolvidos na resposta celular ao hormônio auxina (LOC_Os01g70270.4 e LOC_Os02g04810.1) que codificam as proteínas ARF4 e ARF5 (*Auxin Response Factor 4 and 5*) merecem destaque por serem fatores de transcrição que se ligam nos promotores AuxREs (*Auxin-responsive promoter elements*) para regular a expressão de genes envolvidos no crescimento e desenvolvimento vegetal (Li et al., 2016), podendo indicar que está ocorrendo uma alteração na resposta celular a este hormônio durante o enchimento dos grãos.

O gene que codifica a Celulose sintase 9 (Cesa9, LOC_Os09g25490.1) também foi *down-regulated* nas plantas GM. Esta proteína, em conjunto com a Cesa4 e a Cesa7, formam um complexo proteico funcional para sintetizar a celulose que irá compor a parede celular secundária do arroz (Li et al., 2017; Wang et al., 2012). A parede celular secundária se forma ao redor de algumas células especializadas após o término do crescimento celular para fornecer resistência mecânica e proteção (Lin et al., 2016; Zhong et al., 2019). Ela é composta principalmente por celulose, hemicelulose e lignina (Zhong et al., 2019). Portanto, a alteração da atividade de alguns dos genes que codificam as proteínas celulosas sintases, pode afetar a estrutura das plantas de diferentes formas.

O arroz mutante para esse gene, chamado de Osfc16 (Li et al., 2017), que apresenta dois resíduos de aminoácidos diferentes dos que são originalmente conservados em genes Cesa, apresentou uma redução na deposição de celulose na parede secundária das plantas, o que levou a uma parede celular secundária mais fina, além de terem tido hastes com diâmetro relativamente menor. Porém, apesar das plantas terem apresentado menor altura, tiveram um maior número de perfilhos (Li et al., 2017). Outra mutação no gene Cesa9 levou a um fenótipo chamado de colmo quebradiço (Wang et al., 2012). Já no presente trabalho, apesar de a regulação negativa do gene Cesa9 nas plantas GM ter sido detectada na análise de RNAseq, não foi observada diferença significativa na massa seca e na altura entre o evento OVP1-E4 e as plantas NGM (dados não mostrados).

6 CONCLUSÕES

A superexpressão do gene *OVP1* em arroz de terras altas resultou no aumento do número de grãos cheios e do número de grãos totais produzidos por panícula, o que pode resultar no aumento do potencial produtivo da cultura quando em cultivo comercial. Com base na expressão relativa, observou-se que a expressão do gene *OVP1* foi significativamente maior nas plantas GM apenas no estágio R6. Logo, é possível concluir que este gene exerce uma função importante durante o enchimento dos grãos.

Entre os mecanismos identificados como relacionados ao incremento no número de grãos cheios, estão o aumento da taxa fotossintética, da eficiência de carboxilação e do uso eficiente da água, durante a fase do enchimento de grãos (R6). A alta taxa fotossintética e alta eficiência de carboxilação provavelmente foram ocasionadas pela indução da expressão de dois genes envolvidos na fase fotoquímica e bioquímica da fotossíntese.

Adicionalmente, segundo a análise do transcrito via RNAseq, o evento OVP1-E4 aumentou a expressão de genes envolvidos na biossíntese de flavonoides, na via de transdução de sinais que responde à falta de P_i e no transporte de íons e moléculas tóxicas, ao mesmo tempo em que reduziu a expressão de fatores de transcrição responsivos a auxina. Esses resultados abrem caminho para o desenvolvimento de plantas de arroz mais produtivas e fisiologicamente mais eficientes.

7 PERSPECTIVAS

Apesar de as plantas OVP1-E4 terem apresentado maior quantidade de grãos por panícula, a produção de grãos não foi estatisticamente diferente das plantas NGM. Por isso, outros experimentos serão realizados, em parcelas com maior número de plantas, para avaliar realisticamente a produtividade de grãos do evento OVP1-E4 em comparação ao *wild-type* BRSMG Curinga. Além disso, estas plantas GM serão avaliadas em solos com baixo teor de fósforo, com o objetivo de analisar se irá ocorrer a regulação de genes envolvidos na homeostase celular do P_i , que já foi observada em condições normais de fertilidade do solo, e se esta regulação será capaz de fazer com que as plantas OVP1-E4 sejam mais tolerantes ao solo com baixa disponibilidade de P_i .

REFERÊNCIAS

- ABIARROZ. Perfil da Produção. Disponível em: <<http://brazilianrice.com.br/br/sobre-o-brasil/>>. Acesso em: 11 abr. 2019.
- AHMED, N. et al. Effect of high temperature on grain filling period, yield, amylose content and activity of starch biosynthesis enzymes in endosperm of basmati rice. **Journal of the Science of Food and Agriculture**, v. 95, n. 11, p. 2237–2243, 2015.
- AL-TAHIR, M. M. F. Flag leaf characteristics and relationship with grain yield and grain protein percentage for three cereals. **Journal of Medicinal Plants Studies JMPS**, v. 1, n. 25, p. 1–7, 2014.
- ALTPETER, F. et al. Advancing crop transformation in the era of genome editing. **The Plant Cell**, v. 28, 2016.
- ARNHOLD, E. Easyanova package. Analysis of variance and other important complementary analyses. 2019.
- ASAOKA, M. et al. Contribution of PP_i-hydrolyzing function of vacuolar H⁺-pyrophosphatase in vegetative growth of *Arabidopsis*: Evidenced by expression of uncoupling mutated enzymes. **Frontiers in Plant Science**, v. 7, p. 1–12, 2016.
- ASHIKARI, M. Cytokinin oxidase regulates rice grain production. **Science**, v. 309, n.5735, 2005.
- BATEMAN, A. UniProt: A worldwide hub of protein knowledge. **Nucleic Acids Research**, v. 47, n. D1, p. D506–D515, 2019. Oxford University Press. Disponível em: <<https://www.ebi.ac.uk/proteins/api/doc/>>. Acesso em: 25 agost. 2020.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014. Oxford University Press. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/24568602/>>. Acesso em: 25 agost. 2020.
- BUSTIN, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. **Clinical Chemistry**, v. 55, n. 4, p. 611–622, 2009.
- CARBON, S. et al. AmiGO: online access to ontology and annotation data. **Bioinformatics Applications Note**, v. 25, n. 2, p. 288–289, 2009. Disponível em: <<http://www.geneontology.org>>. Acesso em: 25 agost. 2020.

CARVALHO, M. T. M. et al. O Arroz de terras altas como estratégia para segurança alimentar, intensificação ecológica e adaptação à mudança do clima : rumo aos objetivos. **Boletim de Pesquisa e Desenvolvimento**, v. 1, n. 252, p. 14, 2020.

CHAUDHARY, G; KUMAR, S. S. Global Status of Genetically Modified Crops and its commercialization. Springer Nature Switzerland. **Biotechnol. Prod. everyday life**. New Delhi: Springer International Publishing, 2019. p. 119–129. Disponível em: <<http://link.springer.com/10.1007/978-3-319-92399-4>>.

CHOURA, M.; REBAĬ, A. Identification and characterization of new members of vacuolar H⁺-pyrophosphatase family from *Oryza sativa* genome. **Russian Journal of Plant Physiology**, v. 52, n. 6, p. 821–825, 2005.

CONAB. A cultura do arroz. Brasília, 2015.

CONAB. Acompanhamento da safra brasileira de grãos. Brasília, 2019. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 9 jun. 2019.

CONAB. Acompanhamento da safra brasileira - sétimo levantamento. v.7, n.7, 2020.

COUNCE, P. A.; KEISLING, T. C.; MITCHELL, A. J. A uniform, objectives, and adaptive system for expressing rice development. **Crop Science**, v. 40, n. 2, p. 436–443, 2000.

DEDICOVA, B. et al. High-throughput transformation pipeline for a Brazilian japonica rice with bar gene selection. **Protoplasma**, v. 252, n. 4, p. 1071–1083, 2015.

DESSIMOZ, C. et al. Toward community standards in the quest for orthologs. **Bioinformatics**, v. 28, n. 6, p. 900–904, 2012.

DOBIN, Alexander et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**, v. 29, n. 1, p. 15–21, 2013.

ESMAELI, N. et al. Co-overexpression of *AVP1* and *OsSIZ1* in *Arabidopsis* substantially enhances plant tolerance to drought, salt, and heat stresses. **Scientific Reports**, v. 9, n. 1, p. 1–15, 2019.

FAGERIA, N. K.; WANDER, A. E.; SILVA, S. C. Rice (*Oryza sativa*) cultivation in Brazil. **Indian Journal of Agronomy**, v. 59, n. 3, p. 350–358, 2014.

FAN, C. et al. *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. **Theor Appl. Genet**, v.112, n.6, 2006.

FAO. **The state of food security and nutrition in the world**. Rome, 2020.

FAOSTAT. Top 10 Country Production of Rice. Disponível em: <http://www.fao.org/faostat/en/#rankings/countries_by_commodity>. Acesso em: 24 mar. 2019.

FERREIRA, C. M. et al. Qualidade do arroz no Brasil: evolução e padronização. **Embrapa Arroz e Feijão**, 2005.

FERREIRA, E. B.; CAVALCANTI, P. P.; NOGUEIRA, D. A. **ExpDes.pt: An R Package for ANOVA and Experimental Designs**. 2018.

FIAZ, S. et al. Applications of the CRISPR/Cas9 System for rice grain quality improvement: perspectives and opportunities. **International journal of molecular sciences**, v. 20, n. 4, p. 1–18, 2019.

FINN, R. D. et al. Pfam: The protein families database. **Nucleic Acids Research**, 2014. Oxford University Press. Disponível em: <[/pmc/articles/PMC3965110/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/24762582/)>. Acesso em: 25 agost. 2020.

GARRIS, A. J. et al. Genetic structure and diversity in *Oryza sativa* L. **Genetics**, v. 1638, p. 1631–1638, 2005.

GAXIOLA, R. A. et al. Drought- and salt-tolerant plants result from overexpression of the AVP1 H⁺-pump. **Proceedings of the National Academy of Sciences**, v. 98, n. 20, p. 11444–11449, 2001.

GAXIOLA, R. A. et al. Genetic manipulation of a “vacuolar” H⁺-PPase: from salt tolerance to yield enhancement under phosphorus-deficient soils. **Plant Physiology**, v. 159, n. 1, p. 3–11, 2012.

GAXIOLA, R. A.; REGMI, K.; HIRSCHI, K. D. Moving on up: H⁺-PPase mediated crop improvement. **Trends in Biotechnology**, v. 34, n. 5, p. 347–349, 2016.

GITHUB. TransDecoder/TransDecoder Wiki. GitHub. Disponível em: <<https://github.com/TransDecoder/TransDecoder/wiki>>. Acesso em: 25 agost. 2020.

GITHUB. Trinotate Functional Annotation · griffithlab/rnaseq_tutorial Wiki · GitHub. Disponível em: <https://github.com/griffithlab/rnaseq_tutorial/wiki/Trinotate-Functional-Annotation>. Acesso em: 25 agost. 2020.

GLOVER, N. M.; REDESTIG, H.; DESSIMOZ, C. Homoeologs: what are they and how do we infer them? **Trends in Plant Science**, v. 21, n. 7, p. 609–621, 2016.

GRABHERR, M. G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. **Nature Biotechnology**, v. 29, n. 7, p. 644–652, 2013.

HAAS, B. J. et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. **Nature Protocols**, v. 8, n. 8, p. 1494–1512, 2013. Nature Publishing Group. Disponível em: <<https://www.nature.com/articles/nprot.2013.084>>. Acesso em: 25 agost. 2020.

HAMMERBACHER, A. et al. Flavanone-3-hydroxylase plays an important role in the biosynthesis of spruce phenolic defenses against bark beetles and their fungal associates. **Frontiers in Plant Science**, v. 10, n. February, p. 1–15, 2019.

HAN, Y. et al. Functional analysis of two flavanone-3-hydroxylase genes from *Camellia sinensis*: a critical role in flavonoid accumulation. **Genes**, v. 8, n. 11, 1 nov. 2017. Disponível em: <[pmc/articles/PMC5704213/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/30111113/)>. Acesso em: 6 jan. 2021.

HARRELL, F. E. J. Package “Hmisc”. Disponível em: <<https://hbiostat.org/R/Hmisc/>>. Acesso em: 24 set. 2020.

HENG, Y. et al. *OsALMT7* maintains panicle size and grain yield in rice by mediating malate transport. **The Plant Cell**, v.30: 889-906, 2018.

HUANG, J. J. et al. Expression profile analysis of *MATE* gene family in rice. **Biologia Plantarum**, v. 63, n. January, p. 556–564, 2019.

HUO, X., et al. *NOG1* increases grain production in rice. **Nature Communications**, v.8(1), 2017.

IBGE. Área plantada, área colhida e produção de arroz no Brasil. Disponível em: <<https://sidra.ibge.gov.br/tabela/1618#resultado>>. Acesso em: 27 jun. 2018.

IMIN, N. et al. Low temperature treatment at the young microspore stage induces protein changes in rice anthers. **Molecular & Cellular Proteomics**, Rockville, v. 5, n. 2, p. 274-230, 2006.

ISAAA. **Rice(Oryza sativa L.) GM Events | GM Approval Database**. Disponível em: <<https://www.isaaa.org/gmapprovaldatabase/crop/default.asp?CropID=17&Crop=Rice>>. Acesso em: 18 jan. 2021.

JAIN, M. et al. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. **Biochemical and Biophysical Research Communications**, Philadelphia, v. 345, n.2, p. 646-651, 2006.

JAIN, M. et al. F-Box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. **Plant Physiology**, v. 143, n. 4, p. 1467–1483, 2007.

KANEHISA, M. Toward understanding the origin and evolution of cellular organisms.

Protein Science, 1. 2019. Blackwell Publishing Ltd. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31441146/>>. Acesso em: 25 agost. 2020.

KANEHISA, M.; GOTO, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, 1. 2000. Oxford University Press. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/10592173/>>. Acesso em: 25 agost 2020.

KANEHISA, M. et al. New approach for understanding genome variations in KEGG. **Nucleic Acids Research**, v. 47, n. D1, p. D590–D595, 2019. Oxford University Press. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30321428/>>. Acesso em: 25 agost. 2020.

KHADILKAR, A. S. et al. Constitutive and companion cell-specific overexpression of *AVPI*, encoding a proton-pumping pyrophosphatase, enhances biomass accumulation, phloem loading, and long-distance transport. **Plant Physiology**, v. 170, n. 1, p. 401–414, 2016.

KIM, Y. S. et al. Overexpression of the *Arabidopsis* vacuolar H⁺-pyrophosphatase *AVPI* gene in rice plants improves grain yield under paddy field conditions. **Journal of Agricultural Science**, v. 152, n. 6, p. 941–953, 2014.

KIM, B.R. et a. Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. **Biotechnology Letters**, Heidelberg, v. 25, n. 21, p. 1869-1872, 2003.

KOZERA, B.; RAPACZ, M. Reference genes in real-time PCR. **Plant Genetics Review**, v. 2013, 2013.

LANGMEAD, B. et al. Scaling read aligners to hundreds of threads on general-purpose processors. **Bioinformatics**, v. 35, n. 3, p. 421–432, 2019. Oxford University Press. Disponível em: <<http://bowtie-bio.sourceforge.net/bowtie2.HISAT:http://www.ccb.jhu.edu/software/hisat>>. Acesso em: 25 agost. 2020.

LAWSON, T.; BLATT, M. R. Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. **Plant Physiology**, v. 164, n. 4, p. 1556–1570, 2014.

LI, B.; DEWEY, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 1, p. 323, 2011. BioMed Central. Disponível em: <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-323>>. Acesso em: 25 agost. 2020.

- LI, D. et al. Integrated analysis of phenome, genome, and transcriptome of hybrid rice uncovered multiple heterosis-related loci for yield increase. **Proceedings of the National Academy of Sciences**, v. 113, n. 41, 2016.
- LI, F. et al. *OsCESA9* conserved-site mutation leads to largely enhanced plant lodging resistance and biomass enzymatic saccharification by reducing cellulose DP and crystallinity in rice. **Plant Biotechnology Journal**, v. 15, n. 9, p. 1093–1104, 1 set. 2017.
- LI, J. et al. *Arabidopsis* H⁺-PPase AVP1 regulates auxin-mediated organ development. **Science**, v. 310, n. 5745, p. 121–125, 2005.
- LI, N. et al. Control of grain size in rice. **Plant Reproduction**, v. 31, n. 3, p. 237–251, 2018.
- LI, N. et al. Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in *Populus*. **Journal of Experimental Botany**, v. 68, n. 20, p. 5669–5683, 2017.
- LI, S. B. et al. A review of auxin response factors (ARFs) in plants. **Frontiers in Plant Science**, v. 7, p. 1–7, 2016.
- LIN, F. et al. Cell wall composition and candidate biosynthesis gene expression during rice development. **Plant and Cell Physiology**, v. 57, n. 10, p. 2058–2075, 2016.
- LIN, S. M. et al. Crystal structure of a membrane-embedded H⁺-translocating pyrophosphatase. **Nature**, v. 484, n. 7394, p. 399–403, 2012.
- LIU, J. et al. Genome-Scale Sequence Disruption Following Biolistic Transformation in Rice and Maize. **The Plant Cell**, v. 31, n. 2, p. 368–383, 2019.
- LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. **Methods San Diego Calif.** v. 25, n. 4, p. 402–8, 2001.
- MA'ARUP, R. et al. Emmer wheat (*Triticum dicoccon* Schrank) improves water use efficiency and yield of hexaploid bread wheat. **Plant Science**, v. 295, p. 1–10, 2020.
- MAPA. Sumário Executivo Arroz setembro - 2020. **Ministério da Agricultura, Pecuária e Abastecimento**, 2020a.
- MAPA. Arroz vai continuar nas prateleiras, diz Tereza Cristina. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/noticias/arroz-vai-continuar-nas-prateleiras-diz-tereza-cristina>>. Acesso em: 19 set. 2020b.
- MARTIN, J. A.; WANG, Z. Next-generation transcriptome assembly. **Nature Reviews Genetics**, v. 12, n. 10, p. 671–682, 2011.

MELLO et al. Estimativa da capacidade de campo baseada no ponto de inflexão da curva característica. **Ciência e Agrotecnologia**, v. 26, n. 4, p. 836–841, 2002.

MENDIBURU, F. Package “agricolae” Statistical Procedures for Agricultural Research. 2020.

MI, H. et al. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. **Nucleic Acids Research**, v. 47, n. D1, p. D419–D426, 2019. Oxford University Press. Disponível em: <<http://geneontology>>. Acesso em: 25 agost. 2020.

MIURA, K. et al. *OsPL14* promotes panicle branching and higher grain productivity in rice. **Nature Genetics**, v.42, n.6, 2010.

MORAIS, O. P. De et al. BRSMG Curinga: cultivar de arroz de terras altas de ampla adaptacão para o Brasil. **Embrapa Arroz e Feijão. Comunicado Técnico**, v. 114, 2005.

MORGANTE, C. V.; BLAWID, R. Análise da expressão gênica pela técnica de PCR quantitativa em tempo real: Princípios e Fundamentos. Petrolina, 2016.

MSU RICE GENOME ANNOTATION PROJECT TEAM. Rice genome annotation project: LOC_Os08g40160.1. **Http://Rice.Plantbiology.Msu.Edu/**, 2015. Disponível em: <http://rice.plantbiology.msu.edu/cgi-bin/ORF_infopage.cgi>. Acesso em: 12 set. 2020.

MUELLER, O.; LIGHTFOOT, S.; SCHROEDER, A. RNA Integrity Number (RIN) – Standardization of RNA Quality Control Application. **Agilent Tchenologies**, p. 1–8, 2004.

MÜLLER, M.; MUNNÉ-BOSCH, S. Ethylene response factors: A key regulatory hub in hormone and stress signaling. **Plant Physiology**, v. 169, n. 1, p. 32–41, 2015.

MURCHIE, E. H. et al. Are there associations between grain-filling rate and photosynthesis in the flag leaves of field-grown rice? **Journal of Experimental Botany**, v. 53, n. 378, p. 2217–2224, 2002.

MUTHAYYA, S. et al. An overview of global rice production, supply, trade, and consumption. **Annals of the New York Academy of Sciences**, v. 1324, n. 1, p. 7–14, 2014.

NADIR, S. et al. An overview on reproductive isolation in *Oryza sativa* complex . **AoB PLANTS**, v. 10, n. 6, p. 1–13, 2018.

NELSON, D. L.; COX, M. M. **Lehninger Principles of Biochemistry**. 6. ed. New York: W.H.FREEMAN AND COMPANY, v. 17, 2013.

NEUWIRTH, Erich. ColorBrewer Palettes [R package RColorBrewer version 1.1-2], 2014. Disponível em: <<https://cran.r-project.org/package=RColorBrewer>>. Acesso em: 7 jan. 2021.

- NOLAN, T.; HUGGETT, J.; SANCHEZ, E. Good practice guide for the application of quantitative PCR (qPCR). **National Measurement System**, 2013.
- OKAMURA, M. et al. Characterization of high-yielding rice cultivars with different grain-filling properties to clarify limiting factors for improving grain yield. **Field Crops Research**, v. 219, p. 139–147, 2018.
- PAN, Y. J. et al. The transmembrane domain 6 of vacuolar H⁺-pyrophosphatase mediates protein targeting and proton transport. **Biochimica et Biophysica Acta - Bioenergetics**, v. 1807, n. 1, p. 59–67, 2011.
- PARK, Sangkyu et al. Molecular and functional characterization of *Oryza sativa* Flavonol Synthase (*OsFLS*), a bifunctional dioxygenase. **Journal of Agricultural and Food Chemistry**, v. 67, n. 26, p. 7399–7409, 2019.
- PASAPULA, V. et al. Expression of an *Arabidopsis* vacuolar H⁺-pyrophosphatase gene (*AVPI*) in cotton improves drought- and salt tolerance and increases fibre yield in the field conditions. **Plant Biotechnology Journal**, v. 9, n. 1, p. 88–99, 2011.
- PEGORARO, C. et al. Ethylene response factors gene regulation and expression profiles under different stresses in rice. **Theoretical and Experimental Plant Physiology**, v. 25, n. 4, p. 261–274, 2013. Disponível em: <http://microsatellite.org/cis_input.html>.
- PEI, L. et al. Overexpression of *Thellungiella halophila* H⁺-pyrophosphatase gene improves low phosphate tolerance in maize. **PLoS ONE**, v. 7, n. 8, p. e43501, 30 ago. 2012. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0043501>>. Acesso em: 7 jan. 2021.
- PIZZIO, G. A. et al. *Arabidopsis* type I proton-pumping pyrophosphatase expresses strongly in phloem, where it is required for pyrophosphate metabolism and photosynthate partitioning. **Plant Physiology**, v. 167, n. 4, p. 1541–1553, 2015.
- PRIMO, C. et al. Plant proton pumping pyrophosphatase: the potential for its pyrophosphate synthesis activity to modulate plant growth. **Plant Biology**, v. 21, n. 6, p. 989–996, 2019.
- QUERO, G. et al. Genetic architecture of photosynthesis energy partitioning as revealed by a genome-wide association approach. **Photosynthesis Research**, , n. 0123456789, 2020. Springer Netherlands. Disponível em: <<https://doi.org/10.1007/s11120-020-00721-2>>.
- RAHMAN, M. A. et al. Correlation Analysis of Flag Leaf with Yield in Several Rice Cultivars. **Journal of Life and Earth Science**, v. 8, p. 49–54, 2013.
- RAMIREZ-VILLEGAS, J. et al. Breeding implications of drought stress under future climate for upland rice in Brazil. **Global Change Biology**, v. 24, n. 5, p. 2035–2050, 2018.

RAZA, G. et al. Overexpression of an H⁺-PPase gene from *Arabidopsis* in sugarcane improves drought tolerance, plant growth, and photosynthetic responses. **Turkish Journal of Biology**, v. 40, n. 1, p. 109–119, 2016.

REGMI, K. C.; YOGENDRA, K.; FARIAS, J. G.; et al. Improved Yield and Photosynthate Partitioning in *AVP1* Expressing Wheat (*Triticum aestivum*) Plants. **Frontiers in Plant Science**, v. 11, p. 273, 2020.

REIFSCHNEIDER F.J.B. et al. Uma pitada de biodiversidade na mesa dos brasileiros, ed.1, ISBN: 978-85-917453-0-2. 156 p, 2015.

RGAP. Introduction to the rice genome annotation project. Disponível em: <<http://rice.plantbiology.msu.edu/>>. Acessado em: 29 jun. 2019.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 2010.

ROCHA, A. G. et al. Phosphorylation of *Arabidopsis* transketolase at Ser 428 provides a potential paradigm for the metabolic control of chloroplast carbon metabolism. **Biochem. J**, v. 458, p. 313–322, 2014.

RUIZ-VILLALBA, A. et al. Amplification of nonspecific products in quantitative polymerase chain reactions (qPCR). **Biomolecular Detection and Quantification**, v. 14, p. 7–18, 2017.

SAFITRI, F. A.; UBAIDILLAH, M.; KIM, K. M. Efficiency of transformation mediated by *Agrobacterium tumefaciens* using vacuum infiltration in rice (*Oryza sativa* L.). **Journal of Plant Biotechnology**, v. 43, n. 1, p. 66–75, 2016.

SAKAI, H.; ITOH, T. Massive gene losses in Asian cultivated rice unveiled by comparative genome analysis. **BMC Genomics**, v. 11, n. 1, 2010.

SARKAR, F. et al. **Package “lattice”**. 2020.

SCHILLING, R. K. et al. AVP1: One Protein, Many Roles. **Trends in Plant Science**, v. 22, n. 2, p. 154–162, 2017.

SCHOLZ-STARKE, J. et al. The flip side of the *Arabidopsis* type I proton-pumping pyrophosphatase (*AVP1*): using a transmembrane H⁺ gradient to synthesize pyrophosphate. **Journal of Biological Chemistry**, v. 294, n. 4, p. 1290–1299, 2019.

SCHOUTEN, H.; JACOBSEN, E. Cisgenic plants are similar to traditionally bred plants. **Science e Society**, v. 7, p. 750–753, 2006.

SECCO, D. et al. The emerging importance of the SPX domain-containing proteins in

phosphate homeostasis. **New Phytologist**, v. 193, n. 4, p. 842–851, 2012.

SEGAMI, S. et al. Vacuolar H⁺-pyrophosphatase and cytosolic soluble pyrophosphatases cooperatively regulate pyrophosphate levels in *Arabidopsis thaliana*. **Plant Cell**, v. 30, n. 5, p. 1040–1061, 2018.

SHI, J. et al. The paralogous SPX3 and SPX5 genes redundantly modulate P_i homeostasis in rice. **Journal of Experimental Botany**, v. 65, n. 3, p. 859–870, 2014.

SINGH, R. K.; PRASAD, M. Advances in *Agrobacterium tumefaciens*-mediated genetic transformation of graminaceous crops. **Protoplasma**, v. 253, n. 3, p. 691–707, 2016.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nature Reviews Genetics**, 2019.

SUI, N. et al. Identification and transcriptomic profiling of genes involved in increasing sugar content during salt stress in sweet sorghum leaves. **BMC Genomics**, v. 16, n. 1, 2015.

TUONG, T.P.; BOUMAN, A.A. Rice Production in water-scarce environments. **CAB International**, v. 12, n. 2, p. 186–191, 2003.

UN. **World population prospects 2019 highlights**. New York, 2019. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12283219>>.

UNIPROT. 134P10.11 - Transketolase - *Oryza sativa* subsp. japonica (Rice) 134P10.11 gene & protein. Disponível em: <<https://www.uniprot.org/uniprot/Q84ZY2>>. Acesso em: 12/10/2020.

UPADHYAY, N. et al. The multitasking abilities of MATE transporters in plants. **Journal of Experimental Botany**, v. 70, n. 18, p. 4643–4656, 2019.

USADEL B. et al. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. **Plant Physiol.** 138::1195-204, 2012.

VINECKY, G. F. et al. Análises da expressão gênica por PCR quantitativo em tempo real (qPCR) de genes candidatos para a tolerância à seca em cafeeiro. **VII Simpósio de Pesquisa dos Cafés do Brasil**, 2011.

WARNES, Gregory R. et al. CRAN - Package gplots. 2020. Disponível em: <<https://cran.r-project.org/web/packages/gplots/index.html>>. Acesso em: 7 jan. 2021.

WANG, D. et al. A missense mutation in the transmembrane domain of *CESA9* affects cell wall biosynthesis and plant growth in rice. **Plant Science**, v. 196, p. 117–124, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.plantsci.2012.08.002>>.

WANG, L. et al. The similar and different evolutionary trends of MATE family occurred

between rice and *Arabidopsis thaliana*. **BMC Plant Biology**, v. 16, n. 1, p. 1–19, 2016. Disponível em: <<http://dx.doi.org/10.1186/s12870-016-0895-0>>.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, p. 57–63, 2009.

WEI T & SIMKO V. R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). 2017. Available from <https://github.com/taiyun/corrplot>.

WICKHAM, H. et al. **Package “readxl”**. 2019.

WING, R. A. et al. The *Oryza* map alignment project: the golden path to unlocking the genetic potential of wild rice species. **Plant Molecular Biology**, v. 59, n. 1, p. 53–62, 2005.

XIA et al. Adaptive epigenetic differentiation between upland and lowland rice ecotypes revealed by methylation-sensitive amplified polymorphism. **PLoS ONE** 11(7): e0157810, 2016.

XING, Y.; ZHANG, Q. Genetic and molecular bases of rice yield. **Annual Review of Plant Biology**, v. 61, n. 1, p. 421–442, 2010.

YANG, H. et al. Enhanced phosphorus nutrition in monocots and dicots over-expressing a phosphorus-responsive type I H⁺-pyrophosphatase. **Plant Biotechnology Journal**, v. 5, n. 6, p. 735–745, 2007.

YANG, H. et al. Over-expression of the *Arabidopsis* proton-pyrophosphatase AVP1 enhances transplant survival, root mass, and fruit development under limiting phosphorus conditions. **Journal of Experimental Botany**, v. 65, n. 12, p. 3045–3053, 2014.

ZHANG, J. et al. OVP1, a Vacuolar H⁺-translocating inorganic pyrophosphatase (V-PPase), overexpression improved rice cold tolerance. **Plant Physiology and Biochemistry**, v. 49, n. 1, p. 33–38, 2011.

ZHANG, J. L. et al. Expression pattern of *GS3* during panicle development in rice under drought stress: quantification normalized against selected housekeeping genes in real-time PCR. **Asian Journal of Plant Sciences**, Wuhan, v. 8, n. 4, p. 285-292, 2009.

ZHANG, J. et al. Plastid ribosomal protein S5 is involved in photosynthesis, plant development, and cold stress tolerance in *Arabidopsis*. **Journal of Experimental Botany**, v. 67, n. 9, p. 2731–2744, 2016a.

ZHANG, W. et al. Grain filling characteristics and their relations with endogenous hormones in large- and small-grain mutants of rice. **PLoS ONE**, v. 11, n. 10, p. 1–20, 2016b.

ZHANG, X. et al. Effects of drought stress and water recovery on physiological responses and gene expression in maize seedlings. **BMC Plant Biology**, v. 18, n. 1, p. 68, 2018.

ZHAO, Y. et al. A Literature Review of Gene Function Prediction by Modeling Gene Ontology. **Frontiers in Genetics**, v. 11, 2020.

ZHONG, R.; CUI, D. ; YE, Z. H.. Secondary cell wall biosynthesis. **New Phytologist**, v. 221, n. 4, p. 1703–1723, 2019.

ZHU, T. et al. Transcriptional control of nutrient partitioning during rice grain filling. **Plant Biotechnology Journal**, n. 1, p. 59–70, 2003.

APÊNDICES

Apêndice A. Resultado da quantificação ($\text{ng} \times \mu\text{L}^{-1}$) e análise da qualidade das amostras de RNA provenientes do ensaio 2018/2019.

Genótipo^a	Estádio^b	Repetição^c	RNA ($\text{ng} \times \mu\text{L}^{-1}$)^d	ABS (260/280)^e	ABS (260/230)^e
NGM	R3	112	402	2,03	1,98
NGM	R3	309	334	2,01	2,27
NGM	R3	509	330	2,04	2,24
NGM	R3	712	358	2,01	2,19
OVP1-E4	R3	116	390	2,05	2,10
OVP1-E4	R3	323	184	2,03	2,02
OVP1-E4	R3	503	286	2,05	2,10
OVP1-E4	R3	701	193	2,02	1,82
NGM	R6	112	246	2,03	2,25
NGM	R6	309	246	2,02	2,17
NGM	R6	509	338	2,03	2,00
NGM	R6	712	114	2,00	2,22
OVP1-E4	R6	116	242	2,02	1,98
OVP1-E4	R6	323	404	2,03	2,20
OVP1-E4	R6	503	149	2,04	1,90
OVP1-E4	R6	701	304	2,20	2,03
NGM	R8	112	167	2,01	1,6
NGM	R8	309	135	2,01	1,80
NGM	R8	509	266	2,03	2,05
NGM	R8	712	114	2,02	2,16
OVP1-E4	R8	116	162	2,02	2,21
OVP1-E4	R8	323	167	2,02	2,16
OVP1-E4	R8	503	82	1,85	2,03
OVP1-E4	R8	701	189	2,03	2,06

^a O genótipo NGM equivale as plantas não geneticamente modificadas (BRSMG Curinga controle). ^b O estágio são as fases do desenvolvimento em que o tecido foliar foi coletado para análise de expressão gênica. ^c A repetição são os números das repetições biológicas usadas para cada genótipo. ^d A quantificação foi feita via Qubit[®] 2.0 Fluorometer (Invitrogen). ^e A pureza de cada amostra foi avaliada no espectrofotômetro Nanovue Plus[™] (GE Healthcare) usando a relação entre as absorvâncias (ABS) de ácidos nucleicos (260nm), proteínas (280nm) e reagentes residuais (230nm).

Apêndice B. Resultado da quantificação ($\text{ng} \times \mu\text{L}^{-1}$) e análise da qualidade das amostras de RNA provenientes do ensaio 2019/2020.

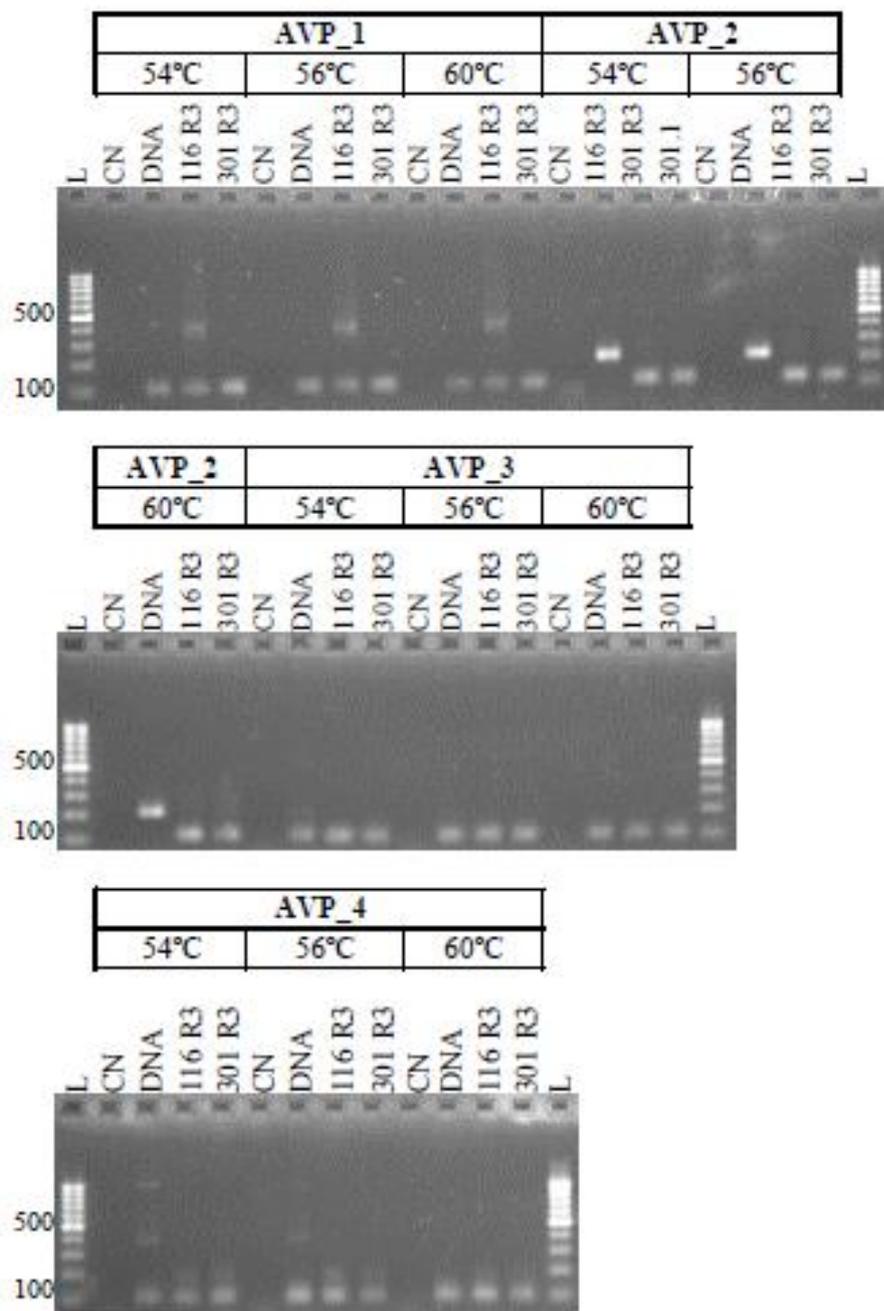
Genótipo^a	Estádio^b	Repetição^c	RNA ($\text{ng} \times \mu\text{L}^{-1}$)^d	ABS (260/280)^e	ABS (260/230)^e
NGM	R2	104	142	2.02	1.73
NGM	R2	307	454	2.01	2.28
NGM	R2	502	314	2.00	1.90
NGM	R2	702	280	2.18	2.35
OVP1-E4	R2	105	248	2.02	2.24
OVP1-E4	R2	304	212	2.02	1.71
OVP1-E4	R2	504	276	2.01	2.11
OVP1-E4	R2	704	236	2.00	2.05
NGM	R6	104	272	2.01	1.96
NGM	R6	307	234	2.00	2.37
NGM	R6	502	186	2.00	1.80
NGM	R6	702	220	2.00	2.31
OVP1-E4	R6	105	328	1.99	2.02
OVP1-E4	R6	304	164	1.92	1.50
OVP1-E4	R6	504	173	2.01	1.60
OVP1-E4	R6	704	214	1.98	1.86
NGM	R7	104	228	2.00	2.39
NGM	R7	307	123	2.00	1.60
NGM	R7	502	169	2.00	2.14
NGM	R7	702	246	2.02	2.05
OVP1-E4	R7	105	200	2.00	1.80
OVP1-E4	R7	304	145	1.99	2.17
OVP1-E4	R7	504	173	2.01	1.99
OVP1-E4	R7	704	131	1.96	1.87

^aO genótipo NGM equivale às plantas não geneticamente modificadas (BRSMG Curinga controle). ^bO estágio são as fases do desenvolvimento em que o tecido foliar foi coletado para análise de expressão gênica. ^cA repetição são os números das repetições biológicas usadas para cada genótipo. ^dA quantificação foi feita via Qubit® 2.0 Fluorometer (Invitogen). ^eA pureza de cada amostra foi avaliada no espectrofotômetro Nanovue Plus™ (GE Healthcare) usando a relação entre as absorvâncias (ABS) de ácidos nucleicos (260nm), proteínas (280nm) e reagentes residuais (230nm).

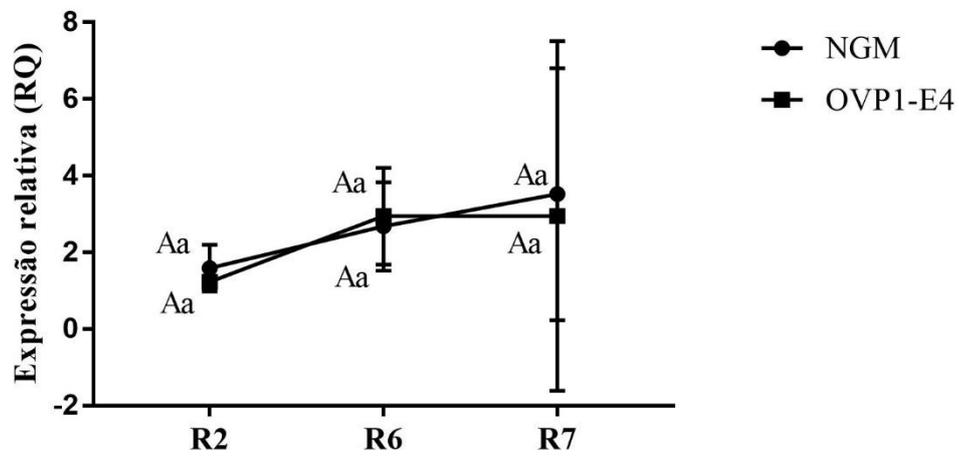
Apêndice C. Resultado da análise de integridade do RNA via Bioanalyzer.

Experimento	Genótipo	Estádio	Número	RIN RNA^a
2018/2019	NGM	R3	701	7,8
2018/2019	NGM	R6	309	6,9
2018/2019	NGM	R8	112	8,0
2018/2019	OVP1-E4	R3	323	8,0
2018/2019	OVP1-E4	R6	503	8,2
2018/2019	OVP1-E4	R8	503	8
2019/2020	NGM	R2	104	7,5
2019/2020	NGM	R6	307	8,0
2019/2020	NGM	R7	502	8,1
2019/2020	OVP1-E4	R2	105	7,5
2019/2020	OVP1-E4	R6	304	7,5
2019/2020	OVP1-E4	R7	504	7,7

^a O número de integridade do RNA (RIN) foi analisado para as amostras pela técnica de microfluidos Bioanalyzer 2100 (Agilent).



Apêndice D. Otimização da temperatura de anelamento dos *primers* desenhados para o gene *OVPI*. Duas amostras de cDNA foram usadas e uma de gDNA. Gel de agarose 1% corado com brometo de etídio. L) Marcador de peso molecular Ladder 100 Thermo Scientific. CN) Controle negativo com água substituindo o cDNA. DNA) Controle positivo com DNA genômico.



Apêndice E. Expressão relativa do gene *OVP1* no ensaio 2019/2020. As letras minúsculas indicam diferença estatística entre os genótipos e as letras maiúsculas indicam diferença estatística entre as fases de desenvolvimento (n=4, p<0,05).

Apêndice F. Análise conjunta da expressão relativa do gene *OVPI* durante o estágio R6.

	Df	Sum	Mean	F value	Pr (>F)
Exp ^a	1	0,09	0,09	3,31	0,11
Gen ^b	1	0,89	0,89	31,84	0,001
Exp:bloco	6	1,4	0,23	8,35	0,01
Exp:gen	1	0,63	0,63	22,53	0,003
Resíduos	6	0,17	0,02	-	-

^aExp são os dois experimentos: 2018/2019 e 2019/2020.

^bGen são os dois genótipos: OVPI-E4 e NGM.