

AMANDA BONOTO GONÇALVES

**A SUPEREXPRESSÃO DE GmbZIPE2 EM PROTOPLASTOS DE SOJA
SUGERE UM PAPEL REGULADOR DA EXPRESSÃO DE GENES DA
FOTOSSÍNTESE E DA VIA DE BIOSÍNTESE DE FENILPROPANÓIDES**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS-BRASIL
2018

Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa

T

G635s
2018
Gonçalves, Amanda Bonoto, 1988-
A superexpressão de GmbZIPE2 em protoplastos de soja
sugere um papel regulador da expressão de genes da fotossíntese
e da via de biossíntese de fenilpropanóides / Amanda Bonoto
Gonçalves. – Viçosa, MG, 2018.
x, 74 f. : il. (algumas color.) ; 29 cm.

Orientador: Luciano Gomes Fietto.

Tese (doutorado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 42-48.

1. Zíper de leucina. 2. Transcriptoma. 3. Soja.
I. Universidade Federal de Viçosa. Departamento de Bioquímica
e Biologia Molecular. Programa de Pós-Graduação em
Bioquímica Aplicada. II. Título.

CDD 22. ed. 572.65

AMANDA BONOTO GONÇALVES

**A SUPEREXPRESSÃO DE GmbZIPE2 EM PROTOPLASTOS DE SOJA
SUGERE UM PAPEL REGULADOR DA EXPRESSÃO DE GENES DA
FOTOSSÍNTESE E DA VIA DE BIOSSÍNTESE DE FENILPROPANÓIDES**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

Pedro Augusto Braga dos Reis

Maximiller Da-Bianco Lamas Costa

Pedro Marcus Pereira Vidigal

Antônio Helvécio Totola

Luciano Gomes Fietto
(Orientador)

AGRADECIMENTOS

Agradeço...

A Deus e nossa Senhora por todas bênçãos e vitórias alcançadas. A Eles agradeço à minha vida, saúde física e mental e sabedoria que me possibilitaram conseguir atravessar mais uma etapa da minha vida acadêmica.

À minha família. Meus pais, Amélia e Toninho, pelo amor incondicional, incentivo e dedicação. Essa vitória é para vocês. Aos meus irmãos Daniel, Lucas e Marco e cunhadas pelo companheirismo e torcida. Ao meu sobrinho Mateus pelo amor e carinho. Aos meus tios Luís e Íris e primas Cássia, Flávia e Artur pelo apoio.

À Universidade Federal de Viçosa, pela oportunidade de estudar em uma instituição de excelência. A UFV possibilitou que eu ficasse por 11 anos na universidade que eu sempre sonhei em estudar. Serei sempre grata e orgulhosa por todos ensinamentos adquiridos.

Ao Luciano, pela orientação, paciência e compreensão. Foram sete anos de orientação e sou muito grata a você. Obrigada por ser um ótimo orientador e por ser uma ótima pessoa. Me espelho em você para ser uma profissional correta, prestativa e humana.

À Paty, pela amizade e companheirismo. Obrigada por toda a ajuda, no laboratório e na vida. Você foi essencial para esse trabalho, minha companheira de bZIP. Obrigada por todas as risadas e choros compartilhados, porque só assim a gente se fortalece, e nós nos fortalecemos juntas.

À Gilza, minha companheira de experimentos desde o mestrado. Agradeço por toda a ajuda e pela amizade que construímos.

À Silvana pela ajuda e amizade.

Ao Murilo pela Coorientação e ajuda.

Ao Laboratório de Biotecnologia Molecular, pelo aprendizado nesses anos. E a todos que fazem parte do LBM, obrigada pela convivência e pela ajuda. Em especial à Fernanda, Ananda, Renato, Higor e Thays.

Aos amigos que passaram pelo LBM em todos esses anos, em especial à Ayla, Vanessa e Gêssica.

À minha amiga Janaína, por estar ao meu lado desde a graduação, seja nas horas boas ou nos perrengues. Obrigada pelo ombro amigo, pela disposição de me escutar e pelos valiosos conselhos.

Aos meus amigos, que sempre estiveram do meu lado nessa jornada: Mayara, Lídia, Fernanda, Thalita, Mayra, Aninha, Gêssica e Lucas.

Ao Du, pela amizade e por toda a ajuda burocrática na pós.

Ao Professor Maximiller pela ajuda no sequenciamento, ao Pedro Vidigal pela ajuda nas análises do transcriptoma, e ao Tiago pela ajuda na bioinformática.

Ao Núcleo de Análise de Biomoléculas da Universidade Federal de Viçosa por fornecer as facilidades para a realização dos experimentos.

Ao Guilherme, funcionário do DFP, pela ajuda e ao professor Murilo Zerbini por permitir o uso de sua casa de vegetação.

Aos professores do Departamento de Bioquímica e Biologia Molecular, pelos ensinamentos.

Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pela concessão da bolsa, à Fundação de Apoio à Pesquisa de Minas Gerais (FAPEMIG) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

A todos que ajudaram, direta ou indiretamente, obrigada.

SUMÁRIO

LISTA DE FIGURAS, GRÁFICOS E TABELAS	vi
LISTA DE ABREVIATURAS E SIGLAS	vii
RESUMO	ix
ABSTRACT	x
1 - INTRODUÇÃO.....	1
2 - REVISÃO DA LITERATURA.....	3
2.1. Sistema imune da planta e sua relação com fatores de transcrição	3
2.2. Família bZIP	4
2.3. GmbZIPE2	6
2.4. Relação de <i>GmbZIPE2</i> com fotossíntese e estresse oxidativo	7
3 - OBJETIVO.....	10
4 - MATERIAIS E MÉTODOS	11
4.1 - Obtenção de protoplasto de soja e superexpressão de <i>GmbZIPE2</i>	11
4.3 - Extração, quantificação e qualidade do RNA	13
4.4 - Tratamento com DNase I e síntese de cDNA	13
4.5 - PCR em Tempo Real (RT-qPCR)	14
4.5.1 - Desenho dos oligonucleotídeos	14
4.5.2 - Eficiência dos oligonucleotídeos	14
4.5.3 - Reação do RT-qPCR	15
4.5.4 - Quantificação da expressão gênica	15
Tabela 1: Oligonucleotídeos utilizados no qRT-PCR	15
Tabela 2: Oligonucleotídeos referentes à genes endógenos utilizados no qRT-PCR	16
4.6 - Construção, quantificação das bibliotecas	17
4.7 - Sequenciamento das bibliotecas	17
4.8 - Tratamento dos dados, análise de qualidade e montagem do transcriptoma	18
4.9 - Validação do RNA-seq	19
4.10 - Silenciamento do gene <i>GmbZIPE2</i> através do ensaio VIGS	19
4.11 - Técnicas de Biologia Molecular	21
5 - RESULTADOS	22
5.1 - Qualidade do RNA e superexpressão de <i>GmbZIPE2</i>	22
5.2 - Qualidade e validação das bibliotecas	24
5.3 - Tratamento de dados do sequenciamento	25
5.4 - Transcriptoma de protoplastos de soja superexpressando <i>GmbZIPE2</i>	27

5.5 - Validação do transcriptoma.....	28
5.6 - Expressão da provável LPA em protoplastos superexpressando <i>GmbZIPE2</i>	31
5.7 - O silenciamento de <i>GmbZIPE2</i> através de VIGS	31
6 - DISCUSSÃO.....	33
6.1. <i>GmbZIPE2</i> ativa genes de fotossíntese	33
6.2. <i>GmbZIPE2</i> ativa fatores de transcrição WRKY	36
6.3. Silenciamento de <i>GmbZIPE2</i>.....	39
7 - CONCLUSÃO.....	41
8 - REFERÊNCIAS	42
9 - MATERIAL SUPLEMENTAR	49

LISTA DE FIGURAS, GRÁFICOS E TABELAS

Tabela 1: Oligonucleotídeos utilizados no RT-qPCR

Tabela 2: Oligonucleotídeos referentes à genes endógenos utilizados no qRT-PCR

Tabela 3: Oligonucleotídeos específicos para a clonagem do gene *GmbZIPE2*.

Tabela 4: Informações geradas pelo FastQC referentes à trimagem das *reads*.

Figura 1: Árvore filogenética com sequências de proteínas bZIPs de *G.max* e *A. arabidopsis*.

Figura 2: Modelo esquemático da via inferida para a possível atuação do fator de transcrição *GmbZIPE2* em resposta a estresse.

Figura 3: Qualidade do RNA extraído das amostras controle e de protoplastos superexpressando *GmbZIPE2*.

Figura 4: Superexpressão de *GmbZIPE2* em protoplastos de soja.

Figura 5: Análise da qualidade das bibliotecas.

Figura 6: Análise das *reads* referentes a *GmbZIPE2*.

Figura 7: Análise da sobreposição das amostras.

Figura 8: Análise da sobreposição das amostras após a deleção de *reads*.

Figura 9: Distribuição dos genes ativados pela superexpressão de *GmbZIPE2* em relação a processos biológicos, função molecular e componentes celulares.

Figura 10: Validação do transcriptoma de protoplastos de soja superexpressando *GmbZIPE2*.

Figura 11: Expressão da LPA em protoplastos superexpressando *GmbZIPE2*.

Figura 12: Expressão de *GmbZIPE2*, genes da via de biossíntese de fenilpropanóides e genes validados como ativados pela superexpressão de *GmbZIPE2* em folhas de soja com o gene *GmbZIPE2* supostamente silenciado.

Figura 13: Representação esquemática da atuação de WKRY33 em situação de estresse.

LISTA DE ABREVIATURAS E SIGLAS

ABA - Ácido abscísico

bZIP - *Basic leucine zipper containing domain proteins*

CDPKs - Quinases dependentes de cálcio

CDS - Sequência codificante

CHS - Chalcona sintase

Cq - Ciclo de quantificação

DEPC - *Diethyl pyrocarbonate*

ET - Etileno

ETI - *Effector-triggered immunity*

ETS - *Effector-triggered susceptibility*

Fe^{3+} - Zn^{2+} PAP - *Fe^{3+} - Zn^{2+} purple acid phosphatase 12*

FLA15 - *Fasciclin-like arabinogalactan protein 15-related*

G/HBF1 - *G-box/H-box binding factor 1*

GLDC - *Glycine dehydrogenase (aminomethyl-transferring)/Glycine-cleavage complex P-protein*

HR - *Hypersensitive response*

JA - *Jasmonic acid*

LPA - *Predicted Low Psii Accumulation*

MAMPs - *Microbial-associated molecular patterns,*

MAPK - Quinase ativadas por mitógenos

MED 37e - *Mediator of RNA polymerase II transcription subunit 37e-related.*

MeJA - Metil Jasmonato

MS - Murashige e Skoog

NB-LRR - *Nucleotide-binding site leucine-rich repeat*

NPR1 - *Non-Expresser of Pathogen-Related 1*

OE - *overexpression*

PAMPS - *Pathogen-associated molecular patterns*

PRR - *Pattern recognition receptors*
PSII - Fotossistema II
PTI - *PAMP-triggered immunity*
Rbcx - *Chaperonin-like Rbcx Protein*
RNAseq - *RNA Sequencing*
ROS - Espécies reativas de oxigênio
SA - Ácido salicílico
SMV - *Soybean Mosaic Virus*
ssDNA - DNA fita simples de esperma de salmão
TGA - proteínas de ligação ao *cis*-elemento TGACGTCA
TMV - *Tobacco mosaic vírus*
TRP - *TPR repeat containing protein*
TRX - Tiorredoxinas
VIGS - *Virus-induced gene silencing*

RESUMO

GONÇALVES, Amanda Bonoto, D.Sc., Universidade Federal de Viçosa, dezembro de 2018. **A superexpressão de GmbZIPE2 em protoplastos de soja sugere um papel regulador da expressão de genes da fotossíntese e da via de biossíntese de fenilpropanóides.** Orientador: Luciano Gomes Fietto. Coorientador: Murilo Siqueira Alves.

A modulação da transcrição gênica é fundamental para a formação de uma resposta de defesa eficaz em plantas sob estresses. Os mecanismos de modulação são regulados principalmente por fatores de transcrição. A família de fatores de transcrição com domínio básico de zíper de leucina (bZIP) é relacionada com vários processos na planta, dentre eles a fotossíntese e estresse abiótico e biótico, e são divididos em 12 grupos em soja. GmbZIPE2 (Glyma05g168100) é um membro do grupo E da família bZIP de *Glycine max*, e membros desse grupo ainda não foram caracterizados nesta espécie. O gene *GmbZIPE2* de soja tem sua expressão alterada em resposta à infecção fúngica e ao tratamento com fitormônios de defesa. A proteína GmbZIPE2 é capaz de se ligar ao *cis*-elemento H-box *in vitro* e ativar a transcrição da chalcona sintase *CHS8*, importante enzima da via de biossíntese de fenilpropanóides. Além disso, GmbZIPE2 interage com duas proteínas relacionadas à fotossíntese e estresse oxidativo. Neste estudo foi realizado o transcriptoma de protoplastos de folhas de soja superexpressando *GmbZIPE2* a fim de se obter mais informações sobre a função desse fator de transcrição. A maior porcentagem de genes ativados pela superexpressão de *GmbZIPE2* estão relacionados com fotossíntese. Um número expressivo de genes ativados são fatores de transcrição, dentre eles dois *WRKY33*, gene que é relacionado a indução da produção de fitoalexinas na defesa contra patógenos. Nossos resultados sugerem que GmbZIPE2 possui um papel regulador na via de biossíntese de fitoalexinas e na fotossíntese, mais especificamente, no controle de estresse oxidativo.

ABSTRACT

GONÇALVES, Amanda Bonoto, D.Sc., Universidade Federal de Viçosa, December, 2018. **Overexpression of GmbZIPE2 in soybean protoplasts suggests a regulatory role for the expression of photosynthesis genes and the phenylpropanoid biosynthesis pathway.** Adviser: Luciano Gomes Fietto. Co-adviser: Murilo Siqueira Alves.

Gene expression modulation is critical for the formation of an effective defense response in stressed plants. Modulation mechanisms are mainly regulated by transcription factors. The family of transcription factors with basic leucine zipper domain (bZIP) is related to several processes in the plant, including photosynthesis and abiotic/biotic stress, and are divided into 12 groups in soybean. GmbZIPE2 (Glyma05g168100) is a member of the E group of the bZIP family in *Glycine max*, and members of this group have not been characterized in this species. The gene *GmbZIPE2* of soybean has altered expression in response to the fungal infection and the treatment with phytohormones related to defense. The GmbZIPE2 protein is capable of binding to the H-box cis-element *in vitro* and activates the transcription of *CHS8* chalcone synthase, an important enzyme in phenylpropanoid biosynthesis pathway. In addition, GmbZIPE2 interact with two proteins related to photosynthesis and oxidative stress. In this study the transcriptome of soybean leaves protoplasts overexpressing *GmbZIPE2* was performed in order to obtain more information on the function of this transcription factor. The highest percentage of genes activated by the overexpression of *GmbZIPE2* are related to photosynthesis. An expressive number of activated genes are transcription factors, among them two *WRKY33*, a gene that is related to production of phytoalexins in defense against pathogens. Our results suggest that GmbZIPE2 has a regulatory role in the biosynthesis pathway of phytoalexins and in photosynthesis, more specifically in the control of oxidative stress. The results generated in this work increase the information about the GmbZIPE2 gene and its possible functions in plant and consequently helps the characterization of the E family of soybean bZIPs.

1 – INTRODUÇÃO

A soja (*Glycine max* (L.) Merr.) é uma das cultivares mais importantes na economia mundial. A doença mais severa que acomete essa cultura é ferrugem asiática, causada pelo fungo biotrófico *Phakopsora pachyrhizi*. A ferrugem asiática ameaça as principais áreas de cultivo de soja na América do Sul, e especificamente no Brasil, essa doença causou perdas de mais de US\$ 10 bilhões desde seu primeiro surto endêmico no ano de 2001 (Yorinori et al., 2005; Da Silva et al., 2014; Langenbach et al., 2016).

As plantas, por serem organismos sésseis, requerem um sistema imunológico complexo para a proteção contra patógenos que comprometem seu crescimento e desenvolvimento (Kapus et al, 2019). O sistema imune vegetal é acionado inicialmente pela percepção de padrões moleculares associados a patógenos (*pathogen-associated molecular patterns*, PAMPs ou *microbial-associated molecular patterns*, MAMPs) na superfície da célula da planta. O reconhecimento de PAMPs acontece através de receptores da planta chamados PRRs (*pattern-recognition receptors*) (Spoel and Dong, 2012) e a resposta aos PAMPs é chamada de imunidade desencadeada por PAMP (*PAMP-triggered immunity*, PTI). Muitos patógenos conseguem contornar essa primeira barreira produzindo efetores que inibem a PTI. Porém as plantas podem reconhecer esses efetores através de receptores adicionais para montar uma segunda fase de defesa, chamada imunidade desencadeada por efetores (*effector-triggered immunity*, ETI) (Boller and He, 2010). Geralmente PTI e ETI geram respostas semelhantes, porém, qualitativamente, a ETI é mais forte e mais rápida (Dodds and Rathjen, 2010). A ETI leva a uma resposta amplificada, e muitas vezes envolve uma forma de morte celular chamada Resposta de hipersensibilidade (HR) no local da infecção, na tentativa de evitar a disseminação do patógeno para outras células da planta (Jones and Dangly, 2006).

A modulação da transcrição de alguns genes é fundamental para montar uma resposta de defesa eficaz nas células do hospedeiro. A reprogramação transcricional em resposta a patógenos busca favorecer a defesa da planta em detrimento de outros processos celulares (Buscaill and Rivas, 2014). Essa regulação é dirigida principalmente pela ação de fatores de transcrição, que são proteínas que modulam a expressão de um grupo específico de genes através de ligação em regiões específicas do DNA. Por controlarem a expressão de vários genes simultaneamente, os fatores de transcrição

estão entre os principais alvos para o aumento da tolerância de plantas a estresses (Latchman, 2003).

As famílias de fatores de transcrição são conservadas entre as espécies vegetais, sendo que vários de seus membros estão ligados a respostas a estresses. Entre estas famílias encontram-se os fatores de transcrição com domínio básico de zíper de leucina bZIP (Alves *et al.*, 2013). Vários transfatores bZIP são caracterizados por possuir um domínio básico de zíper de leucina, responsável pela ligação ao DNA, através da formação de duas estruturas em forma de alfa hélice (Schütze *et al.*, 2008). De acordo com a similaridade de sequência de aminoácidos, a família bZIP foi subdividida em 12 grupos em soja (Zhang *et al.*, 2018). Membros da família bZIP podem regular vários processos na planta, como maturação de sementes, desenvolvimento de flores, resposta a estresse abióticos e a defesa contra patógenos (Jacob *et al.*, 2002).

GmbZIPE2 (Glyma05g168100) é um membro do grupo E da família bZIP de soja, grupo em que nenhum membro foi caracterizado. GmbZIPE2 é capaz de ligar ao *cis*-elemento H-box e ativar a transcrição da chalcona sintase *CHS8*, importante enzima da via de biossíntese de fenilpropanoides. Na infecção de plantas susceptíveis a *P. pachyrhizi*, *GmbZIPE2* é ativado. Esses dados sugerem uma relação desse gene com a defesa da soja contra patógenos. Para entender como a regulação realizada por *GmbZIPE2* acontece, Fontes (2016) identificou dois parceiros de interação desse fator. As proteínas encontradas, *Rbcx* (*Chaperonin-like Rbcx Protein*) e uma possível LPA (*Predicted Low Psii Accumulatio*), têm funções relacionadas à fotossíntese e ao estresse oxidativo na planta, respetivamente.

Este trabalho tem como objetivo principal analisar o efeito da superexpressão (*OE*) do gene *GmbZIPE2* em protoplasto de soja, através da obtenção de um transcriptoma e também analisar as modificações resultantes do silenciamento desse gene, através de VIGS (silenciamento gênico induzido por vírus), em folhas de soja. Os resultados gerados nesse trabalho auxiliam na caracterização do gene *GmbZIPE2* e consequentemente a caracterização da família E de bZIPs de soja.

2 - REVISÃO DA LITERATURA

2.1. Sistema imune da planta e sua relação com fatores de transcrição

As plantas não têm células de defesa móveis e não possuem um sistema imunológico somático adaptativo. A defesa de plantas está associada com a imunidade inata de cada célula e em sinais sistêmicos originados de infecções locais (Jones and Dangly, 2006). O sistema de imunidade inata de plantas é acionado inicialmente pela percepção de PAMPs ou MAMPs através de receptores de reconhecimento de padrões (*pattern recognition receptors*, PRR) na superfície da célula. A resposta aos PAMPs é chamada de PTI, e muitos patógenos conseguem contornar essa primeira barreira da planta produzindo efetores que inibem a PTI. Esses patógenos implantam efetores na planta que podem interferir com a PTI resultando em uma suscetibilidade desencadeada por efector (*effector-triggered susceptibility*, ETS). Porém as plantas podem reconhecer esses efetores através de receptores adicionais, geralmente por receptores NB-LRR (*Nucleotide-binding site leucine-rich repeat*), caracterizados por um domínio de ligação a nucleotídeo e um domínio rico em leucina, para montar uma segunda fase de defesa, a ETI (Boller and He, 2010). O reconhecimento dos efetores pode ser direto através do reconhecimento de um efector pelo NB-LRR, mas também pode ser indireto. ETI é uma resposta acelerada e amplificada da PTI, resultando em resistência à doença e, geralmente, uma resposta de hipersensibilidade (HR), que é um tipo de morte celular programada no local da infecção. Os patógenos evoluíram de maneira a evitar a ETI, diversificando o gene efector, ou através da aquisição de efetores adicionais que suprimem a ETI. Por outro lado, as plantas, podem evoluir e adquirir novos genes R que codificam proteínas de resistência capazes de reconhecer os efetores, desta forma a ETI pode ser acionada novamente (Jones and Dangly, 2006).

As respostas das plantas à infecção envolvem a despolarização da membrana plasmática, a modificação da atividade de canais iônicos, a produção de compostos antimicrobianos e espécies reativas de oxigênio (ROS), fosforilação reversível de proteínas através da ativação de cascatas de proteínas homólogas à quinase ativadas por mitógenos (MAPK) ou por proteínas quinases dependentes de cálcio (CDPKs), modulação da transcrição de genes, ativação de cascatas de sinalização mediada por hormônios e deposição de compostos como a lignina gerando calose na parede celular vegetal (Meng and Zhang, 2013). Os hormônios vegetais, incluindo jasmonato (JA),

ácido salicílico (SA), ácido abscísico (ABA) e etileno (ET), são moléculas de sinalização importantes para uma resposta eficiente a estímulos bióticos. Normalmente, a sinalização por SA está relacionada com a resistência a patógenos biotróficos e hemibiotróficos, e as vias de ET e JA ativam a resistência contra patógenos necrotróficos (Pieterse *et al.*, 2012).

A modulação da transcrição de genes é fundamental para montar uma resposta de defesa eficaz nas células do hospedeiro. A reprogramação transcricional da célula vegetal ocorre de modo a favorecer a defesa da planta em relação a outros processos celulares, como crescimento e desenvolvimento. Este processo envolve os fatores de transcrição que modulam a expressão de um grande número de genes (Buscaill and Rivas, 2014). Os fatores de transcrição regulam a expressão de genes através de ligação sequência-específica com o DNA e interações proteína-proteína. Esses fatores interagem com a maquinaria geral de transcrição, com proteínas de remodelamento de cromatina e/ou com outros fatores de transcrição atuando como ativadores ou repressores da expressão gênica (Latchman, 2003).

Os fatores de transcrição são classificados em famílias de acordo com os motivos conservados presentes nos domínios de ligação ao DNA. A grande maioria dos fatores de transcrição contém domínios característicos de ligação ao DNA, de oligomerização, de regulação da transcrição e de localização nuclear (Liu *et al.*, 2001).

2.2. Família bZIP

A família de transfatores bZIP possui como característica principal a presença de um domínio básico de zíper de leucina (*Basic Leucine Zipper*). O domínio bZIP é formado por duas estruturas em forma de alfa hélice, sendo a primeira constituída de uma região básica de aproximadamente 16 resíduos de aminoácidos contendo um sinal de localização nuclear seguido de um motivo N-x7-R/K. A segunda região é formada por alguns resíduos de leucina espaçados regularmente em intervalos de seis aminoácidos, posicionados exatamente nove aminoácidos em direção ao C-terminal, criando uma hélice anfipática. Para a ligação ao DNA, as duas subunidades se ligam através de interações hidrofóbicas de suas hélices, formando homo ou hetero-dímeros, dando origem ao zíper de leucina. Alguns representantes da família bZIP possuem

isoleucinas, valinas, fenilalaninas ou metioninas no lugar das leucinas (Jacob *et al.*, 2002; Schütze *et al.*, 2008).

As respostas mediadas por bZIP são reguladas pela homo- ou heterodimerização, mecanismo essencial para a ligação desses fatores ao DNA e modulação da afinidade por *cis*-elementos, para assim determinar a via de sinalização a ser desencadeada em resposta um determinado estresse. (Sornaraj *et al.*, 2016). Membros da família bZIP podem regular vários processos na planta, como maturação de sementes, desenvolvimento de flores, fotossíntese, resposta a estresse abióticos e a defesa contra patógenos, dentre outros (Jacob *et al.*, 2002; Zhang *et al.*, 2016).

O genoma de *Arabidopsis thaliana* apresenta 116 genes que codificam fatores bZIP, enquanto que no genoma da soja há 176 genes, podendo estes números variar de acordo com a metodologia utilizada para seleção das sequências (Wang *et al.*, 2010). Em termos de proteína, 127 já foram identificadas em *A. thaliana* e 160 em soja (Ali *et al.*, 2016; Zhang *et al.*, 2018). Os bZIPs de *A. thaliana* foram subdivididos em dez grupos (A, B, C, D, E, F, G, H, I, S) de acordo com a similaridade de sequência e motivos conservados (Jackoby *et al.*, 2002). Posteriormente os bZIPs de soja foram subdivididos nestes mesmos subgrupos, com adição do subgrupo U (Wang *et al.*, 2015; Liao *et al.*, 2008). Porém em 2018, Zhang *et al* realizou outra divisão dos bZIPs de soja, adicionando os subgrupos K e J aos subgrupos de *A. thaliana* (Figura 1). Nesta nova divisão, 160 bZIPs de soja foram identificados através da comparação com sequências de bZIPs de *Arabidopsis*. Os 160 genes então distribuídos de forma desigual entre os 20 cromossomos de soja, o tamanho total das sequências das proteínas varia de 95 a 853 resíduos de aminoácidos, com CDS (*Coding Sequence*) de 288 a 2562 nucleotídeos, e massa molecular relativa de 10,93 a 157,28 kDa.

Dentro da família bZIP existe um sub-grupo bem caracterizado denominado TGA (proteínas de ligação ao *cis*-elemento TGACGTCA) (Singh *et al.*, 2002). Os TGAs atuam como reguladores da sinalização de SA e estão envolvidos na resposta a estresses bióticos. Os membros da família TGA interagem com uma proteína com repetições de anquirina (*ankyrin repeat protein, non-expresser of Pathogen Related* (PR) genes, NPR1), que é um componente chave da via de defesa sinalizada por SA (Pieterse *et al.*, 2012). Sob condições normais, a maior parte do NPR1 é retido no citoplasma como um oligômero através de ligações dissulfeto intermoleculares. Após o ataque de patógenos, SA é sintetizado e induz mudanças no estado redox da célula

através da produção de ROS. O aumento de ROS induz a monomerização de NPR1 através da atividade das tiorredoxinas H3 e H5 (TRX-H3/H5). Nas células induzidas por SA, os monômeros de NPR1 são translocados para o núcleo onde interagem com membros da família TGA que se ligam a promotores de genes responsivos a SA (Pieterse *et al.*, 2012).

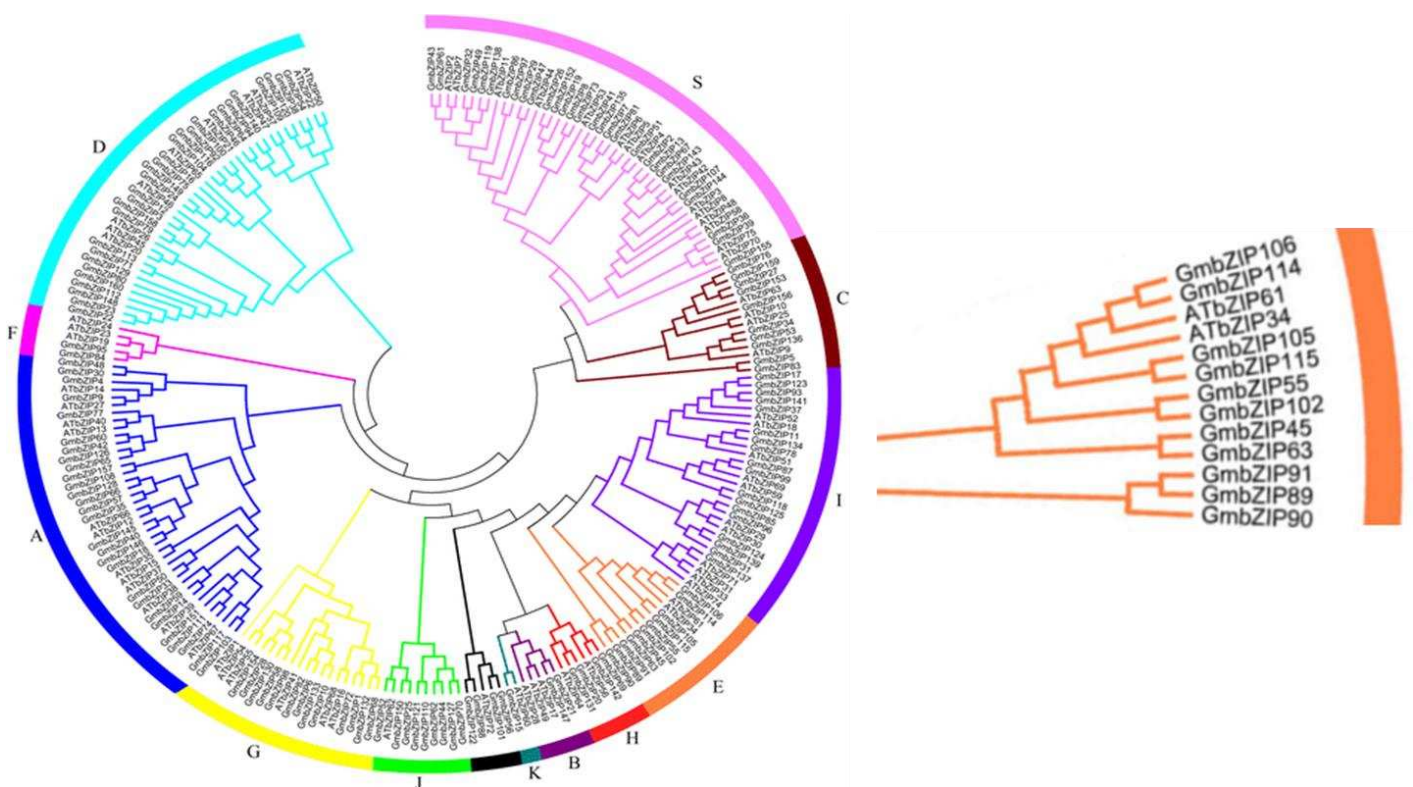


Figura 1. Classificação funcional das proteínas bZIPs de *G.max* e *A. arabidopsis*. Os bZIPs de soja foram classificados em 12 grupos (demonstrado em diferentes cores na árvore) e os bZIPs que não se encaixavam em nenhum grupo estão agrupados na parte preta da árvore. O grupo E está evidenciado, e o GmbZIPE2 é nomeado GmbZIP45 nesse trabalho (Zhang *et al.*, 2018).

2.3. GmbZIPE2

Nosso grupo de pesquisa identificou o fator de transcrição Glyma05G168100 (www.phytozome.net/soybean.php) da família bZIP, que foi posteriormente denominado GmbZIPE2. O GmbZIPE2 é classificado como membro do grupo E de

bZIPs (Alves *et al.*, 2015; Zhang *et al.*, 2018), um dos grupos com menos informações na literatura. Alguns membros desse grupo já foram identificados em soja (Liao *et al.*, 2008; Wang *et al.*, 2015; Zhang *et al.*, 2018), mas nenhum foi caracterizado.

Alves *et al.* 2015 identificou que *GmbZIPE2* apresentava expressão diferenciada em resposta à infecção por *P. pachyrhizi*, e que a proteína GmbZIPE2 apresenta similaridade de sequência de 66% com o transfator TabZIP1 de trigo (Zhang *et al.*, 2009; Alves *et al.*, 2015). TabZIP1 é responsivo a infecções por fungos de ferrugem (Zhang *et al.*, 2009), o que corrobora com a hipótese da função deste bZIP estar relacionado a defesa da soja contra patógenos. Além disso, já foi demonstrado que *GmbZIPE2* é expresso em diferentes tecidos da soja, é induzido por hormônios de defesa como SA e JA em alguns pontos do tratamento, e não tem sua expressão diferenciada em resposta a seca e inundação (Alves *et al.*, 2015, Zhang *et al.*, 2018).

Gonçalves (2014) mostrou que GmbZIPE2 é capaz de se ligar especificamente ao *cis*-elemento H-box *in vitro*, e, durante sua superexpressão em protoplastos de soja, um gene chalcona sintase é ativado. Além disso, *GmbZIPE2* é ativado nos estágios iniciais da infecção por *P. pachyrhizi* em plantas suscetíveis (Alves *et al.*, 2014), concomitantemente com a indução da expressão de enzimas relacionadas à via biossintética dos fenilpropanóides.

2.4. Relação de *GmbZIPE2* com fotossíntese e estresse oxidativo

Geralmente a ação combinada de fatores de transcrição em uma via de sinalização depende de interações destes com outras proteínas. Os fatores de transcrição bZIP são regulados por interações proteína-proteína, seja por homo ou heterodímeros, que ocorrem para que haja a formação do zíper de leucina e então a ligação do bZIP no promotor do gene que ele irá modular. A compreensão do mecanismo da interação é essencial para a elucidação da função do fator de transcrição em diferentes processos biológicos. Na busca de parceiros de interação do GmbZIPE2, Fontes (2016) realizou um *screening* de bibliotecas e identificou por duplo-híbrido duas proteínas que interagem com GmbZIPE2. A primeira (Glyma08g335500) é identificada, no Phytozome (<https://phytozome.jgi.doe.gov>), como Chaperonin-like Rbcx Protein, e a segunda (Glyma09g23260) não é identificada no Phytozome, porém no UNIPROT a

proteína é relacionada à montagem do fotossistema II (PSII). Ambas as proteínas identificadas são aparentemente proteínas plastidiais, sendo a segunda uma potencial proteína integral da membrana. Fontes (2016) identificou por BLASTX que a segunda proteína apresenta 66% de identidade com a proteína Predicted Low Psii Accumulation 2, chloroplastic de *Vigna radiata var. radiata*. No Phytozome cinco genes no genoma da soja são classificados como “low PSII accumulation”. No alinhamento das sequências de resíduos de aminoácidos das proteínas de soja e *A. thaliana*, a proteína não caracterizada de *G. max* é mais similar à LPA2 de *Arabidopsis*. Com a união dos dados de Gonçalves (2014), Fontes (2016) propôs um modelo de atuação de GmbZIPE2 em condições de estresse (Figura 2). Em condições normais, a GmbZIPE2 estaria interagindo com a LPA, que está ancorada à membrana. A fotossíntese ocorre normalmente, com tradução de Rubisco, e então RbcX estaria atuando como chaperona da Rubisco, interagindo com as RbcLs. Em uma situação de estresse, a produção de ROS pode induzir inibição dos fotossistemas e mecanismos de reparo. Nessas situações, LPA deixa de interagir com o GmbZIPE2, liberando-o no estroma, para auxiliar na montagem do PSII. Com a diminuição da fotossíntese (para priorizar a defesa ao estresse) e consequente diminuição da Rubisco, RbcX pode estar livre para interagir com o GmbZIPE2, que agora se encontra no estroma do cloroplasto. Essa interação pode desencadear a sinalização da resposta a estresse, fazendo com que haja a translocação de GmbZIPE2 para o núcleo e então a ativação de genes de defesa, como chalcona sintase.

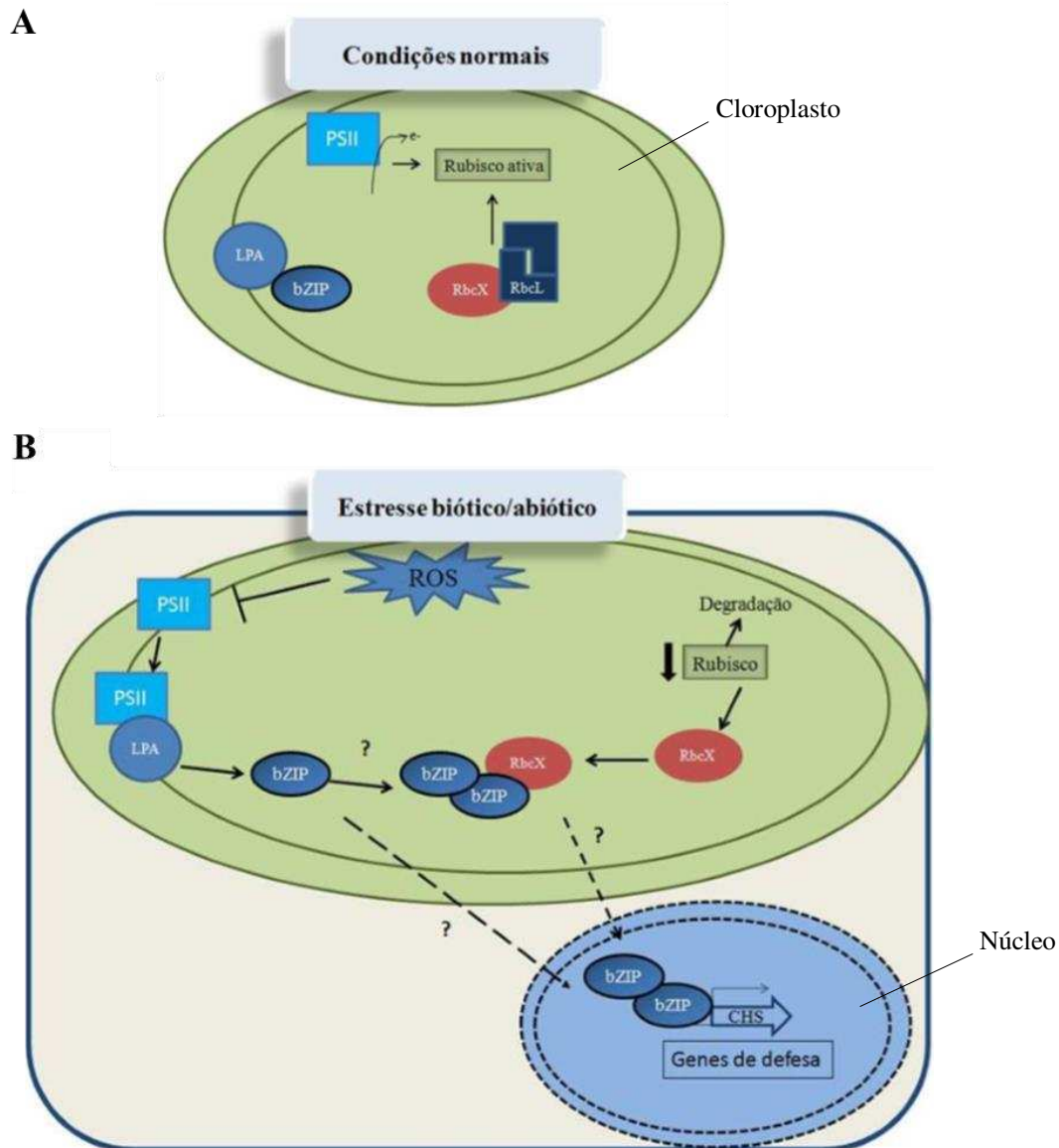


Figura 2: Modelo esquemático da atuação inferida do fator de transcrição GmbZIPE2 em resposta a estresse. A. Possível atuação de GmbZIPE2 em condições normais da célula. GmbZIPE2 interage com a LPA. A fotossíntese ocorre normalmente, com tradução de Rubisco, e então RbcX (interagindo com RbcLs) estaria atuando como chaperona da Rubisco. **B** Possível atuação de GmbZIPE2 em uma situação de estresse. O PSII pode ser inibido pela produção de ROS, e então, mecanismos de reparo são ativados. A LPA deixa de interagir com GmbZIPE2, liberando-o no estroma, para auxiliar na montagem do PSII. Com a diminuição da fotossíntese e consequente diminuição da Rubisco, RbcX pode estar livre para interagir com o GmbZIPE2 no estroma do cloroplasto. Essa interação pode desencadear a sinalização da resposta a estresse, fazendo com GmbZIPE2 seja translocado para o núcleo, desempenhando seu papel na ativação de genes de defesa, como chalcona sintase.

3 - OBJETIVO

O principal objetivo desse trabalho foi analisar o efeito da superexpressão do gene GmbZIPE2 em protoplastos de soja, através de um transcriptoma, e, analisar o silenciamento deste gene, através do silenciamento gênico induzido por vírus (VIGS), em folhas de soja.

4 - MATERIAIS E MÉTODOS

4.1 - Obtenção de protoplasto de soja e superexpressão de *GmbZIPE2*

O gene *GmbZIPE2* foi clonado inicialmente no vetor pDONR201 através do sistema *gateway*. A clonagem no vetor CaMV-35S-pCAMBIA-YFP-pnos1300 foi realizada por recombinação, utilizando a enzima *LR* clonase (*Invitrogen*), obtendo a construção *pYFP-GmbZIPE2*. As clonagens foram realizadas segundo as instruções do fabricante.

Para a obtenção dos protoplastos, folhas de soja da cultivar CD206 em estágio de crescimento V5/V6 foram coletadas e higienizadas através de borrifação, utilizando álcool 70%, seguido de água destilada, água sanitária 1% e novamente água destilada. Em seguida, as folhas foram picotas com estilete em pequenos pedaços (aproximadamente 1 cm²) e agrupadas em placas de petri de 140×15 mm, onde se adicionou 15 mL da solução enzimática (Celulase 2%, Pectoliase 0,1%, Manitol 0,4M, MES pH 5,8, 5mM) em cada placa. As placas foram incubadas sob agitação a 30°C, 60 rpm por 4 horas no escuro. Após esse período as placas foram lavadas com tampão de lavagem (Manitol 0,4 mM, HEPES-KOH 5 mM, pH 5,5) e a solução filtrada em uma peneira de 65 µm. O filtrado foi dividido em tubos falcons de 15 mL e centrifugado a 700xg, 25°C por 3 minutos e o *pellet* lavado duas vezes com tampão de eletroporação (HEPES-KOH 25 mM, KCl 10 mM, MgCl₂ 15 mM, Manitol 0,6 mM, pH 7,2 ajustado com KOH). A solução foi sendo concentrada nas lavagens para o volume usado na eletroporação (aproximadamente 400 µL por amostra) e deixada no gelo por aproximadamente 60 minutos. Após esse período as células foram eletroporadas (solução de protoplasto diluída em tampão de eletroporação acrescida de 10 µg da construção *pYFP-GmbZIPE2* ou pYFP vazio e 40 µg de ssDNA, *single-stranded DNA*, em um total de 800 µL) na condição de 25µF de capacitância e 2500V (*Gene Pulser-BIORAD*). As células foram mantidas por 15 minutos no gelo e em seguida adicionadas a 8mL de solução estéril de MS soja pH5,5 suplementado com 0,6 M de Manitol e Ampicilina 100 µg/mL (Sigma) e vertidas em placas de petri estéreis. As células foram induzidas por 36 horas em um ambiente úmido e escuro, e ao fim da indução, as células foram centrifugadas a 700 xg por 10 minutos e os protoplastos foram coletados. O material foi imediatamente destinado a extração de RNA. Cada amostra é representa um pool de protoplastos derivados de um pool de folhas de soja advindas de diversas

plantas, todas no mesmo estágio de crescimento. Foram gerados quatro amostras finais referentes aos controles e quatro amostras referentes à OE de GmbZIPE2. Esse experimento foi realizado por diversas vezes, e sempre a validação é feita por RT-qPCR.

4.3 - Extração, quantificação e qualidade do RNA

O RNA foi extraído utilizando o kit *RNeasy Plant Mini* (Qiagen) segundo recomendações do fabricante. A qualidade do RNA foi analisada em gel de agarose 1,5 % (p/v), corado com brometo de etídeo 0,1 µg/mL. Os RNAs com boa qualidade e quantidade adequada foram utilizados para produzir *pools*, e cada *pool* contia RNAs obtidos de duas transformações independentes. Ao final foram analisados dois *pools* de amostras superexpressando *GmbZIPE2* e dois *pools* com o controle (vetor vazio). Cada *pool* foi gerado a partir da junção de duas amostras iniciais. Os RNAs foram quantificados por espectrofluorometria através do aparelho Qubit 3.0 (ThermoFisher) utilizando o kit *Qubit RNA Broad Range Assay* (Invitrogen), seguindo as especificações do fabricante. Com os valores da quantificação, os RNAs foram analisados no Agilent 2100 Bioanalyser System, a partir do *kit Agilent RNA 6000 Nano* (Agilent) seguindo as especificações do fabricante.

4.4 - Tratamento com DNase I e síntese de cDNA

Para eliminação de possível contaminação com DNA, o RNA foi tratado com *DNase I, Amplification Grade* (Invitrogen) e quantificado novamente no *Qubit*. A eficácia do tratamento com DNase I foi confirmada a partir de uma reação de PCR utilizando o oligonucleotídeos do gene endógeno ELF1A (Tabela 1) e posterior verificação em gel de agarose 1,5%. A síntese de cDNA foi realizada utilizando 2 µg de RNA total livre de DNA e o *kit High-Capacity cDNA Reverse Transcription* (Applied Biosystems) foi utilizado segundo especificações do fabricante. A eficácia da síntese de cDNA foi confirmada por reação de PCR utilizando o par de primer referente ao gene endógeno ELF1A (Tabela 1) seguido de verificação em gel de agarose 1,5%.

4.5 - PCR em Tempo Real (RT-qPCR)

4.5.1 - Desenho dos oligonucleotídeos

Os oligonucleotídeos foram desenhados através do programa *PrimerPlus3*, utilizando a sequência de codificação de cada gene. As principais configurações utilizadas foram: tamanho de fragmento gerado entre 80 e 120 pares de base, tamanho do primer entre 18 e 23 nucleotídeos (sendo 20 o número ótimo), temperatura de *melting* de 59 a 61°C (sendo 60°C o ótimo), porcentagem de CG de 40 a 80 (sendo 50% ótimo). O programa *Oligo analyser1.0.3* foi utilizado para verificação de formação de *hairpins* e anelamento entre os primers. A verificação do alinhamento dos primers somente no gene de interesse foi realizado pelo e-PCR.

4.5.2 - Eficiência dos oligonucleotídeos

Os oligonucleotídeos utilizados para análise de expressão gênica tiveram a eficiência da reação determinada através de uma curva padrão realizada por meio de RT-qPCR. Para determinar a condição ideal de eficiência da reação, diferentes concentrações dos oligonucleotídeos foram analisadas para par, utilizando em cada reação 100 ng de cDNA de planta sadia ou 50 ng de DNA genômico (quando a expressão do gene era baixa). A curva padrão foi realizada com a concentração do oligonucleotídeo que resultou em maior amplificação, desde que sua curva de *melting* estivesse adequada. Desta forma, para obter curva padrão foram realizadas reações utilizando-se cDNA de plantas de soja sadia previamente obtido em 5 diluições seriadas de 1:10 partindo de 100 ng de cDNA ou 1:5 partindo 50 ng de DNA genômico. O *slope* (coeficiente angular da regressão linear) foi obtido a partir da curva padrão e foi utilizado para calcular a eficiência de amplificação da reação a partir da equação $E = (10^{-1/slope} - 1)$. Para toda diluição de um primer, um branco era testado para evitar contaminações. As eficiências e as concentrações de todos os oligonucleotídeos utilizados estão descritas na Tabela Suplementar 2. Para a escolha do gene endógeno mais adequado, foram testados oligonucleotídeos referentes a genes endógenos para cada experimento (Tabela 2).

4.5.3 - Reação do RT-qPCR

As reações de RT-qPCR foram realizadas segundo os manuais da *Applied Biosystems* e conduzidas no modelo de termociclador *Step one (Applied Biosystems)*. As reações foram realizadas em um volume total de 10 μL , sendo 5 μL de *Power SYBR Green Master Mix (Thermo Fisher Scientific)*, 1 μL de cDNA (concentração variou entre tratamentos) e 4 μL de oligonucleotídeos (concentração variou entre os pares de oligonucleotídeos). Para a reação de amplificação foi condicionado: 95 °C por 10 minutos, 40 ciclos de 94 °C por 15 segundos e 60 °C por 1 minuto. As amostras eram pipetadas em triplicatas biológicas e em duplicatas técnicas.

4.5.4 - Quantificação da expressão gênica

Para a quantificação da expressão gênica foi utilizado o *software REST (Relative Expression Software Tool)* (www.REST.de.com). Esse *software* compara os dois grupos (controle e tratamento), utilizando o gene endógeno como normalizador dos dados. O modelo matemático baseia-se nas eficiências dos *primers* e no desvio médio do ponto de cruzamento (ΔCq) entre a amostra e o grupo de controle. Os métodos integrados de randomização e *bootstrapping* usados nesse *software* testam a significância estatística das razões de expressão calculadas. Como resultado, o *software* informa se o tratamento induziu ou reprimiu determinado gene e informa uma faixa de erro padrão (Pfaffl, 2001; Pfaffl *et al*, 2002; Vandesompele *et al*, 2002). Os gráficos foram confeccionados com o GraphPad Prism 8.

Tabela 1: Oligonucleotídeos utilizados no qRT-PCR

ID	Nome do gene	Orientação do oligonucleotídeo	Sequência (5'-3')
Glyma.05G168100	GmbZIPE2 (VIGS)	Fw	CAACCACAGACACAGCCACAA
		Rv	GCGAACTGAGATTCGAGGTCTCT
Glyma.06G028100	Fe ⁺³ -Zn ⁺² PAP	Fw	TATCATTTACCGTACACAGC
		Rv	ACCAAAAAGCTGAGTAAGAAG
Glyma.11G048300	ATPase transportadora de Ca ⁺²	Fw	TGTAACAGGTGATGGAACATA
		Rv	CAGAATTATGACATCTGCAC
Glyma.03G168000	Transportador ABC	Fw	GCTGTTATATCAGAGGAACC
		Rv	

Glyma.19G201600	FLA15	Rv	CGTTGTATGTTCTTCAACAC
		Fw	CTCTCCATCTATGGTGTCT
Glyma.14G094700	GLDC	Rv	TTATGAGAGAATGCAGAGTG
		Fw	CCATTCTGAGGCGTCTGGTT
Glyma.13G272000	EF-hand ligante de Ca ⁺²	Rv	AACAGAGGACATGCACCTGG
		Fw	TTGGCCCTTGTCTATGCGAG
Glyma.11G140500	MED37e	Rv	TAGAGGGAAGACAGTGCGGA
		Fw	GAGAAACCTATGATTGTGGT
Glyma.16G165700	TRP	Rv	CCTTCATCTTCATAAGAACC
		Fw	GATCTGCATTAGTAGGGAAC
Glyma.01G153300	ABA 8' hidrolase	Rv	ATATGTTGCTTTCTTGGAGT
		Fw	CTAACTTGAGAAGGCTTGTT
Glyma.14G200200	WRKY33	Rv	AAGACTTTAAGCAATCTTGG
		Fw	GTTGAGAGGTCTTTAGATGG
Glyma.06G217900	D2 do PSII	Rv	GAGAAGAGGAGTTTGATGAG
		Fw	FGGTACGACCTTTGTAACCTC
Glyma.09G232600	Provável LPA	Rv	ACTATTAGCAGGAGTGGAGA
		Fw	TGGATCATCGGACACTGCTG
Glyma.02G309300	PAL-1	Rv	ATGCAGGTTTTTCAAGCGGG
		Fw	CAACAGCCAGATCCATTGCA
Glyma.01G228700	CHS7	Rv	ATCTGTGGACCAAGCCACTGA
		Fw	TGTCAAGTGCTTGTGTTCTTTTCA
Glyma.11G011500	CHS8	Rv	TTCACCTGTGTTTTATGTCCATT
		Fw	CAGGCACAAAGGGCAGAAG
		Rv	GTGCTCTGATCAACACAGTTTGGT

Tabela 2: Oligonucleotídeos referentes à genes endógenos utilizados no qRT-PCR

ID	Nome do gene	Orientação do oligonucleotídeo	Sequência (5'-3')
Glyma.12G02790	CYP2 <i>Cyclophilin 2</i>	Fw	CGGGACCAGTGTGCTTCTTCA
		Rv	CCCCTCCACTACAAAGGCTCG
Glyma.05G24110	ELF1A <i>Translation elongation factor 1-alpha</i>	Fw	GACCTTCTTCGTTTCTCGCA
		Rv	CGAACCTCTCAATCACACGC
Glyma.02G44460	ELF1B <i>Translation elongation factor 1-beta</i>	Fw	GTTGAAAAGCCAGGGGACA
		Rv	TCTTACCCCTTGAGCGTGG
Glyma.18G290800	ACT <i>Actin</i>	Fw	GTCTTTCGCTTCAATAACCTA
		Rv	CGGTGGTTCTATCTTGGCATC
Glyma.05G29000	TUA <i>Tubulin alpha-5</i>	Fw	AGGTCGGAAACTCCTGCTGG
		Rv	AAGGTGTTGAAGGCGTCGTG

4.6 - Construção, quantificação das bibliotecas

A construção das bibliotecas de DNA complementar dupla fita (cDNAs) a partir de 1 µg RNA total livre de DNA foi realizada utilizando o guia *Illumina Truseq strand total RNA sample preparation Guide (Low sample protocol)*. As etapas para a construção das bibliotecas através do kit foram: depleção do RNA ribossomal, purificação e fragmentação dos RNAs, síntese da primeira fita de cDNA, remoção da fita de RNA molde e síntese da segunda fita de cDNA, adição de uma adenina 3' no cDNAs, ligação dos adaptadores e, por fim, PCR para amplificação dos fragmentos.

Após a construção das bibliotecas, determinou-se a quantificação e a verificação da qualidade destas. A primeira quantificação das bibliotecas foi realizada no *Qubit* utilizando o *kit dsDNA High Sensitivity Assay (Invitrogen)*, seguindo as instruções do fabricante. Uma segunda quantificação das bibliotecas foi realizada utilizando o *kit Kapa Library Quantification Illumina Platforms (Kapa Biosystems)*, seguindo as instruções do fabricante, e a reação foi realizada no *Step one (Applied)*. Para essa quantificação foi feita uma curva padrão na faixa de 0,0002-20 pM e as amostras foram diluídas em Tris-HCl 10mM e 0,1 % de tween 20 nas proporções de 1:5000, 1:50000 e 1:500000. Esse método se baseia na amplificação dos fragmentos de biblioteca por RT-qPCR utilizando oligonucleotídeos (Tabela 1) que se ligam em uma sequência consenso existente em todos os adaptadores. A quantificação é realizada por comparação entre os valores de Cq (ciclo de quantificação) dos padrões e os valores de Cq das bibliotecas. A análise do tamanho médio dos fragmentos de biblioteca gerados foi feita no *Bioanalyser* utilizando o *kit Agilent High Sensitivity DNA (Agilent)*, seguindo as instruções do fabricante. Juntos, os dados da quantificação e tamanho médio dos fragmentos foram utilizados para aferir a qualidade e refinar a quantificação das bibliotecas.

4.7 - Sequenciamento das bibliotecas

As bibliotecas validadas controle 1, controle 2, GmbZIPE2.1 e GmbZIPE2.2 foram diluídas com Tris-HCl 10mM acrescido de 0,1 % de *tween* 20 para uma concentração de 10mM. Horas antes do sequenciamento as bibliotecas foram diluídas a uma concentração de 2 nM (também com *Tris e Tween*), que foi utilizada para a formação de um *pool*. O *pool* foi desnaturado e diluído para uma concentração de 13 pM de acordo com o manual *Denaturing and Diluting Libraries for the NextSeq™ 500*

(Illumina). O *pool* foi aplicado em uma *flowcell* que foi inserida no aparelho *Hiseq 2500* da Illumina do Núcleo de Análises de Biomoléculas – UFV, seguindo as recomendações dos manuais *Truseq Rapid SBS Kit (200 cycles)* e *Hiseq 2500 System User Guide (Chapter 4 Perform a Rapid Run)*. A primeira etapa é a de clusterização, utilizando o *kit TruSeq® PE Cluster v3* seguindo a recomendação dos fabricantes. O sequenciamento foi realizado utilizando o *Kits TruSeq Rapid SBS (200 Cycles)*, tipo de corrida *paired end*, no modo *Rapid*.

4.8 - Tratamento dos dados, análise de qualidade e montagem do transcriptoma

O sequenciamento gerou dados brutos no formato FASTq, que foram transferidos para o computador Júpiter da Diretoria de Tecnologia e Informação da UFV. Para a análise da qualidade desses dados utilizou-se o programa FastQC (versão 0.11.5, Andrews *et al.*, 2010). O FastQC analisa parâmetros das *reads* como quantidade total, comprimento, pontuação de qualidade e porcentagem de GC das sequências, conteúdo de adaptadores, conteúdo de N (nucleotídeos ambíguos) entre as bases, pontuação de qualidade de todas as bases.

Para a trimagem das *reads* foi utilizado o programa *Trimmomatic 0.36* (Bolger *et al.*, 2014). A trimagem é realizada para eliminação das sequências de baixa qualidade, sequências indesejadas (como os adaptadores) e *reads* de tamanho reduzido. O parâmetro mais importante do *Trimmomatic* é o valor *Phred*, que eliminam sequências de baixa qualidade. O valor de *Phred* representa a probabilidade de que o sequenciamento de uma determinada base esteja incorreta e deve ser selecionado após a verificação da qualidade dos dados brutos, quanto mais sequências de baixa de qualidade, maior o valor de *Phred* será utilizado. Para as análises foi utilizado o valor de *Phred* 20, isto é, esse parâmetro considera que existe um erro a cada 100 bases sequenciadas ou seja, 99% de precisão (Delhomme *et al.*, 2014).

Após a trimagem os dados foram novamente analisados no FastQC, para verificar se a qualidade das *reads* melhoraram e se era preciso mudar algum parâmetro do *Trimmomatic*, a fim de selecionar *reads* com alta qualidade para as próximas etapas.

O *TopHat* versão 2.1.1 é um algoritmo de mapeamento de *reads* que é capaz de alinhar *reads* de um RNA-seq contra um genoma de referência (Trapnell, 2009). No

trabalho, esse programa foi utilizado para analisar a quantidade de *reads* do transcriptoma que alinharam no genoma de soja. Para a utilização desse programa, inicialmente, o programa *Bowtie2* versão 2.2.8 (Langmead *et al.*, 2012), a partir da ferramenta *bowtie2-build*, foi utilizado para indexar o genoma de referência da soja (*G.max* GCA 000004515.3 *Glycine-max-v2.0*) depositado no NCBI (Benson *et al.*, 2014). Com o genoma de referência indexado, o *TopHat* realizou o mapeamento das *reads* no genoma da soja, e estas foram contabilizadas.

O mapeamento do transcriptoma foi realizado através do programa *Kallisto* (Bray *et al.*, 2016). O *Kallisto* é um programa que quantifica a abundância de transcritos a partir de dados de RNA-seq. O programa se baseia em pseudoalinhamentos para determinar a compatibilidade das *reads* com os alvos, sem a necessidade de alinhamento (Bray *et al.*, 2016).

Os GOs (*Gene Ontology*) foram confeccionados no site Soybase (soybase.org).

4.9 - Validação do RNA-seq

A validação do transcriptoma de protoplasto de soja superexpressando *GmbZIPE2* foi realizada através RT-qPCR (como descrito no 4.1) e os genes selecionados e seus oligonucleotídeos correspondentes estão listados na Tabela 1. A seleção dos genes foi direcionada a genes UP-regulados e que tinham anotação.

4.10 - Silenciamento do gene *GmbZIPE2* através do ensaio VIGS

O ensaio de silenciamento gênico induzido por vírus (VIGS) foi realizado na Embrapa Soja em Londrina. A sequência de 252 pares de bases do gene *GmbZIPE2* foi amplificada por PCR utilizando os primers descritos na tabela 3 e DNA extraído da cultivar Williams 82. Os primers foram desenhados no programa Primer3plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). O vetor viral BPMV IA-1033 e os fragmentos de PCR foram previamente digeridos com enzimas de restrição XhoI e BamHI. O inserto foi inserido no vetor utilizando a enzima a T4 ligase, e a reação foi incubada por 16 horas a 4°C. A reação foi confirmada por PCR. A transformação ocorreu por choque térmico utilizando a cepa DH5α de *Escherichia coli*, e os transformantes foram confirmados por PCR de colônia. Os clones positivos foram utilizados para a extração de DNA plasmidial utilizando o kit *QIAprep® Miniprep*

(Qiagen) segundo as recomendações do fabricante. As amostras foram quantificadas no espectrofotômetro *NanoDrop ND-1000* e analisadas em gel de agarose 1%.

A inoculação do vetor viral foi realizada pelo método de biobalística utilizando o bombardeador de micropartículas (*Gene gun*) do Laboratório de Biotecnologia Vegetal da Embrapa Soja. Inicialmente, as partículas de ouro foram higienizadas com etanol 100%, em seguida, ressuspensas em água e armazenadas a -20°C até o momento da preparação destas. Para a infecção eficiente do vetor viral é necessária a presença dos dois componentes RNA1 e o RNA2 (que está com o inserto), e do SMV (*Soybean Mosaic Virus*) para que a visualização dos sintomas seja clara. As partículas de ouro foram acrescidas do DNA plasmidial, RNA1 e SMV para a preparação destas para o bombardeamento. Em seguida, essa preparação foi sonificada, e nela adicionada uma mistura de glicerol 50%, CaCl₂ 2,5M e espermidina 0,1M. Essa suspensão foi lavada com isopropanol e aplicadas no centro da membrana macrocarreadora.

Tabela 3. Oligonucleotídeos específicos para a clonagem do gene *GmbZIPE2*

Nome	Sequência (5' - 3')
Fw Glyma.05G168100.1	ATATCTCGAGAGCACTCAAGCAACGTTTAG
Rv Glyma.05G168100.1	AAATGGATCCATTCTGAGAGCAGGGTCTTG

A membrana foi integrada no aparelho *Gene gun*, que utiliza uma pressão de 1100 psi para o bombardeamento da amostra. Plantas de soja da cultivar Williams 82 com aproximadamente 10 dias após a germinação foram bombardeadas. Em seguida, as plantas foram borrifadas com água destilada e mantidas por 2 dias em câmara de crescimento a 22°C, cobertas com saco plástico para manter a umidade. As plantas foram transferidas para vasos de 5 Kg e para casa de vegetação a 25°C. Após 3 semanas o enrugamento das folhas (sintoma característico do SMV) foi perceptível e assim folhas sintomáticas foram coletas e liofilizadas no Liotop *L101* por aproximadamente 10 horas e em seguida foram transferidas para tubos contendo sílica gel e serão mantidas a -20°C até a inoculação (*Rub-inoculation*).

Plantas do mesmo genótipo usado para produção de inóculo foram utilizadas no experimento de VIGS. O experimento foi conduzido em câmara climatizada (Fitotron) com temperatura controlada (22 °C durante o dia e 20 °C durante a noite), fotoperíodo

de 16 h de luz e aproximadamente 50 % de umidade relativa do ar. Foram realizadas três repetições. Duas semanas após a semeadura, foi realizada a inoculação mecânica (Rub-inoculation). Foram preparadas as soluções partindo de cerca de 100mg de folhas liofilizadas para 5mL de tampão fosfato K_2PO_4 50 mM, pH 7,0. As folhas liofilizadas foram maceradas juntamente com o tampão fosfato. Todas as plantas foram previamente pulverizadas com Carborundum (carbureto de silício), em seguida foi aplicada a solução (construção ou vetor vazio e tampão fosfato) nas duas folhas unifoliadas friccionando levemente com os dedos para ajudar a penetração da construção nas células. Por fim, as plantas foram borrifadas com água destilada e os vasos cobertos com plástico para manter a umidade por dois dias. Aproximadamente três semanas após a inoculação viral, os sintomas começaram a aparecer e uma folha do terceiro trifólio foi coletada para confirmação do silenciamento por RT-qPCR e imediatamente congeladas e armazenadas à -80 °C para posterior extração de RNA.

4.11 - Técnicas de Biologia Molecular

Todas as técnicas de Biologia Molecular que não estão descritas, como preparo de géis de agarose, transformação de bactérias, preparo de tampões e soluções estoques foram conduzidas de acordo com Sambrook *et al* (1989).

5 - RESULTADOS

5.1 - Qualidade do RNA e superexpressão de *GmbZIPE2*

Análises de transcriptoma requerem RNAs de alta qualidade. Primeiramente, verificou-se uma boa qualidade dos RNAs extraídos de protoplasto de soja no gel de agarose na Figura 1A. De um a quatro estão as amostras do controle e de quatro a oito estão as amostras de RNA extraídas de protoplastos superexpressando *GmbZIPE2*. É possível observar a integridade das bandas dos rRNAs das subunidades 28 S (banda superior) e 18 S do ribossomo (banda inferior). Para as etapas seguintes, foi feito um *pool*, concentrando as amostras em duas amostras de controle e duas amostras da superexpressão e o DNA contaminante foi removido.

Para uma verificação mais refinada da qualidade dos RNAs, foi realizada uma análise pelo *Bioanalyser*. Os resultados gerados possibilitaram a análise da qualidade dos RNAs através da intensidade dos picos correspondentes aos rRNAs 28 S e 18 S e através da definição desses picos (Figuras 1B e 1C). Na figura 1B pode-se observar que o gel do capilar referente à amostra E2.2 está mais escuro e com bandas inespecíficas. Na figura 1C é possível observar que somente na amostra E2.2 houve uma sobreposição dos picos, porém mesmo com a sobreposição foi possível diferenciar os dois picos, então esta replicata foi utilizada.

O experimento de superexpressão de *GmbZIPE2* foi confirmado através de qRT-PCR. *GmbZIPE2* apresentou uma expressão aproximadamente 200 vezes maior que o controle (Figura 4).

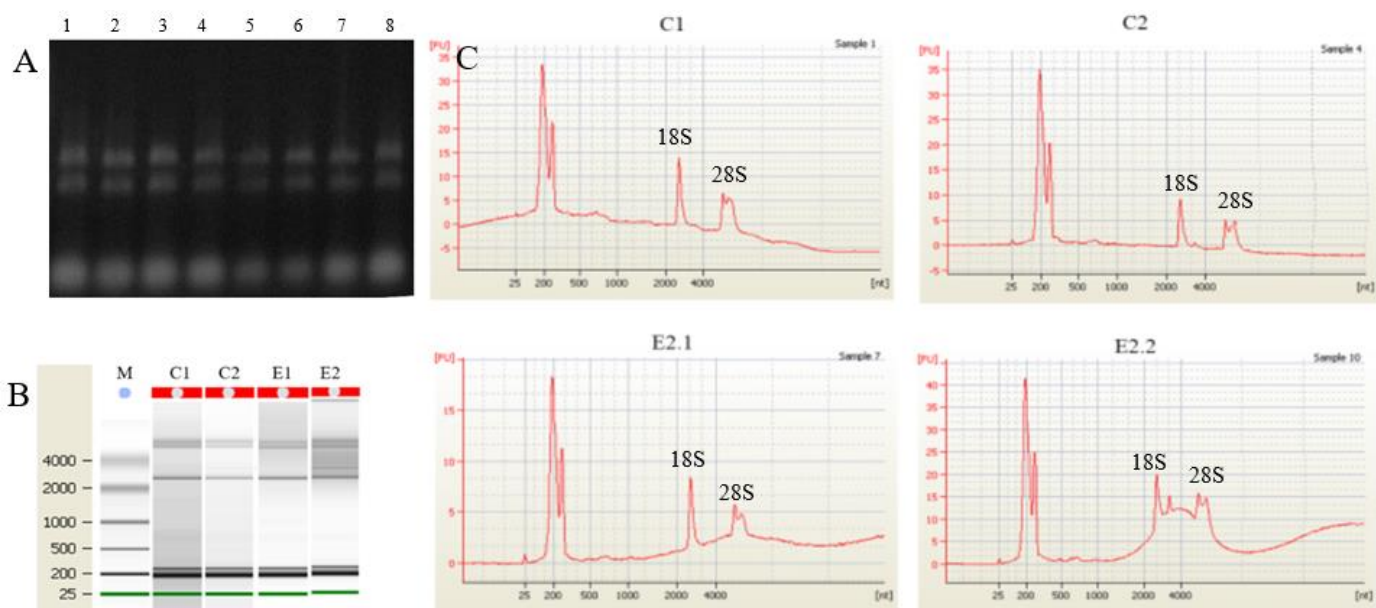


Figura 3. Qualidade do RNA extraído das amostras controle e de protoplastos superexpressando *GmbZIPE2*. A. Gel de agarose, 1-4 amostras controle; 4-8 amostras da *OE* de *GmbZIPE2*. B. Gel do chip de bioanalyser, onde M é o marcador e as amostras são referentes aos *pools* C1, C2, E2.1 e E2.2, C. Picos gerados no Bioanalyser referente aos *pools* C1, C2, E2.1 e E2.2.

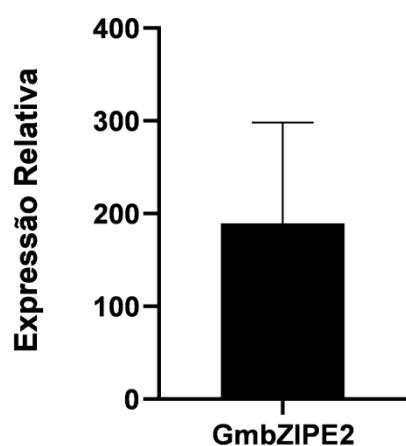


Figura 4: Superexpressão de *GmbZIPE2* em protoplastos de soja. A análise foi realizada pelo *software* REST, usando como endógeno o gene *ELF1A*. A expressão de *GmbZIPE2* foi estatisticamente diferente do controle.

5.2 - Qualidade e validação das bibliotecas

A qualidade das bibliotecas foi verificada por Bioanalyser e a Figura 5 mostra os gráficos de distribuição dos fragmentos de acordo com o tamanho em pares de bases (pb). O tamanho médio das amostras foi obtido através da análise manual (ajustamos as listas em azul para cada amostra) dos gráficos gerados pelo Bioanalyser, que foi 318 e 299pb para C1 e C2 e 285 e 336pb para E2.1 e E2.2 respectivamente. As concentrações foram calculadas através dos dados gerados pelo Bioanalyser e através da curva-padrão, e foi 257,205 e 86,155 nM para C1 e C2 e 242,195 e 251,105 nM para E2.1 e E2.2 respectivamente.

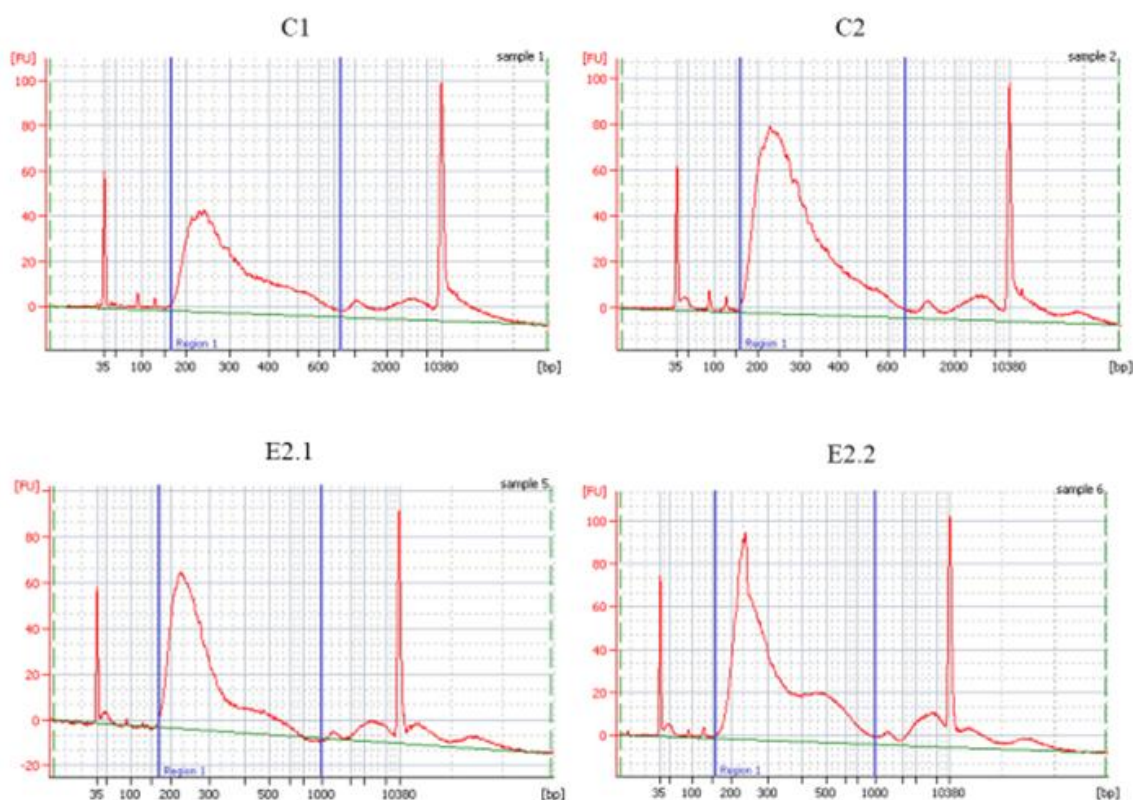


Figura 5. Análise da qualidade das bibliotecas. Gráficos gerados pelo Bioanalyser. O eixo X representa os tamanhos dos fragmentos em pares de pares (bp) e o eixo Y representa a intensidade da fluorescência.

5.3 - Tratamento de dados do sequenciamento

A trimagem com o valor de Phred 20 foi eficiente e as *reads* com baixa qualidade foram excluídas. O resultado da análise da qualidade das *reads* por FastQC está demonstrado na tabela 4 e nas figuras Suplementar 1 e 2.

Após o mapeamento das *reads* trimadas através do *TopHat*, foi observada uma contaminação das amostras. Em relação aos controles 60,7% de C1 e 32,1% das *reads* de C2 mapearam no genoma referência da soja. Em relação as amostras superexpressando *GmbZIPE2*, 20% em E2.1 e 12,2% em E2.2. As *reads* que não mapearam no genoma da soja alinharam com genomas de isolados de bactérias, majoritariamente da espécie *Pantoea ananatis*. Essa bactéria causa a mancha branca do milho, e provavelmente estava na superfície das folhas de soja. A soja utilizada no experimento foi crescida em casa de vegetação, que abrigava outras espécies de plantas, inclusive milho. Provavelmente a desinfestação realizada não foi eficiente para matar todas as bactérias, e então elas proliferaram no meio MS com sacarose. As análises posteriores foram realizadas somente com as *reads* que mapearam no genoma de soja.

Na primeira tentativa da realização do transcriptoma pelo *Kallisto*, o *Heatmap* e o *PCA* (análise dos componentes principais) demonstraram que havia uma dispersão das *reads* referentes aos controles, isto é, uma grande separação física dos replicatas dos controles no *PCA* (Figura 6). O valor de variância também reflete essa dispersão. Então para contornar esse problema, cada *lane* foi dividida em duas, na tentativa de uma melhora na dispersão dos dados. Após essa divisão, as *lanes* que estavam mais dispersas foram deletadas, e os novos *Heatmap* e *PCA* gerados demonstraram um agrupamento melhor das *reads*, assim como um melhor valor de variância (Figura 7).

ID	%CG	Dados brutos		Dados trimados	
		Total de sequências	Tamanho da sequência	Total de sequências	Tamanho da sequência
Controle1_L1	51	28217611	35-101	24054571	50-101
Controle1_L2	51	28122719	35-101	24179798	50-101
Controle2_L1	50	27844976	35-101	23639967	50-101
Controle2_L2	50	27647594	35-101	23704010	50-101
bZIPE2-1_L1	49	26534311	35-101	22721049	50-101
bZIPE2-1_L2	49	26406767	35-101	22831877	50-101
bZIPE2-2_L1	53	34350964	35-101	28688004	50-101
bZIPE2-2_L2	53	34211858	35-101	28842470	50-101

Tabela 4: Informações geradas pelo FastQC referentes à trimagem das *reads*.

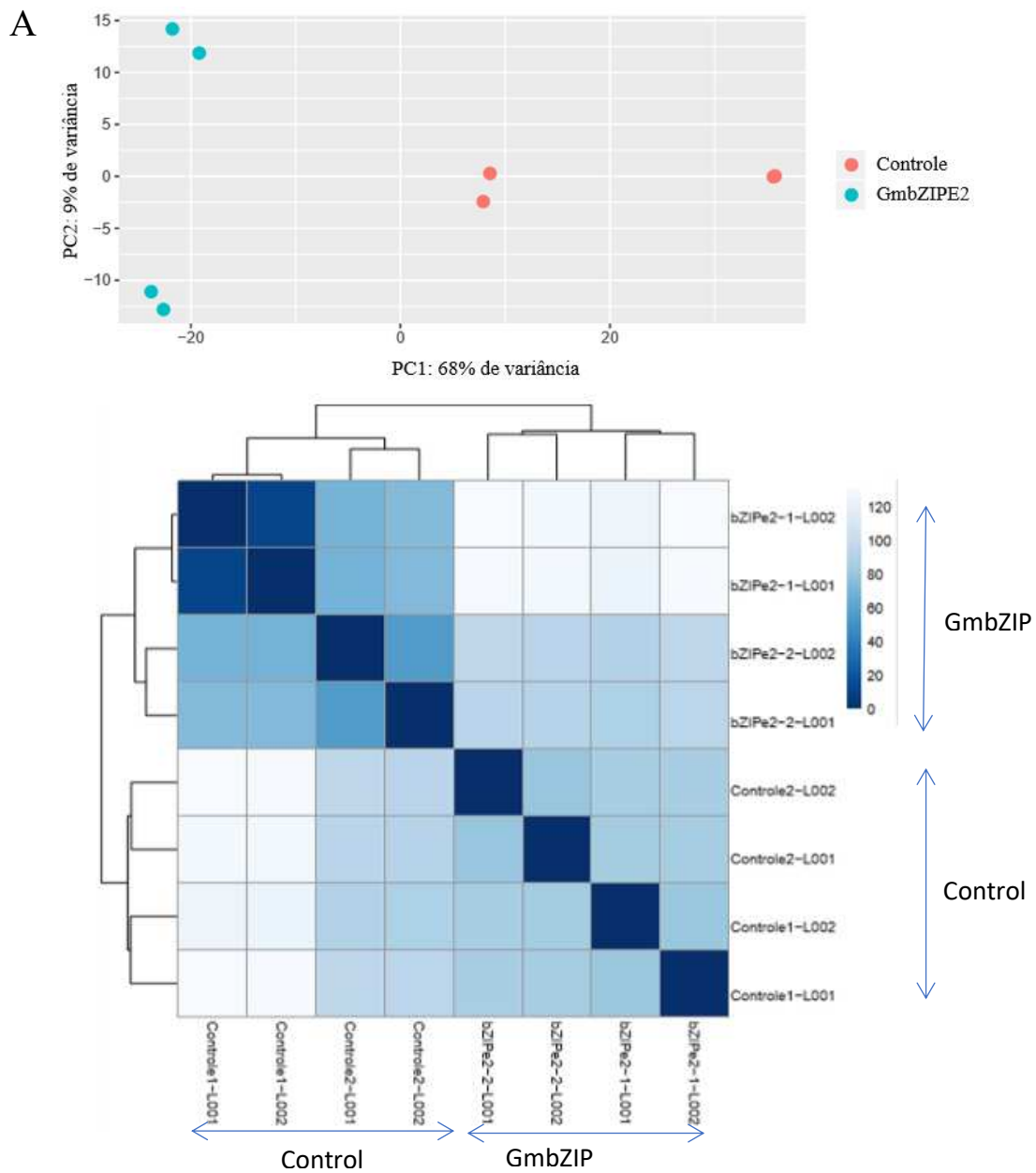


Figura 6: Análise da sobreposição das amostras. A. PCA e B. Heatmap iniciais.

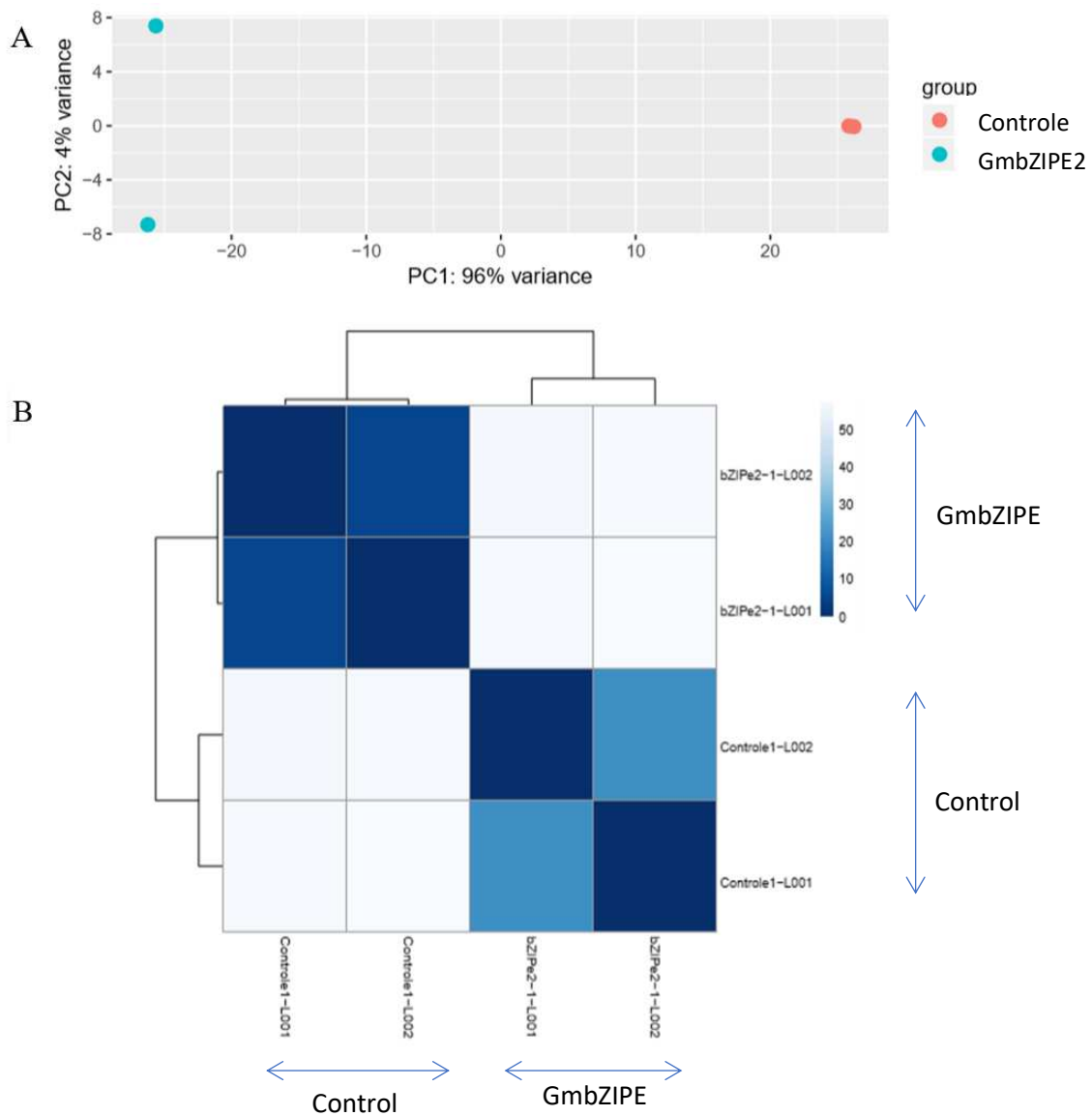


Figura 7. Análise da sobreposição das amostras após a deleção de *reads*. **A.** *PCA* e **B.** *Heatmap*.

5.4 - Transcritoma de protoplastos de soja superexpressando *GmbZIPE2*

Obtivemos o resultado de 671 genes *Up*-regulados e 47 genes *Down*-regulados. Uma grande parte dos genes *Up*-regulados estava anotado, e nos genes *Down*-regulados

somente um gene tinha anotação. Os gráficos de GOs dos genes ativados foram gerados no *SoyBase*, mostrando como esses genes são classificados em relação a processos biológicos, função molecular e componentes celulares (Figura 8). Em relação aos processos biológicos, 34,8% dos genes estão relacionados à fotossíntese, 16,1% à transdução de sinal e 8,7% à tradução. Em relação a função molecular, 19,4% dos genes são classificados como tendo a função de ligação a proteínas, 13,9% ligação à nucleotídeos; 13,4% atividade catalítica e 11,5% tem atividade de fator de transcrição, se ligando a sequencias específicas de DNA. Em relação a componentes celulares 21,4% estão localizados no núcleo, 13% na membrana plasmática, 12,1% em membranas (sem especificação) e 10,5% no citoplasma.

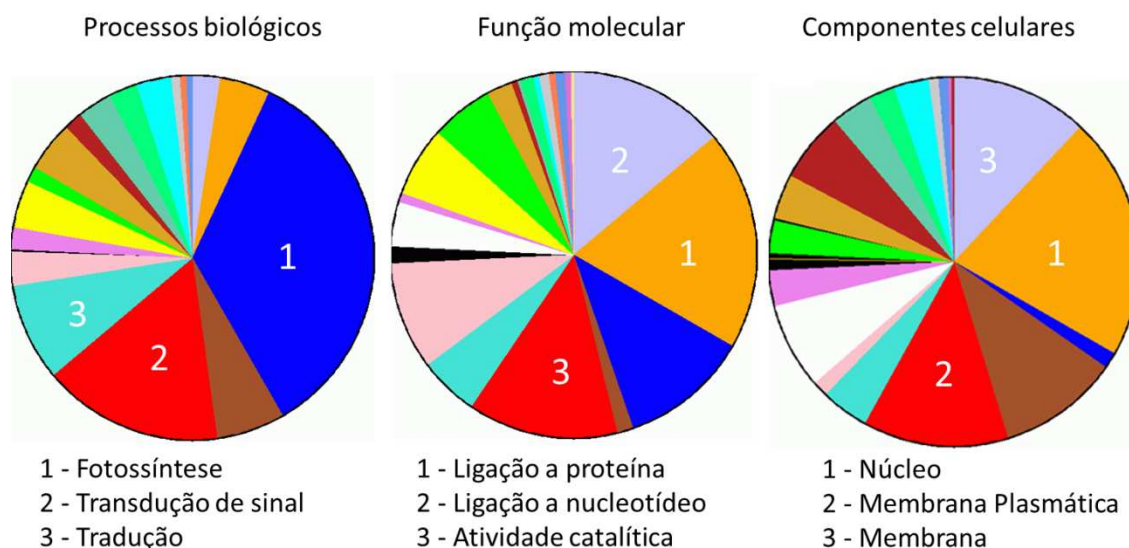


Figura 8. Distribuição dos genes ativados pela superexpressão de *GmbZIPE2* em relação a processos biológicos, função molecular e componentes celulares. Os GOs foram gerados pelo SoyBase (*GO term enrichment tool*) e as três maiores porcentagens de cada gráfico está representada através da numeração.

5.5 - Validação do transcriptoma

Para a validação do transcriptoma, um total de 20 pares de oligonucleotídeos foram sintetizados e a expressão avaliada. Todos esses oligonucleotídeos são relativos a

genes *up*-regulados, pois na relação de genes *down*-regulados somente um gene era anotado. Além disso, a expressão da maior parte dos genes ativados avaliados foi muito baixa nos controles, então o experimento seria inviável para os genes inibidos pela *OE* de *GmbZIPE2*. Os oligonucleotídeos selecionados estavam entre os 100 mais expressos da lista de *up*-regulados e todos tinham anotação. Desses 20, em nove não foi possível calcular a eficiência, ou a eficiência não foi satisfatória (Tabela Suplementar 1). Dos 11 genes avaliados, um não teve a expressão alterada (Glyma.03G168000) e 10 obtiveram um aumento da expressão nas amostras de protoplasto de soja superexpressando *GmbZIPE2* em comparação ao controle.

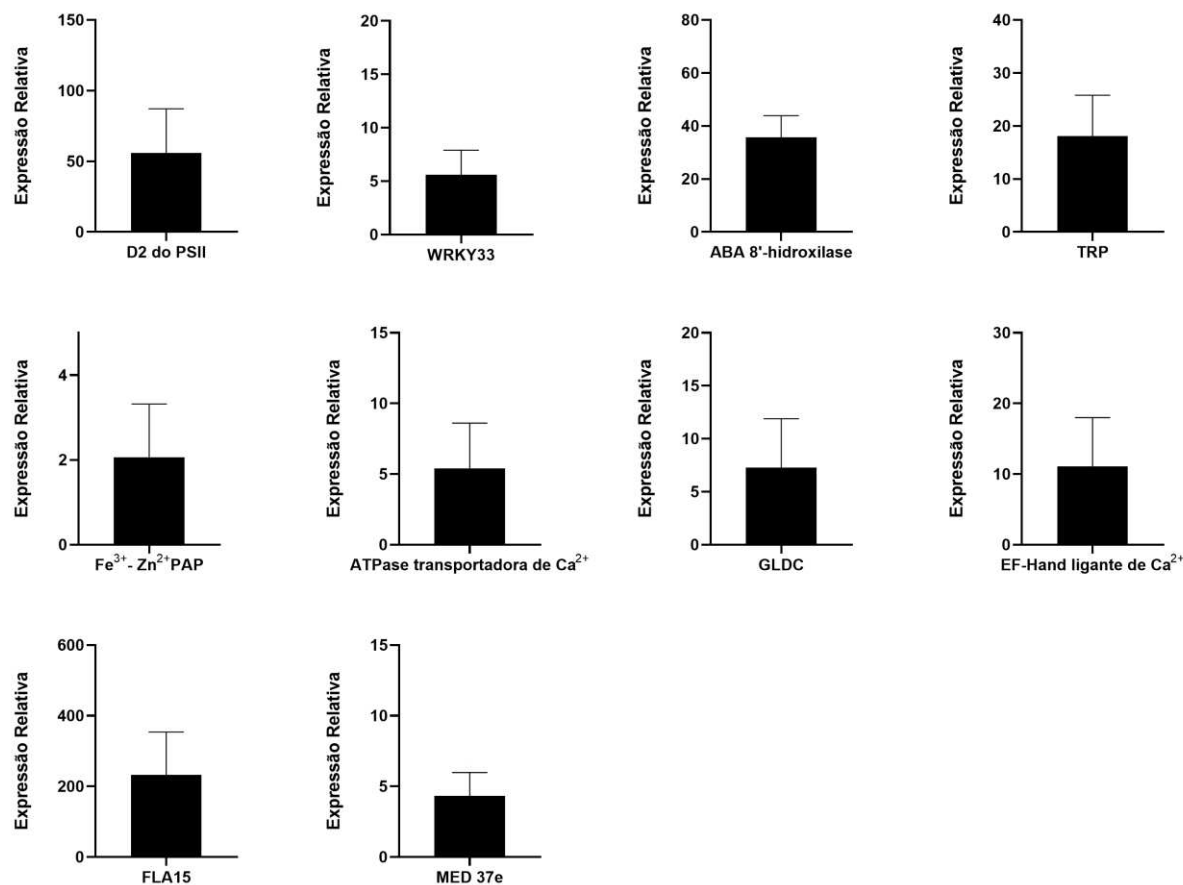


Figura 9: Validação do transcriptoma de protoplastos de soja superexpressando *GmbZIPE2*. A análise foi realizada pelo *software REST*, usando como endógeno o gene *ELF1A*. A expressão de todos os genes foi estatisticamente diferente do controle. Abreviações: TRP: *TPR repeat containing protein*; Fe³⁺-Zn²⁺ PAP: *Fe³⁺-Zn²⁺ purple acid phosphatase 12*; GLDC: *Glycine dehydrogenase (aminomethyl-transferring)/Glycine-cleavage complex P-protein*; FLA15: *Fasciclin-like arabinogalactan protein 15-related*; MED 37e: *Mediator of RNA polymerase II transcription subunit 37e-related*.

5.6 - Expressão da provável LPA em protoplastos superexpressando *GmbZIPE2*

Fontes (2016) identificou por meio do ensaio de duplo-híbrido que *GmbZIPE2* interage com duas proteínas. A provável LPA de soja tem sua expressão induzida em protoplastos de soja superexpressando *GmbZIPE2*.

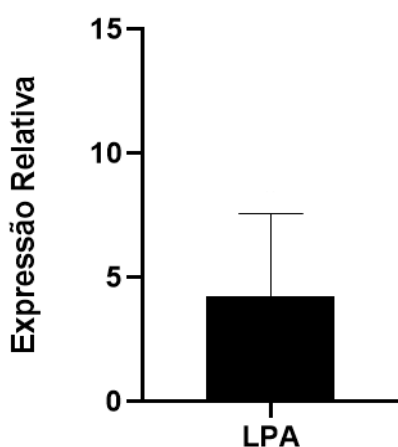


Figura 10: Expressão da LPA em protoplastos superexpressando *GmbZIPE2*. A análise foi realizada pelo *software REST*, usando como endógeno o gene *ELF1A*. A expressão de LPA foi estatisticamente diferente do controle.

5.7 - O silenciamento de *GmbZIPE2* através de VIGS

O silenciamento de *GmbZIPE2* através do experimento *VIGS* não foi eficiente. Inicialmente foi realizado RT-qPCR com *primer* referente ao *GmbZIPE2* que alinhava fora da porção clonada no vetor de silenciamento, e o resultado foi uma expressão aproximadamente 280 vezes maior de *GmbZIPE2* em relação ao controle (Figura 11). O vetor utilizado no experimento de silenciamento é diferente do vetor utilizado na superexpressão induzida por vírus, então esse resultado não poderia ser de uma superexpressão. Na tentativa de elucidar a alta expressão de *GmbZIPE2* no material supostamente silenciado, foi realizada a análise da expressão de genes já avaliados na *OE* desse gene. Na Figura 12B foram analisados os genes *CHS8*, que é ativado e *CHS7* e *PAL-1*, que não tem a expressão alterada na *OE* de *GmbZIPE*. Nenhum desses genes teve sua expressão alterada. Na Figura 12 C encontram-se genes que foram validados na

OE de *GmbZIPE2* e a *LPA*. Somente o gene *EF-Hand* ligante de Ca^{2+} teve a expressão alterada, sendo ativado nesse experimento.

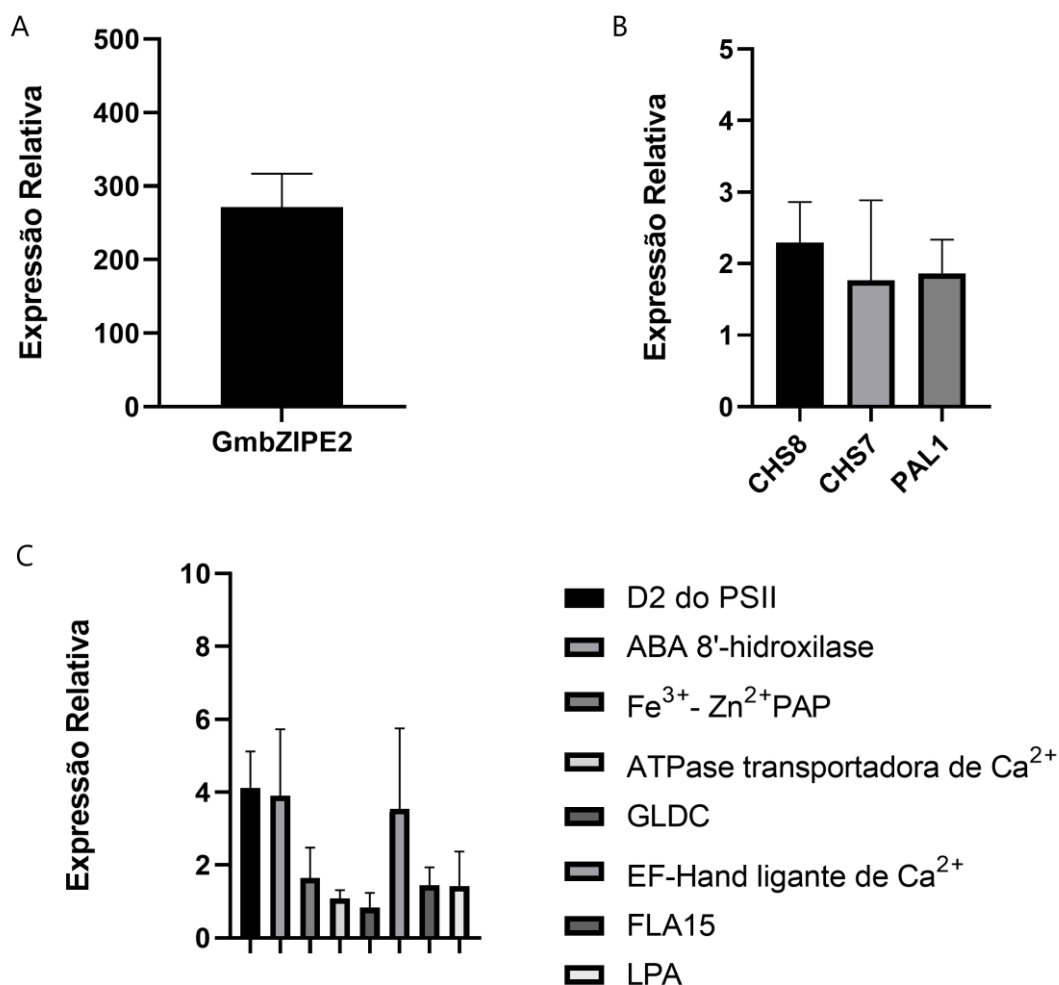


Figura 11: Expressão de *GmbZIPE2*, genes da via de biossíntese de fenilpropanóides e genes validados como ativados pela superexpressão de *GmbZIPE2* em folhas de soja com o gene *GbZIPE2* supostamente silenciado. **A. Validação do silenciamento de *GmbZIPE2*. A expressão de *GmbZIPE2* foi estatisticamente diferente do controle. **B.** Análise de *CHS8*, *CHS7* e *PAL*. **C.** Análise dos genes ativados por *GmbZIPE2* na sua superexpressão. A análise foi realizada pelo software *REST*, usando o gene *CYP* como endógeno para a análise de *GmbZIPE2* e *ELF1A* para os genes restantes. Somente a expressão *EF-Hand* ligante de Ca^{2+} foi estatisticamente diferente do grupo controle.**

6 - DISCUSSÃO

A partir da dificuldade de transformação estável da soja, a nossa estratégia foi utilizar a superexpressão transiente em protoplastos de soja. A extração e transformação dos protoplastos são eficientes. A nossa maior dificuldade foi adaptar o protocolo para uma extração de RNA de boa qualidade. No sequenciamento foi identificada a contaminação de bactérias, que não esperávamos ser tão expressiva, já que foi utilizado antibiótico no meio em que os protoplastos foram induzidos. Genes que são ativados pela superexpressão de *GmbZIPE2* (como *CHS8* e a provável *LPA*) não estão na lista de genes *Up*-regulados do transcriptoma, e essa ausência provavelmente é relacionada a diminuição da cobertura do sequenciamento devido a contaminação. Entretanto, mesmo com a contaminação e diminuição da cobertura, o transcriptoma de protoplasto de soja gerou valiosas informações sobre o gene *GmbZIPE2* e foi validado através de qRT-PCR.

Transcriptoma de uma superexpressão é uma valiosa ferramenta para o estudo de funções de genes. Devido as indicações da função de *GmbZIPE2* na via de biossíntese de fenilpropanoides e na fotossíntese/estresse oxidativo, buscou-se maiores informações da atuação desse fator em soja. O transcriptoma gerou mais dados relacionados a genes ativados por *GmbZIPE2*, e pode-se especular o motivo dessa diferença relacionada a esses números. A superexpressão desse gene pode ter provocado uma alteração no sistema que culminou majoritariamente em indução de expressão de genes, e não na repressão. Existe também a possibilidade da degradação do RNA, principalmente da amostra E2.2, e da contaminação das *reads* com material genético de bactérias ter interferido no número de genes reprimidos nesse transcriptoma.

Além de produzir informações em relação ao *GmbZIPE2*, genes descritos nesse transcriptoma podem servir como alvos para a caracterização de outros membros bZIP do grupo E, já que existe pouca informação na literatura sobre esse grupo em plantas, e nenhuma caracterização em soja.

6.1. *GmbZIPE2* ativa genes de fotossíntese

Dentre as várias funções atribuídas aos fatores de transcrição bZIP em diferentes espécies, está a fotossíntese. Em *Arabidopsis* a proteína bZIP AtHY5 ativa a transcrição do gene *Phytoene Synthase* (PSY) através da ligação ao *cis*-elemento G-box presente em seu promotor, promovendo então a acumulação de clorofila e fotopigmentos carotenoides (Toledo-Ortiz et al., 2014). Também em *Arabidopsis* a expressão constitutiva do bZIP *ABP9* melhora a capacidade fotossintética de plantas sob estresse, ajustando a composição do pigmento fotossintético, dissipando o excesso de luz e elevando a eficiência do uso de carbono (Zhang et al., 2018). Em soja, os fatores de transcrição GmbZIP04g e GmbZIP07g se ligam ao promotor do gene da Rubisco ativase (GmRCA α), promovendo a ativação de sua transcrição (Zhang et al., 2016). A maior proporção (aproximadamente 35%) dos genes ativados no transcriptoma de superexpressão de *GmbZIPE2* estão relacionados com a fotossíntese, indicando uma forte relação desse gene com este processo.

Na lista de genes ativados pela OE de *GmbZIPE2*, 21 são relacionados com os fotossistemas. Dentre eles, dois genes que codificam a proteína D2 do PSII (Glyma.06G217900 e Glyma.01G153500), sendo que a primeira foi validada por RT-qPCR. A provável proteína LPA que interage com GmbIPE2 (Fontes, 2016) também tem sua expressão aumentada na OE de *GmbZIPE2* e sua atuação na planta provavelmente está relacionada ao PSII.

A fotoinibição é um fenômeno que ocorre com organismo fotossintéticos, caracterizado pela inibição do fotossistema II resultante de uma exposição a luz intensa. Como a fotoinibição é inevitável em plantas, estas em condições normais, são capazes de superar os efeitos tóxicos dessa inibição através do reparo rápido e eficiente do PSII (Murata et al 2007). No esquema clássico de fotoinibição, a intensa iluminação produz ROS por redução excessiva de Q_A, o principal acceptor de elétrons de PSII, ou recombinação das cargas entre o lado receptor e o lado doador de PSII. As ROS geradas atacam diretamente o centro de reação fotoquímico de PSII (Murata et al 2007). O centro de reação PSII é composto por um heterodímero D1-D2 que transporta todos os componentes essenciais para o transporte de elétrons (Mattoo et al.1989; Baker 1991). Em situações de estresse, os danos oxidativos podem induzir inibição dos fotossistemas e de mecanismos de reparo, que envolvem a substituição da proteína do centro de reação D1, principal alvo dos danos oxidativos. A fosforilação dos polipeptídeos do PSII e o turnover da proteína D1 foram propostos como dois mecanismos de proteção

da integridade estrutural e funcional do PSII. O turnover de D1 leva a uma regulação negativa da atividade de PSII sob alto estresse oxidativo, constituindo um ciclo de reparação de danos. (Giardi *et al*, 1996)

As proteínas LPA participam da montagem e reparo do PSII. LPA1 atua como chaperona, auxiliando no dobramento e integração de D1 na membrana tilacoide e promovendo a interação de D1 com D2. A LPA3 interage com CP43 (β -caroteno-Chla, proteína do PSII) e LPA2, auxiliando na montagem de CP43 no complexo PSII (Peng *et al.*, 2006; Cai *et al.*, 2010; Järvi *et al.*, 2015)

No trabalho uma LPA (mais parecida com a LPA2 de *A.thaliana*) e a D2 do fotossistema II foram *up*-reguladas em protoplastos superexpressando *GmbZIPE2*, e Fontes (2016) identificou a interação dessa LPA à *GmbZIPE2*. Fontes (2016) sugeriu um esquema de regulação de LPA e *GmbZIPE2* em situações de estresse (Figura 2). Em condições normais a *GmbZIPE2* estaria interagindo com a LPA, que está ancorada à membrana. Em uma situação de estresse, a produção de ROS pode induzir inibição dos fotossistemas e mecanismos de reparo, então a LPA deixaria de interagir com *GmbZIPE2* para auxiliar na montagem do PSII. Através de um outro mecanismo, *GmbZIPE2* entraria no núcleo e desempenharia seu papel como fator de transcrição.

A ativação desses genes por *GmbZIPE2* na OE juntamente com a interação entre *GmbZIPE2* e a provável LPA reforça a hipótese de Fontes (2016) de que este gene está envolvido em processos da soja que tentam aliviar o estresse oxidativo. Uma análise de 1000pb do promotor da possível LPA (Figura suplementar 7) revela a presença do *cis*-elemento H-box, seguido de quatro bases que também existem após o H-box no promotor de *CHS8* (Gonçalves *et al*, 2018, não publicado), indicando a possibilidade de *GmbZIPE2*, além de ligar a proteína, regule a transcrição da provável LPA em soja. A localização predita da proteína *GmbZIPE2* (Fontes, 2016) é no cloroplasto, devido à presença de um possível peptídeo de trânsito de cloroplasto. A partir dos dados da interação desse fator com proteínas do cloroplasto relacionadas com estresse oxidativo e do alto número de genes relacionados a fotossíntese no transcriptoma, pode-se sugerir um papel de *GmbZIPE2* como um sensor de estresse no cloroplasto. A partir do estímulo do estresse, o fator iria para o núcleo regulando genes de defesa.

6.2. *GmbZIPE2* ativa fatores de transcrição WRKY

Os fatores de transcrição regulam a expressão de genes através de ligação sequência-específica com o DNA, ação que geralmente ocorre através de interações proteína-proteína (Latchman, 2003). Nas análises dos GOs dos genes com expressão aumentada na OE de *GmbZIPE2*, em relação a função molecular as maiores porcentagens foram 19,4% dos genes que possuem a função de ligação a proteínas e 13,9% ligação à nucleotídeos e em relação a componentes celulares, observa-se que a 21,4% dos genes estão localizados no núcleo, e provavelmente alguns desses genes podem ser fatores de transcrição. O maior número de fatores de transcrição de uma mesma família identificados entre os genes ativados foram da família WRKY.

Fatores de transcrição WRKY interagem com TGAs, que são bZIPs participantes da via mais elucidada da família (Van Verk *et al*, 2011). A família WRKY é exclusiva de plantas, caracterizados por possuírem um ou dois domínios que contêm a sequência WRKYGQK, também por se ligarem ao *cis*-elemento W box (Rushton *et al.*, 2010). As proteínas WRKY estão envolvidas em vários processos nas plantas, como desenvolvimento de sementes, senescência, dormência, germinação, com um destaque na atuação de fatores de transcrição na modulação da resposta a estresses abióticos e bióticos (Yu *et al* 2016; Chen *et al.* 2012).

Na lista de genes ativados através da OE de *GmbZIPE2*, identificamos nove WRKYs. O fator WRKY33 é um dos fatores mais estudados dessa família em *A. thaliana*. Alguns WRKY de soja são nomeados como WRKY33 pelo Phytozome. Uma comparação da sequência da proteína WRKY33 de *A.thaliana* com as proteínas de *G.max* foi realizada no *Phytozome*. A proteína de soja mais similar a WRKY33 de *A. thaliana* é a Glyma.14G200200 (48%), que é nomeada WRKY49 pelo NCBI. A segunda proteína mais similar (47%), Glyma.02G232600, é nomeada de WRKY39 pelo NCBI. Esses dois genes aparecem na lista de genes *Up*-regulados na OE de *GmbZIPE2* e são nomeadas WRKY33 pelo Phytozome. A expressão de Glyma.14G200200 foi validada por RT-qPCR (Figura 10)

A proteína AtWRKY33 em *Arabidopsis* desempenha um papel importante durante a infecção por patógenos necrotróficos. A interação de AtWRKY33 com as proteínas *SIGMA FACTOR-INTERACTING PROTEIN* 1 e 2 (SIB1 e SIB2) após a contato da planta com o fungo *Botrytis cinerea*, culmina na ligação desse WRKY no

promotor de genes de defesa, ativando a transcrição destes genes (Lai *et al*, 2011) (Figura 12).

O mutante *wrky33* de *Arabidopsis* é altamente susceptível a *B. cinerea*. Liu *et al*, 2015 identificou através de RNA-seq e ChiP-seq que AtWRKY33 pode agir, de uma maneira dependente da ligação do fator ao promotor, como um repressor ou ativador da transcrição de determinados genes. Como ativador, os alvos de WRKY33 são genes envolvidos na sinalização hormonal e na biossíntese de fitoalexinas. E como repressor, AtWRKY33 regula negativamente genes (NCED3 e NCED5) da biossíntese de ácido abscísico (ABA). Plantas mutantes *wrky33* são mais susceptíveis a infecção fúngica e apresentam maiores níveis de expressão de NCED3 e NCED5 em comparação com WT. Plantas com mutação tripla *wrky33 nced3 nced5* tem praticamente o mesmo fenótipo de resistência a *B.cinerea* em relação a WT, e apresentam biomassa fúngica fortemente reduzida, indicando que o aumento dos níveis de expressão de NCED3 e NCED5 no mutante *wrky33* contribui para a suscetibilidade à infecção por *B. cinerea*, e que uma função chave de WRKY33 na imunidade do hospedeiro contra este patógeno consiste em reprimir a biossíntese de ABA. (Liu *et al*, 2015)

Outra via importante descrita que AtWRKY33 participa é relacionada com a transcrição de sinais por MPK cinases. Esta via é desencadeada após a indução por *Pseudomonas syringae* ou flagelina, e então, a proteína MPK4 é ativada e fosforila seu substrato, MSK1. A fosforilação de MSK1 libera o complexo AtWRKY33, permitindo que AtWRKY33 se ligue à região promotora de alguns genes, incluindo o promotor da deficiência de fitoalexina3 (PAD3), que codifica uma enzima que participa da síntese de camalexina, um tipo de fitoalexina que desempenha papel importante na defesa das plantas (Quil *et al*,2008). AtWRKY33 também se liga ao promotor de *CYP71A13*, outra enzima da via de biossíntese de camalexina (Birkenbihl *et al.*, 2012). AtWRKY33 é fosforilada por MPK3/MPK6, e possivelmente através dessa fosforilação, ocorre um aumento da expressão de AtWRKY33, desencadeando um mecanismo de feedback positivo que desencadeia a resposta da planta a patógenos. Além de MPK4, a proteína AtWRKY33 também pode interagir com MPK3 e MPK6. Em plantas mutantes *wrky33*, a expressão de genes envolvidos na produção de camalexina através da cascata MPK3/MPK6 e a própria indução de camalexina, são comprometidas, demonstrando a importância desse gene em mecanismos de defesa (Chen *et al*, 2001) (Figura 12).

Além da associação de WRKY33 com vias relacionadas a estresse biótico, os dois genes GmWRKY33 listados como ativados pela OE de *GmbZIPE2* foram analisados em relação à estresse salino. O Glyma.14G200200 teve sua expressão induzida e Glyma.02G232600 foi inibido por estresse salino (Yu *et al.*, 2016), demonstrando que esses dois genes podem ter funções diferentes em soja sob estresse abiótico.

O fato de WRKY33 ser ativado na OE de *GmbZIPE2* e estar relacionado com a via de biossíntese de fitoalexinas, como a camalexina, é um indício de que essas duas proteínas possam atuar conjuntamente na ativação de genes em resposta a estresse biótico. *GmbZIPE2* atua ligando ao promotor de *CHS8*, ativando sua transcrição e assim aumentando o nível de narigerina chalcona, que pode gerar vários tipos metabólitos relacionados a defesa (também chamados de fitoalexinas). WRKY33 ligaria ao promotor de *PAD3* e *CYP71A13* (e possivelmente outros genes dessa via) para ativar a transcrição desses genes e assim aumentar a produção de camalexina. As vias de produção dos metabólitos secundários gerados através da ativação de genes por *GmbZIPE2* e WRKY33 são diferentes, porém ambas podem resultar na produção de fitoalexinas.

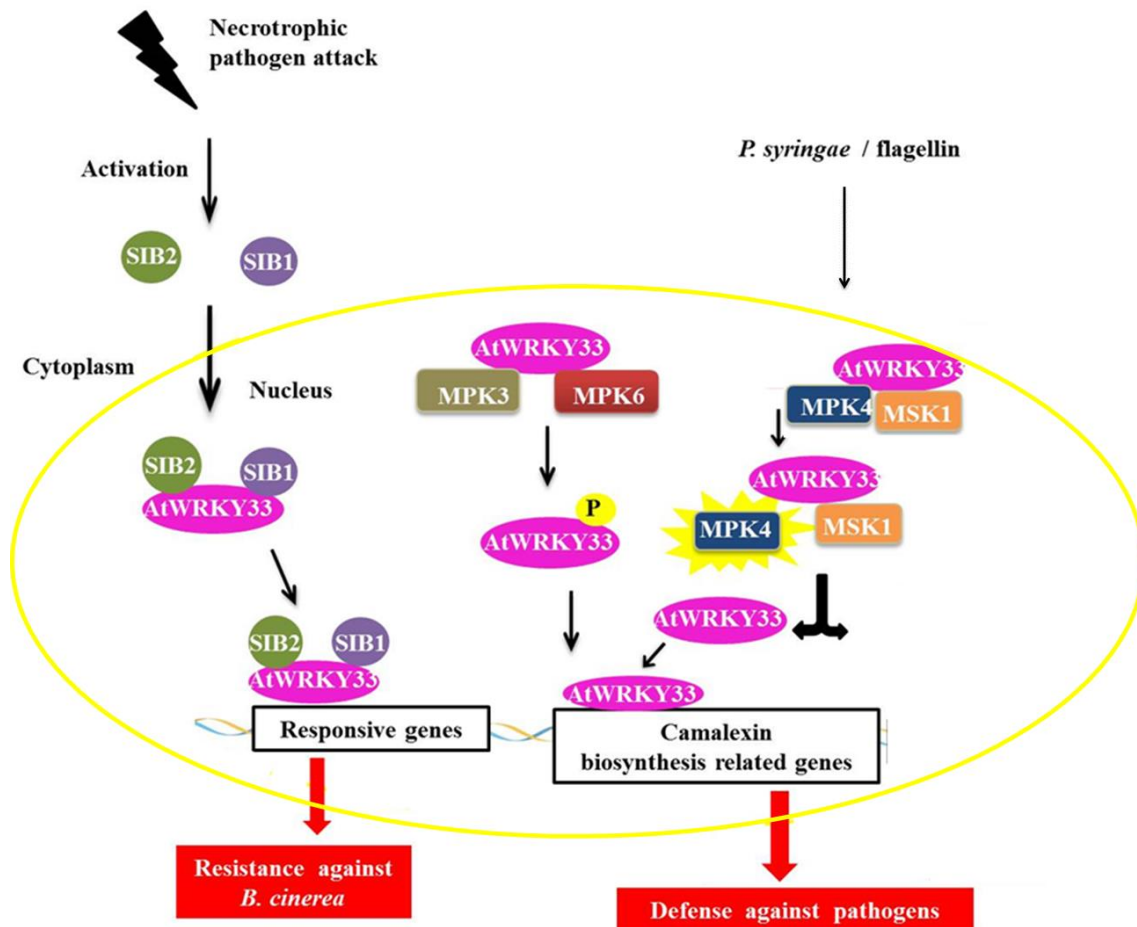


Figura 12: Representação esquemática da atuação de WRKY33 em situação de estresse. Figura adaptada de Alves *et al*, 2014.

6.3. Silenciamento de *GmbZIPE2*

O silenciamento de *GmbZIPE2* através do VIGS não foi eficiente. Após o resultado de um aumento da expressão de *GmbZIPE2*, ao invés de esperada diminuição, investigou-se o possível motivo desse resultado. O par de *primer* de RT-qPCR utilizado anelava fora da porção clonada no vetor, não existindo a possibilidade de detecção do vetor viral de RNA ao invés do transcrito. Na tentativa de extrair maiores informações sobre o material supostamente silenciado, a expressão genes cujo o padrão da expressão já era conhecida na OE de *GmbZIPE2* foi analisada. A expressão desses genes majoritariamente não foi alterada nas folhas de soja supostamente silenciadas em relação ao controle, demonstrando que esse experimento de silenciamento de

GmbZIPE2 por VIGS não funcionou. Provavelmente houve um acúmulo de transcritos de *GmbZIPE2*, que não foram traduzidos em proteínas.

7 - CONCLUSÃO

O transcriptoma de protoplasto de soja superexpressando *GmbZIPE2* gerou importantes informações sobre este gene, e que apesar de da contaminação das amostras, gerou uma boa quantidade de genes induzidos. O transcriptoma foi validado através de RT-qPCR através do aumento significativo da expressão de dez genes tidos como ativados no transcriptoma.

A maior parte dos genes anotados do transcriptoma são relacionados à fotossíntese. Com essa informação relacionada aos resultados anteriores do nosso grupo de pesquisa, acredita-se que *GmbZIPE2* tem um papel na fotossíntese, mais especificamente, no controle do estresse oxidativo realizado nos fotossistemas.

A ativação do gene WRKY33 na superexpressão de *GmbZIPE2* pode estar relacionada a um mecanismo de defesa da soja, podendo WRKY33 e *GmbZIPE2* agirem em uma ação conjunta na via de biossíntese de fitoalexinas e/ou outras vias de defesa.

Os dados gerados enriqueceram as informações sobre *GmbZIPE2* e abrem possibilidades para novos experimentos utilizando essas informações tanto para a continuação da caracterização de *GmbZIPE2*, quanto para a caracterização do grupo E de soja.

8 – REFERÊNCIAS

Alves, M.S., Dadalto, S.P., Gonçalves, A.B., De Souza, G.B., Barros, V.A., Fietto, L.G. 2013. Plant bZIP Transcription Factors Responsive to Pathogens: A Review. *Int J. Mol. Sci.* 14,7814-7828.

Alves, M.S., Dadalto, S.P., Gonçalves, A.B., De Souza, G.B., Barros, V.A., Fietto, L.G. 2014. Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. *Proteomes.* 2, 85-106

Alves, M.S., Soares, Z.G., Vidigal, P.M.P., Barros E.G. , Poddanosqui, A.M., Aoyagi, L.N., Abdelnoor, R.V., Marcelino-Guimarães, F.C., Fietto, L.G. 2015. Differential expression of four soybean bZIP genes during *Phakopsora pachyrhizi* infection. *Funct Integr Genomics.* 15(6), 685-96

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data.

Baker, N.R.1991. Possible role of photosystem II in environmental perturbations of photosynthesis. *Physiol Plant* 81:563-570

Benson, D. A., Clark, K., Karsch-Mizrachi,I., Lipman, D. J., Ostell,J., Eric W. Sayers, E. W.2014. GenBank. *Nucleic Acids Research.* 42, 30–35.

Birkenbihl, R.P., Diezel, C., Somssich, I.E. 2012. Arabidopsis WRKY33 Is a Key Transcriptional Regulator of Hormonal and Metabolic Responses toward *Botrytis cinerea* Infection. *Plant Physiol*May; 159(1): 266–285.

Bolger, A. M., Lohse, M., Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30 (15), 2114-2120.

Boller T. and He, S. Y. 2009. Innate immunity in plants: An arms race between pattern recognition receptors in plants and effectors in microbial pathogen. *Science.* 8; 324(5928): 742–744.

Buscaill, P., Rivas, S. 2014. Transcriptional control of plant defence responses. *Current Opinion in Plant Biology*, 20:35–46.

Cai, W., Ma, J., Chi, W., Zou, M., Guo, J., Lu, C., Zhang, L. 2010. Cooperation of LPA3 and LPA2 is essential for photosystem II assembly in Arabidopsis. *Plant physiology*, v. 154, n. 1, p. 109-120.

Chern, M.S.; Fitzgerald, H.A.; Yadav, R.C.; Canlas, P.E.; Dong, X.; Ronald, P.C. 2001. Evidence for a disease-resistance pathway in rice similar to the NPR1-mediated signaling pathway in Arabidopsis. *Plant J.* 2001, 27, 101–113.

Chen, L.G., Song, Y., Li SJ, Zhang, L.P., Zou, C.S., Yu, D.Q. 2012. The role of WRKY transcription factors in plant abiotic stresses. *BBA-Gene Regul Mech.* 819:120–128

da Silva A. C., de Souza P. E., Amaral D. C., Zeviani W. M., Pinto J. E. B. P. 2014. Essential oils from *Hyptis marrubioides*, *Aloysia gratissima* and *Cordia verbenacea* reduce the progress of Asian soybean rust. *Acta Sci. Agron.* 36 159–166.

Dodds, P. N. and Rathjen, J.P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics.* 11, 539-548.

Fontes, P.R. 2016. Identificação de interações proteína-proteína do fator de transcrição GmbZIPE2 de soja (*Glycine max*). Dissertação de mestrado em Bioquímica Aplicada. UFV. Viçosa

Giardi, M. T., Cona, A., Geiken, B., Kučera, T., Masojídek, J., & Mattoo, A. K. 1996. Long-term drought stress induces structural and functional reorganization of photosystem II. *Planta*, 199(1), 118–125. doi:10.1007/bf00196888

Gonçalves, A. B. 2014. Caracterização funcional do fator de transcrição GmbZIPE2 de soja (*Glycine max*). Dissertação de mestrado em Bioquímica Agrícola. UFV. Viçosa

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D. S. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*.40, 1178-1186.

Hartman, G.L., Miles, M.R., Frederick, R.D. 2005. Breeding for resistance to soybean rust. *Plant Dis* 89:664–666.

Jackoby, M., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Parcy, T.K.F. 2002. bZIP transcription factors in Arabidopsis. *Trends in Plant Science* Vol 7, N 03.

Järvi, S., Suorsa, M., Aro, E. 2015. Photosystem II repair in plant chloroplasts—regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, v. 1847, n. 9, p. 900-909.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. *Nature*: 444: 323-329.

Kapos, P., Devendrakumar K.T., Li X. 2019. Plant NLRs: From discovery to application. *Plant Sci*.279:3-18

Langenbach, C., Campe, R., Beyer, S. F., Mueller, A. N., Conrath, U. 2016. Fighting Asian Soybean Rust. *Front Plant Sci*. 7: 797.

Latchman, D. S. 2003. Eucaryotic transcription factors, edn4. San Diego: Academic Press.

Liao, Y., Zou, H., Wei, W., Hao, Y., Tian, A., Huang, J., Liu, Y., Zhang, J., Chen, S. 2008. Soybean GmbZIP44, GmbZIP62 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic *Arabidopsis*. *Planta*. 228:225-240

Liu L., White, M. J., Macrae, T. H. 2001. Transcription factors and their genes in higher plants. *European Journal of Biochemistry*, v.262, i.2, p.247–257.

Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R.P., Somssich, I.E. 2015. Negative regulation of ABA signaling by WRKY33 is critical for Arabidopsis immunity towards *Botrytis cinerea* 2100. *Elife*.15;4:e07295

Mattoo, A.K., Marder, J.B., Edelman, M. 1989. Dynamics of the photosystem II reaction center. *Cell* 56:241-246

Meng X and Zhang, S. 2013. MAPK cascades in plant disease resistance signaling. *Annu Rev Phytopathol*, 51, pp. 245–266.

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., Marshall, D. 2010. Tablet- Next generation sequence assembly visualization. *Bioinformatics*, v. 26, p. 401-402, 201

Murata, N., Takahashi. S., Nishiyama, N., Allakhverdiev S.I. 2007. Photoinhibition of photosystem II under environmental stress. *Biochimica et Biophysica Acta (BBA) – Bioenergetics*.1767, 6, 414-421

Peng, L.; Ma, J.; Chi, W.; Guo, J.; Zhu, S.; Lu, Q.; Lu, C.; Zhang, L. 2006. LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. *The Plant Cell*, v. 18, n. 4, p. 955-969.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *NAR* 29, e45

Pfaffl, M.W., Horgan, G.W., Dempfle, L. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *NAR* 30, e36

Pieterse, C.M.J.; van der Does, D.; Zamioudis, C.; Leon-Reyes, A.; van Wees, S.C.M. 2012. Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* 28, 489– 521.

Qiu, J.L.; Fiil, B.K.; Petersen, K.; Nielsen, H.B.; Botanga, C.J.; Thorgrimsen, S.; Palma, K.; Suarez-Rodriguez, M.C.; Sandbech-Clausen, S.; Lichota, J. 2008. Arabidopsis MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. *EMBO J.* 27, 2214–2221.

Rushton, P.J., Somssich, I.E., Ringler, P., Shen, Q.J. 2010. WRKY transcription factors. *Trends in Plant Science* 15: 247–258.

Sambrook, J, Fritsch, EF, Maniats, T. 1989. *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor.

Schütze, K., Harter, K., Chaban, C. 2008. Post-translational regulation of plant bZIP factors. *TRENDS in Plant Science Vol 13, N 05*.

Singh, K. B., Foley, R. C., Oñate-sánchez, L. 2002. Transcription factors in plant defense and stress responses. *Plant Biology*, 5:430, p.430-436.

Sornaraj, P., Luang, S., Lopato, S., Hrmova, M. Basic leucine zipper (bZIP) transcription factors involved in abiotic stresses: a molecular model of a wheat bZIP factor and implications of its structure in function. 2016. *Biochimica et Biophysica Acta (BBA)-General Subjects*, v. 1860, n. 1, p. 46-56.

Spoel, S. H. and Dong, X. 2012. How do plants achieve immunity? Defence without specialized immune cells. *Nat Ver Immunol*, 12(2):89-100.

Toledo-Ortiz, G., Johansson, H., Lee, K.P., Bou-Torrent, J., Stewart, K., Steel, G., Rodríguez-Concepción, M., Halliday, K. J. 2014. The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. *PLOS Genetics*. 10 (6).

Van Verk, M.C., Neeleman, L., Bol, J.F., Linthorst, H.J. Tobacco transcription factor NtWRKY12 interacts with TGA2.2 in vitro and in vivo. *Front. Plant Sci.* 2011, 2, 32.

Vandesompele J., Preter, K., Pattyn, F., Poppe, B., Roy, N.V., Paepe, A., Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, research0034.1.

Wang, Z., Cheng, K., Wan, L., Yan, L., Jiang, L., Liu, S., Lei, Y., Liao, B. 2015. Genome-wide analysis of the basic leucine zipper (bZIP) transcription factor gene family in six legume genomes. *BMC Genomics* 10;16:1053.

Wang, Z., Libault, M., Joshi, T., Valliyodan, B., Nguyen, H. T., XU, D., Stacey, G., Cheng, J. 2010. SoyDB: a knowledge database of soybean transcription factors. *Plant Biology*, 10:14.

Yorinori J. T., Paiva W. M., Frederick R. D., Costamilan L. M., Bertagnolli P. F., Hartman G. E. 2005. Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Dis.* 89 675–677.

Yu Y., Wang, N., Hu, R. and Xiang, F. 2016. Genome-wide identification of soybean WRKY transcription factors in response to salt stress. *SpringerPlus.* 5:920

Zhang X, Wollenweber B, Jiang D, Liu F, Zhao J. 2008. Water deficits and heat shock effects on photosynthesis of a transgenic *Arabidopsis thaliana* constitutively expressing ABP9, a bZIP transcription factor. *Journal of Experimental Botany* 59: 839–848.

Zhang, M., Liu, Y., Shi, H., Guo, M., Chai, M., He, Q., Yan, M., Cao, D., Zhao, L., Cai, H., Qin, Y. 2018. Evolutionary and expression analyses of soybean basic Leucine zipper transcription factor family. *BMC Genomics*, 19:159.

Zhang, Y., Zhang G., Dong Y., Guo J., Huang L., Kang Z. 2009. Cloning and Characterization of a MBF1 Transcriptional Coactivator Factor in Wheat Induced by Stripe Rust Pathogen. *ACTA Agronomica Sinica*, v.35:1, p.11-17.

Zhang. J., Du, H., Chao, M., Yin, Z., Yang, H., Li, Y., Huang, F., Yu, D. 2016. Identification of two bZIP transcription factors interacting with the promoter of soybean rubisco activase gene (GmRCAa). *Frontiers in Plant Science*. 7 (628).

9 - MATERIAL SUPLEMENTAR

ID	Orientação do oligonucleotídeo	Sequência (5'-3')
Glyma.02G015500	Fw	CCATTTACTAATTCCCTGTC
	Rv	CAAGCATACTCAAAGAGTGA
Glyma.15G011900	Fw	ATGATATCCACTCACCTCAT
	Rv	GACTTCCTCAATGAACATCT
Glyma.13G285400	Fw	TGCAACTCCATCACACAGCT
	Rv	CAGTGTGGTCACCTTGGGAG
Glyma.05G092000	Fw	CTACTAGGTAGGATCGGATG
	Rv	CGTTATGACAAAGAACAAGA
Glyma.18G124800	Fw	GAACACCTTAGCCTCTACAC
	Rv	CCTTTGTAATGGAGTAACC
Glyma.08G285900	Fw	TCCTTCCAAAACCTTATTGAT
	Rv	GAGGTTTGATCTCTGTCACT
Glyma.05G187300	Fw	TAACAATAGCAAGGCACTTA
	Rv	TGTCTATCAATTCCATCAA
Glyma.07G122200	Fw	ATCGCTTACTACTAGGTTGG
	Rv	AACAAGAAGAAAACAAAAGG
Glyma.11G162200	Fw	GGTTTGCACCTTTTATCCTAT
	Rv	TAGCAAGCTACACCAAGTAA

Tabela Suplementar 1: Oligonucleotídeos usados para RT-qPCR, que não funcionaram.

ID	Nome do gene	Eficiência	Concentração μM
Glyma.05G168100	GmbZIPE2 (VIGS)	1	1,5
Glyma.06G028100	Fe ⁺³ -Zn ⁺² PAP	0,9	2,5
Glyma.11G048300	ATPase transportadora de Ca ⁺²	1	2,5
Glyma.03G168000	Transportador ABC	0,71	2,5
Glyma.19G201600	FLA15	0,89	2,5
Glyma.14G094700	GLDC	0,86	2,5
Glyma.13G272000	EF-hand ligante de Ca ⁺²	1	2,5
Glyma.11G140500	MED37e	0,94	2,5
Glyma.16G165700	TRP	1	2,5
Glyma.01G153300	ABA 8'hidrolase	1	2,5
Glyma.14G200200	WRKY33	1	2,5
Glyma.06G217900	D2 do PSII	0,92	2,5
Glyma.09G232600	Provável LPA	1	2,5
Glyma.02G309300	PAL-1	1	2,5
Glyma.01G228700	CHS7	1	2,5
Glyma.11G011500	CHS8	1	2,0
Glyma.12G02790	CYP2	0,92	2,5
	<i>Cyclophilin 2</i>		
Glyma.05G24110	ELF1A Translation elongation factor 1-alpha	0,86	2,0

Tabela Suplementar 2: Eficiência dos oligonucleotídeos usados no RT-qPCR.

9.1 Análise de qualidade das *reads* através do FastQC

O padrão de cores dos gráficos representa um perfil de qualidade (verde, laranja e rosa), onde verde são bases de boa qualidade, a região laranja de qualidade razoável e a faixa rosa representa bases de baixa qualidade. Antes da trimagem algumas *reads* se encontravam na parte laranja e rosa do gráfico, e após a trimagem todas as amostras se encontram na parte verde do gráfico, demonstrando a deleção das *reads* de baixa qualidade.

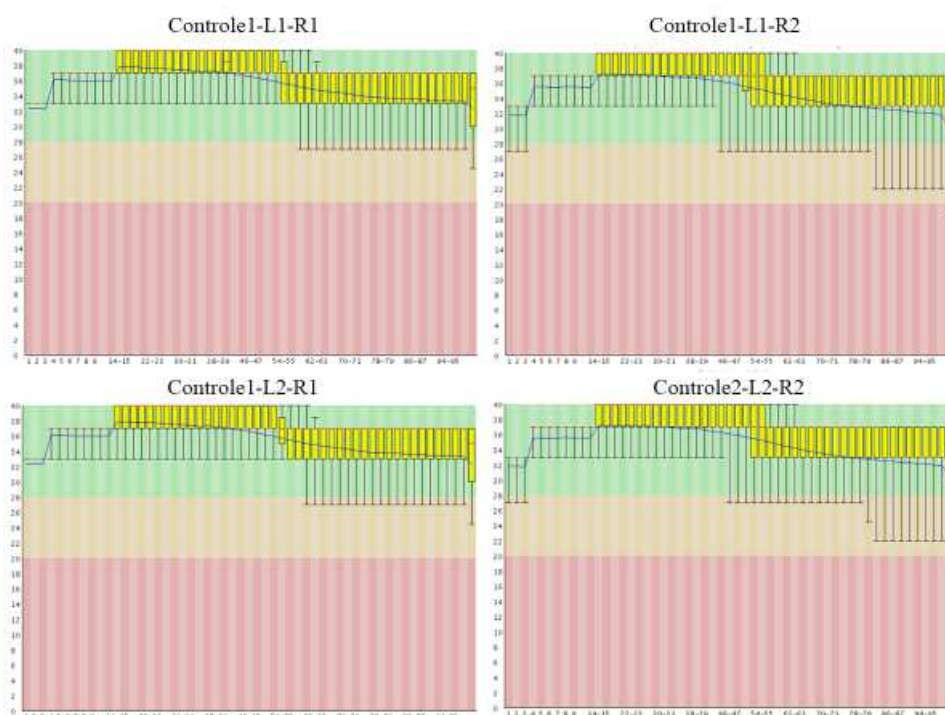


Figura Suplementar 1: Gráficos gerados pelo FastQC referentes aos dados brutos da amostra C1.

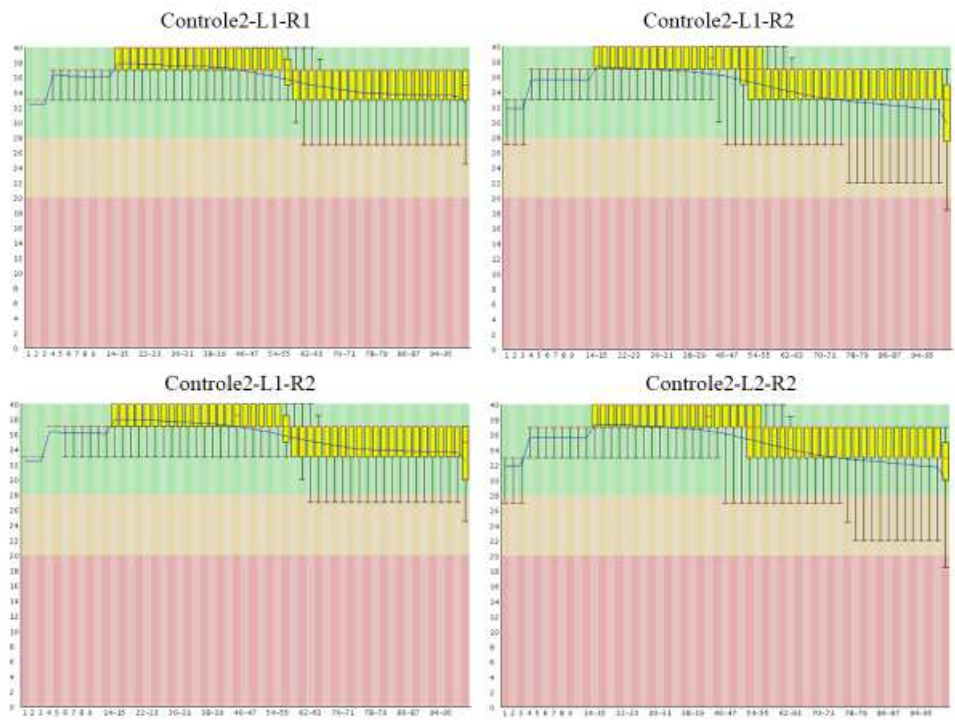


Figura Suplementar 2: Gráficos gerados pelo FastQC referentes aos dados brutos da amostra C2 .

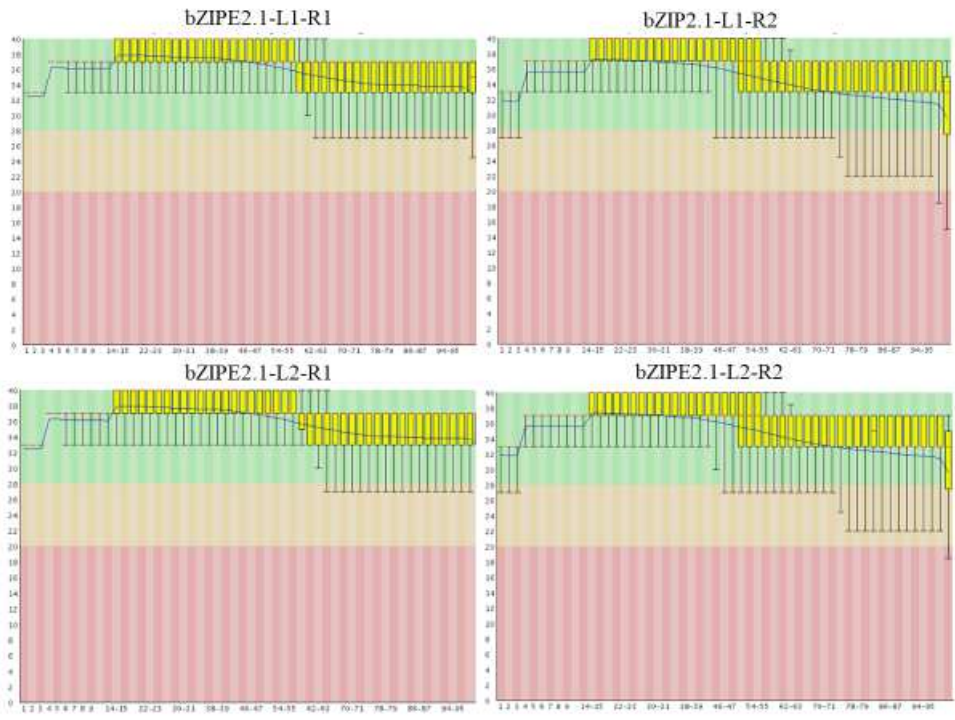


Figura Suplementar 3: Gráficos gerados pelo FastQC referentes aos dados brutos da amostra E2.1.

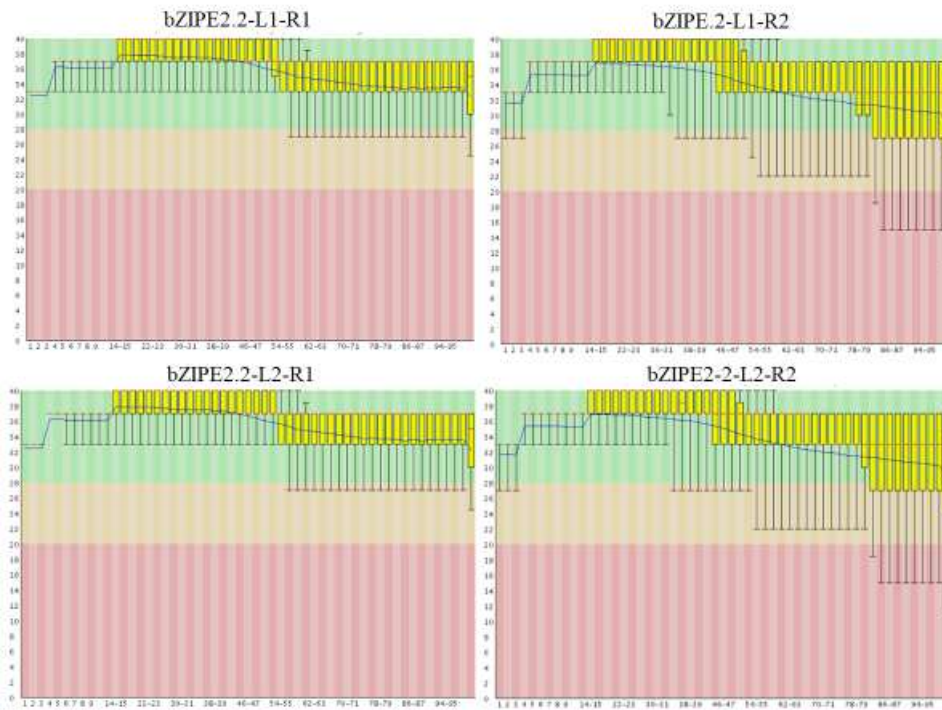


Figura Suplementar 4: Gráficos gerados pelo FastQC referentes aos dados brutos da amostra E2.2.

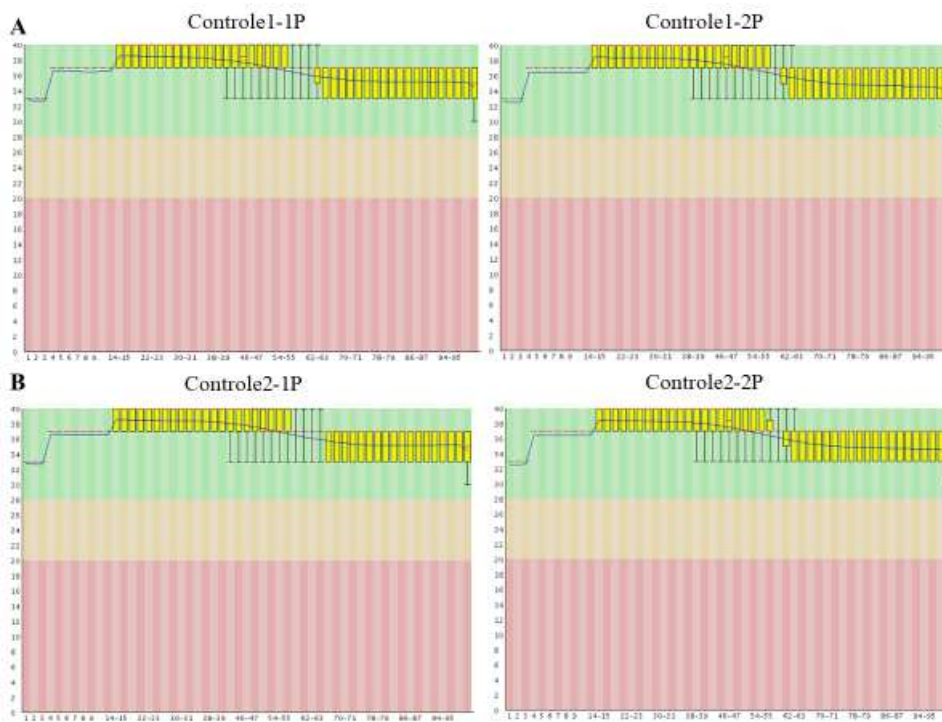


Figura Suplementar 5: Gráficos gerados pelo FastQC referentes às *reads* trimadas das amostras controles.

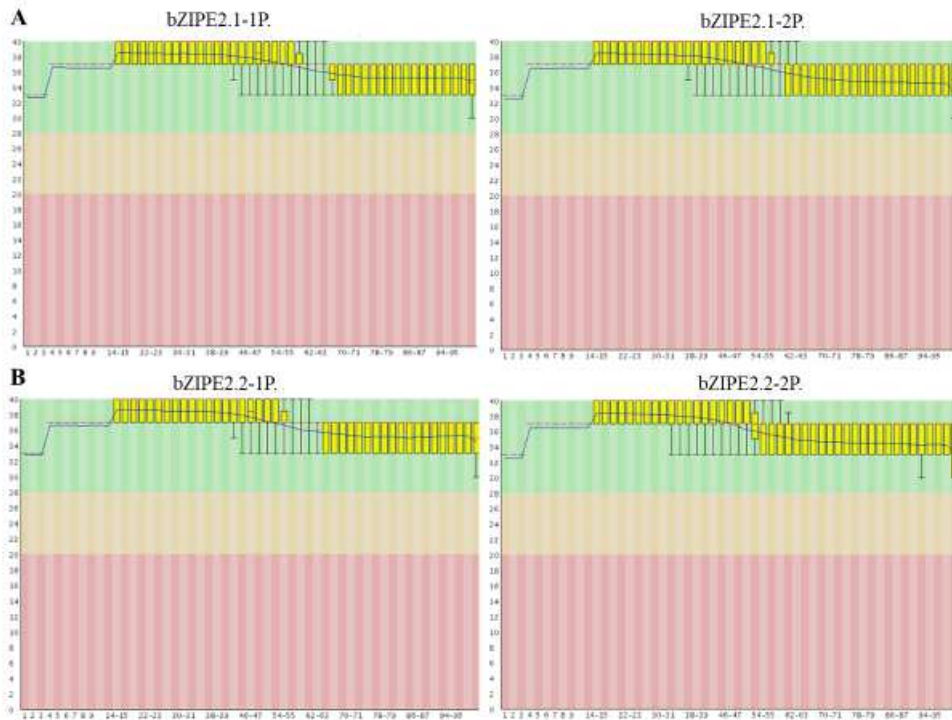


Figura Suplementar 6: Gráficos gerados pelo FastQC referentes às *reads* trimadas das amostras *OE GmbZIPE2*.

```

>Glyma.09G232600
CATTGAGAAACGGATCTAGCTGTAGAAGAACCTAATCAAATCGGAAAACAGCAAGCAGTTGGTGTCA
GAATCGTTCAAAAACATCCATTCCTCTACTCACAAATTTTTTTAGATCACGATAACATAAGCTGAAT
CCCAGTCCAAACTATTAATAAATAAATACTAGGGAAATGACATTACGATTGTCCTGAGGAAATTTACGT
AACTAGATGGTGAATTTTGGAACTACTAGTGTGAATTATTAAGGGTAGATTATTATTTCCCGCAAC
GATCTAATCTACATTTAAAAAACGCGGAGGAGCAAGGGAATGGAAGAACGCAGAAAAAGTAGGGAAA
AATAATTAAGAGAAGAGATTACCATAGTTTTGGGTGCTGAAACTGAAAACGCAGTGTTCGTCTTTA
GGTTGAGTTGAAGGGGAAAAGAAAAGCCACGAGAAATTAACAATTGATAACGAGTTAACTGTTGAAC
TGAAGTGAAGGCCAAAGCCCAAATAAATTGCTATTTGCACTACCCTTTGTTATTCATACGTCCCC
ATGTTCTTTTTTATTCCTAG-CCTACC-CTTC-CCCCTCAAATTCATGGATGGCACATTAATTAT
CGGTGAAAGTCTCTAACATATTCCTTTGAACATATTTCTTTTATTAATTGAAATTTATTGAAAATC
ACAAAATTATAGATCTTATTCAATAAGTCTTTCTTATGATTTTGTAAATTTTAAATATTTTAACTA
ATAAAAAAATGTATATGAAGAAATCTTAAGAACCTATTAATAAATAATGTTACTAAAACATGTTT
CTAGAATAAAGAGTCTCGAACTCAAACACGAGGACTTTAAAAATTTTATATGAGTTAATTTTTTTT
AATTATTGTATATACTACTATATTAGGAGTTGTTTTATTAATAATAAATAATGGTATTGTATTATA
TTATAGAGCCATTACTTTATGTATGTGTAATAATTGCAAAGTAGAAGAAAAGAGGGGATGTGACTATA
TAACATTATCAGAAACAAAGGATGTGGTGAAGCTGGAGCCACGTTTGAGGTGTGAGCAATAACTCTG
TTACGGAAGAAGAAAGTGAGAGTGAAAATG

```

Figura Suplementar 7: Análise da região promotora da possível LPA de soja que interage com *GmbZIPE2*. Na figura está representada 1000pb da região *upstream* do gene. Em verde escuro está representado o *cis*-elemento H-box, seguido das quatro bases existentes também após o H-box da região promotora de *CHS8*. A região verde clara representa a UTR e azul o ATG do início da transcrição.

Identificador	Gene	log2 FoldChange
Glyma.05G168100.3	GmbZIPE2	13,26711
Glyma.05G168100.1	GmbZIPE2	9,732443
Glyma.06G028100.1	FE(3+)-ZN(2+) PURPLE ACID PHOSPHATASE 12	8,63327
Glyma.U016600.1	PHOTOSYSTEM I P700 CHLOROPHYLL A APOPROTEIN A1	8,546709
Glyma.12G061600.1	RIBULOSE BISPHTOSPHATE CARBOXYLASE LARGE CHAIN-RELATED	8,144652
Glyma.02G015500.1	Kinesin motor domain (Kinesin) // Calponin homology (CH) domain (CH)	8,009271
Glyma.04G095000.1	Photosynthetic reaction centre protein (Photo_RC)	7,995272
Glyma.11G048300.1	Calcium-transporting ATPase / Calcium- translocating P-type ATPase	7,826394
Glyma.15G114600.1	CYTOCHROME B	7,814509
Glyma.12G231900.1	ATP synthase B/B' CF(0) (ATP-synt_B)	7,808237
Glyma.01G095900.1	PHOTOSYSTEM II REACTION CENTER PROTEIN Z	7,75426
Glyma.07G122200.1	CYTOCHROME B	7,696461
Glyma.18G124800.1	SINGLE-STRANDED DNA-BINDING PROTEIN WHY1, CHLOROPLASTIC-RELATED	7,686284
Glyma.03G168000.1	ABC TRANSPORTER G FAMILY MEMBER 38	7,641826
Glyma.05G092000.1	Cytochrome b	7,564028
Glyma.13G028200.1	PHOTOSYSTEM Q(B) PROTEIN	7,537813
Glyma.07G143800.1	PHOTOSYSTEM I P700 CHLOROPHYLL A APOPROTEIN A1	7,537664
Glyma.15G188400.1	PROTEIN PSBN	7,53045
Glyma.05G074900.1	PHOTOSYSTEM II CP43 CHLOROPHYLL APOPROTEIN	7,476747
Glyma.12G232900.1	APOCYTOCHROME F	7,462822
Glyma.12G232100.1	39S RIBOSOMAL PROTEIN L20	7,435288
Glyma.12G232300.1	50S RIBOSOMAL PROTEIN L33, CHLOROPLASTIC	7,430779
Glyma.06G217900.1	PHOTOSYSTEM II D2 PROTEIN	7,415285
Glyma.05G073600.1	Photosynthetic reaction centre protein (Photo_RC)	7,378984
Glyma.05G092300.1	ATP SYNTHASE SUBUNIT ALPHA, MITOCHONDRIAL	7,285925
Glyma.16G127900.2	WRKY TRANSCRIPTION FACTOR 1-RELATED	7,245823
Glyma.12G232200.1	30S RIBOSOMAL PROTEIN S18	7,232726
Glyma.06G265000.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	7,229424
Glyma.11G140500.1	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED	7,209269
Glyma.19G201600.1	FASCICLIN-LIKE ARABINOGLACTAN PROTEIN 15-RELATED	7,206836
Glyma.13G088500.1	4Fe-4S binding domain (Fer4)	7,198771
Glyma.08G285900.1	NUCLEOREDOXIN 1-RELATED	7,184903
Glyma.15G011900.1	ABC TRANSPORTER G FAMILY MEMBER 38	7,162784
Glyma.16G165700.1	TPR REPEAT CONTAINING PROTEIN	7,15688
Glyma.01G058600.1	CYTOCHROME B6	7,133661
Glyma.15G274900.1	Protein of unknown function (DUF2647) (DUF2647)	7,104992

Glyma.13G272000.1	EF-HAND CALCIUM-BINDING DOMAIN CONTAINING PROTEIN	7,084989
Glyma.16G195300.1	(M=305) PF00646 - F-box domain (F-box)	7,084645
Glyma.14G094700.1	Glycine dehydrogenase (aminomethyl-transferring) / Glycine-cleavage complex P-protein	7,070944
Glyma.11G162200.1	PHOTOSYSTEM Q(B) PROTEIN	7,059922
Glyma.08G363500.1	PHOTOSYSTEM II REACTION CENTER PROTEIN Z	7,053307
Glyma.02G189000.1	Ycf1 (Ycf1)	7,048887
Glyma.19G054200.1	ATP SYNTHASE SUBUNIT A	7,039862
Glyma.04G208300.1	GRAB1-LIKE PROTEIN	7,001719
Glyma.13G285400.1	GRAS domain family (GRAS)	6,978723
Glyma.U040900.1	NADH-UBIQUINONE OXIDOREDUCTASE 49 KDA SUBUNIT	6,95718
Glyma.13G361900.1	ABC TRANSPORTER G FAMILY MEMBER 38	6,951262
Glyma.12G232000.1	ATP SYNTHASE SUBUNIT ALPHA, CHLOROPLASTIC	6,950518
Glyma.17G185300.1	Mitovirus RNA-dependent RNA polymerase (Mitovir_RNA_pol)	6,941773
Glyma.05G187300.1	Cellulose synthase (GDP-forming) / GDP-glucose-cellulose glucosyltransferase	6,891086
Glyma.13G046200.1	RIBULOSE BISPHTHOSPHATE CARBOXYLASE SMALL CHAIN 1A, CHLOROPLASTIC-RELATED	6,886803
Glyma.U041500.1	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5	6,884603
Glyma.09G144500.1	PHOTOSYSTEM I P700 CHLOROPHYLL A APOPROTEIN A1	6,858968
Glyma.07G028200.1	CemA family (CemA)	6,832955
Glyma.02G296600.1	ZINC FINGER CCCH DOMAIN-CONTAINING PROTEIN 29-RELATED	6,829123
Glyma.17G185500.1	40S RIBOSOMAL PROTEIN S10, MITOCHONDRIAL	6,826215
Glyma.09G131500.1	HEAT SHOCK PROTEIN 90	6,813992
Glyma.04G145900.1	ATP SYNTHASE PROTEIN YMF19	6,811049
Glyma.17G042100.1	POLYUBIQUITIN 4	6,799259
Glyma.14G200200.1	WRKY TRANSCRIPTION FACTOR 1-RELATED	6,795813
Glyma.20G042300.1	NADH DEHYDROGENASE UBIQUINONE IRON-SULFUR PROTEIN 8, MITOCHONDRIAL	6,781749
Glyma.17G121600.1	TRANSCRIPTION INITIATION FACTOR TFIID	6,761545
Glyma.19G159900.1	E1-E2 ATPase (E1-E2_ATPase) // Cation transporting ATPase, C-terminus (Cation_ATPase_C) // Cation transporter/ATPase, N-terminus (Cation_ATPase_N) // haloacid dehalogenase-like hydrolase (HAD)	6,729065
Glyma.11G114700.1	PHOTOSYSTEM II CP43 CHLOROPHYLL APOPROTEIN	6,700609
Glyma.06G138100.1	F11O4.3-RELATED	6,692158
Glyma.14G016300.1	ZINC FINGER CCCH DOMAIN-CONTAINING PROTEIN 29-RELATED	6,690226
Glyma.01G153300.1	ABSCISIC ACID 8'-HYDROXYLASE 1-RELATED	6,685735
Glyma.06G263900.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	6,655883

Glyma.07G222600.1	Domain of unknown function (DUF4283) (DUF4283)	6,617195
Glyma.03G098600.1	2-alkenal reductase (NADP(+)) / NADPH- dependent alkenal/one oxidoreductase // 2-alkenal reductase (NAD(P)(+)) / NADPH:2-alkenal alpha,beta-hydrogenase	6,591359
Glyma.12G064000.1	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED	6,576937
Glyma.11G179300.1	PHENYLPROPIONATE/CINNAMIC ACID DIOXYGENASE FERREDOXIN--NAD(+) REDUCTASE COMPONENT	6,575595
Glyma.04G009400.1	Dehydrin (Dehydrin)	6,574099
Glyma.15G248600.1	PHOTOSYSTEM II REACTION CENTER PROTEIN K	6,560966
Glyma.03G065000.1	Photosystem II 10 kDa phosphoprotein (PsbH)	6,542171
Glyma.07G139800.1	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING	6,537064
Glyma.01G153500.1	PHOTOSYSTEM II D2 PROTEIN	6,536654
Glyma.18G155700.1	NADH:ubiquinone oxidoreductase, NDUFS2/49 kDa subunit	6,534573238
Glyma.14G149900.1	40S RIBOSOMAL PROTEIN S14/30S RIBOSOMAL PROTEIN S11	6,515835442
Glyma.02G222300.1	F11O4.3-RELATED	6,497480631
Glyma.08G042700.1	GB	6,495808386
Glyma.08G162200.1	METHYLATED RNA-BINDING PROTEIN 1	6,495271816
Glyma.11G000500.1	GLUCOSYL/GLUCURONOSYL TRANSFERASES	6,494073501
Glyma.12G232500.1	CYTOCHROME B6-F COMPLEX SUBUNIT 5	6,491488113
Glyma.05G091900.1	60S RIBOSOMAL PROTEIN L11-RELATED	6,488339255
Glyma.15G177800.1	Cathepsin H / N-benzoylarginine-beta- naphthylamide hydrolase	6,452016571
Glyma.13G095000.1	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE PROTEIN 25-RELATED	6,449487766
Glyma.13G095000.1	PHOTOSYSTEM II REACTION CENTER PROTEIN J	6,445121018
Glyma.03G157700.1	E1-E2 ATPase (E1-E2_ATPase) // Cation transporting ATPase, C-terminus (Cation_ATPase_C) // Cation transporter/ATPase, N-terminus (Cation_ATPase_N) // haloacid dehalogenase-like hydrolase (HAD)	6,407205291
Glyma.02G093200.1	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED	6,404677707
Glyma.15G251600.1	GLUTATHIONE S-TRANSFERASE U21- RELATED	6,40349239
Glyma.07G023300.1	WRKY TRANSCRIPTION FACTOR 40- RELATED	6,385199916
Glyma.01G071300.1	S-ADENOSYLMETHIONINE DECARBOXYLASE PROENZYME-RELATED	6,3685605
Glyma.17G185600.1	ATP SYNTHASE PROTEIN MI25	6,341252038
Glyma.16G133700.1	ATP SYNTHASE SUBUNIT A	6,328456561
Glyma.03G236300.1	RESPIRATORY BURST OXIDASE HOMOLOG PROTEIN B	6,322966231
Glyma.08G200500.1	POLY A POLYMERASE	6,310582944

Glyma.15G110300.1	WRKY DNA -binding domain (WRKY)	6,309528648
Glyma.05G149700.1	KINESIN-LIKE PROTEIN KIN12A-RELATED	6,309464622
Glyma.05G195000.1	NAC DOMAIN-CONTAINING PROTEIN 2	6,30189095
Glyma.07G227300.1	Type II intron maturase (Intron_maturas2)	6,277674052
Glyma.03G171100.1	HEAT SHOCK PROTEIN 70KDA	6,274341989
Glyma.15G208300.1	Photosystem II reaction centre M protein (PsbM) (PsbM)	6,27289445
Glyma.15G238900.1	30S RIBOSOMAL PROTEIN S18	6,272584781
Glyma.06G069600.1	CELLULOSE SYNTHASE A CATALYTIC SUBUNIT 1 [UDP-FORMING]	6,258690787
Glyma.13G070800.1	GENOMIC DNA, CHROMOSOME 3, P1 CLONE:MMG15	6,257430944
Glyma.13G261200.1	CALCIUM-DEPENDENT LIPID-BINDING DOMAIN-CONTAINING PROTEIN	6,253224414
Glyma.15G251500.1	GLUTATHIONE S-TRANSFERASE U21-RELATED	6,252433042
Glyma.09G090100.1	OUTER MEMBRANE PROTEIN, OMPA-RELATED	6,249084238
Glyma.14G132100.1	MatK/TrnK amino terminal region (MatK_N)	6,246916885
Glyma.11G187700.1	Acid phosphatase / Phosphomonoesterase // Protein-tyrosine-phosphatase / PTPase	6,212371096
Glyma.12G232700.1	CYTOCHROME B559 SUBUNIT ALPHA	6,210514444
Glyma.02G298500.1	Xenobiotic-transporting ATPase / Steroid-transporting ATPase	6,204263107
Glyma.13G042700.1	Integral membrane protein	6,201615641
Glyma.06G052900.1	Mitochondrial solute carrier protein	6,201435241
Glyma.12G155300.1	CYTOCHROME C OXIDASE SUBUNIT III	6,19598465
Glyma.15G132800.1	F22M8.11 PROTEIN-RELATED	6,189520548
Glyma.03G248400.1	ALDO/KETO REDUCTASE	6,177168889
Glyma.01G193600.1	CALCIUM-TRANSPORTING ATPASE 2, PLASMA MEMBRANE-TYPE-RELATED	6,152890653
Glyma.06G191300.1	Lipid exporter ABCA1 and related proteins, ABC superfamily	6,145748028
Glyma.18G018600.1	Inositol-3-phosphate synthase / Myo-inositol-1-phosphate synthase	6,14228419
Glyma.02G203800.1	WRKY TRANSCRIPTION FACTOR 32-RELATED	6,141738151
Glyma.03G005800.1	MATE EFFLUX FAMILY PROTEIN	6,139154722
Glyma.14G195200.1	PROTEIN PHOSPHATASE 2C-LIKE PROTEIN-RELATED	6,138505643
Glyma.09G076300.1	CLUSTERED MITOCHONDRIA PROTEIN HOMOLOG	6,127861523
Glyma.02G128000.1	S-ADENOSYLMETHIONINE DECARBOXYLASE PROENZYME-RELATED	6,127002214
Glyma.20G048700.1	ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1	6,102014208
Glyma.07G039900.1	LIPOXYGENASE 3, CHLOROPLASTIC-RELATED	6,097678422
Glyma.08G281300.1	Photosystem II protein (PSII)	6,096221864
Glyma.20G150600.1	EARLY LIGHT-INDUCED PROTEIN 1, CHLOROPLASTIC-RELATED	6,095443434
Glyma.19G046800.1	RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1A, CHLOROPLASTIC-	6,094992776

	RELATED	
Glyma.14G151100.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	6,085534017
Glyma.06G295700.1	O-METHYLTRANSFERASE FAMILY PROTEIN	6,085284923
Glyma.12G204400.1	VOLTAGE-GATED CATION CHANNEL CALCIUM AND SODIUM	6,084857232
Glyma.02G292800.1	GDP-D-GLUCOSE PHOSPHORYLASE 1	6,084530083
Glyma.12G237000.1	X-BOX TRANSCRIPTION FACTOR-RELATED	6,079505222
Glyma.14G120700.1	EXOSTOSIN FAMILY PROTEIN	6,074975668
Glyma.08G301000.1	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 6	6,068705353
Glyma.13G042200.1	OXIDOREDUCTASE, 2OG-FE II OXYGENASE FAMILY PROTEIN	6,068663233
Glyma.12G096200.1	ATP SYNTHASE SUBUNIT A	6,06054277
Glyma.01G084700.1	ATP-dependent Clp protease adaptor protein ClpS (ClpS)	6,050800714
Glyma.04G008300.1	FRUCTOSE-BISPHOSPHATE ALDOLASE 1, CHLOROPLASTIC-RELATED	6,018501697
Glyma.15G026500.1	Linoleate 9S-lipoxygenase / Linoleate 9-lipoxygenase	6,018296614
Glyma.10G179400.1	Very-long-chain 3-oxoacyl-CoA synthase / Very-long-chain beta-ketoacyl-CoA synthase	6,017612864
Glyma.05G068600.1	LON PROTEASE HOMOLOG 2, PEROXISOMAL	6,016788557
Glyma.15G173500.1	CHARGED MULTIVESICULAR BODY PROTEIN	6,016154827
Glyma.03G094400.1	CYTOCHROME C OXIDASE SUBUNIT 1	6,011631487
Glyma.15G015100.1	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase // Chitinase / Poly-beta-glucosaminidase	6,011505794
Glyma.09G058100.1	30S RIBOSOMAL PROTEIN S12 FAMILY MEMBER	6,010127754
Glyma.13G043800.1	ATP-BINDING CASSETTE TRANSPORTER	6,008121141
Glyma.17G256100.1	3-isopropylmalate dehydratase (aconitase superfamily)	6,008116339
Glyma.12G173500.1	Kelch motif (Kelch_1)	6,007089528
Glyma.03G031100.1	30S RIBOSOMAL PROTEIN S2 PROKARYOTIC AND ORGANELLAR	6,006173493
Glyma.15G188000.1	NADH-UBIQUINONE OXIDOREDUCTASE 49 KDA SUBUNIT	6,002800518
Glyma.13G112000.1	PROTEIN TIFY 10A-RELATED	5,981535899
Glyma.07G252100.1	SWI/SNF-RELATED MATRIX-ASSOCIATED ACTIN-DEPENDENT REGULATOR OF CHROMATIN SUBFAMILY-RELATED	5,952941594
Glyma.10G293500.1	Transketolase / Glycoaldehyde transferase // Transaldolase / Glycerone transferase	5,952889613
Glyma.07G147700.1	TUBBY-LIKE F-BOX PROTEIN 1-RELATED	5,951375062
Glyma.08G243000.1	UDP-GLUCOSE DEHYDROGENASE/UDP-MANNAC DEHYDROGENASE	5,950502189
Glyma.13G255900.1	CLASS I GLUTAMINE AMIDOTRANSFERASE DOMAIN-CONTAINING PROTEIN-RELATED	5,948886598
Glyma.09G044200.1	TRANSLATIONALLY CONTROLLED TUMOR PROTEIN-RELATED	5,948680799
Glyma.05G095400.1	SF44 - FERRITIN	5,948277047
Glyma.15G187800.1	NADH-UBIQUINONE OXIDOREDUCTASE	5,948252371

	CHAIN 5	
Glyma.12G054700.1	PROTEIN C13C4.7	5,948123304
Glyma.10G035700.1	DNA REPLICATION FACTOR CDT1	5,923355397
Glyma.13G279900.1	NAC DOMAIN-CONTAINING PROTEIN 19-RELATED	5,923169042
Glyma.20G066800.1	ERD (EARLY-RESPONSIVE TO DEHYDRATION STRESS) FAMILY PROTEIN-RELATED	5,881019183
Glyma.18G045100.1	Spermine oxidase / SMO // Non-specific polyamine oxidase / Polyamine oxidase	5,877289324
Glyma.02G195600.1	PROTEIN VIRILIZER HOMOLOG	5,876540334
Glyma.18G180800.1	Metallothionein (Metallothio_2)	5,870130905
Glyma.01G238800.1	SUGAR TRANSPORT PROTEIN 13	5,869138503
Glyma.17G063000.1	FLAVODOXIN-RELATED	5,865390876
Glyma.06G114300.1	ALDEHYDE DEHYDROGENASE FAMILY 3 MEMBER I1, CHLOROPLASTIC	5,864745014
Glyma.17G186600.1	TRANSLATION FACTOR	5,859900109
Glyma.14G213500.1	30S RIBOSOMAL PROTEIN S12, CHLOROPLASTIC	5,856369541
Glyma.09G178600.1	PHOTOSYSTEM I ASSEMBLY PROTEIN YCF3	5,843744525
Glyma.U009100.1	30S RIBOSOMAL PROTEIN S3, CHLOROPLASTIC	5,82543188
Glyma.11G039900.1	PLECKSTRIN HOMOLOGY (PH) DOMAIN-CONTAINING PROTEIN	5,80476981
Glyma.19G044900.1	ALLENE OXIDE CYCLASE 1, CHLOROPLASTIC-RELATED	5,802132794
Glyma.08G042900.1	CDC48-INTERACTING UBX-DOMAIN PROTEIN	5,801363351
Glyma.14G216300.1	CYSTEINE PROTEASE FAMILY C1-RELATED	5,801154357
Glyma.09G011200.1	AUXIN-INDUCED IN ROOT CULTURES PROTEIN 12	5,800919759
Glyma.10G002200.1	CALCIUM BINDING PROTEIN	5,799380387
Glyma.13G220900.1	EARLY-RESPONSIVE TO DEHYDRATION STRESS PROTEIN (ERD4)	5,797846216
Glyma.15G199800.1	TRANSMEMBRANE PROTEIN DDB_G0292058	5,797113946
Glyma.16G065300.1	Ycf1 (Ycf1)	5,794199904
Glyma.04G150600.1	DNA-directed RNA polymerase / RNA polymerase III	5,794173707
Glyma.10G098500.1	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4	5,787799479
Glyma.15G019300.1	MALIC ENZYME-RELATED	5,781976876
Glyma.08G050600.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT K, CHLOROPLASTIC	5,77450711
Glyma.13G215000.1	INACTIVE BETA-AMYLASE 9	5,769585227
Glyma.16G008700.1	LIPOXYGENASE 3, CHLOROPLASTIC-RELATED	5,721528577
Glyma.14G116800.1	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE FAF-RELATED	5,721443992
Glyma.09G138100.1	Oxalate--CoA ligase / Oxalyl-CoA synthetase	5,720383913
Glyma.06G157400.1	GRAB1-LIKE PROTEIN	5,718942069
Glyma.20G180100.1	GATA TRANSCRIPTION FACTOR 8	5,717872834
Glyma.07G233800.1	S-ADENOSYLMETHIONINE SYNTHASE 4	5,716636217
Glyma.13G047300.1	ALLENE OXIDE CYCLASE 1,	5,71309476

	CHLOROPLASTIC-RELATED	
Glyma.07G261500.1	Adenylyl-sulfate reductase (thioredoxin) / Thioredoxin-dependent 5'-adenylylsulfate reductase	5,710987729
Glyma.11G096700.1	EXPANSIN-LIKE A1-RELATED	5,704995722
Glyma.01G114300.1	MEMBRANE PROTEIN YCF1-RELATED	5,704162705
Glyma.11G043700.1	TRANSCRIPTION FACTOR NAI1	5,701811919
Glyma.18G155300.1	4Fe-4S dicluster domain (Fer4_7)	5,698202268
Glyma.01G101600.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 2 A, CHLOROPLASTIC-RELATED	5,696007571
Glyma.05G132900.1	E1-E2 ATPase (E1-E2_ATPase) // Heavy-metal-associated domain (HMA) // haloacid dehalogenase-like hydrolase (HAD)	5,691921524
Glyma.09G218600.1	ABSCISIC ACID 8'-HYDROXYLASE 1-RELATED	5,690724815
Glyma.14G179200.1	OVARIAN CANCER GENE-2 PROTEIN-RELATED	5,670899319
Glyma.13G354900.1	MALIC ENZYME-RELATED	5,631778992
Glyma.13G003200.1	COPINE	5,631314614
Glyma.08G170000.1	TRANSLATION FACTOR	5,630867136
Glyma.06G108100.1	BETA-1,3-GALACTOSYLTRANSFERASE 7-RELATED	5,629696899
Glyma.07G262700.1	WRKY DNA -binding domain (WRKY)	5,627865413
Glyma.07G015000.1	POLY A POLYMERASE	5,625975658
Glyma.18G259000.1	60S RIBOSOMAL PROTEIN L34	5,62596518
Glyma.17G162200.1	F-box associated (FBA_1)	5,622680456
Glyma.11G121900.1	Domain of unknown function (DUF4228) (DUF4228)	5,622464182
Glyma.09G027900.1	WALL-ASSOCIATED RECEPTOR KINASE-LIKE 1-RELATED	5,621391304
Glyma.19G081000.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 6, CHLOROPLASTIC	5,616353456
Glyma.20G141600.1	Ubiquitin and ubiquitin-like proteins // Ubiquitin-like protein	5,610094455
Glyma.20G131500.1	PROTEIN PHOSPHATASE 2C 63-RELATED	5,605954809
Glyma.19G051900.1	50S RIBOSOMAL PROTEIN L2, CHLOROPLASTIC	5,591057651
Glyma.15G022500.1	Jacalin-like lectin domain (Jacalin)	5,570683596
Glyma.03G248300.1	MuDR family transposase (DBD_Tnp_Mut) // SWIM zinc finger (SWIM) // MULE transposase domain (MULE)	5,53755808
Glyma.13G296900.1	VOLTAGE-GATED CATION CHANNEL CALCIUM AND SODIUM	5,536887024
Glyma.01G082400.1	AGENET DOMAIN-CONTAINING PROTEIN / BROMO-ADJACENT HOMOLOGY (BAH) DOMAIN-CONTAINING PROTEIN	5,536426561
Glyma.04G249000.1	GRAB1-LIKE PROTEIN	5,536348229
Glyma.12G095700.1	DNAJ HOMOLOG SUBFAMILY C MEMBER	5,535697102
Glyma.11G077000.1	TRANSFORMATION/TRANSCRIPTION DOMAIN-ASSOCIATED PROTEIN	5,534440361
Glyma.05G133100.1	E3 UBIQUITIN-PROTEIN LIGASE HUWE1	5,534417982
Glyma.05G138700.1	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE PROTEIN 27-RELATED	5,533291607

Glyma.03G202600.1	E3 UBIQUITIN-PROTEIN LIGASE PUB22-RELATED	5,533193144
Glyma.02G008000.1	ABC TRANSPORTER B FAMILY MEMBER 11-RELATED	5,53127676
Glyma.18G067300.1	TRANSFORMER-SR RIBONUCLEOPROTEIN-RELATED	5,531097206
Glyma.01G010500.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	5,529423172
Glyma.01G208600.1	G2484-1 PROTEIN	5,526920688
Glyma.05G236300.1	PLAC8 family (PLAC8)	5,526333374
Glyma.06G186300.1	Betaine-aldehyde dehydrogenase / Betaine aldehyde oxidase	5,524661824
Glyma.02G047600.1	NADPH dehydrogenase / NADPH diaphorase	5,523763655
Glyma.16G089800.1	NADH DEHYDROGENASE UBIQUINONE IRON-SULFUR PROTEIN 3	5,519365529
Glyma.17G186100.1	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4L	5,516656332
Glyma.08G316500.1	MAP KINASE-ACTIVATED PROTEIN KINASE 5	5,515066612
Glyma.10G047700.1	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING	5,512233587
Glyma.10G208900.1	Photosystem I reaction centre subunit VIII (PSI_8)	5,504073053
Glyma.14G021800.1	GDP-D-GLUCOSE PHOSPHORYLASE 1	5,465182322
Glyma.19G106800.1	Glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) / Triosephosphate dehydrogenase // Glyceraldehyde-3-phosphate dehydrogenase (NADP(+)) (phosphorylating) / Triosephosphate dehydrogenase (NADP+)	5,457081079
Glyma.15G003300.1	WRKY TRANSCRIPTION FACTOR 40-RELATED	5,435429983
Glyma.20G165000.1	ABC TRANSPORTER C FAMILY MEMBER 10	5,434078416
Glyma.17G202700.1	CHY AND CTCHY AND RING-TYPE ZINC FINGER PROTEIN	5,43367406
Glyma.09G273300.1	RECEPTOR-LIKE PROTEIN KINASE FERONIA	5,433647069
Glyma.08G332900.1	HEAT SHOCK PROTEIN 90	5,43283965
Glyma.06G114000.1	GRAB1-LIKE PROTEIN	5,432299176
Glyma.05G055500.1	RIBOSOME-RELEASING FACTOR 2, MITOCHONDRIAL	5,431514998
Glyma.02G008500.1	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase	5,430380818
Glyma.07G139700.1	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING	5,429287307
Glyma.05G045700.1	TRANSCRIPTION FACTOR GTE10	5,429231412
Glyma.20G151500.1	Sulfate adenylyltransferase / Sulfurylase	5,424442828
Glyma.13G295200.1	ZINC FINGER CCCH DOMAIN-CONTAINING PROTEIN 2-RELATED	5,424264741
Glyma.03G066800.1	ALDO-KETO REDUCTASE FAMILY 4 MEMBER C10	5,423100578
Glyma.U009000.1	60S RIBOSOMAL PROTEIN L23A	5,421553989
Glyma.01G166400.1	TRANSFORMATION/TRANSCRIPTION DOMAIN-ASSOCIATED PROTEIN	5,417378406
Glyma.07G072500.1	SYNAPTOTAGMIN	5,416348738
Glyma.10G174800.1	ATP SYNTHASE EPSILON CHAIN, CHLOROPLASTIC	5,410733688

Glyma.16G008400.1	GTP diphosphokinase / Stringent factor // Xylan 1,4-beta-xylosidase / Xylobiase // Non-reducing end alpha-L-arabinofuranosidase / Arabinosidase	5,40060676
Glyma.10G068800.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 2 A, CHLOROPLASTIC-RELATED	5,394277625
Glyma.11G194600.1	HEAVY METAL TRANSPORT/DETOXIFICATION DOMAIN-CONTAINING PROTEIN-RELATED	5,372523011
Glyma.14G103300.1	NADH DEHYDROGENASE [UBIQUINONE] IRON-SULFUR PROTEIN 2, MITOCHONDRIAL	5,368131633
Glyma.13G370100.1	WRKY TRANSCRIPTION FACTOR 40-RELATED	5,343683972
Glyma.15G248500.1	ACETYL-COENZYME A CARBOXYLASE CARBOXYL TRANSFERASE SUBUNIT BETA, CHLOROPLASTIC	5,328313853
Glyma.08G252600.1	PUSHOVER/RETINOBLASTOMA-ASSOCIATED FACTOR 600	5,322483745
Glyma.20G040700.1	TELOMERASE-BINDING PROTEIN EST1A	5,322192946
Glyma.04G200600.1	AUXIN RESPONSE FACTOR 2	5,32196506
Glyma.17G020600.1	Pyruvate, phosphate dikinase / Pyruvate, phosphate dikinase	5,321877796
Glyma.10G097800.1	SPOT2-RELATED	5,32159739
Glyma.01G219000.1	E1-E2 ATPase (E1-E2_ATPase) // Heavy-metal-associated domain (HMA)	5,321378393
Glyma.17G174500.1	ZEAXANTHIN EPOXIDASE, CHLOROPLASTIC	5,32137278
Glyma.18G281600.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	5,321281544
Glyma.15G171000.1	CALMODULIN-BINDING PROTEIN	5,321257702
Glyma.17G072400.1	HEAT SHOCK 70 KDA PROTEIN 5	5,321050938
Glyma.04G066900.1	F-box domain (F-box) // Leucine Rich repeat (LRR_6)	5,320840594
Glyma.08G039600.1	ARF-GAP DOMAIN AND FG REPEAT-CONTAINING PROTEIN 1	5,320813561
Glyma.10G224500.1	Protein of unknown function (DUF793) (DUF793)	5,320790551
Glyma.10G055300.1	HEAVY-METAL-ASSOCIATED DOMAIN-CONTAINING PROTEIN-RELATED	5,319837417
Glyma.02G009100.1	PROTEIN PHOSPHATASE 2C 13-RELATED	5,319087487
Glyma.12G102400.1	PROTEIN MET-1, ISOFORM A	5,317354541
Glyma.03G230000.1	dTDP-glucose 4-6-dehydratase/UDP-glucuronic acid decarboxylase // C-3 sterol dehydrogenase/3-beta-hydroxysteroid dehydrogenase and related dehydrogenases	5,31699004
Glyma.19G199500.1	BEL1-LIKE HOMEODOMAIN PROTEIN 1-RELATED	5,316159613
Glyma.19G018600.1	AAA-TYPE ATPASE FAMILY PROTEIN-RELATED	5,315216133
Glyma.02G202500.1	Glyceraldehyde-3-phosphate dehydrogenase (NADP(+)) / Triosephosphate dehydrogenase	5,315150869
Glyma.13G205600.1	POLYADENYLATE-BINDING PROTEIN RBP45B-RELATED	5,314164596
Glyma.08G336400.1	PROTEIN T28B8.3, ISOFORM A-RELATED	5,308235586
Glyma.08G299900.1	Delta(8)-fatty-acid desaturase / SLD	5,307202828
Glyma.19G053400.1	50S/60S RIBOSOMAL PROTEIN L16	5,298374385
Glyma.14G135900.1	Structural maintenance of chromosome protein 3	5,295324846

	(sister chromatid cohesion complex Cohesin, subunit SMC3) // E3 ubiquitin ligase involved in syntaxin degradation // Uncharacterized conserved protein // Uncharacterized conserved coiled-coil protein	
Glyma.19G045000.1	PNAS-3 RELATED	5,294188502
Glyma.09G065000.1	CALMODULIN-BINDING PROTEIN	5,293571215
Glyma.06G224500.1	PHOTOSYSTEM I P700 CHLOROPHYLL A APOPROTEIN A1	5,292601173
Glyma.11G235100.1	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 3	5,290478125
Glyma.18G155500.1	NADH DEHYDROGENASE UBIQUINONE IRON-SULFUR PROTEIN 8, MITOCHONDRIAL	5,289682497
Glyma.09G171300.1	CYTOCHROME B6-F COMPLEX SUBUNIT 8	5,286279291
Glyma.13G068600.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 3, CHLOROPLASTIC	5,243962077
Glyma.20G145800.1	SERINE/THREONINE-PROTEIN KINASE WNK WITH NO LYSINE -RELATED	5,236385652
Glyma.20G191700.1	PASILLA, ISOFORM D	5,209531154
Glyma.05G082500.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	5,19925787
Glyma.12G071300.1	E3 SUMO-PROTEIN LIGASE SIZ1	5,199236796
Glyma.20G170000.1	PROTEIN GIGANTEA	5,19916265
Glyma.06G098700.1	HECT DOMAIN UBIQUITIN-PROTEIN LIGASE	5,19905205
Glyma.16G025900.1	ABSCISIC ACID RECEPTOR PYL7-RELATED	5,198332351
Glyma.06G315100.1	PHOSPHOLIPID-TRANSPORTING ATPASE 1	5,19815577
Glyma.06G045400.1	C2H2-type zinc finger (zf-C2H2_6)	5,198123572
Glyma.06G076100.1	Proton-exporting ATPase / Proton-translocating P-type ATPase	5,197773666
Glyma.17G096400.1	PROTEIN NRT1/ PTR FAMILY 2.10-RELATED	5,197305716
Glyma.10G067500.1	MONODEHYDROASCORBATE REDUCTASE, CYTOPLASMIC ISOFORM 1-RELATED	5,197201256
Glyma.05G191700.1	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 3	5,196998466
Glyma.18G030400.1	Sedoheptulose-bisphosphatase / Sedoheptulose-1,7-bisphosphatase	5,196634266
Glyma.10G219600.1	F23M19.14 PROTEIN-RELATED	5,1954011
Glyma.07G132000.1	Metallothionein (Metallothio_2)	5,191203018
Glyma.07G199800.1	MACPF DOMAIN-CONTAINING PROTEIN NSL1	5,191013239
Glyma.08G188000.1	TESTIS-SPECIFIC PROTEIN PBS13 T-COMPLEX 11	5,188781054
Glyma.14G176400.1	(PHI-1) PROTEIN, PUTATIVE-RELATED	5,185628284
Glyma.08G259000.1	DISEASE RESISTANCE PROTEIN RPP13-RELATED	5,175532649
Glyma.11G157900.1	ALPHA/BETA-HYDROLASES SUPERFAMILY PROTEIN	5,175176903
Glyma.20G045200.1	Plant protein of unknown function (DUF825) (DUF825)	5,175164653
Glyma.07G028100.1	CemA family (CemA)	5,17242657
Glyma.10G078900.1	50S RIBOSOMAL PROTEIN L2, CHLOROPLASTIC	5,1712674
Glyma.10G238400.1	30S RIBOSOMAL PROTEIN S7, CHLOROPLASTIC	5,166435327

Glyma.06G265400.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	5,158440829
Glyma.17G030700.1	COBRA-LIKE PROTEIN 7-RELATED	5,126670286
Glyma.07G043900.1	PHOSPHOINOSITIDE PHOSPHATASE SAC3-RELATED	5,087453231
Glyma.04G017500.1	CATALASE	5,064663562
Glyma.02G089000.1	LISH DOMAIN AND HEAT REPEAT-CONTAINING PROTEIN KIAA1468	5,064323377
Glyma.12G073900.1	TWO-COMPONENT RESPONSE REGULATOR-LIKE APRR3	5,063958581
Glyma.13G239100.1	UBIQUITIN-CONJUGATING ENZYME E2 24-RELATED	5,063377035
Glyma.13G347700.1	Linoleate 9S-lipoxygenase / Linoleate 9-lipoxygenase	5,062958375
Glyma.06G259100.1	NB-ARC domain (NB-ARC) // Leucine Rich Repeat (LRR_3) // TIR domain (TIR_2)	5,062893749
Glyma.02G296400.1	GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE-LIKE PROTEIN-RELATED	5,062580666
Glyma.10G051100.1	Ubiquitin and ubiquitin-like proteins // Ubiquitin-like protein	5,062498976
Glyma.07G046800.1	Non-specific protein-tyrosine kinase / Cytoplasmic protein tyrosine kinase	5,0624913
Glyma.07G041400.1	PROTEIN TIFY 10A-RELATED	5,062159394
Glyma.05G215900.1	WRKY DNA -binding domain (WRKY)	5,061263552
Glyma.08G168700.1	TPR REPEAT CONTAINING PROTEIN	5,061106348
Glyma.08G024600.1	DCD (DEVELOPMENT AND CELL DEATH) DOMAIN PROTEIN	5,06098555
Glyma.16G035100.1	PROTEIN TAF-3, ISOFORM A	5,060805407
Glyma.06G261500.1	NB-ARC domain (NB-ARC) // Leucine Rich Repeat (LRR_3) // TIR domain (TIR_2)	5,060015585
Glyma.05G120100.1	EMB	5,059815551
Glyma.03G131900.1	PROTEIN EARLY RESPONSIVE TO DEHYDRATION 15	5,059740934
Glyma.04G019100.1	ATP-DEPENDENT ZINC METALLOPROTEASE FTSH 1, CHLOROPLASTIC-RELATED	5,059423947
Glyma.17G255800.1	BASIC-LEUCINE ZIPPER (BZIP) TRANSCRIPTION FACTOR FAMILY PROTEIN	5,05914572
Glyma.17G048900.1	SER/THR PROTEIN KINASE ACIK1B	5,058478056
Glyma.10G242600.1	Sulfate adenylyltransferase / Sulfurylase	5,058439408
Glyma.08G173700.1	PHOTOSYSTEM II 10 KDA POLYPEPTIDE, CHLOROPLASTIC	5,057982746
Glyma.04G004000.1	E3 UBIQUITIN-PROTEIN LIGASE TRIP12	5,057835614
Glyma.06G312300.1	ABC TRANSPORTER C FAMILY MEMBER 11-RELATED	5,057736783
Glyma.13G259400.1	GIBBERELLIN 2-BETA-DIOXYGENASE 1	5,057133606
Glyma.14G085800.1	CYSTEINE PROTEASE FAMILY C1-RELATED	5,055394322
Glyma.04G136400.1	Conserved hypothetical protein (Lin0512_fam) (Lin0512_fam)	5,051730587
Glyma.15G040700.1	C2H2-TYPE ZINC FINGER FAMILY PROTEIN-RELATED	5,051232762
Glyma.05G094500.1	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 2 A, CHLOROPLASTIC-RELATED	5,051188371
Glyma.03G121000.1	TRANSCRIPTION ELONGATION FACTOR	5,05079158

	SPT6	
Glyma.02G184000.1	ATP synthase, Delta/Epsilon chain, long alpha-helix domain (ATP-synt_DE)	5,049142113
Glyma.02G221500.1	30S RIBOSOMAL PROTEIN S2 PROKARYOTIC AND ORGANELLAR	5,048625805
Glyma.13G070500.1	TRANSCRIPTION ELONGATION FACTOR B POLYPEPTIDE 3	5,046280161
Glyma.18G019700.1	Acinus (induces apoptotic chromatin condensation)	5,04557544
Glyma.09G212900.1	MITOGEN-ACTIVATED PROTEIN KINASE	5,045024655
Glyma.13G117700.1	POLYUBIQUITIN 4	5,044816493
Glyma.01G235600.1	NADH OXIDOREDUCTASE-RELATED	5,043562778
Glyma.17G047700.1	PROTEIN TIFY 10A-RELATED	5,042081588
Glyma.14G121000.1	GTPASE CONTAINING FAMILY OF SIGNAL RECOGNITION PARTICLE PROTEINS	5,040588705
Glyma.07G002400.1	40S RIBOSOMAL PROTEIN S10, MITOCHONDRIAL	5,038003891
Glyma.16G096600.1	PHOTOTROPIN-2	5,028940396
Glyma.15G190500.1	S-ADENOSYLMETHIONINE SYNTHASE 4	5,027609876
Glyma.08G045200.1	NUDIX HYDROLASE 17, MITOCHONDRIAL-RELATED	5,025169842
Glyma.13G176400.1	POLYUBIQUITIN 3	5,000950399
Glyma.13G109800.1	12-OXOPHYTODIENOATE REDUCTASE 3	4,999648063
Glyma.19G220200.1	S-ADENOSYLMETHIONINE SYNTHASE 2	4,998583576
Glyma.18G155400.1	Methanosarcina-phenazine hydrogenase / Methylviologen-reducing hydrogenase	4,9954575
Glyma.08G055900.1	PROTEIN SHUTTLE CRAFT	4,971759621
Glyma.05G114900.1	TRANSLATION FACTOR	4,967849487
Glyma.01G102600.1	ALDO-KETO REDUCTASE FAMILY 4 MEMBER C10	4,940148043
Glyma.01G204400.1	PROTEIN TIFY 10A-RELATED	4,917241915
Glyma.03G075300.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	4,915872094
Glyma.13G145000.1	HISTIDINE AMMONIA-LYASE	4,914541433
Glyma.01G225200.1	CDP-diacylglycerol--inositol 3-phosphatidyltransferase / Phosphatidylinositol synthase // Long-chain-fatty-acid--CoA ligase / Lignoceroyl-CoA synthase	4,914500677
Glyma.20G244900.1	CALLOSE SYNTHASE 9	4,914276336
Glyma.04G040000.1	Transferase family (Transferase)	4,914136948
Glyma.01G216600.1	MATE EFFLUX FAMILY PROTEIN	4,913784082
Glyma.14G074700.1	Leucine Rich Repeat (LRR_1) // Protein tyrosine kinase (Pkinase_Tyr) // Leucine rich repeat (LRR_8)	4,91371257
Glyma.16G179900.1	SCARECROW-LIKE PROTEIN 8	4,913276994
Glyma.11G111400.1	FRUCTOSE-BISPHOSPHATE ALDOLASE 1, CHLOROPLASTIC-RELATED	4,913192802
Glyma.13G283600.1	CALCIUM BINDING PROTEIN	4,913072207
Glyma.08G226600.1	NAD DEPENDENT EPIMERASE/DEHYDRATASE	4,912417983
Glyma.02G268000.1	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE 3-RELATED	4,911831541
Glyma.09G020800.1	REGULATORY PROTEIN NPR3-RELATED	4,911067253
Glyma.11G022200.1	HOMEODOMAIN PROTEIN TRANSCRIPTION	4,910868834

	FACTORS	
Glyma.12G109800.1	O-METHYLTRANSFERASE FAMILY PROTEIN	4,910619718
Glyma.01G166100.1	CALCIUM-DEPENDENT PROTEIN KINASE 16-RELATED	4,910309223
Glyma.18G150000.1	SERINE HYDROXYMETHYLTRANSFERASE 1, MITOCHONDRIAL	4,910295454
Glyma.17G230700.1	HEMOGLOBINASE FAMILY MEMBER	4,910053294
Glyma.11G119500.1	POLYOL TRANSPORTER 5	4,909978574
Glyma.08G152500.1	Helix-loop-helix DNA-binding domain (HLH)	4,909667135
Glyma.08G072200.1	Ubiquinol oxidase (non-electrogenic) / Ubiquinol oxidase	4,908694638
Glyma.05G079700.1	CALMODULIN-1-RELATED	4,905918784
Glyma.06G265500.1	GRAS domain family (GRAS)	4,904891679
Glyma.13G094800.1	CALMODULIN-BINDING PROTEIN	4,90458636
Glyma.15G017000.1	PROTEIN NRT1/ PTR FAMILY 1.1-RELATED	4,904571941
Glyma.14G046100.1	UBIQUITIN FUSION DEGRADATION PROTEIN 1	4,904323071
Glyma.01G026700.1	Glycine transaminase / Glycine aminotransferase	4,904166976
Glyma.16G044900.1	Glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) / Triosephosphate dehydrogenase // Glyceraldehyde-3-phosphate dehydrogenase (NADP(+)) (phosphorylating) / Triosephosphate dehydrogenase (NADP+)	4,903088336
Glyma.08G145600.1	Cellulose synthase (GDP-forming) / GDP-glucose-cellulose glucosyltransferase	4,901946984
Glyma.13G330400.1	FLIPPASE KINASE 1-RELATED	4,901803375
Glyma.09G203500.1	Phosphoenolpyruvate carboxykinase (GTP) / Phosphopyruvate carboxylase // Phosphoenolpyruvate carboxykinase (ATP) / Phosphopyruvate carboxylase	4,900875343
Glyma.02G297700.1	GRAS domain family (GRAS)	4,900545346
Glyma.14G093100.1	Metallothionein (Metallothio_2)	4,900482681
Glyma.12G087200.1	CYTOCHROME P450 94C1	4,900316087
Glyma.08G165500.1	PHOSPHOGLYCERATE KINASE 1, CHLOROPLASTIC-RELATED	4,90011581
Glyma.04G059000.1	PROTEIN PHOSPHATASE 2C 59-RELATED	4,899538573
Glyma.05G194100.1	CGI-141-RELATED/LIPASE CONTAINING PROTEIN	4,899295734
Glyma.15G246000.1	Polyphosphate kinase / Polyphosphoric acid kinase	4,898716386
Glyma.15G166800.1	Helix-loop-helix DNA-binding domain (HLH)	4,898488712
Glyma.01G091900.1	DNA-directed RNA polymerase / RNA polymerase III	4,892561086
Glyma.12G090600.1	Protein of unknown function (DUF2647) (DUF2647)	4,889329235
Glyma.14G044400.1	Leucine Rich Repeat (LRR_1) // Leucine rich repeat N-terminal domain (LRRNT_2) // Leucine rich repeat (LRR_8)	4,885373092
Glyma.14G143700.1	LA PROTEIN 1-RELATED	4,884719328
Glyma.13G326800.1	Kelch motif (Kelch_1)	4,884493706
Glyma.13G101100.1	Helix-loop-helix DNA-binding domain (HLH)	4,882456234
Glyma.18G155600.1	NADH DEHYDROGENASE SUBUNIT 1	4,87324397
Glyma.02G155000.1	AAA ATPASE	4,862286918
Glyma.12G116800.1	PROTEIN PHOSPHATASE 2C 20-RELATED	4,861617452

Glyma.12G027400.1	Nuclear localization sequence binding protein	4,861249191
Glyma.12G073000.1	MITOGEN-ACTIVATED PROTEIN KINASE 3	4,860512643
Glyma.18G050600.1	60S RIBOSOMAL PROTEIN L10	4,856696809
Glyma.15G187700.1	DNA polymerase type B, organellar and viral (DNA_pol_B_2)	4,855299245
Glyma.11G215500.1	GLUTAMINE SYNTHETASE	4,767798666
Glyma.14G209100.1	Ycf4 (Ycf4)	4,757738215
Glyma.08G214000.1	TETRATRICOPEPTIDE REPEAT-CONTAINING PROTEIN	4,750386762
Glyma.10G019400.1	AAA ATPASE	4,750320009
Glyma.04G075100.1	Proton-exporting ATPase / Proton-translocating P-type ATPase	4,749910013
Glyma.03G182800.1	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG D	4,749485847
Glyma.10G225800.1	EUKARYOTIC TRANSLATION INITIATION FACTOR 5B	4,749126231
Glyma.04G071900.1	PROTEIN PARTNER OF SNF	4,748797961
Glyma.12G193800.1	Aconitate hydratase / Citrate(isocitrate) hydro-lyase	4,748619443
Glyma.10G143900.1	EMB	4,748518939
Glyma.08G181700.1	Histidine decarboxylase / L-histidine carboxy-lyase	4,748402368
Glyma.18G123800.1	PROTEIN KINASE 2A, CHLOROPLASTIC-RELATED	4,748283933
Glyma.01G068000.1	HEAT SHOCK PROTEIN 89.1	4,748254722
Glyma.18G204800.1	AMMONIUM TRANSPORTER 2	4,7479606
Glyma.01G203500.1	PROTEIN KINASE FAMILY PROTEIN	4,74667182
Glyma.11G126200.1	Hydroxyacylglutathione hydrolase / Glyoxalase II	4,746639259
Glyma.06G264300.1	LEUCINE-RICH REPEAT-CONTAINING PROTEIN	4,74623744
Glyma.13G134200.1	Pectate lyase / PPase-N	4,746214395
Glyma.10G169300.1	F-box-like (F-box-like) // Phloem protein 2 (PP2)	4,745738137
Glyma.02G144200.1	WD domain, G-beta repeat (WD40) // Beige/BEACH domain (Beach) // PH domain associated with Beige/BEACH (PH_BEACH)	4,745626248
Glyma.18G158400.1	Xenobiotic-transporting ATPase / Steroid-transporting ATPase	4,744951561
Glyma.17G065200.1	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE PROTEIN 25-RELATED	4,744739263
Glyma.10G152200.1	RESPIRATORY BURST OXIDASE HOMOLOG PROTEIN B	4,744303945
Glyma.10G200700.1	GLYCOSYLTRANSFERASE FAMILY PROTEIN 47	4,743889834
Glyma.19G046600.1	RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1A, CHLOROPLASTIC-RELATED	4,743511645
Glyma.19G172200.1	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37E-RELATED	4,743275163
Glyma.05G185300.1	SPATACSIN	4,742077159
Glyma.15G007100.1	PRE-MRNA SPLICING FACTOR PRP8	4,741855947
Glyma.20G144400.1	PROTEIN TIME FOR COFFEE	4,741759189
Glyma.15G070100.1	SNF2 , HELICASE AND ZINC-FINGER DOMAIN-CONTAINING PROTEIN-RELATED	4,741377792
Glyma.19G233900.1	RESPIRATORY BURST OXIDASE HOMOLOG	4,740663415

	PROTEIN B	
Glyma.15G186400.1	ASPARTYL PROTEASE FAMILY PROTEIN	4,740253787
Glyma.15G071600.1	Splicing factor 3b, subunit 2	4,740162954
Glyma.05G064700.1	BAX INHIBITOR 1	4,737568178
Glyma.20G124200.1	Domain of unknown function (DUF4228) (DUF4228)	4,737227131
Glyma.13G069800.1	SENESCENCE-ASSOCIATED CARBOXYLESTERASE 101	4,71811355
Glyma.06G119700.1	F24J8.3 PROTEIN	4,707693964
Glyma.12G037400.1	(FRUCTOSE-BISPHOSPHATE ALDOLASE 1, CHLOROPLASTIC-RELATED	4,706979729
Glyma.18G099900.1	Lipoyl(octanoyl) transferase / Octanoyl-[acyl- carrier-protein]-protein N-octanoyltransferase	4,700351311
Glyma.06G290200.1	PROTEIN PHOSPHATASE 2C 20-RELATED	4,676618452
Glyma.11G010400.1	CELLULOSE SYNTHASE-LIKE PROTEIN D3	4,586237648
Glyma.01G244900.1	PRESEQUENCE PROTEASE 1, CHLOROPLASTIC/MITOCHONDRIAL- RELATED	4,530310763
Glyma.08G086600.1	METALLOPROTEASE M41 FTSH	4,515338446
Glyma.03G092600.1	Plant protein of unknown function (DUF825) (DUF825)	4,51255983
Glyma.08G300900.1	Mitovirus RNA-dependent RNA polymerase (Mitovir_RNA_pol)	4,424513345
Glyma.07G039600.1	GTP diphosphokinase / Stringent factor // Xylan 1,4-beta-xylosidase / Xylobiase // Non-reducing end alpha-L-arabinofuranosidase / Arabinosidase	4,419544347
Glyma.20G221000.1	UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N- ACETYLGLUCOSAMINYLTRANSFERASE 110 KDA SUBUNIT	4,322916453
Glyma.07G236000.1	Alkaline and neutral invertase (Glyco_hydro_100)	4,322374537
Glyma.19G100000.1	GRAM domain (GRAM)	4,318062951
Glyma.06G068400.1	F-box domain (F-box) // Leucine Rich repeat (LRR_6)	4,314635852
Glyma.15G115600.1	AUXIN-INDUCED IN ROOT CULTURES PROTEIN 12	4,306486966
Glyma.18G043700.1	GLUTATHIONE S-TRANSFERASE F11- RELATED	4,257389075
Glyma.01G201600.1	PHOTOSYSTEM I ASSEMBLY PROTEIN YCF3	4,213791851
Glyma.01G168900.1	PHD FINGER, SWIB/MDM2 AND GYF DOMAIN-CONTAINING PROTEIN	4,207139769
Glyma.U021800.1	MITOGEN-ACTIVATED PROTEIN KINASE 3	4,203514084
Glyma.14G155900.1	ZINC FINGER CLIPPER -RELATED	4,198715581
Glyma.20G243500.1	Transketolase / Glycoaldehyde transferase // Transaldolase / Glycerone transferase	4,197233777
Glyma.12G076700.1	CYTOKININ RIBOSIDE 5'-MONOPHOSPHATE PHOSPHORIBOHYDROLASE LOG1	4,180629872
Glyma.17G230500.1	Metallothionein (Metallothio_2)	4,172870668
Glyma.15G250800.1	(S)-2-hydroxy-acid oxidase / Hydroxy-acid oxidase B	4,17064811
Glyma.19G099200.1	PPR repeat (PPR) // PPR repeat family (PPR_2)	4,102108114
Glyma.02G249600.1	RIBULOSE BISPHOSPHATE CARBOXYLASE/OXYGENASE ACTIVASE, CHLOROPLASTIC	4,004884085
Glyma.07G015800.1	ABC TRANSPORTER G FAMILY MEMBER 29	3,988216291

Glyma.04G156200.1	TETRATRICOPEPTIDE REPEAT DOMAIN PROTEIN	3,827200216
-------------------	--	-------------

Tabela suplementar 3: Genes *Up*-regulados pela OE de *GmbZIPE2*. Genes anotados em ordem decrescente de *Log2FoldChange*.

Identificador	log₂FoldChange
Glyma.10G232100.1	8,592482
Glyma.12G194800.2	8,333089
Glyma.17G185800.1	7,902656
Glyma.02G267000.1	7,743696
Glyma.19G099100.1	7,618942
Glyma.11G221000.2	7,552292
Glyma.14G078300.1	7,304979
Glyma.05G092200.1	7,29091
Glyma.16G127900.2	7,278913
Glyma.11G068000.1	7,1352
Glyma.10G067600.1	7,059451
Glyma.17G180100.2	6,950945
Glyma.05G082400.2	6,948495
Glyma.15G091600.2	6,844274
Glyma.14G050500.1	6,791751
Glyma.07G140400.2	6,724474
Glyma.06G177600.1	6,719024
Glyma.08G302100.1	6,693136
Glyma.05G077300.3	6,64843
Glyma.13G361900.3	6,618639
Glyma.10G287400.1	6,586618
Glyma.19G130800.2	6,540231
Glyma.03G100500.1	6,531309
Glyma.08G138200.2	6,496133
Glyma.13G111500.3	6,495595
Glyma.15G163900.6	6,435022
Glyma.14G038100.1	6,427016
Glyma.18G036400.2	6,426407
Glyma.02G310000.6	6,359886
Glyma.18G155800.1	6,352236
Glyma.10G221500.3	6,310076
Glyma.08G041700.3	6,274085
Glyma.15G108700.2	6,144893
Glyma.16G213700.2	6,143674
Glyma.19G168300.1	6,14235
Glyma.09G011600.1	6,136368
Glyma.06G157400.2	6,117775
Glyma.06G165200.2	6,103784
Glyma.17G022300.2	6,084146
Glyma.06G242900.1	6,076489

Glyma.08G328300.1	6,071417
Glyma.20G036000.1	6,066072
Glyma.08G229300.1	6,061773
Glyma.14G158500.1	6,054958
Glyma.20G182700.1	6,036508
Glyma.08G124600.2	6,028475
Glyma.02G244800.1	6,018921
Glyma.02G041800.2	6,018659
Glyma.08G041700.2	6,018235
Glyma.10G249800.2	6,014995
Glyma.17G259300.2	6,014181
Glyma.07G168700.2	6,007494
Glyma.20G066100.3	6,005418
Glyma.11G166300.2	5,951878
Glyma.19G030000.1	5,949839
Glyma.10G292400.2	5,946375
Glyma.U013900.1	5,913496
Glyma.08G017100.1	5,902322
Glyma.19G148400.1	5,896107
Glyma.04G200400.2	5,880609
Glyma.03G128300.3	5,879802
Glyma.14G136900.2	5,876077
Glyma.03G232400.2	5,801512
Glyma.08G068600.1	5,800717
Glyma.16G073300.1	5,800642
Glyma.02G038100.2	5,794324
Glyma.07G121900.5	5,721668
Glyma.07G047800.2	5,719885
Glyma.15G228400.1	5,716318
Glyma.05G082200.3	5,703009
Glyma.05G084200.5	5,692333
Glyma.03G244100.1	5,668782
Glyma.16G106700.3	5,622259
Glyma.07G226600.1	5,618243
Glyma.13G011800.1	5,617666
Glyma.03G100400.1	5,586438
Glyma.13G300800.1	5,534258
Glyma.05G080300.4	5,533107
Glyma.12G155100.1	5,496805
Glyma.06G089800.3	5,43709
Glyma.09G046100.1	5,434178
Glyma.18G096500.3	5,428273
Glyma.11G048300.3	5,423025
Glyma.10G014400.2	5,402032
Glyma.01G014700.4	5,38748

Glyma.09G069800.2	5,344561
Glyma.04G069500.3	5,343286
Glyma.06G201900.1	5,321882
Glyma.05G039400.4	5,321774
Glyma.20G198000.2	5,318715
Glyma.01G040900.2	5,307949
Glyma.01G124600.1	5,307915
Glyma.14G098100.2	5,307087
Glyma.04G026600.3	5,305788
Glyma.16G030500.1	5,301736
Glyma.U034500.11	5,199251
Glyma.07G127800.3	5,199185
Glyma.16G095000.2	5,198228
Glyma.15G078300.2	5,195861
Glyma.19G260900.3	5,195736
Glyma.02G106100.2	5,193879
Glyma.06G118300.3	5,191488
Glyma.14G090000.2	5,188196
Glyma.03G005800.2	5,186167
Glyma.06G223500.2	5,176712
Glyma.07G037400.2	5,176585
Glyma.14G080800.3	5,176111
Glyma.06G218700.1	5,176012
Glyma.16G004800.2	5,159601
Glyma.17G039400.1	5,151549
Glyma.13G220000.2	5,146796
Glyma.17G187600.1	5,121811
Glyma.09G071600.4	5,088329
Glyma.01G143000.1	5,087246
Glyma.01G234300.2	5,063601
Glyma.05G178200.3	5,063161
Glyma.16G012300.2	5,063014
Glyma.15G232000.2	5,062781
Glyma.08G028200.2	5,061644
Glyma.19G052400.2	5,06136
Glyma.16G016300.2	5,058753
Glyma.19G101100.3	5,056102
Glyma.16G155700.3	5,05524
Glyma.06G259400.4	5,029089
Glyma.16G015400.3	5,003044
Glyma.16G178800.2	4,931714
Glyma.11G008700.1	4,930604
Glyma.03G070500.2	4,914923
Glyma.05G089000.2	4,913907
Glyma.01G038300.2	4,913814

Glyma.16G017400.3	4,913122
Glyma.14G000800.3	4,912673
Glyma.19G239700.2	4,911564
Glyma.20G204500.2	4,910511
Glyma.06G292500.3	4,910453
Glyma.06G265100.1	4,910213
Glyma.04G082900.2	4,909572
Glyma.08G302600.5	4,903612
Glyma.16G018500.2	4,900679
Glyma.14G223200.3	4,900384
Glyma.11G096700.2	4,899589
Glyma.19G141200.2	4,885399
Glyma.04G067600.2	4,883746
Glyma.17G112500.2	4,862072
Glyma.01G085200.2	4,859846
Glyma.02G184600.2	4,775047
Glyma.15G206100.2	4,749057
Glyma.17G188300.2	4,748465
Glyma.06G065900.1	4,74842
Glyma.04G020700.1	4,746881
Glyma.13G112100.2	4,746283
Glyma.13G002200.2	4,744122
Glyma.16G182500.2	4,741244
Glyma.10G143500.3	4,739566
Glyma.17G187600.2	4,738602
Glyma.20G125500.1	4,735443
Glyma.04G096900.2	4,709992
Glyma.14G219500.4	4,704453
Glyma.17G040600.1	4,629755
Glyma.20G242400.2	4,616691
Glyma.06G224900.1	4,50409
Glyma.09G278700.1	4,425604
Glyma.U017900.1	4,424753
Glyma.19G218900.1	4,421267
Glyma.18G195200.8	4,419888
Glyma.07G054300.1	4,375796
Glyma.07G207000.1	4,316739
Glyma.18G216800.2	4,242075
Glyma.07G063600.2	4,20827
Glyma.01G168900.1	4,20714
Glyma.07G065600.3	4,202653
Glyma.13G013100.1	2,950787
Glyma.01G163400.1	2,624593
Glyma.13G019500.1	1,405041
Glyma.13G012300.1	1,1925

Glyma.13G011400.1	1,013793
-------------------	----------

Tabela suplementar 4: Genes *Up*-regulados pela OE de *GmbZIPE2*. Genes identificados não anotados em ordem decrescente de *Log2FoldChange*.

Identificador	log₂FoldChange
Glyma.13G021600.1	-4,81414
Glyma.13G022700.1	-4,52061
Glyma.13G018800.1	-3,5216
Glyma.13G014400.1	-3,49229
Glyma.13G012200.1	-3,24804
Glyma.U005800.1	-3,04249
Glyma.13G021200.1	-2,53704
Glyma.13G015300.1	-2,15809
Glyma.13G018700.1	-2,14999
Glyma.18G068500.1	-1,99912
Glyma.13G018200.1	-1,9672
Glyma.13G011500.1	-1,87403
Glyma.05G084400.1	-1,65799
Glyma.13G012100.1	-1,6407
Glyma.13G017500.1	-1,63987
Glyma.13G022300.1	-1,60728
Glyma.13G021500.1	-1,46227
Glyma.13G021900.1	-1,44157
Glyma.U007200.1	-1,35321
Glyma.20G208700.1	-1,31895
Glyma.13G021100.1	-1,18515
Glyma.13G021700.1	-1,0701
Glyma.13G016600.1	-1,06452
Glyma.13G011700.1	-0,99798
Glyma.09G133900.1	-0,82673
Glyma.13G011300.1	-0,73468
Glyma.13G015400.1	-0,72943
Glyma.13G014100.1	-0,68802
Glyma.13G023700.1	-0,67844
Glyma.13G022200.1	-0,6556
Glyma.09G164000.1	-0,60493
Glyma.19G187300.1	-0,5843
Glyma.13G014500.1	-0,57062
Glyma.05G084600.1	-0,56777
Glyma.13G023300.1	-0,51033
Glyma.13G016700.1	-0,50065
Glyma.13G020100.1	-0,46349
Glyma.07G079200.1	-0,36434
Glyma.07G079100.1	-0,3095

Glyma.19G247400.1	-0,30684
Glyma.05G084500.1	-0,28432
Glyma.11G218300.1	-0,23311
Glyma.08G269900.1	0,171213
Glyma.13G024200.1	0,437838
Glyma.13G014600.1	0,448694
Glyma.13G015700.1	0,854469

Tabela suplementar 5: Genes Down-regulados pela OE de *GmbZIPE2*. Genes identificados em ordem decrescente de *Log2FoldChange*.