THIAGO JONAS NAKAYAMA

ESTRATÉGIAS DE PROSPECÇÃO GÊNICA E TRANSFORMAÇÃO EM PLANTAS PARA TOLERÂNCIA À HIPÓXIA

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

VIÇOSA
MINAS GERAIS – BRASIL
2016

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

T
N163e
2016

Nakayama, Thiago Jonas, 1984-
Estratégias de prospecção gênica e transformação em plantas para tolerância à hipóxia / Thiago Jonas Nakayama. – Viçosa, MG, 2016.
vii, 93f. : il. (algumas color.) ; 29 cm.

Inclui anexo.
Orientador: Aluízio Borém de Oliveira.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Plantas - Melhoramento genético. 2. Transformação genética vegetal. 3. G*lycine max*. 4. *Setaria viridis*. 5. Hipóxia. I. Universidade Federal de Viçosa. Departamento de Fitotecnia. Programa de Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22 ed. 632.233

THIAGO JONAS NAKAYAMA

ESTRATÉGIAS DE PROSPECÇÃO GÊNICA E TRANSFORMAÇÃO EM PLANTAS PARA TOLERÂNCIA À HIPÓXIA

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

APROVADA: 19 de julho de 2016.

André Luiz de Lima Passianotto

Fernando Augusto Henning

Rogério Fernandes de Souza

Luciano do Amarante

Aluízio Borém de Oliveira
(Orientador)

AGRADECIMENTOS

À Universidade Federal de Viçosa (UFV) e ao Programa de Pós-Graduação em Genética e Melhoramento.

À Embrapa Soja, Embrapa LABEX e ao USDA/ARS-PGEC.

Ao Dr. Aluízio Borém de Oliveira, Dr. Alexandre Lima Nepomuceno, Drª. Liliane Marcia Mertz Henning e Dr. Frank Harmon pelos ensinamentos e confiança em mim depositada.

Ao Dr. Rogério Fernandes de Souza, Dr. Luciano do Amarante, Dr. Fernando Augusto Henning, Dr. André Luiz de Lima Passianotto, Dr. José Renato Bouças Farias e Drª. Larissa Alexandra Cardoso Moraes pela disposição em avaliar este trabalho.

A todos do Laboratório de Biotecnologia Vegetal e da Ecofisiologia Vegetal da Embrapa Soja e do Harmon Lab (PGEC).

A todos que contribuíram para o desenvolvimento deste trabalho.

À FAPEMIG pela concessão da bolsa de doutorado e doutorado sanduíche.

Aos amigos.

À minha família, por tudo.

SUMÁRIO

RESUMO

NAKAYAMA, Thiago Jonas, D.Sc., Universidade Federal de Viçosa, julho de 2016. **Estratégias de prospecção gênica e transformação em plantas para tolerância à hipóxia**. Orientador: Aluízio Borém de Oliveira. Coorientadores: Alexandre Lima Nepomuceno e Liliane Marcia Mertz Henning.

A disponibilidade de gás oxigênio (O_2) é uma das principais forças que moldam a evolução dos organismos vivos e é essencial para a sobrevivência das plantas. A deficiência parcial (hipóxia) ou total (anóxia) de gás oxigênio é componente importante do estresse de alagamento. Para o aprofundamento do conhecimento genético, a transcriptômica possibilita identificar genes envolvidos nos mecanismos de respostas, e a transformação desses genes em plantas permite avaliar seus efeitos. Com o objetivo de caracterizar genes responsivos à deficiência de oxigênio e identificar genes candidatos para tolerância ao alagamento, analisamos dados de RNA-seq de raízes das cultivares de soja Embrapa 45 (tolerante ao encharcamento do solo) e BR 4 (sensível ao encharcamento) sob hipóxia. Das diferenças entre as duas cultivares, Embrapa 45 apresentou menor número de genes *up*-, maior quantidade de genes *down*-regulados, e maior expressão dos genes que codificam *phosphoglucomutase* (Glyma05g34790), *unknown protein related to N-terminal protein myristoylation* (Glyma06g03430), *protein suppressor of phyA-105* (Glyma06g37080) e *fibrillin* (Glyma10g32620). Dados de RNA-seq e qRT-PCR da hemoglobina não simbionte Glyma11g12980 indicaram divergência estrutural desse gene entre as cultivares. Em comum entre as duas cultivares, mudanças na abundância de transcritos envolvidos no metabolismo de aminoácidos e derivados sugerem perturbação destes em modificações do tRNA, acurácia/eficiência traducional e estresse do retículo endoplasmático (RE) sob hipóxia. Também observamos que grupos de genes (estáveis, *up*- e *down*-regulados) diferem quanto à frequência de elementos *cis* TATA box, ABREs (*ABA-responsive elements*) e CRT/DREs (*C-repeat/dehydration-responsive elements*), assim como em estrutura, composição e uso de códons sinônimos. Especulamos que *cis* elementos ABRE podem mediar expressão gênica independentemente de ABA em raízes de soja sob hipóxia. Por fim, a nosso conhecimento, somos os primeiros a demonstrar viabilidade de se transformar *Setaria viridis* (espécie modelo de

monocotiledôneas C4) mediada por *Agrobacterium* por meio da técnica *floral dip*, que não demanda cultura de tecidos nem regeneração de plantas transgênicas. Assim como em arabidopsis, espécie modelo de dicotiledôneas C3, *floral dip* em *S. viridis* pode acelerar estudos genéticos e de genômica funcional (inclusive ao estresse de alagamento) de gramíneas importantes para alimentação humana e animal, fibras e biocombustíveis, tais como cana-de-açúcar, *Miscanthus*, capim-elefante, *Brachiaria*, sorgo e milho.

ABSTRACT

NAKAYAMA, Thiago Jonas, D.Sc., Universidade Federal de Viçosa, July, 2016. **Strategies for gene prospecting and transformation in plants to hypoxia tolerance**. Adviser: Aluízio Borém de Oliveira. Co-advisers: Alexandre Lima Nepomuceno and Liliane Marcia Mertz Henning.

Oxygen gas (O_2) availability is one of primary forces shaping the evolution of living organisms and essential for land plants survival. Partial (hypoxic) or total (anoxic) oxygen gas deficiency is important component of flooding stress. For deepening the genetic knowledge, transcriptomics identify genes involved in stress response mechanisms, and transformation of these genes in plants allows to evaluate its effects. In order to characterize genes responsive to oxygen deficiency and identify candidate flooding tolerance genes we analyzed root RNA-seq data from flood-tolerant Embrapa 45 and flood-sensitive BR 4 soybean cultivars under hypoxic stress. From differences between the two cultivars, Embrapa 45 showed less up- and more down-regulated genes, and stronger induction of phosphoglucomutase (Glyma05g34790), unknown protein related to protein N-terminal myristoylation (Glyma06g03430), protein suppressor of Phya-105 (Glyma06g37080), and fibrillin (Glyma10g32620). RNA-seq and qRT-PCR data of non-symbiotic hemoglobin Glyma11g12980 indicated structural divergence of this gene between cultivars. In common between the two cultivars, transcriptional changes in amino acids and derivative metabolic process suggest its disturbance in tRNA modifications, translation accuracy/efficiency, and endoplasmic reticulum (ER) stress under hypoxia. In addition, genes groups (stable, up-, and down-regulated genes) differed in promoter TATA box, ABREs (ABA-responsive elements), and CRT/DREs (C-repeat/dehydration-responsive elements) frequency, as well as in structure, composition, and codon usage. We speculate that *cis*-acting ABREs can mediate gene expression independent of ABA in hypoxic soybean roots. Finally, to our knowledge, we first report *Agrobacterium*-mediated transformation method in *Setaria viridis* (C4 monocotyledonous model species) using floral dip, which does not require tissue culture nor transgenic plant regeneration. As in arabidopsis, C3 dicotyledonous model species, floral dip in *S. viridis* may accelerate genetic studies and functional genomics (including flooding stress) of important food–feed–fiber–fuel

grass crops such as sugarcane, *Miscanthus*, elephant grass, *Brachiaria*, *Sorghum* and maize.

Introdução Geral

O alagamento é relativamente comum em áreas cultivadas, sendo um dos estresses abióticos de grande importância para a agricultura mundial. Frequente em solos de várzea (solos hidromórficos) (Silva and Parfitt, 2008), também ocorre devido ao excesso de chuva, ao rápido derretimento da neve (Banach et al., 2009) e em regiões áridas quando irrigadas em excesso (Dennis et al., 2000). Além disso, sua frequência vem aumentando ao longo das últimas seis décadas associada às mudanças climáticas, agravando os desafios globais para a segurança alimentar da crescente população humana (Bailey-Serres et al., 2012a). Entre as quatro grandes culturas - soja, trigo, milho e arroz - apenas plantas de arroz são adaptadas ao encharcamento do solo, e todas são sensíveis à submersão total (Bailey-Serres et al., 2012b). A deficiência parcial (hipóxia) ou total (anóxia) do gás oxigênio é componente importante do estresse de alagamento, pois o oxigênio é necessário para produção de energia via respiração aeróbica (Bailey-Serres and Voesenek, 2008).

Orquestrados por uma complexa rede de regulação gênica, plantas e outros organismos precisam perceber, sinalizar e promover mudanças bioquímicas, fisiológicas e morfoanatômicas adequadas para sobreviverem e prosperarem aos estresses abióticos. Para o aprofundamento do conhecimento genético, a transcriptômica possibilita identificar genes envolvidos nos mecanismos de respostas, e a transformação desses genes em plantas permite avaliar seus efeitos sob estresses.

Ainda são poucas as análises de transcriptomas de plantas de soja sob alagamento (Komatsu et al., 2009; Nanjo et al., 2011; Tamang et al., 2014), e até o momento há apenas um trabalho que utilizou RNA-seq (Chen et al., 2016), em que avaliaram o transcriptoma de folhas de plantas de soja após sete dias de encharcamento do solo. Comparada a outras tecnologias, como as baseadas em hibridização (macro e microarranjo) ou em sequenciamento Sanger (SAGE, CAGE e MPSS), RNA-seq permite mensurações mais precisas e acuradas numa ampla faixa de expressão dos transcritos e suas variantes (Wang et al., 2009). Outra limitação dos estudos de transcriptômica para alagamento é o uso de apenas um genótipo. Não comparam diferenças nos

níveis de expressão entre genótipos tolerante e sensível ao alagamento, o que dificulta selecionar genes candidatos para tolerância.

A transgenia tem auxiliado obter plantas mais tolerantes a estresses abióticos (Kasuga et al., 1999; Chen et al., 2007; Cheng et al., 2009; Valente et al., 2009; Xie et al., 2009; Hao et al., 2011; Qin et al., 2012; Ying et al., 2012; Mickelbart et al., 2015), inclusive ao alagamento (Banti et al., 2010; Licausi et al., 2011). Porém, há limitações das técnicas de transformação e regeneração de plantas de soja (He et al., 2011), o que demanda muito mais tempo, esforço de trabalho e recurso financeiro. Primeiro desenvolvido para arabidopsis (*Arabidopsis thaliana*), o método de transformação de inflorescências mediada por *Agrobacterium* (*floral dip*) é simples e rápido, pois não demanda cultura de tecidos nem regeneração de plantas transgênicas (Clough and Bent, 1998; Zhang et al., 2006). *Floral dip* também é aplicável em *Medicago* (Trieu et al., 2000), rabanete (Curtis and Nam, 2001), tomate (Yasmeen et al., 2009), trigo (Zale et al., 2009), canola (Li et al., 2010), *Melilotus* (Hirsch et al., 2010), *Camelina* (Liu et al., 2012), milho (Mu et al., 2012) e arroz (Rod-in et al., 2014). Esse método possibilitou acelerar estudos genéticos e de genômica funcional em arabidopsis, espécie modelo de dicotiledôneas C3, estendendo nossa compreensão dos mecanismos conservados com espécies agrícolas dicotiledôneas. Já na espécie *Setaria viridis*, modelo de gramíneas C4, a nosso conhecimento *floral dip* ainda não foi testado. Assim como em arabidopsis, se o *floral dip* também for viável em *S. viridis* ele pode acelerar estudos genéticos e de genômica funcional (inclusive ao estresse de alagamento) de gramíneas importantes para alimentação humana e animal, fibras e biocombustíveis, tais como cana-de-açúcar, *Miscanthus*, capim-elefante, *Brachiaria*, sorgo e milho. Neste contexto, caracterizamos o transcriptoma de raízes das cultivares de soja Embrapa 45 (tolerante ao encharcamento do solo) e BR 4 (sensível ao encharcamento) sob hipóxia fazendo uso de RNA-seq. Também demonstramos a viabilidade de transformar *S. viridis* por meio da técnica *floral dip*.

Referências

Bailey-Serres J, Fukao T, Gibbs DJ, Holdsworth MJ, Lee SC, Licausi F,

Perata P, Voesenek LACJ, van Dongen JT (2012a) Making sense of low oxygen sensing. Trends Plant Sci **17**: 129–138

Bailey-Serres J, Lee SC, Brinton E (2012b) Waterproofing Crops: Effective Flooding Survival Strategies. Plant Physiol **160**: 1698–1709

Bailey-Serres J, Voesenek LACJ (2008) Flooding Stress: Acclimations and Genetic Diversity. Annu Rev Plant Biol **59**: 313–339

Banach K, Banach AM, Lamers LPM, De Kroon H, Bennicelli RP, Smits AJM, Visser EJW (2009) Differences in flooding tolerance between species from two wetland habitats with contrasting hydrology: Implications for vegetation development in future floodwater retention areas. Ann Bot **103**: 341–351

Banti V, Mafessoni F, Loreti E, Alpi A, Perata P (2010) The heat-inducible transcription factor HsfA2 enhances anoxia tolerance in Arabidopsis. Plant Physiol **152**: 1471–83

Chen M, Wang Q, Cheng X, Xu Z, Li L, Ye X, Xia L, Ma Y (2007) GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. Biochem Biophys Res Commun **353**: 299–305

Chen W, Yao Q, Patil GB, Agarwal G, Deshmukh RK, Lin L, Wang B, Wang Y, Prince SJ, Song L, et al (2016) Identification and Comparative Analysis of Differential Gene Expression in Soybean Leaf Tissue under Drought and Flooding Stress Revealed by RNA-seq. Front Plant Sci **7**: 1–19

Cheng L, Huan S, Sheng Y, Hua X, Shu Q, Song S, Jing X (2009) GMCHI, cloned from soybean [*Glycine max* (L.) Meer.], enhances survival in transgenic Arabidopsis under abiotic stress. Plant Cell Rep **28**: 145–153

Clough SJ, Bent AF (1998) Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735–743

Curtis IS, Nam HG (2001) Transgenic radish (*Raphanus sativus* L. *longipinnatus* Bailey) by floral-dip method - Plant development and surfactant are important in optimizing transformation efficiency. Transgenic Res **10**: 363–371

Dennis ES, Dolferus R, Ellis M, Rahman M, Wu Y, Hoeren FU, Grover A, Ismond KP, Good AG, Peacock WJ (2000) Molecular strategies for improving waterlogging tolerance in plants. J Exp Bot **51**: 89–97

Hao Y, Wei W, Song Q, Chen H-W, Zhang Y-Q, Wang F, Zou H, Lei G, Tian A, Zhang W-K, et al (2011) Soybean NAC transcription factors promote

abiotic stress tolerance and lateral root formation in transgenic plants. Plant J **68**: 302–313

He L, Zhao J, Zhao M, He C (2011) Current development and application of soybean genomics. **6**: 337–348

Hirsch AM, Lee A, Deng W, Tucker SC (2010) An open-flower mutant of *Melilotus Alba*: Potential for floral-dip transformation of a papilionoid legume with a short life cycle? Am J Bot **97**: 395–404

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol **17**: 287–291

Komatsu S, Yamamoto R, Nanjo Y, Mikami Y, Yunokawa H, Sakata K (2009) A Comprehensive Analysis of the Soybean Genes and Proteins Expressed under Flooding Stress using Transcriptome and Proteome Techniques. J Proteome Res **8**: 4766–4778

Li J, Tan X, Zhu F, Guo J (2010) A Rapid and Simple Method for *Brassica napus* Floral-Dip Transformation and Selection of Transgenic Plantlets. Int J Biol **2**: 127–131

Licausi F, Kosmacz M, Weits D a., Giuntoli B, Giorgi FM, Voesenek L a. CJ, Perata P, van Dongen JT (2011) Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. Nature **479**: 419–422

Liu X, Brost J, Hutcheon C, Guilfoil R, Wilson AK, Leung S, Shewmaker CK, Rooke S, Nguyen T, Kiser J, et al (2012) Transformation of the oilseed crop *Camelina sativa* by *Agrobacterium*-mediated floral dip and simple large-scale screening of transformants. Vitr Cell Dev Biol - Plant **48**: 462–468

Mickelbart M V, Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nat Rev Genet **16**: 237–251

Mu G, Chang N, Xiang K, Sheng Y, Zhang Z, Pan G (2012) Genetic Transformation of Maize Female Inflorescence Following Floral Dip Method Mediated by *Agrobacterium*. Biotechnology **11**: 178–183

Nanjo Y, Maruyama K, Yasue H, Yamaguchi-Shinozaki K, Shinozaki K, Komatsu S (2011) Transcriptional responses to flooding stress in roots including hypocotyl of soybean seedlings. Plant Mol Biol **77**: 129–144

Qin Y, Wang M, Tian Y, He W, Han L, Xia G (2012) Over-expression of

TaMYB33 encoding a novel wheat MYB transcription factor increases salt and drought tolerance in Arabidopsis. Mol Biol Rep **39**: 7183–92

Rod-in W, Sujipuli K, Ratanasut K (2014) The floral-dip method for rice (*Oryza sativa*) transformation. J Agric Technol **10**: 467–474

Silva CAS, Parfitt JMB (2004) Drenagem Superficial para Diversificação do Uso dos Solos de Várzea do Rio Grande do Sul. Pelotas: EMBRAPA-CPACT (Circular Técnica, 40) 1–10

Tamang BG, Magliozzi JO, Maroof MAS, Fukao T (2014) Physiological and transcriptomic characterization of submergence and reoxygenation responses in soybean seedlings. Plant, Cell Environ **37**: 2350–2365

Trieu a T, Burleigh SH, Kardailsky I V, Maldonado-Mendoza IE, Versaw WK, Blaylock L a, Shin H, Chiou TJ, Katagi H, Dewbre GR, et al (2000) Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. Plant J **22**: 531–41

Valente MAS, Faria JAQA, Soares-Ramos JRL, Reis PAB, Pinheiro GL, Piovesan ND, Morais AT, Menezes CC, Cano MAO, Fietto LG, et al (2009) The ER luminal binding protein (BiP) mediates an increase in drought tolerance in soybean and delays drought-induced leaf senescence in soybean and tobacco. J Exp Bot **60**: 533–546

Wang Z, Gerstein M, Snyder M (2009) RNA-seq: a revolutionary tool for transcriptomics.

Xie Z, Zou H, Lei G, Wei W, Zhou Q, Niu C, Liao Y, Tian A-G, Ma B, Zhang W-K, et al (2009) Soybean Trihelix Transcription Factors GmGT-2A and GmGT-2B Improve Plant Tolerance to Abiotic Stresses in Transgenic Arabidopsis. PLoS One **4**: e6898

Yasmeen A, Mirza B, Inayatullah S, Safdar N, Jamil M, Ali S, Choudhry MF (2009) In planta transformation of tomato. Plant Mol Biol Report **27**: 20–28

Ying S, Zhang D-F, Fu J, Shi Y-S, Song Y-C, Wang T-Y, Li Y (2012) Cloning and characterization of a maize bZIP transcription factor, ZmbZIP72, confers drought and salt tolerance in transgenic Arabidopsis. Planta **235**: 253–266

Zale JM, Agarwal S, Loar S, Steber CM (2009) Evidence for stable transformation of wheat by floral dip in *Agrobacterium tumefaciens*. Plant Cell Rep **28**: 903–913

Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H (2006) *Agrobacterium-*

mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nat Protoc **1**: 641–646

CAPÍTULO I (REVISÃO)

Avanços biotecnológicos para o desenvolvimento de plantas tolerantes ao encharcamento do solo

Introdução

O estresse por encharcamento é um dos mais importantes estresses abióticos, tendo efeito severo sobre o crescimento e rendimento de plantas cultivadas. A difusão dos gases na água é 10.000 vezes menor que na atmosfera, e plantas submetidas a essa condição sofrem com a deficiência de oxigênio (Armstrong, 1980). O estresse por deficiência de oxigênio em plantas normalmente ocorre quando essas são submetidas a condições de solos encharcados. Entretanto, determinados tecidos vegetais podem sofrer com a hipóxia mesmo em ambientes em que a concentração do oxigênio é próxima a da atmosfera, a exemplo de tecidos e orgãos em que a atividade metabólica é intensa e a entrada do oxigênio via difusão não é suficiente para atender a demanda. Essa situação pode ocorrer, por exemplo, em sementes em fase de desenvolvimento (Borisjuk et al., 2007).

O estresse por encharcamento altera vários aspectos do metabolismo vegetal, prejudicando o crescimento e o desenvolvimento das plantas. A falta do oxigênio reduz a produção de ATP, pois impede a fosforilação oxidativa na mitocôndria. Dessa forma, a produção de energia passa a ocorrer por meio da via fermentativa, a qual apresenta baixa eficiência quando comparada à fosforilação oxidativa (Bailey-Serres e Voesenek, 2008).

As alterações na concentração de oxigênio podem variar da normóxia (concentração normal), ou seja, situação em que a respiração aeróbica procede normalmente, hipóxia (deficiência de oxigênio), que limita a produção de ATPs oriundos da fosforilação oxidativa, ou anóxia (ausência total de oxigênio), quando não ocorre produção de ATPs via fosforilação oxidativa (Sousa e Sodek, 2002).

Na célula, uma resposta comum à deficiência de oxigênio em organismos eucariotos é o efeito Pasteur, em que a glicólise e a fermentação são promovidas e o ciclo do ácido tricarboxílico (TCA) e a respiração mitocondrial

são reprimidos. Enquanto que pela respiração aeróbica são gerados 36 ATPs por molécula de glicose, pela rota fermentativa são produzidos apenas 2 ATPs. Para compensar o déficit em energia, a glicólise é acelerada, levando ao consumo dos carboidratos de reserva (Parent et al., 2008).

Ajustes metabólicos em resposta ao encharcamento do solo envolvem mecanismos de detecção e de sinalização que resultam em complexa reconfiguração da expressão gênica. A identificação e caracterização de genes que conferem tolerância facilita e agiliza o desenvolvimento de culturas adaptadas ao encharcamento. O aprimoramento e a redução dos custos de uso das novas tecnologias de sequenciamento de DNA em larga escala têm potencializado análises de QTLs (*Quantitative Trait Loci*) e de transcriptomas de diferentes espécies para vários estresses ambientais, como por exemplo, a deficiência de oxigênio ocasionada pelo encharcamento.

Sinalização à deficiência de oxigênio e regulação transcricional em plantas

As respostas de aclimatação dos diferentes organismos à deficiência de oxigênio podem ser desencadeadas por múltiplos mecanismos de sinalização diretos ou indiretos (Bailey-Serres e Chang, 2005). Mecanismos de sinalização diretos envolvem proteínas ou substratos que diretamente reagem ou se ligam com o oxigênio, sendo essa forma de sinalização melhor compreendida em metazoários do que em plantas. Em metazoários, a restrição de oxigênio altera o acúmulo da subunidade α do fator de transcrição heterodimérico HIF 1α/β (*Hypoxia Inducible Factor*). O fator HIF 1α é constitutivamente sintetizado, porém, não é acumulado em condições de normóxia, pois na presença do oxigênio ocorre a hidroxilação de dois resíduos de prolina presentes na subunidade α desse fator de transcrição, o que desencadeia sua ubiquitinação e degradação mediada por proteassoma 26S. Essa reação de hidroxilação é catalisada pelas enzimas prolina hidrolases (PHDs), consumidoras de oxigênio. Com decréscimo na concentração de oxigênio, as enzimas PHDs tornam-se menos ativas e consequentemente o fator HIF 1α é acumulado e transportado para o núcleo onde HIF 1α/β pode funcionar na ativação transcricional de vários

genes (Bailey-Serres and Chang, 2005; Licausi e Perata, 2009; Bailey-Serres et al., 2012).

Em plantas, não foi detectada a presença de ortólogo ao fator HIF 1α/β, o que fez com que até recentemente, apenas mecanismos indiretos de sinalização à deficiência de oxigênio fossem relatados (Licausi, 2013). Entretanto, estudos recentes destacam o papel dos fatores de transcrição ERFs (*Etilene Responsive Factor*), pertencentes ao Grupo VII, no mecanismo direto de sinalização à hipóxia em plantas. O acúmulo ou degradação dessas proteínas depende da ausência ou presença de oxigênio, funcionando como sensor de deficiência de oxigênio em vegetais (Gibbs et al., 2011; Licausi et al., 2011; Licausi, 2013).

Mecanismos indiretos de percepção ao estresse envolvem alterações na homeostase celular, a exemplo no estado energético por meio de variações nos níveis de ATP, ADP e AMP, declínio no pH citosólico, mudanças nos níveis do cálcio citosólico (Ca^{2+}) e de metabólitos tais como a sacarose ou pela produção de espécies reativas de oxigênio (ROS) e óxido nítrico (NO) (Bailey-Serres e Chang, 2005; Bailey-Serres, 2012).

O fitohormônio etileno também é associado aos mecanismos de resposta ao encharcamento, promovendo alterações morfológicas como a elongação da parte aérea, presença de aerênquima radicular e raízes adventícias. Plantas submetidas a condições de encharcamento apresentam aumento na regulação do gene que codifica a enzima aminociclopropano carboxilato sintase (ACC sintase), envolvida na biossíntese do etileno (Olson et al., 1995; Dat et al., 2004).

Genes responsivos ao estresse por deficiência de oxigênio ou encharcamento e que normalmente são utilizados para validação da indução do estresse são associados ao metabolismo fermentativo, como o gene que codifica a enzima álcool desidrogenase (ADH) e piruvato descarboxilase (PDC), ao catabolismo da sacarose, com a indução da sacarose sintase (SUSY) em inibição à rota da invertase (INV) e os genes que atuam na rota da respiração alternativa, como o gene que codifica a enzima hemoglobina não simbiótica Classe I (nsHb) (Bailey-Serres e Voesenek, 2008). Em plantas de arabidopsis (*Arabidopsis thaliana*) super-expressando a enzima PDC, o incremento na produção de etanol se correlacionou positivamente com a tolerância à baixa

concentração de oxigênio (Ismond et al., 2003). Por meio de análise das sequências promotoras desses genes comumente induzidos por hipóxia em várias espécies de plantas foi possível identificar elementos de DNA supostamente responsáveis pela indução de genes em resposta ao estresse por hipóxia (Dolferus et al., 1994; Mohanty et al., 2005). Os fatores de transcrição pertencentes as famílias MYC, MYB, AP2/ERF e NAC foram relatados por serem capazes de regular a expressão de ADH em arabidopsis (Abe et al., 2003; Papdi et al., 2008; Christianson et al., 2009).

Em estudo comparativo, utilizando ferramentas transcriptômicas, Mustroph et al. (2010) analisaram alterações no acúmulo de mRNAs de 21 espécies distantes evolutivamente, incluindo animais, fungos, bactérias e espécies vegetais. Foram observadas respostas altamente conservadas entre os organismos pertencentes aos diferentes reinos, dentre esses, genes que codificam proteínas associadas à glicólise, fermentação, respiração alternativa, transporte de metabólitos, também envolvidas na metabolização/detoxificação de espécies reativas de oxigênio, proteínas chaperonas e biogênese dos ribossomos. Entretanto, genes relacionados a regulação transcricional foram pouco conservados entre espécies de diferentes reinos (Mustroph et al., 2010). Maior número de genes diferencialmente expressos em condições de hipóxia foi observado em vegetais comparativamente aos demais reinos, demonstrando a complexidade da rede de sinalização em resposta ao estresse por deficiência de oxigênio em plantas, prevalecendo a indução de fatores de transcrição pertencentes às famílias ERF, MYB, proteínas CCCH zinc finger, WRKY e bZIP (Mustroph et al., 2010).

Além disso, em arabidopsis, muitos genes altamente regulados em resposta à hipóxia e co-regulados com genes associados ao metabolismo anaeróbio são ainda pouco caracterizados, sendo identificados como *HYPOXIA-RESPONSIVE UNKNOW PROTEIN* (HUP) genes. A avaliação fenotípica de mutantes de arabidopsis superexpressando esses genes confirma o envolvimento dos mesmos na sobrevivência de plantas em condições de deficiência de oxigênio (Mustroph, et al., 2010).

A inibição de processos biossintéticos por meio da redução da expressão de genes associados à síntese de parede celular, síntese de cromatina e genes associados ao metabolismo secundário também é uma

estratégia utilizada em vegetais, resultando na conservação de energia e diminuição no consumo de CO_2 (Nanjo et al., 2011).

Fatores de trasncrição ERF (*Etilene responsive factors*) nos mecanismo de resposta ao encharcamento

Os fatores de transcrição da família ERF (*Etilene Responsive Factors*) são pertencentes à superfamília AP2/ERF, a qual pertencem também as famílias AP2 (Apetala 2) e RAV (*Related to* ABI3/VP1). A superfamília AP2/ERF é caracterizada pelo domínio AP2/ERF, que consiste em cerca de 60 a 70 aminoácidos implicados na ligação ao DNA. As proteínas da família ERF contêm um único domínio AP2/ERF, enquanto que as pertencentes à família AP2 contêm dois domínios AP2/ERF repetidos; já a família RAV contém um domínio AP2/ERF e um domínio B3. A família ERF pode ser ainda dividida em duas subfamílias, a subfamília a ERF e a CBF/DREB (Sakuma et al., 2002; Nakano et al., 2006).

O domínio ERF foi identificado pela primeira vez em tabaco (*Nicotiana tabacum*), sendo quatro proteínas de ligação denominadas ethylene-responsive binding proteins 1, 2, 3, e 4 ou “proteína de ligação ao elemento responsivo ao etileno” (EREBP1, 2, 3, e 4, atualmente renomeados ERF1, 2, 3, e 4), e foi comprovado que essas proteínas se ligam a uma região GCC box, reconhecendo o elemento *cis*-acting AGCCGCC, a qual é uma sequência de DNA envolvida na transcrição de genes responsivos ao etileno (Ohme-Takagi e Shinshi, 1995; Nakano et al., 2006). Posteriormente foram identificas várias proteínas da família ERF implicadas em diversas funções, tais como transdução de sinal hormonal, resposta a estresses bióticos e abióticos e regulação do metabolismo (Nakano et al., 2006).

Em arabidopsis e arroz foram identificados respectivamente 122 e 139 membros da familia ERF, classificados em 12 e 15 grupos respectivamente, sendo a maioria desses presentes em ambas as espécies. Dentre esses grupos destaca-se o Grupo VII, da família ERF, o qual é baseado na presença de um motivo conservado na região N terminal (MCGGAI(I/L) (Nakano et al., 2006). A indução hipóxica de alguns genes ERF Grupo VII apontam para o seu

envolvimento na sinalização em condições de baixa concentração de oxigênio nas plantas (Licausi et al., 2010).

As plantas apresentam diversidade genética quanto às respostas relacionadas aos mecanismos de sobrevivência ao estresse de encharcamento, as quais incluem alterações na arquitetura e crescimento de plantas e no metabolismo. Em arroz existem duas estratégias diferentes para sobrevivência ao estresse por submersão, sendo uma definida como escape, e outra, como quiescência. Ambas as estratégias são controladas por genes ERF Grupo VII, sendo o escape controlado pelo locus SNORKEL (escape) e a quiescência pelo locus SUBMERGENCE 1 (Bailey-Serres et al., 2012).

O mecanismo de escape consiste em promover o alongamento de entrenós da planta podendo obter sucesso caso as folhas ultrapassem o nível da água entrando em contato com o ar atmosférico (fonte de gás oxigêno e gás carbônico, essenciais para a respiração celular e para a fotossíntese, respectivamente) para que as plantas possam tolerar a submergência parcial durante meses. Estudos fisiológicos têm mostrado que os fitohormônios etileno, giberelina e ácido abscísico estão envolvidos nessa resposta. Recentemente foi elucidado o mecanismo molecular envolvido por meio da identificação do locus SNORKEL (SK) para os genes SNORKEL1 (SK1) e SNORKEL2 (SK2) (Figura 1). Sob condições de alagamento, ocorre acúmulo de etileno na planta e a indução da expressão destes dois genes SK1 e SK2, os quais promovem o alongamento dos entrenós via giberelina (Hattori et al., 2009).

Outra estratégia de sobrevivência em arroz é regulada pelo locus SUB1, um cluster de três genes, SUB1A, SUB1B e SUB1C. Dois desses genes (SUB1B e SUB1C) são invariáveis entre genótipos sensíveis e tolerantes ao alagamento, já o gene SUB1A pode apresentar dois alelos, o SUB1A-1 e SUB1A-2, sendo o primeiro presente em genótipos tolerantes e o segundo em genótipos sensíveis (Xu et al., 2006). O gene SUB1A mostrou ser altamente induzido em condições de alagamento e na presença de etileno. Esse gene controla respostas metabólicas e de crescimento associadas ao etileno e ao ácido giberélico. Na presença do alelo tolerante, as plantas em condições de estresse apresentaram restrição no crescimento foliar e na elongação dos entrenós, menor degradação de clorofila e menor consumo de carboidratos. Por outro lado, houve incremento nas atividades das enzimas PDC e ADH, e

aumento da expressão de genes envolvidos no controle das espécies reativas de oxigênio, melhorando assim o reestabelecimento de plantas após a submersão (Fukao et al., 2006; Fukao e Bailey-Serres, 2008). A observação de um fenótipo de floração tardia em linhagens superexpressando o gene SUB1A levou à descoberta de que a inibição da indução floral também é um dos componentes da estratégia de quiescência em arroz (Peña-Castro et al., 2011). Dessa forma, constatou-se que o alelo SUB1A-1 promove quiescência no crescimento das plantas submetidas ao estresse por submersão, favorecendo a manutenção dos carboidratos de reserva e a recuperação e reativação do crescimento quando as condições se tornam favoráveis. O gene SUB1A, apesar de ser induzido pelo etileno, também controla a produção desse hormônio em condições de submergência, pois embora o etileno tenha papel fundamental na tolerância ao estresse, seu acúmulo em quantidades excessivas pode comprometer o metabolismo vegetal (Fukao et al., 2006; Bailey-Serres et al., 2012). A introgressão do alelo SUB1A-1 em germoplasma asiático de arroz, por meio de seleção assistida por marcadores moleculares, permitiu o desenvolvimento de cultivares de arroz tolerantes à submersão, garantindo estabilidade de produção (Xu et al., 2006). O SUB1A-1 promove o acúmulo de dois repressores da sinalização ao AG (ácido giberélico), SLENDER RICE 1 (SLR1) e SLENDER RICE-LIKE 1 (SLRL1) e diminui a expressão de genes responsivos a esse hormônio (Fukao e Bailey-Serres, 2008) (Figura 1).

Ambos os genes Sub1A e SK1-2 foram identificados via análise de QTLs combinada com clonagem posicional, correspondendo a 69 e 30% da variação fenotípica respectivamente, e caracterizados molecular e funcionalmente por meio de técnicas de biologia molecular, fisiologia e transgenia (Jackson e Ram, 2003; Xu et al., 2006; Hattori at al., 2009).

Em arabidopsis os genes ERF Grupo VII também estão associados ao mecanismo de sobrevivência à deficiência de oxigênio. Foram identificados cinco membros ERF Grupo VII no genoma de arabidopsis, nomeadamente At1g53910 (RAP2.12), At1g72360, At2g47520, At3g14230 (RAP2.2) e At3g16770 (também conhecido como RAP2.3 AtEREBP). Estes ERFs são caracterizados por um domínio de ligação conservado na extremidade N-terminal (MCGGAV/IISD) e por resíduos de leucina/isoleucina e triptofano na região C terminal (Licausi et al., 2010).

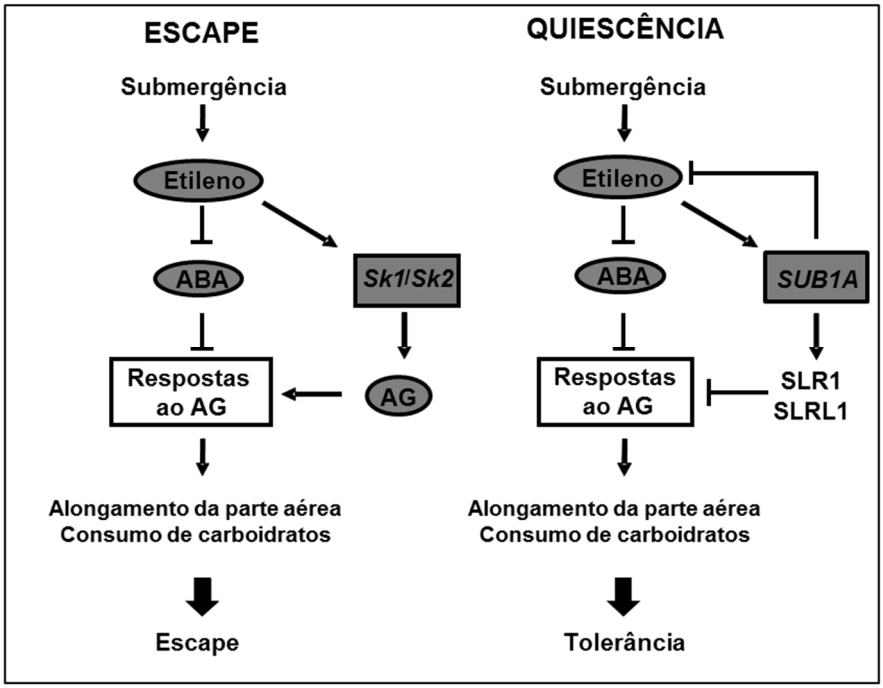


Figura 1 – Genes ERF Grupo VII e as rotas que regulam o crescimento de plantas de arroz submetidas ao estresse por submergência. Ambos os genes, SK1/SK2 e SUB1A são induzidos por etileno. O locus SK1/SK2 controla a estratégia de escape, por meio de acúmulo de ácido giberélico bioativo (AG) promovendo o alongamento de entrenós da planta. O locus SUB1A controla a estratégia de quiescência promovendo a regulação negativa de genes que controlam as respostas ao ácido giberélico SLR1 e SLRL1, resultando na paralização do crescimento e do consumo de energia durante a submergência e manutenção da viabilidade celular. Adaptado de Fukao e Bailey-Serres (2008) e Bailey-Serres et al. (2012).

Em estudo conduzido por Licausi et al. (2010), plantas de arabidopsis submetidas ao estresse por hipóxia apresentaram alteração na regulação dos genes At1g72360 e Atg47520, os quais foram denominados respectivamente HRE1 e HRE2 (*Hypoxia Responsive Factors*). Plantas transgênicas de arabidopsis superexpressando o gene HRE1 apresentaram aumento da tolerância à hipóxia, associado ao incremento nas atividades das enzimas ADH e PDC.

Da mesma forma, obeservou-se que outro gene pertencente a esse grupo, o gene RAP2.2, afeta parte das respostas das plantas à baixa concentração de oxigênio levando a expressão de genes envolvidos no metabolismo dos açúcares e enzimas da rota fermentativa, bem como na biossíntese de etileno. Esse gene é constitutivamente transcrito em altos níveis nas raízes e em menores quantidades na parte aérea, porém, em plantas submetidas a hipóxia no escuro, passa a ser altamente induzido em ambos os

tecidos. A superexpressão do gene RAP2.2 também promoveu aumento na tolerância à hipóxia (Hinz, et al., 2010), da mesma forma que seu homólogo RAP2.12 (Licausi et al., 2011).

Até pouco tempo, mecanismos diretos de sinalização à deficiência de oxigênio em plantas ainda não haviam sido descritos. Entretanto, estudos recentes desenvolvidos independentemente por dois grupos de pesquisa, revelaram os mecanismos pelos quais o oxigênio é percebido em células vegetais por meio de sinal de degradação via extremidade N-terminal de proteínas ERFs do Grupo VII (Gibbs et al., 2011; Licausi et al., 2011). Ambos os grupos de pesquisa identificaram um motivo conservado na região N-terminal (MCGGAI/L) dos transcritos dos genes ERFs do Grupo VII de arabidopsis, como um substrato ideal para degração via resíduos da extremidade N-terminal, os quais são reconhecidos por uma ubiquitina-ligase, mecanismo presente em eucariotos e procariotos (Licausi, 2013).

Um teste *in vitro* indicou que todos os transcritos dos cinco genes *ERF* do Grupo VII de arabidopsis são substratos para essa via de degradação. As evidências sugerem que a estabilidade do ERFs Grupo VII sob hipóxia é provavelmente relacionada com a inibição da oxidação do resíduo de cisteína N-terminal (NH_2–Met_1–Cys_2), o que é necessário para que a proteína possa ser arginilada e degradada (Gibbs et al., 2011; Licausi et al., 2011) (Figura 2). O gene SUB1A, principal responsável pela tolerância de plantas de arroz à submergência, embora contenha o motivo conservado na região N-terminal, não é substrato para essa via de degradação. A maior estabilidade dos transcritos desse gene pode estar relacionada ao desempenho superior desse genótipo mediante condições de estresse (Gibbs et al., 2011).

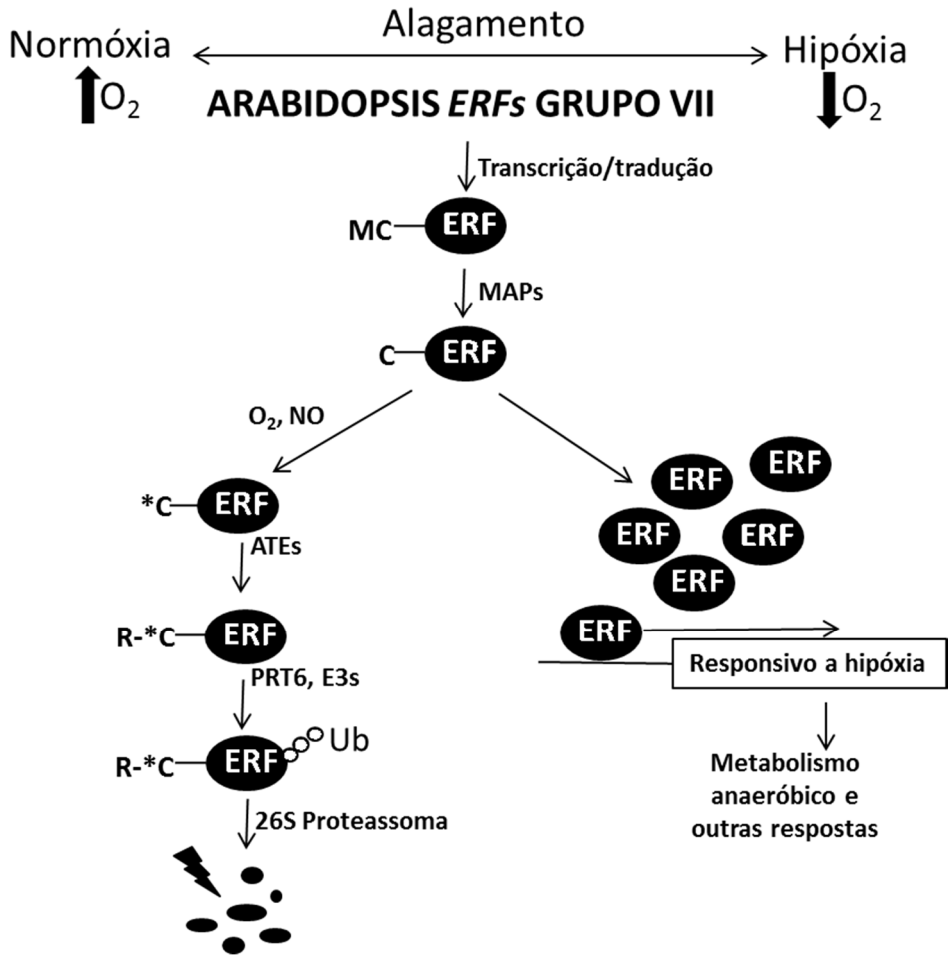


Figura 2 – Mecanismo de degradação via extremidade N-terminal de ERFs Grupo VII de arabidopsis em condição de estresse por hipóxia. Em condições de normóxia, após a transcrição, as proteínas ERFs são degradadas via extremidade N-terminal, envolvendo as seguintes etapas: i) a metionina (Met) N-terminal é constitutivamente clivada por uma aminopeptidase (MAP). ii) A cisteína (C) exposta é convertida a uma forma oxidada (C*) por O_2 ou NO; iii) um resíduo de arginina (R) é adicionado a C* por uma arginil RNAt transferase (ATE); iv) a proteína argininilada é reconhecida pela Proteólise 6 (PRT6) ou outras ligases E3 as quais promovem a poliubiquitinação da proteína tornando-a alvo para degradação (proteassoma 26S). O resultado desse processo é a inibição da transcrição de genes responsivos à hipóxia em condições de normóxia. Em condições de hipóxia, quando o oxigênio se torna limitante, a via de degradação pela extremidade N-terminal é inibida porque na ausência de oxigênio não ocorre a oxidação da cisteína. Os ERFs estabilizados podem ativar a transcrição de genes que incrementam o metabolismo anaeróbio e outras estratégias de sobrevivência em condição de hipóxia. Adaptado de Bailey-Serres et al. (2012).

Em resumo, esses estudos demonstram que o mecanismo de degradação via extremidade N-terminal regula o acúmulo de proteínas ERF do Grupo VII e consequentemente, o acúmulo de transcritos associados com as respostas à baixa concentração de oxigênio em arabidopsis. Propõe-se que as proteínas ERFs pertencentes ao Grupo VII são constitutivamente sintetizadas, porém, em situação de normóxia são degradadas via extremidade N-terminal ou sequestradas por meio de interação com a membrana plasmática por meio da

associação com proteínas de ligação Acyl-CoA (ACBP1 ou ACBP2) (Figura 3). Quando os níveis de oxigênio se tornam limitantes, a degradação dessas proteínas é inibida, sendo as mesmas transportadas para o núcleo onde ativam a transcrição de outros genes responsivos à hipóxia (Gibbs et al., 2011; Licausi et al., 2011; Bailey-Serres et al., 2012).

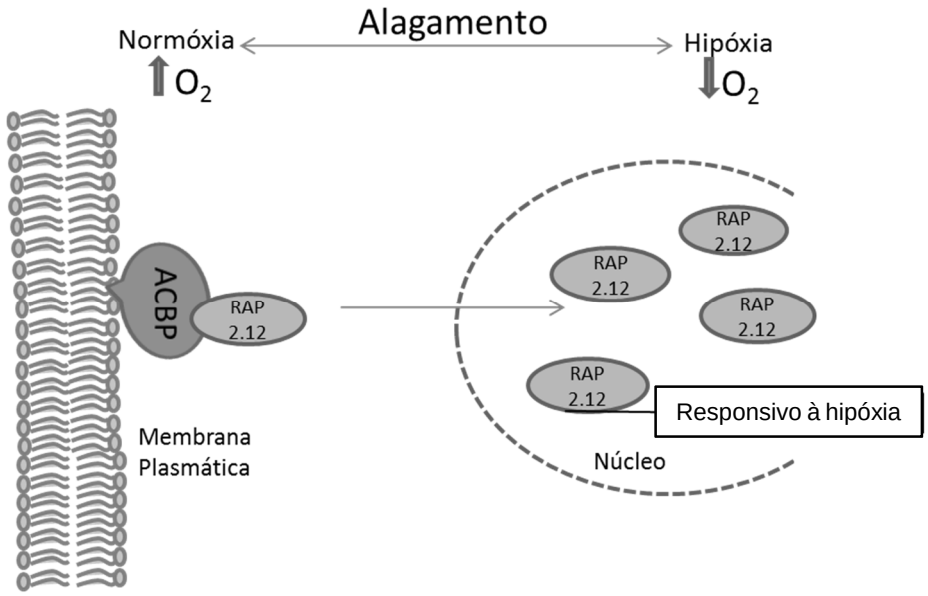


Figura 3 – Em condições de normóxia, transcritos do gene RAP2.12 permanecem associados com a membrana plasmática via interação com ACBP. Durante a hipóxia o RAP2.12 é realocado para o núcleo onde ativa a transcrição de genes responsivos à hipóxia. Adaptado de Bailey-Serres et al. (2012).

Respostas moleculares à deficiência de oxigênio em soja e milho

Até o momento apenas QTLs de pequeno efeito para tolerância ao encharcamento foram preditos em soja (VanToai et al., 2001; Nguyen et al., 2012). A não identificação de QTLs de efeito maior pode ser devido ao desconhecimento de genótipos de soja altamente tolerantes para utilização em estudos de mapeamento. Para o milho foram preditos QTLs para amarelecimento foliar (QTL no cromossomo 4 responsável por 29 a 49% da variação fenotípica) (Mano e Omori, 2013) e formação de aerênquima constitutivo em raízes (um QTL no cromossomo 1 e um no cromossomo 7 juntos responsáveis por 32% da variação fenotípica) (Mano et al., 2012), via análises de cruzamentos entre plantas de milho e o parente selvagem *Zea nicaraguensis*

e em uma população proveniente de um indivíduo de *Z. nicaraguensis*, respectivamente. A espécie *Z. nicaraguensis* é encontrada em áreas alagáveis, possui sistema radicular aerenquimatoso e desenvolve raízes adventícias na superfície de solos encharcados (Mano et al., 2012). A identificação e caracterização de genes referentes aos QTLs acima citados contribuirá para o melhoramento de soja e de milho e para o avanço do conhecimento dos mecanismo moleculares envolvidos na tolerância ao encharcamento.

Em soja, além dos genes comumente associados às respostas em função da deficiência de oxigênio em plantas, como os da rota fermentativa, do metabolismo de açúcares e da respiração alternativa, também foi identificada a presença de genes *ERFs* do Grupo VII, e a alteração da regulação desses em resposta ao estresse. Em trabalho conduzido por Tamang et al. (2014), foram identificados nove genes pertencentes ao Grupo VII da família *ERF*, designados *ERFVII1-9*, todos contendo o motivo conservado na região N-terminal [NH2-MCGGAII(A/S)D] (Figura 4). Por meio de análise filogenética foi estimada a relação evolutiva entre as proteínas desse grupo em arroz, soja e arabidopsis. Aparentemente os genomas da soja e arabidopsis não contêm ortólogos dos genes *SUB1* ou *SK* de arroz. Entretanto, os resultados indicaram uma alta conservação entre as proteínas ERF de arabidopsis e soja. Foi observado aumento dos transcritos desses genes em plantas de soja submetidas ao estresse por submergência.

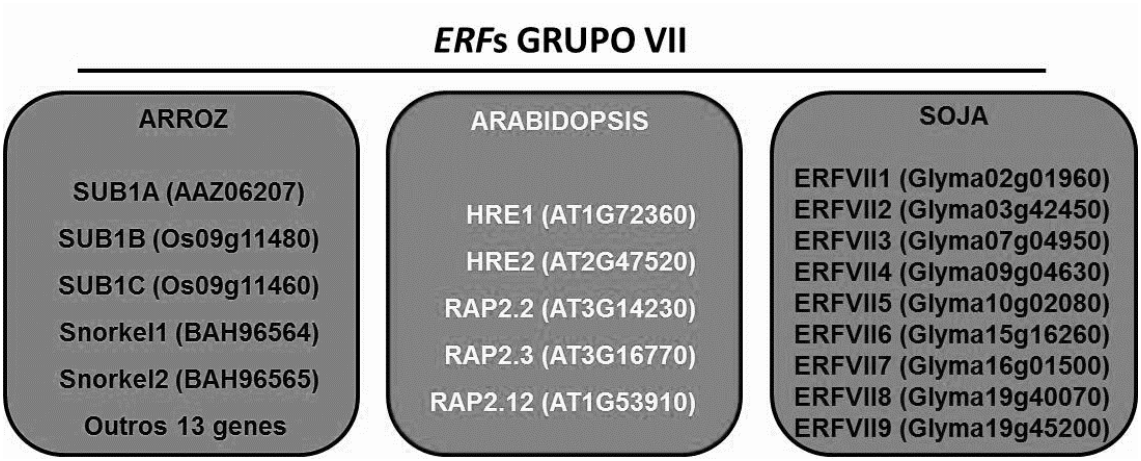


Figura 4 –Genes ERF pertencentes ao Grupo VII em diferentes espécies vegetais. Adaptado de Tamang et al. (2014).

Nesse mesmo estudo foi analisado o transcriptoma em plantas de soja submetidas ao estresse por submergência, por períodos curto e prolongado de submerção, e também, após a drenagem durante reoxigenação. A comparação entre o tratamento controle e as situações de estresse indicaram mais de 5.000 transcritos com regulação diferencial, demonstrando a complexa rede de sinalização e reconfiguração da expressão gênica em decorrência do estresse. Genes comumente induzidos em resposta ao estresse foram os que codificam para PDC (Piruvato Descarboxilase), aldeído desidrogenase (ALDH), aspartato aminotransferase (AST), inibidores da invertase, hemoglobinas não simbiontes, ACC oxidase e um grande número de genes associados com a respiração anaeróbia e estresse responsivos, muitos dos quais já observados como sendo altamente induzidos em plantas de arabidopsis submetidas à submergência, além de fatores de transcrição, fosfatases e quinases (Mustroph et al., 2009; Lee et al., 2011, Tamang et al., 2014). Essas alterações também foram observadas em trabalhos de proteômica em soja, principalmente na fase de germinação, indicando a inibição de proteínas associadas ao crescimento e o aumento de proteínas associadas à fermentação (Nanjo, et al., 2013; Komatsu et al., 2013).

Outra via que tem sido estudada em soja entre os mecanismos de tolerância ao encharcamento é a via da respiração alternativa. Muitos pesquisadores têm reportado que a aplicação exógena de nitrato em soja melhora o desempenho da cultura em condições de deficiência de oxigênio (Oliveira et al., 2013). Estudos sugerem que na medida em que o nitrato é reduzido pela enzima nitrato redutase (NR), há incremento na produção de nitrito, o qual é reduzido a óxido nítrico. A redução do nitrato em nitrito sob condições de hipóxia é associada ao mecanismo de respiração alternativa, onde o nitrito atua como aceptor de elétrons alternativo na cadeia respiratória, em um processo acoplado à regeneração do NAD^+ e à síntese de ATP (Igamberdiev e Hill, 2004). O nitrito ao ser reduzido, gera óxido nítrico, uma molécula sinalizadora presente em diferentes tecidos nas plantas e que participa de uma ampla variedade de processos fisiológicos, bem como na ativação dos mecanismos de defesa contra estresses bióticos e abióticos, incluindo o estresse por deficiência de oxigênio (Stohr e Stremlau, 2006; Borisjuk e Rolletschek, 2008). A enzima nsHb Classe 1 atua nessa rota, modulando a quantidade de

óxido nítrico produzido e convertendo o mesmo em nitrato. Existem várias classes de hemoglobinas em plantas, porém, com propriedades distintas. As hemoglobinas simbiontes são encontradas em nódulos de raízes enquanto as hemoglobinas não simbiontes podem ser expressas em vários tecidos vegetais, presentes tanto em monocotiledôneas como dicotiledôneas. A categoria das nsHb possui duas classes, sendo a Classe 1 induzida em condições de baixa concentração de oxigênio (Igamberdiev et al., 2005; Sairam et al., 2009).

Em milho e em outras espécies, houve aumento na expressão de genes envolvidos no metabolismo anaeróbio, glicólise e outros processos relacionados, como os genes codificadores para álcool desidrogenase, aldolase, enolase, isomerase da glucose-fosfato, gliceraldeído-3-fosfato-desidrogenase, piruvato-descarboxilase e sacarose sintase (Sachs et al., 1996). Alguns estudos sugerem que os níveis de Ca^{2+} citosólico, bem como um rápido estabelecimento da homeostase iônica podem ser essenciais na sinalização para que essas mudanças adaptativas em resposta à hipóxia ocorram (Subbaiah e Sachs, 2003). A participação de miRNAs responsivos ao encharcamento, na regulação do metabolismo e nas adaptações morfológicas e fisiológicas em raízes de milho em nível pós-transcricional, foi destacada em estudo conduzido por Zhang et al. (2008).

No Brasil, a Embrapa - Centro Nacional de Pesquisa de Milho e Sorgo (CNPMS) desenvolveu a cultivar de milho 'Saracura BRS-4154', uma cultivar com uma ampla base genética que, em comparação com a cultivar sensível ao encharcamento (BR 107), apresentou maior sobrevivência ao estresse e retardo no aparecimento de danos no mesocótilo cuja causa é a degradação da parede celular (Vitorino et al., 2001). Em estudo avaliando os transcritos de genes que codificam para enzimas do sistema antioxidante, ascorbato peroxidase (APX), catalase (CAT) e superóxido dismutase (SOD), observou-se uma diminuição no acúmulo desses transcritos em resposta ao estresse por encharcamento (Campos, 2010), sugerindo aumento na formação de espécies reativas de oxigênio e ativação das alterações decorrentes do mecanismo de sobrevivência ao estresse.

Avaliando a capacidade combinatória de genótipos de milho para tolerância ao encharcamento do solo, Silva et al. (2006) demonstraram a existência de variabilidade genética, além de um efeito materno pronunciado

para o caráter, sugerindo um possível papel para genomas organelares na resposta anaeróbica.

Considerações Finais

Até o momento, as informações geradas dos genes identificados, caracterizados e envolvidos nas respostas ao encharcamento e a outras características complexas estão longe de serem suficientes para o entendimento do *pool* gênico disponível e de como é a conexão genótipo-fenótipo. Para diminuir essa distância, avanços estão ocorrendo no desenvolvimento de sequenciadores de alto desempenho, métodos estatísticos e técnicas de transformação. Um maior aprofundamento do conhecimento genético auxiliará na formulação de estratégias mais eficientes no melhoramento genético, seja utilizando técnicas de engenharia genética, ou seleção via marcadores para otimização da predição fenotípica e consequente obtenção de genótipos com maior tolerância.

REFERÊNCIAS

ARMSTRONG, W. Aeration in higher plants. Advances in Botanical Research, v.7, p.225-332, 1980.

ABE, H.; URAO, T.; ITO, T.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. The Plant Cell, v.15, p.63–78, 2003.

BAILEY-SERRES, J.; CHANG, R. Sensing and signaling in response to oxygen deprivation in plants and other organisms. Annals of Botany, v. 96, p.507–518, 2005.

BAILEY-SERRES, J.; FUKAO, T.; GIBBS, D.J.; HOLDSWORTH, M.J.; LEE, S.C.; LICAUSI, F.; PERATA, P.; VOESENEK, L.A.; VAN DONGEN, J.T. Making sense of low oxygen sensing. Trends in Plant Science, v.17, p.129-138, 2012.

BAILEY-SERRES, J.; VOESENEK, L.A. Flooding stress: acclimations and genetic diversity. Annual Review of Plant Biology, v.59, p.313–339, 2008.

BORISJUK, L.; MACHEREL, D.; BENAMAR, A.; WOBUS, U.; ROLLETSCHEK, H. Low oxygen sensing and balancing in plant seeds: a role for nitric oxide. New Phytologist, v.176, p.813–823, 2007.

BORISJUK, L.; ROLLETSCHEK, H. Nitric oxide is a versatile sensor of low oxygen stress in plants. Plant Signaling and Behavior, v.3, p.391-393, 2008.

CAMPOS, N.A. Expressão gênica associada a formação de aerênquimas em raízes de milho Saracura BRS 4154 sob hipóxia. 2010. 54p. Dissertação (Mestrado) - Programa de Pós-Graduação em Biotecnologia Vegetal, Universidade Federal de Lavras, Lavras, MG.

CHRISTIANSON, J.A.; WILSON, I.W.; LLEWELLYN, D.J.; DENNIS, E.S. The low-oxygen-induced NAC domain transcription factor ANAC102 affects viability of Arabidopsis seeds following low-oxygen treatment. Plant Physiology, v.149, p.1724–1738, 2009.

DAT, J.F.; CAPELLI, N.; FOLZER, H.; BOURGEADE, P.; BADOT, P-M. Sensing and signalling during plant flooding. Plant Physiology and Biochemistry, v. 42, p. 273–282, 2004.

DOLFERUS, R.; JACOBS, M.; PEACOCK, W.J.; DENNIS, E. S. Differential interactions of promoter elements in stress responses of the Arabidopsis adh gene. Plant Physiology, n.105, p.1075-1087, 1994.

FUKAO, T.; BAILEY-SERRES, J. Submergence tolerance conferred by Sub1A is mediated by SLR1 and SLRL1 restriction of gibberellin responses in rice. Proceedings of the National Academy of Sciences of the United States of America, v.105, p.16814–16819, 2008.

FUKAO, T.; XU, K.; RONALD, P.C.; BAILEY-SERRES J. A variable cluster of ethylene response factor-like genes regulates metabolic and developmental acclimation responses to submergence in rice. Plant Cell, v.18, p.2021–34, 2006.

GIBBS, D.J.; LEE, S.C.; ISA, N.M.; GRAMUGLIA, S.; FUKAO,T.; BASSEL, G.W.; CORREIA, C.S.; CORBINEAU, F.; THEODOULOU, F.L.; BAILEY-SERRES, J.; HOLDSWORTH, M.J. Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. Nature, v.479, p.415–418, 2011.

HATTORI, Y.; NAGAI, K.; FURUKAWA, S.; SONG, X-J.; KAWANO, R.; SAKAKIBARA, H.; WU, J.; MATSUMOTO, T.; YOSHIMURA, A.; KITANO, H.;

MATSUOKA, M.; MORI, H.; ASHIKARI, M. The ethylene response factors snorkel1 and snorkel2 allow rice to adapt to deep water. Nature, v.460, p.1026–1030, 2009.

HINZ, M.; WILSON, I.W.; YANG, J.; BUERSTENBINDER, K.; LLEWELLYN, D.; DENNIS, E.S.; SAUTER, M.; DOLFERUS, R. Arabidopsis RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. Plant Physiology, v.153, p.757–772, 2010.

IGAMBERDIEV, A.U.; HILL, R.D. Nitrate, NO and haemoglobin in plant adaptation to hypoxia: an alternative to classic fermentation pathways. Journal of Experimental Botany, v.55, p.2473-2482, 2004.

IGAMBERDIEV, A.U.; BARON, K.; MANAC'H-LITTLE, N.; STOIMENOVA, M.; HILL, R.D. The haemoglobin/nitric oxide cycle: involvement in flooding stress and effects on hormone signalling. Annals of Botany, v. 96, p.557–564, 2005.

ISMOND, K.P.; DOLFERUS, R.; DE PAUW, M.; DENNIS, E.S.; GOOD, A.G. Enhanced low oxygen survival in Arabidopsis through increased metabolic flux in the fermentative pathway. Plant Physiology, v.132, p.1292–1302, 2003.

JACKSON, M.B.; RAM, P. Physiological and molecular basis of susceptibility and tolerance of rice plants to complete submergence. Annals of Botany, v. 91, p. 227–241, 2003.

KOMATSU, S.; NANJO, Y.; NISHIMURA, M. Proteomic analysis of the flooding tolerance mechanism in mutant soybean. Journal of Proteomics, v. 79, p.231-250, 2013.

LEE, S.C.; MUSTROPH, A.; SASIDHARAN, R.; VASHISHT, D.; PEDERSEN, O.; OOSUMI, T.; VOESENEK, L. A.C.J.; BAILEY-SERRES, J. Molecular characterization of the submergence response of the *Arabidopsis thaliana* ecotype Columbia. The New Phytologist, v.190, p.457–471, 2011.

LICAUSI, F.; PERATA, P. Low oxygen signaling and tolerance in plants. Advances in Botanical Research, v.50, p.139-198, 2009.

LICAUSI, F.; VAN DONGEN, J.T.; GIUNTOLI, B.; NOVI, G.; SANTANIELLO, A.; GEIGENBERGER, P.; PERATA, P. HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. The Plant Journal, v.62, p.302–315, 2010.

LICAUSI, F.; KOSMACZ, M.; WEITS, D.A.; GIUNTOLI, B.; GIORGI, F.M.; VOESENEK, L.A.C.J.; PERATA, P.; VAN DONGEN, J.T. Oxygen sensing in

plants is mediated by an N end rule pathway for protein destabilization. Nature, v.479, p.419–422, 2011.

LICAUSI, F. Molecular elements of low-oxygen signaling in plants. Physiologia Plantarum, v.148, p.1-8, 2013.

MANO, Y.; OMORI, F. Flooding tolerance in interspecific introgression lines containing chromosome segments from teosinte (*Zea nicaraguensis*) in maize (Zea mays subsp. mays). Annals of botany, v. 112, p. 1125–39, 2013.

MANO, Y.; OMORI, F.; TAKEDA, K. Construction of intraspecific linkage maps, detection of a chromosome inversion, and mapping of QTL for constitutive root aerenchyma formation in the teosinte *Zea nicaraguensis*. Molecular Breeding, v. 29, p. 137–146, 2012.

MOHANTY, B.; KRISHNAN, S.P.; SWARUP, S.; BAJIC, V.B. Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. Annals of Botany, v.96, p.669-681, 2005.

MUSTROPH, A.; LEE, S.C.; OOSUMI, T.; ZANETTI, M.E.; YANG, H.; MA, K.; YAGHOUBI-MASIHI, A.; FUKAO, T.; BAILEY-SERRES, J. Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. Plant Physiology, 152:1484–1500, 2010.

MUSTROPH, A.; ZANETTI, M.E.; JANG, C.J.; HOLTAN, H.E.; REPETTI, P.P.; GALBRAITH, D.W.; GIRKE, T.; BAILEY-SERRES, J. Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in arabidopsis. Proceedings of the National Academy of Sciences of the United States of America v.106, p.18843–18848, 2009.

NAKANO, T.; SUZUKI, K.; FUJIMURA, T.; SHINSHI, H. Genome-wide analysis of the ERF gene family in arabidopsis and rice. Plant Physiology, v.140, p.411–432, 2006.

NANJO, Y.; MARUYAMA, K.; YASUE, H.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K.; KOMATSU, S. Transcriptional responses to flooding stress in roots including hypocotyl of soybean seedlings. Plant Molecular Biology, v.77, p.129–144, 2011.

NANJO, Y.; NAKAMURA, T.; KOMATSU, S. Identification of indicator proteins associated with flooding injury in soybean seedlings using label-free quantitative proteomics. Journal of Proteome Research, v. 12, p.4785−4798, 2013.

NGUYEN, V.T.; VUONG, T.D.; VANTOAI, T.; LEE, J.D.; WU, X.; ROUF MIAN, M.A.; DORRANCE, A.E.; SHANNON, J.G.; NGUYEN, H.T. Mapping of quantitative trait loci associated with resistance to *Phytophthora sojae* and flooding tolerance in soybean. Crop Science, v. 52, p. 2481, 2012.

OHME-TAKAGI, M.; SHINSHI, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. The Plant Cell, v.7, p.173-182, 1995.

OLIVEIRA, H.C.; SALGADO, I.; SODEK, L. Involvement of nitrite in the nitrite-mediated modulation of fermentative metabolismo and nitric oxide production of soybean roots during hypoxia. Planta, v. 237, p. 255-264, 2013.

OLSON, D.C.; OETIKER, J. H.; YANG, S.F. Analysis of LE-ACS3, a 1-aminocyclopropane-1-carboxylic acid synthase gene expressed during flooding in the roots of tomato plants. The Journal of Biological Chemistry, v.270, p.14056-14061, 1995.

PAPDI, C.; ABRAHÁM, E.; JOSEPH, M.P.; POPESCU, C.; KONCZ, C.; SZABADOS, L. Functional identification of Arabidopsis stress regulatory genes using the controlled cDNA overexpression system. Plant Physiology, v.147, p.528–542, 2008.

PARENT, C.; CAPELLI, N.; BERGER, A.; CREVECOEUR, M.; DAT, J.F. An overview of plant responses to soil waterlogging. Plant Stress, v.2. p.20-27, 2008.

PEÑA-CASTRO, J.M.; VAN ZANTEN, M.; LEE, S.C.; PATEL, M.R.; VOESENEK, L.A.; FUKAO, T.; BAILEY-SERRES, J. Expression of rice *SUB1A* and *SUB1C* transcription factors in Arabidopsis uncovers flowering inhibition as a submergence tolerance mechanism. The Plant Journal, v.67, p.434–446, 2011.

SACHS, M. M.; SUBBAIAH, C.C.; SAAB, I.N. Anaerobic gene expression and flooding tolerance in maize. Journal of Experimental Botany, v. 47, p. 1-15, 1996.

SAIRAM, R.K.; KUMUTHA, D. EZHILMATHI, K. Waterlogging tolerance: nonsymbiotic haemoglobin-nitric oxide homeostasis and antioxidants. Current Science, v.96, p.674-682, 2009.

SAKUMA, Y.; LIU, Q.; DUBOUZET, J.G.; ABE, H.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochemical and Biophysical Research Communications, v. 290, p.998–1009, 2002.

SILVA, S. D. dos A.; SERENO, M. J. C. de M.; SILVA, C. F. L.; BARBOSA NETO, J. F. Capacidade combinatória de genótipos de milho para tolerância ao encharcamento do solo. Ciência Rural, v.36, p.391-396, 2006.

SOUSA, C.A.F., SODEK, L. The metabolic response of plants to oxygen deficiency. Brazilian Journal of Plant Physiology, v.14, p.83-94, 2002.

STOHR, C.; STREMLAU, S. Formation and possible roles of nitric oxide in plant roots. Journal of Experimental Botany, v. 57, p. 463–470, 2006.

SUBBAIAH, C.C.; SACHS, M.M. Molecular and cellular adaptations of maize to flooding stress. Annals of Botany, v.90, p. 119-127, 2003.

TAMANG, B.G.; MAGLIOZZI, J.O.; MAROOF, M.A.; FUKAO, T. Physiological and transcriptomic characterization of submergence and reoxygenation responses in soybean seedlings. Plant Cell Environment, p.1-16, 2014.

VANTOAI, T.T.; MARTIN, S.K.ST.; CHASE, K.; BORU, G.; SCHNIPKE, V.; SCHMITTHENNER, F.; LARK, K.G. Identification of a QTL associated with tolerance of soybean to soil waterlogging. Crop Science, v. 41, p. 1247–1252, 2001.

VITORINO, P.G.; ALVES, J. D.; MAGALHÃES, P.C.; MAGALHÃES, M.M.; LIMA, L.C.O.; OLIVEIRA, L.E.M. Flooding tolerance and cell wall alterations in maize mesocotyl during hypoxia. Pesquisa Agropecuária Brasileira, v. 36, p.1027-1035, 2001.

XU, K.; XU, X.; FUKAO, T.; CANLAS, P.; MAGHIRANG-RODRIGUEZ, R.; HEUER, S.; ISMAIL, A.M.; BAILEY-SERRES, J.; RONALD, P.C.; MACKILL, D.J. S*ub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature, n.442, p.705-708, 2006.

ZHANG, Z.; WEI,L.; ZOU, X.; TAO, Y.; LIU, Z.; ZHENG, Y. Submergence-responsive MicroRNAs are potentially involved in the regulation of morphological and metabolic adaptations in maize root cells. Annals of Botany, v.102, p.509–519, 2008.

CAPÍTULO II (ARTIGO)

Insights into soybean transcriptome reconfiguration under hypoxic stress: functional, regulatory, and structural/compositional characterization

Abstract Oxygen gas (O_2) availability is one of primary forces shaping the evolution of living organisms and essential for land plants survival. Partial (hypoxic) or total (anoxic) oxygen gas deficiency is important component of flooding stress. Here, we analyzed soybean root RNA-seq data from flood-tolerant Embrapa 45 and flood-sensitive BR 4 cultivars under hypoxic stress. From divergences between cultivars, Embrapa 45 showed less up-, more down-regulated genes, and stronger induction of phosphoglucomutase (Glyma05g34790), unknown protein related to N-terminal protein myristoylation (Glyma06g03430), protein suppressor of phyA-105 (Glyma06g37080), and fibrillin (Glyma10g32620). RNA-seq and qRT-PCR data of non-symbiotic hemoglobin Glyma11g12980 indicated structural divergence of this gene between cultivars. Conserved between the two cultivars, transcriptional changes in amino acids and derivative metabolic process suggest involvement of amino acids metabolism in tRNA modifications, translation accuracy/efficiency and endoplasmic reticulum (ER) stress under hypoxia. Aiming for progress to understand how genes evolve and respond to stresses, stable and differentially expressed genes (DEGs) were structural and compositional characterized, comparing its mechanistic relationship with expression regulation. Genes groups differed in promoter TATA box, ABREs (ABA-responsive elements), and CRT/DREs (C-repeat/dehydration-responsive elements) frequency, as well as in structure, composition, and codon usage. We speculate that *cis*-acting ABRE can mediate gene expression independent of ABA in hypoxic soybean roots.

Keywords: flooding, oxygen deprivation, *Glycine max*

1 Introduction

Regimes of excess water (flooding) influence the distribution and diversity of species in natural ecosystems (Silvertown et al., 1999) and lead productivity losses of many farmland crops (Bailey-Serres et al., 2012b). The increased flooding events over the past six decades is associated with climate change, which threatens food security to the growing human population (Bailey-Serres et al., 2012a). Among the four major crops — soybean, wheat, maize, and rice — only rice is adapted to soil waterlogging and all are sensitive to submergence (Bailey-Serres et al., 2012b). The oxygen diffusion in water is 10.000 times lower than in air (Armstrong, 1980). Thus, water surrounding roots (waterlogging) or entire plants (submergence) can cause plant energy crisis once oxygen is required for energy production through mitochondrial respiration (Bailey-Serres and Voesenek, 2008). Additionally, oxygen is required in several metabolic pathways, including haem, sterol, and fatty-acid biosynthesis (Geigenberger, 2003). Orchestrated by complex gene regulatory network, plants and other organisms need to perceive, signal and promote biochemical, physiological, and morpho-anatomical changes appropriate to survive and thrive under oxygen level fluctuations.

In evolutionary context, paleontological and geological records indicate life on Earth originated prior to 3.4 billion-years-ago (GA) (Wacey et al., 2011; Bontognali et al., 2013) and evolved in anoxic (absence of oxygen) environments (Fischer et al., 2016). Around 2.4 GA, accumulation of oxygen, probably photosynthetic produced from cyanobacteria ancestors, would have led to heavy selective pressure by oxidative stress due to incompatibilities with core cellular biochemical processes (Fischer et al., 2016; Hamilton et al., 2016). After life have been at bacterial level for more than 2 billion years, transition to eukaryotic level would arose around 1.5 GA (Olson, 2012), involving mitochondrial endosymbiosis between an alpha-proteobacteria endosymbiont and a proto-eukaryote archaea host (Olson, 2012; Martin et al., 2015; Blackstone, 2016). Synergistic effects of bioenergetics and population genetics shows strong involvement with eukaryotic cellular and genomic complexity (Koonin, 2015). In general, compared to prokaryotes, larger and more energetic demanding eukaryotic cell has higher total energy budget (Lane and Martin, 2010; Lynch and

Marinov, 2015) mainly supplied from up to hundreds of thousands mitochondria (Okie et al., 2016) synthesizing ATPs by aerobic oxidative phosphorylation. The drop of effective population size accompanying organism enlargement from prokaryotes to multicellular aerobic eukaryotes magnify the power of random genetic drift, which allows eukaryotic genome restructure, providing novel substrates for secondary evolution of phenotypic complexity by natural selection (Lynch, 2003). Expansions in sizes and numbers of introns (Koonin, 2009), numbers of mobile elements, lengths of intergenic regions and numbers of genes are observed in nuclear genomes along phylogenetic lineages (Lynch, 2005). Mechanisms of duplication and divergence of genes (Zhang, 2003; Innan and Kondrashov, 2010) show to have been heavily important in the origin and radiation of eukaryotes. Comparative genomics indicate that substantial duplication and divergence of inherited ancient prokaryotic genes would have involved in early evolution of eukaryotes and subsequent transitions (Koonin et al., 2004; Makarova et al., 2005). Further duplication events (Cronk, 2001; Bowman et al., 2007; Jiao et al., 2011; Sørensen et al., 2012) would contributed for new genes involved in different plant hormone signaling mechanisms (Komatsu et al., 2013a; Wang et al., 2015a) along the transition from aquatic simple body green algae to land plants consisted of complex organs and tissue systems (Bowman, 2013; Delwiche and Cooper, 2015). Importantly, land plants constitute the basis for agriculture, industrial, and pharmaceutical products (Delwiche and Cooper, 2015), and adaptation capacities validated using crops in field conditions for extremes in water, temperature, and ions have shown association with duplication of genes (Mickelbart et al., 2015). To summarize with an example, the multigenic SNORKEL (SK) and SUBMERGENCE-1 (SUB1) loci are found in deepwater and submergence-tolerant rice varieties, respectively (Xu et al., 2006; Hattori et al., 2009). Both encode tandem-repeated genes, of which SK2 (Hattori et al., 2009) and SUB1A-1 (Xu et al., 2006) are major responsible for the rapid-growth avoidance escape (SK2) and energy saving quiescence (SUB1A-1) strategies against oxygen deprivation. Both loci are members of AP2 (APETALA2)/ERF (Ethylene Responsive Factor) plant-specific transcription factors superfamily, of which other members are associated with abiotic and biotic stress tolerance (Rehman and Mahmood, 2015), response to hormones, control of primary and secondary metabolism as well as growth and development in

diverse plant species (Licausi et al., 2013). Strikingly, although AP2/ERF domain have not been found in animals, it is widespread in transcription factors of plants including green algae and in protist, cyanobacterial, and viral HNH endonucleases (Wessler, 2005). Thus, it has been hypothesized that endonucleases bearing AP2/ERF domain from cyanobacterium by endosymbiosis or from a bacterium or a virus by other lateral gene transfer events have occurred into plants, leading to the origin of the AP2/ERF domain in plants via transposition and homing processes (Magnani, 2004).

Different from rice, soybean genome does not contain SUB1 and SK2 orthologs (Tamang et al., 2014), and only QTLs (Quantitative Trait Loci) with small effect for waterlogging tolerance have already been predicted (Vantoai et al., 2001; Nguyen et al., 2012). Deepening the genetic knowledge, transcriptomics identify genes involved in stress response mechanisms. Compared to other technologies, such as those based on hybridization (macro and microarray) or Sanger sequencing (SAGE, CAGE and MPSS), RNA-seq allows more precise and accurate measurements over a wider expression range of transcripts and their variants (Wang et al., 2009). To our knowledge, only one transcriptome study for flooding stress using RNA-seq technique have already been reported (Chen et al., 2016). The authors evaluated leaves transcriptome of soybean plants seven days after waterlogging. Another few already reported studies analyzed transcriptomes from shoots (cotyledons including hypocotyls) (Tamang et al., 2014) and roots including hypocotyls (Komatsu et al., 2009; Nanjo et al., 2011) of newly germinated soybean seedlings under hypoxic stress. All above studies evaluated one genotype. In the present study, we analyzed conserved and divergent root transcriptional responses between flood-tolerant Embrapa 45 and flood-sensitive BR 4 soybean cultivars under hypoxic stress by using RNA-seq platform. Aiming for progress to understand how genes evolve and respond to hypoxic stress, stable and DEGs were structural and compositional characterized, comparing its mechanistic relationship with expression regulation.

2 Materials and Methods

2.1 Grain yield

Embrapa 45 and BR 4 seeds were sown on December 21th of 2011 at Embrapa Clima Temperado, Pelotas, Brazil (latitude 31°42'S and longitude 52°24'W). We conducted the experiment in randomized block, with four replicates, 75 plants per plot, at a density of 200,000 plants per ha. The seed emergence occurred six days after sowing. The first waterlogging occurred 18 days after sowing due to heavy rain, lasting five days. We observed mild symptoms of yellowing leaves. On February 15th of 2012, we waterlogged the soil beds, maintaining water level 2 cm above the soil surface for 10 days. The plants developed typical reaction features to waterlogging. The harvest took place on May 23th, 2012.

2.2 Plant material for RNA extraction and qRT-PCR

For expression analysis of genes from hypoxic soybean roots, the experiment was set in a randomized block design, composed of twelve treatments: two cultivars (Embrapa 45 and BR 4), two oxygen conditions [fully aerobic state (normoxy) and hypoxic], and three treatment sampling times (0.5h, 4h, and 28h), each with three biological replicates (four plantlets per replicate). We used total RNA samples from biological material described previously (Nakayama et al., 2014). For quantitative real-time PCR (qRT-PCR) analysis, the synthesis of cDNA, design of primers, and expression analysis of genes used to verify reliability of RNA-seq expression data were done as described previously (Nakayama et al., 2014).

2.3 Transcriptome library construction, deep sequencing, and mapping of reads

For each of twelve treatments, equimolar quantities of purified total RNA from roots of twelve plants were pooled to result one library. Then, the twelve

libraries were sent to Fasteris SA (Plan-les-Ouates, Switzerland) for processing and sequencing. After processing the twelve libraries (poly-A purification, fractionation, cDNA synthesis using random primers, and ligation to bar-coded adapters), fragments of 150–250 pb were isolated and multiplexed, resulting one sequencing library (a pooled of equimolar quantities from twelve initial libraries; each library with a specific barcode for further bioinformatic discrimination). Then, sequencing library was used to produce clusters for a 1 x 100 bp single end-sequencing run into one lane on a flow cell for sequencing in a Hi-Seq 2000 (Illumina). The raw data was uploaded to the GeneSifter database (Geospiza, Seattle, WA) for alignment with a reference genome. The mapping of reads and transcripts analysis was done as described previously (Rodrigues et al., 2015).

2.4 Data Analysis

For each time point (0.5, 4, and 28h), the expression ratio (fold-change, fc) of genes was performed by dividing transcript abundance values (in RPM= Reads per Mapped Million) from plants under hypoxic and normoxic conditions. The statistical significance of DEGs were obtained by using Bioconductor package edgeR (Robinson et al., 2010), corrected by Benjamini and Hochberg method (Benjamini and Hochberg, 1995). We only considered as DEGs those showing fold-change ≥ 2 (up), ≤ -2 (down), adj. *p*-value ≤ 0.01, and with more than 20 mapped reads (RPM ≥ 9) in at least one of the two compared libraries.

Gene set enrichment analysis was performed using Singular Enrichment Analysis (SEA) provided by agriGO (http:// bioinfo.cau.edu.cn/agriGO/) (Du et al., 2010). We chose hypergeometric test, corrected by Hochberg FDR method, plant GO slim database. Soybean pathways of DEGs related to amino acids and its derivatives were analyzed using KEGG Mapper - Search&Color Pathway (Kanehisa et al., 2012). The RSCU (Relative Synonymous Codon Usage) was calculated with CodonW 1.4.4 (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::CodonW). The copy numbers of soybean nuclear tRNA genes was extracted from the genomic tRNA database (http://gtrnadb.ucsc.edu/) (Chan and Lowe, 2009). Samples sizes used for structural and compositional analysis among groups of genes are shown in the additional file 1.

The Microsoft Excel (Microsoft, Redmond, WA) was used for TATA-box searching in promoter regions and structural/compositional characterization of genes. For hypothesis testing on binary data, we used Microsoft Excel add-in Real Statistics Resource Pack (ver. 3.3.1, http://www.real-statistics.com/).

3 Results and Discussion

In the present study, we analyzed root RNA-seq data from flood-tolerant Embrapa 45 and flood-sensitive BR 4 cultivars that showed contrasted grain yields in waterlogged soil (Fig. 1). In order to evaluate pairwise RNA-seq data, the relative expression of six common hypoxia-responsive genes were determined by qRT-PCR. The qRT-PCR and RNA-seq expression data were consistent for all genes (Fig. 2) showing a strong positive Pearson correlation ($r = 0.95$; $P < 0.001$), indicating reliability in our transcriptome analysis.

3.1 Transcriptome reconfiguration

From 54,174 predicted protein-coding genes in soybean genome (assembly Glyma 1.1) (Schmutz et al., 2010), 2,505 up- and 4,769 down-regulated genes (URGs, DRGs) were found in Embrapa 45 in common to BR 4 responding to low-oxygen stress (data not shown). After 0.5, 4, and 28h of root hypoxia, 1,144, 5,687, and 3,761 genes differentially expressed in both cultivars, of which 89, 28, and 46% were URGs, respectively (Fig. 3). In thale cress (*Arabidopsis thaliana*), another flood-sensitive species, more URGs compared to DRGs were also observed in roots after 0.5h of oxygen deprivation, but kept URGs number prevalence even at later stress (Van Dongen et al., 2009). In contrast to present soybean data whose DRGs number decreased after 28h, predominance of DRGs remained along 168h of waterlogging in roots of flood-tolerant gray poplar (*Populus* × *canescens*) (Kreuzwieser et al., 2009). Even so, Embrapa 45 showed less URGs and more DRGs than BR 4, of which the highest contrasted DEGs number between cultivars was for more DRGs in Embrapa 45 (4,216 DRGs) than in BR 4 (2,582 DRGs) after 28h (Fig. 3).

Summarizing genes function, more gene ontology (GO) enriched categories were found after 0.5h for URGs in both cultivars and after 4-28h for DRGs mainly in Embrapa 45 (Fig. 3). It is noteworthy GO enrichment for gene expression regulation (GO:0010468), more specifically for transcription (GO:0006350) and protein modification (GO:0006464) involving transcription factors (GO:0003700) and kinases (GO:0016301) activities in URGs. In DRGs, highlight energy-demanding processes such as transport (GO:0005215) and biosynthesis (GO:0009058) including translation (GO:0006412), of which most of the genes encode ribosomal proteins (GO:0005198; structural molecule activity). Further, molecular function enrichment differed DEGs versus stable genes (SGs) in transcription factors (GO:0003700), kinases (GO:0016301) as well as transporters (GO:0005215) for DEGs, as mentioned above, and in a more general functions such as biding (GO:0005488) for stable genes.

As shown above, we observed that induction of controlling/signaling genes and down-regulation of genes related with energy-consuming processes highlight the responses of soybean (and possible other plants species) to oxygen deprivation, and some positive association occur between repression of genes and flooding tolerance.

3.1.1 Transcriptional changes in amino acids and derivative metabolic process.

After 4h, both URGs and DRGs were enriched for reorganization of cellular amino acid and derivative metabolic process (GO:0006519) (Fig. 3). Alterations in amino acid metabolism have been observed in hypoxic root of soybean (Oliveira and Sodek, 2013), *Lotus japonicus* (Rocha et al., 2010), and gray poplar (Kreuzwieser et al., 2009), showing high accumulation of alanine and GABA (Gamma-Amino Butyric Acid) as well as reduction of aspartate level. In agreement, we observed up-regulation of Glyma07g05130 (alanine aminotransferase) and Glyma08g09670 (glutamate decarboxylase) at all time points in both cultivars. Glutamate is directly related to alanine and aspartate metabolism via transamination, and to GABA via decarboxylation (Morot-Gaudry et al., 2001). Interestingly, we observed induction of genes NADH-dependent

glutamate synthase (Glyma04g41540 and Glyma19g16486) and aspartate aminotransferase (Glyma14g13480, Glyma17g33050, Glyma06g08670, and Glyma01g32360) in contrast to the repression of ferredoxin-dependent glutamate synthase (Glyma03g28410 and Glyma19g31120) and ATP-dependent asparagine synthase (Glyma11g27480, Glyma11g27720, Glyma14g37440, Glyma18g06840). It suggests metabolic improvement by NADH oxidation and ATP saving for glutamate synthesis under limiting oxygen. Although the transcriptional induction of aspartate kinase, aspartate semialdehyde dehydrogenase, and homocysteine S-methyltransferase involved in methionine synthesis, we observed repression of polyamines (derived from methionine as well as arginine and ornithine pathways) and phenylpropanoids biosynthesis-related genes, including upstream genes from shikimate pathway (data not shown). Among DRGs was S-adenosyl-L-methionine (SAM) synthetase (EC 2.5.1.6) from which product SAM connects ethylene, polyamines, and phenylpropanoid-derived lignin pathways. In addition, SAM is methyl donor of histones and nucleic acids for regulation of gene expression (at transcriptional and translational levels) (Li et al., 2011; Hori, 2014), further discussed in the last topic of this paper. Similar changes were also found in transcriptome of gray poplar (Kreuzwieser et al., 2009), proteome of soybean (Nanjo et al., 2012) roots under hypoxic stress, and decreased lignification (Komatsu et al., 2010). Nonetheless, opposed to our results, Wang et al., (2014) observed transcriptional and enzymatic activity induction of SAM, arginine, and ornithine decarboxylases (EC 4.1.1.50, EC 4.1.1.19, and EC 4.1.1.17, respectively) as well as increment of polyamine levels in hypoxia-stressed melon (*Cucumis melon* L.) roots. Although divergences in endogenous polyamine metabolism and its irregular association with stress tolerance in conventional plants, studies involving exogenous supplying or endogenous production of polyamines via genetic manipulation have shown increased tolerance to a broad spectrum of abiotic stresses (reviewed in Alcázar et al., 2010; Minocha et al., 2014; Pál et al., 2015; Saha et al., 2015) opening opportunities for improvement of soybean flooding tolerance by genetic engineering approaches.

3.1.2 Same gene, different regulation between cultivars

It is noteworthy that the gene phosphoglucomutase (PGM; Glyma05g34790) were up-regulated in Embrapa 45 and down-regulated in BR 4 after 4h of hypoxia (3 fc in Embrapa 45; -2 fc in BR 4) (Fig. 4). PGM participates in SuSy-UGPase (sucrose synthase, UDP-glucose pyrophosphorylase) pathway, and its up-regulation in Embrapa 45 is in accordance with the sucrose catabolism shift from ATP-dependent invertase-hexokinase to energy-saving SuSy-UGPase pathway (Bailey-Serres and Voesenek, 2008). In both cultivars occurred up-regulation of SuSy (Glyma19g40041, Glyma09g08550, and Glyma15g20180) and UGPase (Glyma11g33160) as well as down-regulation of invertase (Glyma10g35890) and hexokinase (Glyma05g35890, Glyma07g12190, and Glyma17g37720). In addition, we observed stronger induction of an unknown protein related to N-terminal protein myristoylation (Glyma06g03430; 386 fc in Embrapa 45; ns fc in BR 4 after 4h) (ns, no statistically significant), suppressor of phyA-105 (Glyma06g37080; 25 fc in Embrapa 45; ns fc in BR 4 after 28h), and fibrillin (Glyma10g32620; 142 fc in Embrapa 45; 20 fc in BR 4 after 28h) in Embrapa 45 than in BR 4. Fibrillins maintain structural integrity of plastoglobule, a plastid lipoprotein body found in all tissue types, and are involved in diverse plastid metabolism and biotic-abiotic stress tolerance (Lundquist et al., 2012). The thale cress AT3G23400, ortholog of Glyma10g32620, is required for lipophilic antioxidants transport in and out of the plastoglobules and for resistance to multiple stresses (Singh et al., 2010). Moreover, differential expression of fibrillins encoded genes correspond to plastoglobule number in leaves of contrasting soybean genotypes under drought and waterlogging stresses (Mutava et al., 2015), supporting our results in root system. Lastly, although the non-symbiotic hemoglobin Glyma11g12980 exhibits similar relative expression ratio between Embrapa 45 and BR 4, the expression level (in RPM) in BR 4 was higher than in Embrapa 45 under normoxic and hypoxic conditions (Fig. 4), suggesting structural gene divergence between cultivars. In agreement, we observed no qRT-PCR amplification of the Glyma11g12980 3'-end region in Embrapa 45, but a very similar transcriptional profile by qRT-PCR and RNA-seq in BR 4 (Fig. 2).

3.2 Structural and compositional characterization of genes groups

Aiming for progress to understand how genes evolve and respond to stresses, top 40 (all time points in both cultivars) and top 500 (at least one time point in both cultivars; except for S500 keeping all time points as criterion) stable and DEGs were structural/compositional characterized, comparing its mechanistic relationship with expression regulation.

3.2.1 The top genes groups

Previously reported in this work, genes aspartate aminotransferase (Glyma01g32360) and SAM decarboxylase (Glyma17g07830) ranked in top 40 up- (U40) and down- (D40) regulated groups, respectively. Likewise, phenylpropanoid/flavonoid related (D500) and SuSy-UGPase (U500) genes as well as three paralogs of Glyma06g03430 (N-terminal protein myristoylation) (U500) belonged to top 500 groups. Moreover, genes associated with ethylene biosynthesis (ACC oxidase, Glyma05g36310), glycolysis (Pyruvate kinase, Glyma10g34490), ethanol fermentation (pyruvate decarboxylase, Glyma01g29190, Glyma03g07380; alcohol dehydrogenase, Glyma04g39190, Glyma14g27940), biotic stress defense (kunitz trypsin protease inhibitor, Glyma16g33770; polygalacturonase inhibiting protein, Glyma05g25370, Glyma08g08360), and flooding governing acclimation (ERFVII related to thale cress HRE2, Glyma19g40070) were found in U40. In D40, we observed genes involved in antimicrobial defense (cysteine-rich secretory proteins, Glyma13g32500, Glyma13g32510, Glyma13g32530, and Glyma13g32540), gene regulation (bZIP, Glyma05g22860, Glyma17g17100; MYB, Glyma13g20510; NAC, Glyma08g18470), oxygen consuming (2-oxoglutarate oxygenase, Glyma03g23770, Glyma07g12210, Glyma08g18000; cytochrome P450, Glyma03g02410, Glyma04g03790), and transport of dicarboxylate (Glyma07g39580), sulfite (Glyma07g30570), manganese (2+), and iron (2+) (Glyma08g08090, Glyma16g28340, Glyma09g21920), which are ions that can reach toxic levels in waterlogged plants (Laanbroek, 1990; Lamers et al., 2013;

Shabala et al., 2014). Top 40 stable group (S40) has genes associated to protein (Glyma09g08830, Glyma06g01120, Glyma18g10060, Glyma18g32830, and Glyma04g23560) and nucleic acid binding (Glyma03g01920, Glyma18g11950, Glyma18g32190, Glyma06g48070, Glyma09g06750, Glyma12g29270, Glyma04g04880, and Glyma01g44460). The highest fold change through 0.5, 4, and 28h of hypoxic stress varied from 1.1 fc in P-loop containing nucleoside triphosphate hydrolase gene (Glyma18g32190) to 1.2 fc (S40) and -1.4 fc (S500) in the unknown genes Glyma06g17390 and Glyma11g18011, respectively. For flooding stress, this genes are promising candidate references for qRT-PCR normalization, given their higher stability compared to stable genes ELF1B (1.7 fc, Glyma02g44460) and ACTB (1.9 fc,Glyma15g05570), commonly used in the literature (Nakayama et al., 2014).

3.2.2 Top genes groups differed in promoter TATA box, ABRE, and DRE motifs frequency. Does ABRE mediate gene expression independent of ABA in hypoxic soybean roots?

Transcription of protein-coding genes in eukaryotes requires numerous protein factors to recognize specific DNA loci in a chromatin context, since nucleosome interfere the protein factor occupancy (Sadeh and Allis, 2011). The core promoter region, proximal to the transcription start-site (TSS), recruits general transcription factors involved in basal transcription (Nikolov and Burley, 1997) and *cis*-regulatory elements from extended promoter recruits DNA-bound transcription factors (activators or repressors) to fine-tune the transcriptional control (Fessele et al., 2002).

The general regulator TATA-box binding protein (TBP) is required for transcription initiation by all three eukaryotic RNA polymerases (Vannini and Cramer, 2012). TBP can bind to various DNA sequences but has higher affinity for the consensus TATA-box (Bonham et al., 2009). Based on previous works (Kim and Burley, 1994; Tirosh et al., 2006; Tora and Timmers, 2010), we scanned for the core sequence TATA extending 4 bp at 3' direction (coding strand) in 50 bp region upstream the predicted TSS (between -50 and -1), and found more DEGs (from 21% in U500 to 40% in D40) with the consensus TATA box sequence

TATA(T/A)ATA than SGs (at most 6% in S500) (Fig. 5; Additional file 2). Our results are in agreement with the hexamer sequences TATA(T/A)A over-represented in thale cress, rice (*Oryza sativa*), and soybean genomes (Maruyama et al., 2012) as well as with Tirosh et al. (2006), who observed correlation of consensus TATA-containing genes being differentially expressed while TATA less-containing genes stable expressed in yeast, metazoans, and plants. Moreover, higher TBP turnover on consensus TATA- compared to TATA-less promoters is associated with specific coactivators (van Werven et al., 2009; Tora and Timmers, 2010).

Coactivators are multisubunit complexes represented by SAGA (Spt-Ada-Gcn5-Acetyltransferase), TFIID (transcription Factor II D), related with TBP binding on TATA and TATA-less promoters, respectively (van Werven et al., 2009), and Mediator (Paul et al., 2015). The latter is organized into head, middle, tail, and Cdk8 kinase modules to converge and transmit signals from sequence-specific transcription factors to RNA polymerase (Eyboulet et al., 2015; Yang et al., 2016). Here, mediator components Glyma13g16910 (head MED20a) and Glyma13g31480 (tail MED16) were at least three times down- and nine times up-regulated, respectively, after 4 and 28h of hypoxia in both soybean cultivars. The head module MED20a subunit participates in transcription regulation of miRNA and protein-coding genes involved with plant development, time flowering, and fruits size in thale cress (Kim et al., 2011). In contrast, the tail module MED16 component regulates several biotic (Zhang et al., 2012; Wang et al., 2015b) and abiotic (Boyce et al., 2003; Knight et al., 2009; Wathugala et al., 2012) stress responses in plants. For example, not observed for MED20a, MED16 has important role in *Sclerotinia sclerotiorum* basal resistance, recruiting RNA Pol II for transcriptional activation of resistance genes (Wang et al., 2015b). Med16 is also required for transcriptional activation of cold- and dehydration-inducible genes that have C-repeat/dehydration-responsive elements (CRT/DRE) promoter motifs controlled by CRT/DRE-binding transcription factors (CBF/DREB) (Boyce et al., 2003; Knight et al., 2009). Although no protein-protein interaction between Med16 and DREB members has been observed yet, Med16 and DREB2A bind to another mediator component called Med25 (Yang et al., 2016). Here, soybean Med25 genes were not differentially expressed, except Glyma01g21831 (2 fc) after 28h in Embrapa 45. Opposing to Med16, MED25 is

not involved in cold tolerance (Yang et al., 2016), but med25 mutants display ABA hypersensitivity (Chen et al., 2012) and increased tolerance to drought with stronger Dreb2A mRNA up-regulation (Elfving et al., 2011).

As CRT/DRE and CBF/DREB, the *cis*-acting ABA-responsive elements (ABRE) and corresponded *trans*-acting factors ABRE-binding proteins/ABRE-binding factors (AREB/ABF) play important roles in abiotic stress tolerance in plants (Nakashima et al., 2014). From genomatix database (http://www.genomatix.de/) (Cartharius et al., 2005), *cis*-acting ABRE and DRE/CRT motifs were most frequent in the putative promoter region (1000 bp up- and 100 bp down-stream the TSS) of top up- compared to top down-regulated genes, mainly for genes containing both ABRE and DRE/CRT motifs in their promoter (Fig. 5; Additional file 2). The AREB/ABF genes Glyma06g04353 and Glyma19g30230 tended to increase their mRNA abundance (~2 fc) after 28h of hypoxia. The higher differential expression of GmDREB1B;1 (Glyma20g29410; Kidokoro et al., 2015) was observed after 0.5h (>20 fc), and GmDREB2A;2 (Glyma14g06080; Mizoi et al., 2013) after 28h (>24 fc) of hypoxic stress. GmDREB1B;1 and GmDREB2A;2 are also induced by heat, cold, drought, and high salt stresses, and both increase the transcription abundance of ABA receptor GmPYL21 (Glyma13g30210; Bai et al., 2013) (Mizoi et al., 2013; Kidokoro et al., 2015). GmPYL21 and GmPYL14 (Glyma14g06100; transcriptional activated by GmDREB2A;2; Mizoi et al., 2013), both up-regulated under hypoxia (>3 and >9 fc after 4h, respectively), interact with the phosphatase GmPP2C1 [Glyma13g16640, at most 3 fc under hypoxia, and 2 fc in leaves under drought stress (Rodrigues et al., 2015)] inhibiting it in an ABA independent manner (Bai et al., 2013). Interestingly, Kidokoro et al. (2015) observed that transcriptional activation of GmPYL21 by GmDREB1B;1 can enhance ABRE-mediated gene expression in an ABA-independent manner under cold stress, although ABA levels are not increased under such condition.

Here, we speculate ABRE-mediate gene expression independent of ABA in hypoxic soybean roots. Although exogenous application of ABA improves flooding tolerance of plants (Hwang and Vantoai, 1991; Ellis et al., 1999; Komatsu et al., 2013b), endogenous ABA decreases in roots under flooding stress (Arbona and Gómez-Cadenas, 2008; Hsu et al., 2011; Argamasilla et al., 2014). We observed down-regulation of ABA biosynthetic related genes (Glyma19g06540,

Glyma06g08944, Glyma13g27220, Glyma11g21160, and Glyma14g04950) and up-regulation of GmCYP707As (Glyma16g20490, Glyma01g35660, and Glyma09g35250), which are ABA inactivators (Zheng et al., 2012). In addition to GmCYP707As, we also found hypoxic soybean DEGs (Additional file 3) that are orthologs of thale cress genes involved in ABA signaling (ABA receptors Pyl4-6; phosphatases ABI1-2, HAB1-2, HAI1-3, and AHG3; AFP1-4; MAPKKK18) dependent of SRK2D/E/I (SNF1-related kinases SRK2D/SnRK2.2, SRK2E/SnRK2.6, and SRK2I/SnRK2.3) (Fig. 6) (from Genevestigator microarray database; Hruz et al., 2008). SRK2D/E/I are key activators of AREB/ABF proteins (Fujita et al., 2013). Glyma01g39020, which protein similarity is closer to SRK2D/E/I, increased its steady-state mRNA levels two and three times after 4 and 28h of hypoxia. Moreover, Glyma17g08270, Glyma04g06520, and Glyma06g06550 up-regulated at all time points under hypoxia (from 4 to 40 fc). They are orthologs of CIPK-6 and -25, kinases involved in Ca^{+2}-mediated expression of DREB1-2 and ABA signaling (He et al., 2013; Meena et al., 2015). Strikingly, although ABI1-2, PP2CA, and HAI1-2 thale cress genes down-regulated under hypoxic stress (Fig. 6), we observed up-regulation of soybean orthologs, highlighting the 41 to 930 folds of Glyma01g43460 (HAI3) transcripts at all time points (Additional file 3). These thale cress genes are up-regulated under drought, osmotic, and salinic stress (Fig. 6) and soybean genes under drought stress (Rodrigues et al., 2015), stresses associated with ABA production. These results suggest soybean-specific response to hypoxia, as observed by Kidokoro et al. (2015) to cold stress, involving divergent downstream genes responses to GmDREBs compared to AtDREBs, also observed by Mizoi et al. (2013) to thale cress developmental phenotypes. To summarize, in agreement to the important role of DREB and AREB for tolerance to diverse abiotic stresses, our data indicate that although ABA level decreases in hypoxic soybean roots, CRT/DRE and ABRE-mediated gene expression may occur in soybean under hypoxia. In addition, to overcome the reduced ABA level, we speculate that a constitutive active form of AREB/ABF protein (e.g. AREB1ΔQT; Fujita et al., 2005) is potential candidate to increase flooding tolerance.

3.2.3 Top genes groups differed in structure, composition, and codon usage. How does hypoxia may interfere translation?

Compared to SGs, DEGs had smaller primary transcript, CDS (coding sequence) length, and introns number (Table 1; Additional file 4). Similar results were also observed in thale cress (Aceituno et al., 2008), yeasts, and mice (Jeffares et al., 2008). Shorter genes with fewer introns respond economically and quickly. In average, 46 and 2 high-energy phosphate bonds are consumed for each nucleotide synthesis and its polymerization during transcription, and 27/4 high-energy phosphate bonds for each amino acid synthesis/polymerization during translation (Lynch and Marinov, 2015). Moreover, intron splicing is time-consuming process, taking several minutes (Bentley, 2014). Interestingly, SGs also seem to show energetic and time cost improvement. Compared to SGs here analyzed, hypoxia tend to down regulate cognate soybean proteins (Wang et al., 2016), and stable mRNAs from thale cress orthologs tend to be poorly associated with translating ribosomes and sequestrated into stress granules (ribonucleoprotein complexes) under hypoxia, whereas they are rapidly recruited to polyribosomes upon reoxygenation (Branco-Price et al., 2008; Sorenson and Bailey-Serres, 2014). On the other hand, protein abundance of hypoxic DEGs seems mainly governed by its total mRNA abundance.

In addition to higher intron number and splicing variants, SGs have higher gene body (i.e., transcript region) methylation (Aceituno et al., 2008; Tatarinova et al., 2013; Elhaik et al., 2014), which is also involved in pre-mRNA splicing regulation (Luco et al., 2011; Shukla and Oberdoerffer, 2012; Takuno and Gaut, 2013). So far, positive 5’ to 3’ methylation gradient opposite correlates with recombination in thale cress genes (Choi et al., 2013). Likewise, methylation and recombination are involved with AT-biased 5-methylcytosine deamination and GC-biased gene conversion, respectively (Glémin et al., 2014). Gene body of plants contain almost exclusively cytosine methylation in CpG dinucleotide than in CpHpG and CpHpH (p= phosphate bond; H= C, T, or A) trinucleotide targets (Schmitz et al., 2013; Song et al., 2013), as well as methylation are enriched on exons compared to introns (Chodavarapu et al., 2010). Here, whereas SGs exhibited steeper 5’ to 3’ negative G+C and CpG gradient, no decrement from start to middle region of DEGs were observed (Fig. 7; Additional file 4). The low

strength but diverged G+C and CpG patterns among genes groups is possible influenced by gene structure, recombination, and DNA methylation, as proposed by Glémin et al. (2014). Moreover, C3pG1 (C at third position of a codon binding G at first position of a neighbor codon) and G1 content were higher among di- and mono-nucleotides, respectively, and C3 was the mono-nucleotide with more diverged content among genes groups, mainly at CDS middle region.

By using RSCU (Relative Synonymous Codon Usage) analysis for all 59 synonymous codons, C3 content divergences highlighted 2-fold degenerate pyrimidine [cytosine (C), thymine (T), and uracil (U)] ending codons between CDS edges as well as among genes groups at CDS middle region (Fig. 8). It is noteworthy that 2-fold degenerate pyrimidine ending codons seems majority (maybe exclusively) decoded by G34 tRNAs (G at position 34 of the tRNA, the first anticodon position) in all three domains of life (archaea, bacteria, and eukarya) (Grosjean et al., 2010; Novoa et al., 2012). This suggests strong positive selection to discriminate correct cognate C3 and wobble U3 codons from the incorrect near-cognate codons G3 and A3 (e.g., CAC/U histidine versus CAG/A glutamine). On the other hand, A34 tRNAs decode 3 and 4-fold degenerate codon boxes in eukarya, but only for Arg in bacteria and none in archaea, correlating deaminase A34-to-I34 relaxed specificity in eukarya, Arg specificity in bacteria, and no homologue in archaea (Novoa et al., 2012; Paris et al., 2012; Torres et al., 2014). I34 is chemically analogue to guanosine, but the former lacks N2 amino group (Torres et al., 2014) and pairs with C-, U-, and A-ended codons (Wald and Margalit, 2014). C3 over U3 2-fold degenerate ending codon bias also occurs at evolutionarily conserved amino acids sites across 12 fly drosophilid species and correspond to higher levels of G34-to-Q34 substitution (Zaborske et al., 2014). This opens question if Q34 tRNA influences the compositional divergence among soybean genes groups. Interestingly, queuine (q), the free base of Q (queuosine), is only synthesized by bacteria and salvaged by most eukaryotes (Zallot et al., 2014). Example is the legume model *Medicago truncatula*, in which rhizobium Q synthesis is required for effective nitrogen-fixing symbiosis (Marchetti et al., 2013). On the other hand, this is not observed in brassicaceae, including thale cress, given the absence of genes encoding transglycosylases that catalyze q insertion in target G34U35N36 tRNAs (N= any base), found in *Medicago* (Zallot et al., 2014) and soybean (Glyma04g10706,

Glyma01g40041, Glyma13g10281, Glyma11g05250, Glyma06g10555, and Glyma08g48310). Further Q, wybutosine (yW) is another hypermodified nucleoside derived from G, but yW is found exclusively at position 37 (neighboring anticodon sequence) of tRNAs that decode phenylalanine (codons UUU and UUC; biased here among soybean genes groups), important to reduce polyuridine translational frameshift errors (Jackman and Alfonzo, 2013).

Here, DEGs related to modification of I34, Q34, yW37, pseudouridine, and methylation (tRNA and DNA methyltransferases), as well as aminoacylation of tRNAs (amino acid charging) were observed under hypoxia (Additional file 5). Pseudouridine is highly abundant post-transcriptional RNA modification (Ge and Yu, 2013). It is isomer of uridine with an extra hydrogen bond donor at non-Watson-Crick edge that rigidify sugar-phosphate backbone and enhance base stacking, contributing to stabilize local structures like tRNA stem-loops and anticodon (Charette and Gray, 2000). For instance, cytoplasmic tyrosine tRNAs contain a pseudouridine in the middle of their anticodon at position 35 (Szweykowska-Kulinska and Beier, 1992), and their cognate C3 and wobble U3 codons biased here among soybean genes groups. Among RNA species, tRNA molecules remain the most numerous and chemically modifications for proper structure, codon binding, and tRNA aminoacylation (Jackman and Alfonzo, 2013), and methylation types are most present (Jackman and Alfonzo, 2013; Swinehart and Jackman, 2015) or required as obligate precursor (Hori, 2014) for subsequent tRNA modifications, e.g. for yW37 synthesis (Noma et al., 2006). Conversely, considering that methyl-transfer reaction by majority of methyltransferases consumes SAM as methyl-group donor (Swinehart and Jackman, 2015) and its activity is inhibited by SAM immediate catabolic S-adenosyl-L-homocysteine (SAH) (Zhou et al., 2013), our transcriptome data indicate that hypoxia impairs biosynthesis of methionine (discussed in the second topic of this paper), mainly SAM and consequently DNA and tRNA methylation. Although S-methylmethionine cycle seems to alleviate methionine decrement by transcriptional up-regulation of homocysteine S-methyltransferase (ec: 2.1.1.10), which enzymatic activity does not require folate as cofactor (genes related to folate biosynthesis and upstream shikimate pathway were repressed here), genes encoding folate-dependent methionine synthase (ec: 2.1.1.14) were transcriptionally down-regulated as well as SAM synthetase (EC 2.5.1.6) and

SAH hydrolase (EC 3.3.1.1) under hypoxia (data not shown), also at protein level (Wang et al., 2016).

Finally, differential expression of aminoacyl-tRNA synthetases (aaRS) also occur in soybean leaf under drought stress, with more aaRS DEGs (but weaker drought tolerance) in wild-type than in transgenic lines overexpressing BiP (Carvalho et al., 2014). The molecular chaperone BiP (Binding Protein) major regulate the endoplasmic reticulum (ER) stress (Reis et al., 2011). Here, many ER stress related genes (Silva et al., 2015) expressed differentially (Additional file 6), including down-regulation of BIP (Glyma05g36600, Glyma05g36620, Glyma08g02940, and Glyma08g02960; from -2 to -10 fc after 4 and 28h in both cultivars under hypoxia). Among aaRS differentially expressed under hypoxic stress, Glyma14g11711 transcripts (aspartyl-tRNA synthetase) up-regulated at all time points in the two soybean cultivars (from 5 to 16 fc) here, also at protein level (2.3 fc) reported by Oh and Komatsu (2015). Curiously, transgenic overexpression of thale cress AspRS orthologue At4g31180 increase tolerance to biotic stress, and At4g31180 is target of a synthetic isomer of GABA, called BABA (β-Amino Butyric Acid) (Luna et al., 2014). BABA primes plants to enhance tolerance against broad-spectrum of biotic and abiotic stresses (Cohen, 2002), potentiating various hormone-mediated signaling processes (Baccelli and Mauch-Mani, 2015). Inhibition of AspRS activity by BABA accumulates uncharged tRNA Asp, which as others uncharged tRNAs is signaling molecule to attenuate translation by Gcn2-eIF2α system, and subsequently alleviate ER stress (Kørner et al., 2015). Further studies may help to elucidate involvement of amino acids metabolism in tRNA modifications, translation accuracy/efficiency, and ER stress under hypoxia. In addition, BIP and aminoacyl-tRNA synthetases are candidates for biotechnology applications for increment of flooding tolerance.

Literature Cited

Aceituno FF, Moseyko N, Rhee SY, Gutiérrez RA (2008) The rules of gene expression in plants: Organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. BMC

Genomics **9**: 438

Alcázar R, Altabella T, Marco F, Bortolotti C, Reymond M, Koncz C, Carrasco P, Tiburcio AF (2010) Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. Planta **231**: 1237–1249

Arbona V, Gómez-Cadenas A (2008) Hormonal Modulation of Citrus Responses to Flooding. J Plant Growth Regul **27**: 241–250

Argamasilla R, Gómez-Cadenas A, Arbona V (2014) Metabolic and Regulatory Responses in Citrus Rootstocks in Response to Adverse Environmental Conditions. J Plant Growth Regul **33**: 169–180

Armstrong W (1980) Aeration in Higher Plants. Adv Bot Res **7**: 225–332

Baccelli I, Mauch-Mani B (2015) Beta-aminobutyric acid priming of plant defense: the role of ABA and other hormones. Plant Mol Biol 1–9

Bai G, Yang D-H, Zhao Y, Ha S, Yang F, Ma J, Gao X-S, Wang Z-M, Zhu J-K (2013) Interactions between soybean ABA receptors and type 2C protein phosphatases. Plant Mol Biol **83**: 651–664

Bailey-Serres J, Fukao T, Gibbs DJ, Holdsworth MJ, Lee SC, Licausi F, Perata P, Voesenek LACJ, van Dongen JT (2012a) Making sense of low oxygen sensing. Trends Plant Sci **17**: 129–138

Bailey-Serres J, Lee SC, Brinton E (2012b) Waterproofing Crops: Effective Flooding Survival Strategies. Plant Physiol **160**: 1698–1709

Bailey-Serres J, Voesenek LACJ (2008) Flooding Stress: Acclimations and Genetic Diversity. Annu Rev Plant Biol **59**: 313–339

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc **57**: 289–300

Bentley DL (2014) Coupling mRNA processing with transcription in time and space. Nat Rev Genet **15**: 163–175

Blackstone N (2016) An Evolutionary Framework for Understanding the Origin of Eukaryotes. Biology (Basel) **5**: 18

Bonham AJ, Neumann T, Tirrell M, Reich NO (2009) Tracking transcription factor complexes on DNA using total internal reflectance fluorescence protein binding microarrays. Nucleic Acids Res **37**: e94–e94

Bontognali TRR, Fischer WW, Föllmi KB (2013) Siliciclastic associated banded iron formation from the 3.2Ga Moodies Group, Barberton Greenstone Belt, South Africa. Precambrian Res **226**: 116–124

Bowman JL (2013) Walkabout on the long branches of plant evolution. Curr Opin Plant Biol **16**: 70–77

Bowman JL, Floyd SK, Sakakibara K (2007) Green Genes—Comparative Genomics of the Green Branch of Life. Cell **129**: 229–234

Boyce JM, Knight H, Deyholos M, Openshaw MR, Galbraith DW, Warren G, Knight MR (2003) The sfr6 mutant of Arabidopsis is defective in transcriptional activation via CBF/DREB1 and DREB2 and shows sensitivity to osmotic stress. Plant J **34**: 395–406

Branco-Price C, Kaiser KA, Jang CJH, Larive CK, Bailey-Serres J (2008) Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. Plant J **56**: 743–755

Byfield GE, Xue H, Upchurch RG (2006) Two Genes from Soybean Encoding Soluble Δ9 Stearoyl-ACP Desaturases. Crop Sci **46**: 840

Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T (2005) MatInspector and beyond: Promoter analysis based on transcription factor binding sites. Bioinformatics **21**: 2933–2942

Carvalho HH, Brustolini OJB, Pimenta MR, Mendes GC, Gouveia BC, Silva PA, Silva JCF, Mota CS, Soares-Ramos JRL, Fontes EPB (2014) The Molecular Chaperone Binding Protein BiP Prevents Leaf Dehydration-Induced Cellular Homeostasis Disruption. PLoS One **9**: e86661

Chan PP, Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res **37**: D93–D97

Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB Life **49**: 341–51

Chen R, Jiang H, Li L, Zhai Q, Qi L, Zhou W, Liu X, Li H, Zheng W, Sun J, et al (2012) The Arabidopsis Mediator Subunit MED25 Differentially Regulates Jasmonate and Abscisic Acid Signaling through Interacting with the MYC2 and ABI5 Transcription Factors. Plant Cell **24**: 2898–2916

Chen W, Yao Q, Patil GB, Agarwal G, Deshmukh RK, Lin L, Wang B, Wang Y, Prince SJ, Song L, et al (2016) Identification and Comparative Analysis of Differential Gene Expression in Soybean Leaf Tissue under Drought and Flooding Stress Revealed by RNA-seq. Front Plant Sci **7**: 1–19

Chodavarapu RK, Feng S, Bernatavichute Y V, Chen P, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al (2010) Relationship between nucleosome positioning and DNA methylation. Nature **466**: 388–392

Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver GP, Franklin FCH, et al (2013) Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. Nat Genet **45**: 1327–1336

Cohen YR (2002) β-Aminobutyric Acid-Induced Resistance Against Plant Pathogens. Plant Dis **86**: 448–457

Cronk QCB (2001) Plant evolution and development in a post-genomic context. Nat Rev Genet **2**: 607–619

Delwiche CF, Cooper ED (2015) The Evolutionary Origin of a Terrestrial Flora. Curr Biol **25**: R899–R910

Van Dongen JT, Fröhlich A, Ramírez-Aguilar SJ, Schauer N, Fernie AR, Erban A, Kopka J, Clark J, Langer A, Geigenberger P (2009) Transcript and metabolite profiling of the adaptive response to mild decreases in oxygen concentration in the roots of arabidopsis plants. Ann Bot **103**: 269–280

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res **38**: W64–W70

Elfving N, Davoine C, Benlloch R, Blomberg J, Brannstrom K, Muller D, Nilsson A, Ulfstedt M, Ronne H, Wingsle G, et al (2011) The *Arabidopsis thaliana* Med25 mediator subunit integrates environmental cues to control plant development. Proc Natl Acad Sci **108**: 8245–8250

Elhaik E, Pellegrini M, Tatarinova T V (2014) Gene expression and nucleotide composition are associated with genic methylation level in *Oryza sativa*. BMC Bioinformatics **15**: 23

Ellis MH, Dennis ES, Peacock WJ (1999) Arabidopsis roots and shoots have different mechanisms for hypoxic stress tolerance. Plant Physiol **119**: 57–64

Eyboulet F, Wydau-Dematteis S, Eychenne T, Alibert O, Neil H, Boschiero C, Nevers MC, Volland H, Cornu D, Redeker V, et al (2015) Mediator independently orchestrates multiple steps of preinitiation complex assembly in vivo. Nucleic Acids Res **43**: 9214–9231

Fessele S, Maier H, Zischek C, Nelson PJ, Werner T (2002) Regulatory context

is a crucial part of gene function. Trends Genet **18**: 60–63

Fischer WW, Hemp J, Valentine JS (2016) How did life survive Earth's great oxygenation? Curr Opin Chem Biol **31**: 166–178

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, Hiratsu K, Ohme-Takagi M, Shinozaki K, Yamaguchi-Shinozaki K (2005) AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell **17**: 3470–88

Fujita Y, Yoshida T, Yamaguchi-Shinozaki K (2013) Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. Physiol Plant **147**: 15–27

Ge J, Yu Y-T (2013) RNA pseudouridylation: new insights into an old modification. Trends Biochem Sci **38**: 210–218

Geigenberger P (2003) Response of plant metabolism to too little oxygen. Curr Opin Plant Biol **6**: 247–256

Glémin S, Clément Y, David J, Ressayre A (2014) GC content evolution in coding regions of angiosperm genomes: a unifying hypothesis. Trends Genet **30**: 263–270

Grosjean H, de Crécy-Lagard V, Marck C (2010) Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. FEBS Lett **584**: 252–264

Hamilton TL, Bryant DA, Macalady JL (2016) The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans. Environ Microbiol **18**: 325–340

Hattori Y, Nagai K, Furukawa S, Song X, Kawano R, Sakakibara H, Wu J, Matsumoto T, Yoshimura A, Kitano H, et al (2009) The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. Nature **460**: 1026–1030

He L, Yang X, Wang L, Zhu L, Zhou T, Deng J, Zhang X (2013) Molecular cloning and functional characterization of a novel cotton CBL-interacting protein kinase gene (GhCIPK6) reveals its involvement in multiple abiotic stress tolerance in transgenic plants. Biochem Biophys Res Commun **435**: 209–215

Hori H (2014) Methylated nucleosides in tRNA and tRNA methyltransferases. Front Genet **5**: 1–26

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. Adv Bioinformatics **2008**: 1–5

Hsu F-C, Chou M-Y, Peng H-P, Chou S-J, Shih M-C (2011) Insights into hypoxic systemic responses based on analyses of transcriptional regulation in Arabidopsis. PLoS One **6**: e28888

Hwang S, Vantoai TT (1991) Abscisic Acid Induces Anaerobiosis Tolerance in Corn. Plant Physiol **97**: 593–597

Innan H, Kondrashov F (2010) The evolution of gene duplications: classifying and distinguishing between models. Nat Rev Genet **11**: 4

Jackman JE, Alfonzo JD (2013) Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA **4**: 35–48

Jeffares DC, Penkett CJ, Bähler J (2008) Rapidly regulated genes are intron poor. Trends Genet **24**: 375–378

Jian B, Liu B, Bi Y, Hou W, Wu C, Han T (2008) Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol Biol **14**: 1–14

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, et al (2011) Ancestral polyploidy in seed plants and angiosperms. Nature **473**: 97–100

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res **40**: D109–D114

Kidokoro S, Watanabe K, Ohori T, Moriwaki T, Maruyama K, Mizoi J, Myint Phyu Sin Htwe N, Fujita Y, Sekita S, Shinozaki K, et al (2015) Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. Plant J **81**: 505–518

Kim JL, Burley SK (1994) 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. Nat Struct Biol **1**: 638–653

Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X (2011) The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. EMBO J **30**: 814–822

Knight H, Mugford SG, Ülker B, Gao D, Thorlby G, Knight MR (2009)

Identification of SFR6, a key component in cold acclimation acting post-translationally on CBF function. Plant J **58**: 97–108

Komatsu K, Suzuki N, Kuwamura M, Nishikawa Y, Nakatani M, Ohtawa H, Takezawa D, Seki M, Tanaka M, Taji T, et al (2013a) Group A PP2Cs evolved in land plants as key regulators of intrinsic desiccation tolerance. Nat Commun **4**: 2219

Komatsu S, Han C, Nanjo Y, Altaf-Un-Nahar M, Wang K, He D, Yang P (2013b) Label-Free Quantitative Proteomic Analysis of Abscisic Acid Effect in Early-Stage Soybean under Flooding. J Proteome Res **12**: 4769–4784

Komatsu S, Kobayashi Y, Nishizawa K, Nanjo Y, Furukawa K (2010) Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. Amino Acids **39**: 1435–1449

Komatsu S, Yamamoto R, Nanjo Y, Mikami Y, Yunokawa H, Sakata K (2009) A Comprehensive Analysis of the Soybean Genes and Proteins Expressed under Flooding Stress using Transcriptome and Proteome Techniques. J Proteome Res **8**: 4766–4778

Koonin E V (2009) Intron-dominated genomes of early ancestors of eukaryotes. J Hered **100**: 618–623

Koonin E V, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, et al (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biol **5**: R7

Koonin E V. (2015) Energetics and population genetics at the root of eukaryotic cellular and genomic complexity. Proc Natl Acad Sci **112**: 15777–15778

Kørner C, Du X, Vollmer M, Pajerowska-Mukhtar K (2015) Endoplasmic Reticulum Stress Signaling in Plant Immunity—At the Crossroad of Life and Death. Int J Mol Sci **16**: 26582–26598

Kreuzwieser J, Hauberg J, Howell KA, Carroll A, Rennenberg H, Millar AH, Whelan J (2009) Differential response of gray poplar leaves and roots underpins stress adaptation during hypoxia. Plant Physiol **149**: 461–73

Laanbroek HJ (1990) Bacterial cycling of minerals that affect plant growth in waterlogged soils: a review. Aquat Bot **38**: 109–125

Lamers LPM, Govers LL, Janssen ICJM, Geurts JJM, Van der Welle MEW, Van Katwijk MM, Van der Heide T, Roelofs JGM, Smolders AJP (2013)

Sulfide as a soil phytotoxin—a review. Front Plant Sci **4**: 1–14

Lane N, Martin W (2010) The energetics of genome complexity. Nature **467**: 929–934

Li W, Han Y, Tao F, Chong K (2011) Knockdown of SAMS genes encoding S-adenosyl-l-methionine synthetases causes methylation alterations of DNAs and histones and leads to late flowering in rice. J Plant Physiol **168**: 1837–1843

Licausi F, Ohme-Takagi M, Perata P (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. New Phytol **199**: 639–649

Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in Alternative Pre-mRNA Splicing. Cell **144**: 16–26

Luna E, van Hulten M, Zhang Y, Berkowitz O, López A, Pétriacq P, Sellwood M a, Chen B, Burrell M, van de Meene A, et al (2014) Plant perception of β-aminobutyric acid is mediated by an aspartyl-tRNA synthetase. Nat Chem Biol **10**: 450–456

Lundquist PK, Poliakov A, Bhuiyan NH, Zybailov B, Sun Q, van Wijk KJ (2012) The Functional Network of the Arabidopsis Plastoglobule Proteome Based on Quantitative Proteomics and Genome-Wide Coexpression Analysis. Plant Physiol **158**: 1172–1192

Lynch M (2003) The Origins of Genome Complexity. Science (80-) **302**: 1401–1404

Lynch M (2005) The Origins of Eukaryotic Gene Structure. Mol Biol Evol **23**: 450–468

Lynch M, Marinov GK (2015) The bioenergetic costs of a gene. Proc Natl Acad Sci **112**: 201514974

Magnani E, Sjölander K, Hake S (2004) From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. Plant Cell **16**: 2265–77

Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin E V. (2005) Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. Nucleic Acids Res **33**: 4626–4638

Marchetti M, Capela D, Poincloux R, Benmeradi N, Auriac M-C, Le Ru A, Maridonneau-Parini I, Batut J, Masson-Boivin C (2013) Queuosine

Biosynthesis Is Required for Sinorhizobium meliloti-Induced Cytoskeletal Modifications on HeLa Cells and Symbiosis with *Medicago truncatula*. PLoS One **8**: e56043

Martin WF, Garg S, Zimorski V (2015) Endosymbiotic theories for eukaryote origin. Philos Trans R Soc B Biol Sci **370**: 20140330

Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, et al (2012) Identification of Cis-Acting Promoter Elements in Cold- and Dehydration-Induced Transcriptional Pathways in Arabidopsis, Rice, and Soybean. DNA Res **19**: 37–49

Meena MK, Ghawana S, Dwivedi V, Roy A, Chattopadhyay D (2015) Expression of chickpea CIPK25 enhances root growth and tolerance to dehydration and salt stress in transgenic tobacco. Front Plant Sci **6**: 1–11

Mickelbart M V, Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nat Rev Genet **16**: 237–251

Minocha R, Majumdar R, Minocha SC (2014) Polyamines and abiotic stress in plants: a complex relationship1. Front Plant Sci **5**: 175 (11–17)

Mizoi J, Ohori T, Moriwaki T, Kidokoro S, Todaka D, Maruyama K, Kusakabe K, Osakabe Y, Shinozaki K, Yamaguchi-Shinozaki K (2013) GmDREB2A;2, a Canonical DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN2-Type Transcription Factor in Soybean, Is Posttranslationally Regulated and Mediates Dehydration-Responsive Element-Dependent Gene Expression. Plant Physiol **161**: 346–361

Morot-Gaudry J-F, Job D, Lea PJ (2001) Amino Acid Metabolism. *In* J-F Morot-Gaudry, PJ Lea, eds, Plant Nitrogen. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 167–211

Mutava RN, Prince SJK, Syed NH, Song L, Valliyodan B, Chen W, Nguyen HT (2015) Understanding abiotic stress tolerance mechanisms in soybean: A comparative evaluation of soybean response to drought and flooding stress. Plant Physiol Biochem **86**: 109–120

Nakashima K, Yamaguchi-Shinozaki K, Shinozaki K (2014) The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. Front Plant Sci

5: 170

Nakayama TJ, Rodrigues FA, Neumaier N, Marcelino-Guimarães FC, Farias JRB, de Oliveira MCN, Borém A, de Oliveira ACB, Emygdio BM, Nepomuceno AL (2014) Reference genes for quantitative real-time polymerase chain reaction studies in soybean plants under hypoxic conditions. Genet Mol Res **13**: 860–871

Nanjo Y, Maruyama K, Yasue H, Yamaguchi-Shinozaki K, Shinozaki K, Komatsu S (2011) Transcriptional responses to flooding stress in roots including hypocotyl of soybean seedlings. Plant Mol Biol **77**: 129–144

Nanjo Y, Skultety L, Uváčková L, Klubicová K, Hajduch M, Komatsu S (2012) Mass Spectrometry-Based Analysis of Proteomic Changes in the Root Tips of Flooded Soybean Seedlings. J Proteome Res **11**: 372–385

Nguyen VT, Vuong TD, VanToai T, Lee JD, Wu X, Rouf Mian MA, Dorrance AE, Shannon JG, Nguyen HT (2012) Mapping of quantitative trait loci associated with resistance to Phytophthora sojae and flooding tolerance in soybean. Crop Sci **52**: 2481–2493

Nikolov DB, Burley SK (1997) RNA polymerase II transcription initiation: A structural view. Proc Natl Acad Sci **94**: 15–22

Noma A, Kirino Y, Ikeuchi Y, Suzuki T (2006) Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. EMBO J **25**: 2142–2154

Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L (2012) A Role for tRNA Modifications in Genome Structure and Codon Usage. Cell **149**: 202–213

Oh M, Komatsu S (2015) Characterization of proteins in soybean roots under flooding and drought stresses. J Proteomics **114**: 161–181

Okie JG, Smith VH, Martin-Cereceda M (2016) Major evolutionary transitions of life, metabolic scaling and the number and size of mitochondria and chloroplasts. Proc. R. Soc. London B Biol. Sci. 283:

Oliveira HC, Sodek L (2013) Effect of oxygen deficiency on nitrogen assimilation and amino acid metabolism of soybean root segments. Amino Acids **44**: 743–755

Olson KR (2012) Mitochondrial adaptations to utilize hydrogen sulfide for energy and signaling. J Comp Physiol B **182**: 881–897

Pál M, Szalai G, Janda T (2015) Speculation: Polyamines are important in abiotic stress signaling. Plant Sci **237**: 16–23

Paris Z, Fleming IMC, Alfonzo JD (2012) Determinants of tRNA editing and modification: Avoiding conundrums, affecting function. Semin Cell Dev Biol **23**: 269–274

Paul E, Zhu ZI, Landsman D, Morse RH (2015) Genome-wide association of mediator and RNA polymerase II in wild-type and mediator mutant yeast. Mol Cell Biol **35**: 331–42

Rehman S, Mahmood T (2015) Functional role of DREB and ERF transcription factors: regulating stress-responsive network in plants. Acta Physiol Plant **37**: 1–14

Reis P a. a., Rosado GL, Silva L a. C, Oliveira LC, Oliveira LB, Costa MDL, Alvim FC, Fontes EPB (2011) The Binding Protein BiP Attenuates Stress-Induced Cell Death in Soybean via Modulation of the N-Rich Protein-Mediated Signaling Pathway. Plant Physiol **157**: 1853–1865

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**: 139–140

Rocha M, Licausi F, Araújo WL, Nunes-Nesi A, Sodek L, Fernie AR, van Dongen JT (2010) Glycolysis and the tricarboxylic acid cycle are linked by alanine aminotransferase during hypoxia induced by waterlogging of Lotus japonicus. Plant Physiol **152**: 1501–1513

Rodrigues FA, Fuganti-Pagliarini R, Marcolino-Gomes J, Nakayama TJ, Molinari HBC, Lobo FP, Harmon FG, Nepomuceno AL (2015) Daytime soybean transcriptome fluctuations during water deficit stress. BMC Genomics **16**: 505

Sadeh R, Allis CD (2011) Genome-wide “Re”-Modeling of Nucleosome Positions. Cell **147**: 263–266

Saha J, Brauer EK, Sengupta A, Popescu SC, Gupta K, Gupta B (2015) Polyamines as redox homeostasis regulators during salt stress in plants. Front Environ Sci **3**: 1–13

Schmitz RJ, He Y, Valdes-Lopez O, Khan SM, Joshi T, Urich M a., Nery JR, Diers B, Xu D, Stacey G, et al (2013) Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome

Res **23**: 1663–1674

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, et al (2010) Genome sequence of the palaeopolyploid soybean. Nature **463**: 178–183

Shabala S, Shabala L, Barcelo J, Poschenrieder C (2014) Membrane transporters mediating root signalling and adaptive responses to oxygen deprivation and soil flooding. Plant Cell Environ **37**: 2216–33

Shukla S, Oberdoerffer S (2012) Co-transcriptional regulation of alternative pre-mRNA splicing. Biochim Biophys Acta - Gene Regul Mech **1819**: 673–683

Silva PA, Silva JCF, Caetano HDN, Machado JPB, Mendes GC, Reis PAB, Brustolini OJB, Dal-Bianco M, Fontes EPB (2015) Comprehensive analysis of the endoplasmic reticulum stress response in the soybean genome: conserved and plant-specific features. BMC Genomics **16**: 783

Silvertown J, Dodd M, Gowing DJ, Mountford JO (1999) Hydrologically defined niches reveal a basis for species richness in plant communities. Nature **400**: 61–63

Singh DK, Maximova SN, Jensen PJ, Lehman BL, Ngugi HK, McNellis TW (2010) FIBRILLIN4 Is Required for Plastoglobule Development and Stress Resistance in Apple and Arabidopsis. Plant Physiol **154**: 1281–1293

Song Q, Lu X, Li Q, Chen H, Hu X, Ma B, Zhang W, Chen S-Y, Zhang J-S (2013) Genome-Wide Analysis of DNA Methylation in Soybean. Mol Plant **6**: 1961–1974

Sørensen I, Rose JKC, Doyle JJ, Domozych DS, Willats WGT (2012) The Charophycean green algae as model systems to study plant cell walls and other evolutionary adaptations that gave rise to land plants. Plant Signal Behav **7**: 1–3

Sorenson R, Bailey-Serres J (2014) Selective mRNA sequestration by OLIGOURIDYLATE-BINDING PROTEIN 1 contributes to translational control during hypoxia in Arabidopsis. Proc Natl Acad Sci **111**: 2373–2378

Swinehart WE, Jackman JE (2015) Diversity in mechanism and function of tRNA methyltransferases. RNA Biol **12**: 398–411

Szweykowska-Kulinska Z, Beier H (1992) Sequence and structure requirements for the biosynthesis of pseudouridine (psi 35) in plant pre-tRNA(Tyr). EMBO J **11**: 1907–12

Takuno S, Gaut BS (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. Proc Natl Acad Sci **110**: 1797–1802

Tamang BG, Magliozzi JO, Maroof MAS, Fukao T (2014) Physiological and transcriptomic characterization of submergence and reoxygenation responses in soybean seedlings. Plant, Cell Environ **37**: 2350–2365

Tatarinova T, Elhaik E, Pellegrini M (2013) Cross-Species Analysis of Genic GC3 Content and DNA Methylation Patterns. Genome Biol Evol **5**: 1443–1456

Tirosh I, Weinberger A, Carmi M, Barkai N (2006) A genetic signature of interspecies variations in gene expression. Nat Genet **38**: 830–834

Tora L, Timmers HTM (2010) The TATA box regulates TATA-binding protein (TBP) dynamics in vivo. Trends Biochem Sci **35**: 309–314

Torres AG, Piñeyro D, Filonava L, Stracker TH, Batlle E, Ribas de Pouplana L (2014) A-to-I editing on tRNAs: Biochemical, biological and evolutionary implications. FEBS Lett **588**: 4279–4286

Vannini A, Cramer P (2012) Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. Mol Cell **45**: 439–446

Vantoai TT, Martin SKS, Chase K, Boru G, Schnipke V, Schmitthenner AF, Lark KG (2001) Identification of a QTL associated with tolerance of soybean to soil waterlogging. Crop Sci **41**: 1247–1252

Wacey D, Kilburn MR, Saunders M, Cliff J, Brasier MD (2011) Microfossils of sulphur-metabolizing cells in 3.4-billion-year-old rocks of Western Australia. Nat Geosci **4**: 698–702

Wald N, Margalit H (2014) Auxiliary tRNAs: large-scale analysis of tRNA genes reveals patterns of tRNA repertoire dynamics. Nucleic Acids Res **42**: 6552–6566

Wang C, Fan L, Gao H, Wu X, Li J, Lv G, Gong B (2014) Polyamine biosynthesis and degradation are modulated by exogenous gamma-aminobutyric acid in root-zone hypoxia-stressed melon roots. Plant Physiol Biochem **82**: 17–26

Wang C, Liu Y, Li S-S, Han G-Z (2015a) Insights into the Origin and Evolution of the Plant Hormone Signaling Machinery. Plant Physiol **167**: 872–886

Wang C, Yao J, Du X, Zhang Y, Sun Y, Rollins JA, Mou Z (2015b) The

Arabidopsis Mediator Complex Subunit16 Is a Key Component of Basal Resistance against the Necrotrophic Fungal Pathogen Sclerotinia sclerotiorum. Plant Physiol **169**: 856–872

Wang X, Oh M, Sakata K, Komatsu S (2016) Gel-free/label-free proteomic analysis of root tip of soybean over time under flooding and drought stresses. J Proteomics **130**: 42–55

Wang Z, Gerstein M, Snyder M (2009) RNA-seq: a revolutionary tool for transcriptomics.

Wathugala DL, Hemsley PA, Moffat CS, Cremelie P, Knight MR, Knight H (2012) The Mediator subunit SFR6/MED16 controls defence gene expression mediated by salicylic acid and jasmonate responsive pathways. New Phytol **195**: 217–230

van Werven FJ, van Teeffelen HAAM, Holstege FCP, Timmers HTM (2009) Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. Nat Struct Mol Biol **16**: 1043–8

Wessler SR (2005) Homing into the origin of the AP2 DNA binding domain. Trends Plant Sci **10**: 54–56

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature **442**: 705–708

Yang Y, Li L, Qu L-J (2016) Plant Mediator complex and its critical functions in transcription regulation. J Integr Plant Biol **58**: 106–118

Zaborske JM, Bauer DuMont VL, Wallace EWJ, Pan T, Aquadro CF, Drummond DA (2014) A Nutrient-Driven tRNA Modification Alters Translational Fidelity and Genome-wide Protein Coding across an Animal Genus. PLoS Biol **12**: e1002015

Zallot R, Brochier-Armanet C, Gaston KW, Forouhar F, Limbach P a, Hunt JF, de Crécy-Lagard V (2014) Plant, Animal, and Fungal Micronutrient Queuosine Is Salvaged by Members of the DUF2419 Protein Family. ACS Chem Biol **9**: 1812–1825

Zhang J (2003) Evolution by gene duplication: an update. Trends Ecol Evol **18**: 292–298

Zhang X, Wang C, Zhang Y, Sun Y, Mou Z (2012) The Arabidopsis Mediator

Complex Subunit16 Positively Regulates Salicylate-Mediated Systemic Acquired Resistance and Jasmonate/Ethylene-Induced Defense Pathways. Plant Cell **24**: 4294–4309

Zheng Y, Huang Y, Xian W, Wang J, Liao H (2012) Identification and expression analysis of the *Glycine max* CYP707A gene family in response to drought and salt stresses. Ann Bot **110**: 743–756

Zhou H-R, Zhang F-F, Ma Z-Y, Huang H-W, Jiang L, Cai T, Zhu J-K, Zhang C, He X-J (2013) Folate Polyglutamylation Is Involved in Chromatin Silencing by Maintaining Global DNA Methylation and Histone H3K9 Dimethylation in Arabidopsis. Plant Cell **25**: 2545–2559

Table 1. Structural features among genes groups.

	Introns number	Intronless genes (%)	Alternative spliced genes (%)	Gene length (pb)	CDS length (pb)
U40	2	35	13	2,008	839
U500	2	29	20	2,111	864
D40	1	40	10	2,368	906
D500	2	30	14	2,456	1,058
Genome	3	16	22	3,053	987
S40	8	5	55	5,380	1,469
S500	6	3	42	4,926	1,349

Top 40 (all time points in both cultivars) and top 500 (at least one time point in both cultivars; except for S500 keeping all time points as criterion) up- (U), down- (D), stable-genes (S), and genome features were calculated using median, except when using percentage. Statistical significances from pairwise comparisons are provided in additional files 2 and 4.

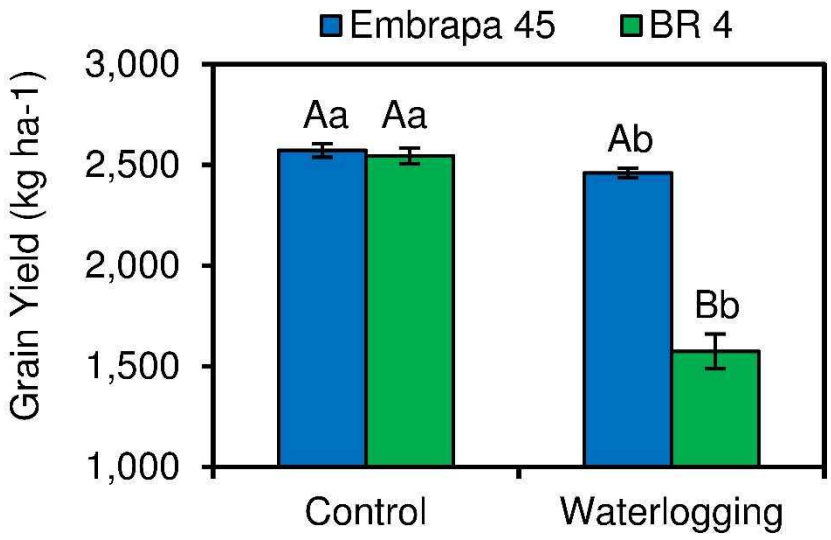

Embrapa 45
BR 4
Grain Yield (kg ha-1)
3,000
2,500
2,000
1,500
1,000
Aa
Aa
Ab
Bb
Control
Waterlogging

Figure 1

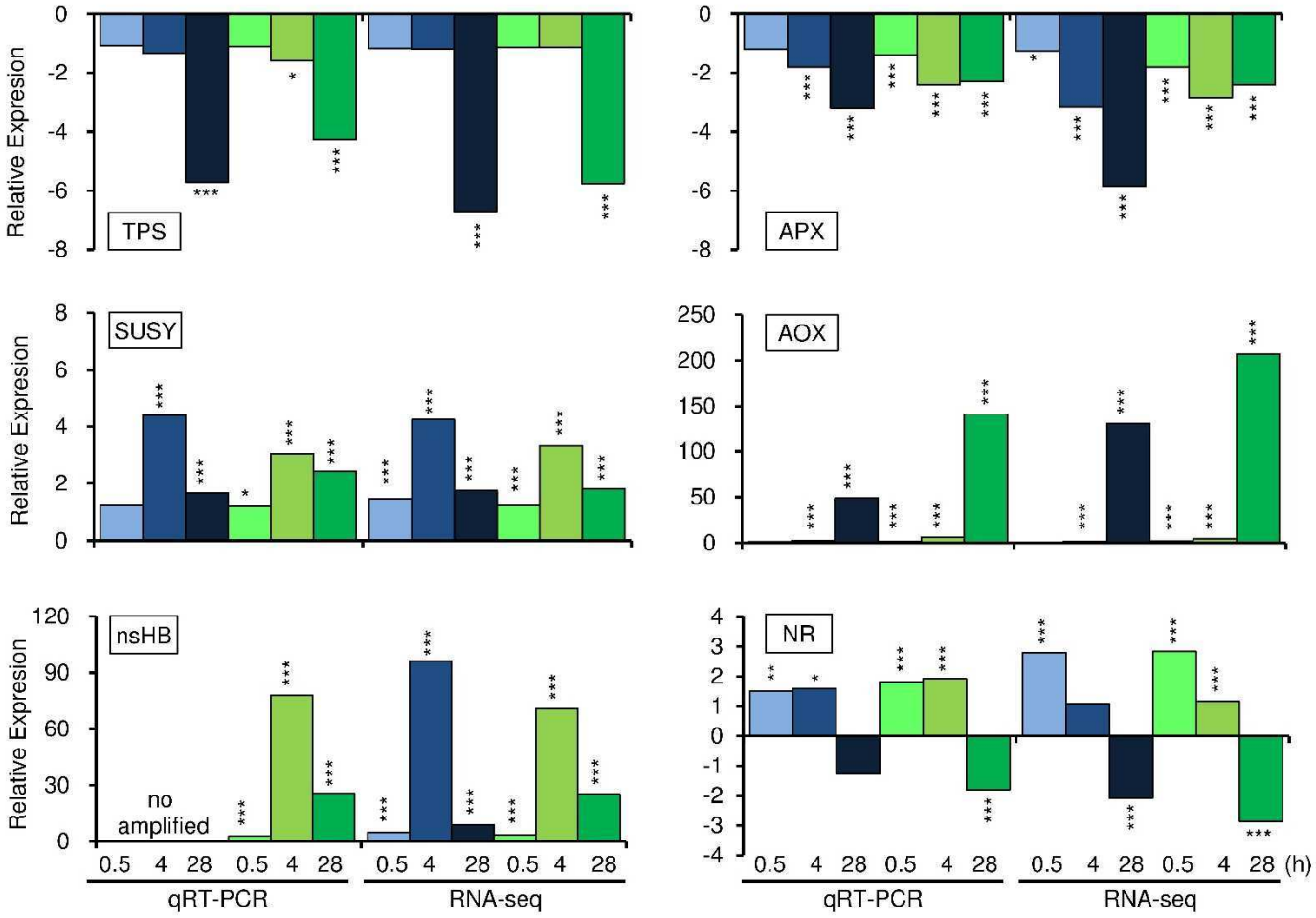
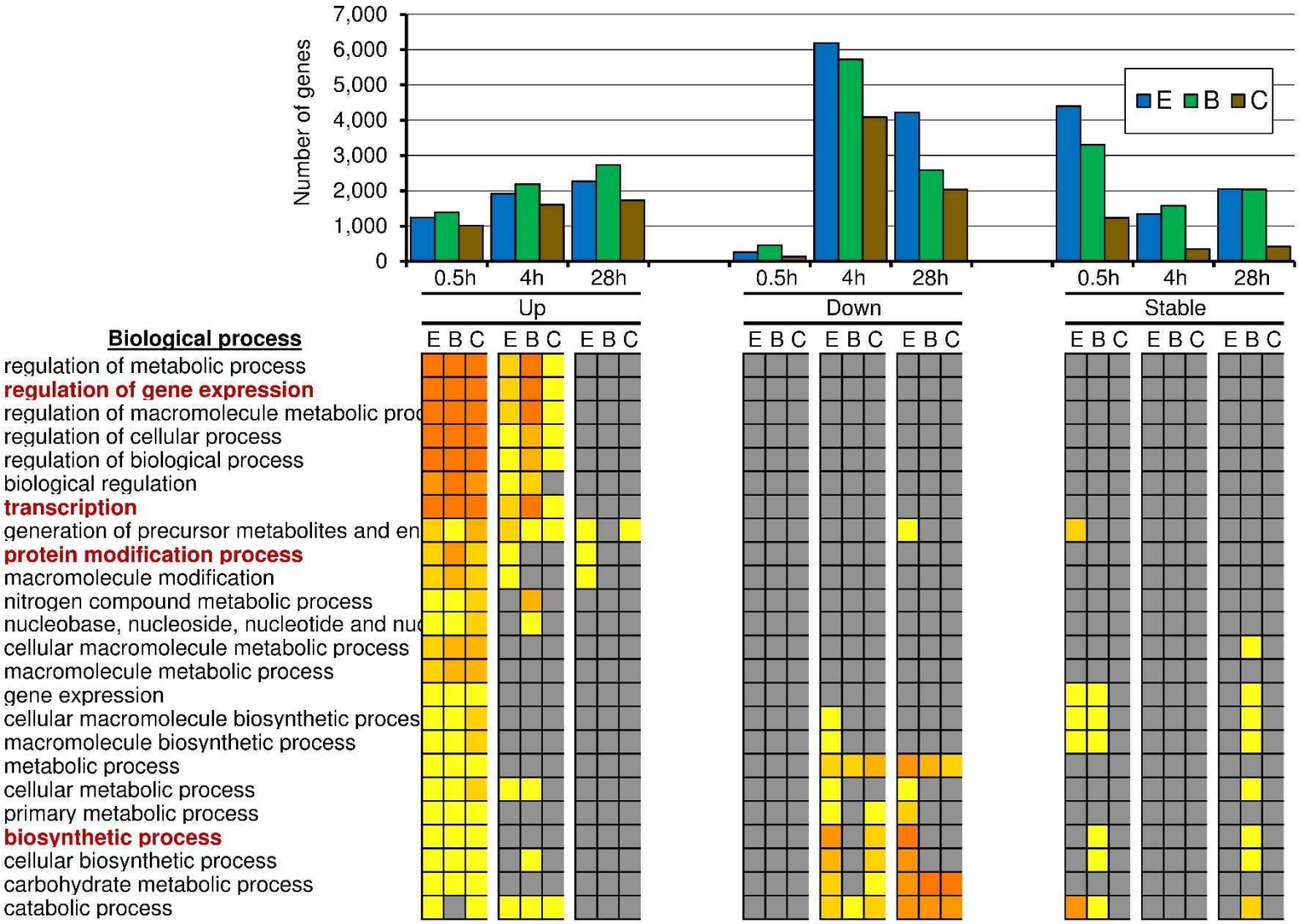

Relative Expresion
TPS
APX
SUSY
AOX
nsHB
NR
no amplified
0.5 4 28 0.5 4 28
qRT-PCR
RNA-seq
(h)

Figure 2


Number of genes
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
E B C
0.5h 4h 28h
Up
Down
Stable
Biological process
E B C
regulation of metabolic process
regulation of gene expression
regulation of macromolecule metabolic proc
regulation of cellular process
regulation of biological process
biological regulation
transcription
generation of precursor metabolites and en
protein modification process
macromolecule modification
nitrogen compound metabolic process
nucleobase, nucleoside, nucleotide and nuc
cellular macromolecule metabolic process
macromolecule metabolic process
gene expression
cellular macromolecule biosynthetic proces
macromolecule biosynthetic process
metabolic process
cellular metabolic process
primary metabolic process
biosynthetic process
cellular biosynthetic process
carbohydrate metabolic process
catabolic process

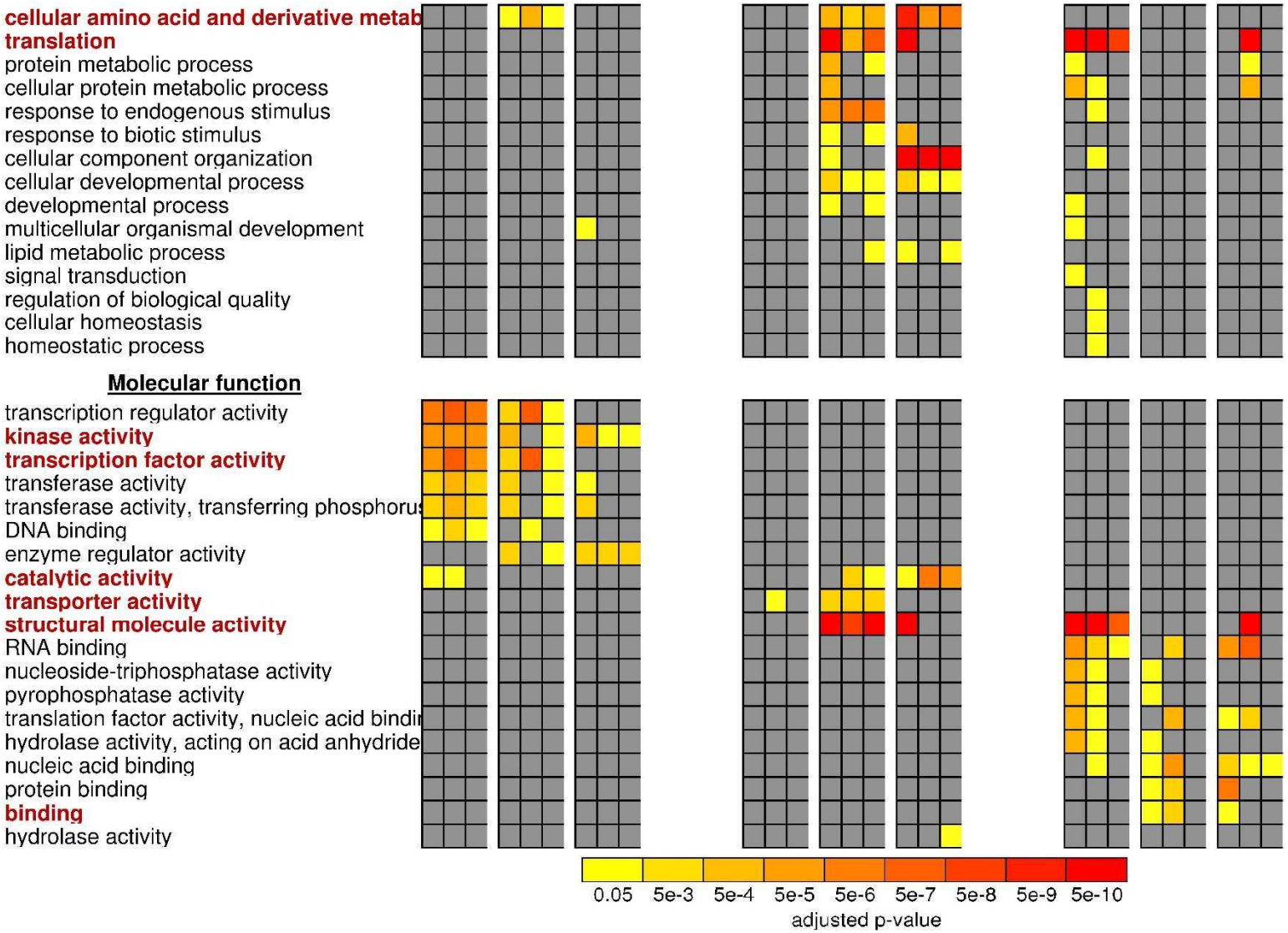

cellular amino acid and derivative metab
translation
protein metabolic process
cellular protein metabolic process
response to endogenous stimulus
response to biotic stimulus
cellular component organization
cellular developmental process
developmental process
multicellular organismal development
lipid metabolic process
signal transduction
regulation of biological quality
cellular homeostasis
homeostatic process
Molecular function
transcription regulator activity
kinase activity
transcription factor activity
transferase activity
transferase activity, transferring phosphoru
DNA binding
enzyme regulator activity
catalytic activity
transporter activity
structural molecule activity
RNA binding
nucleoside-triphosphatase activity
pyrophosphatase activity
translation factor activity, nucleic acid bindi
hydrolase activity, acting on acid anhydride
nucleic acid binding
protein binding
binding
hydrolase activity
0.05
5e-3
5e-4
5e-5
5e-6
5e-7
5e-8
5e-9
5e-10
adjusted p-value

Figure 3

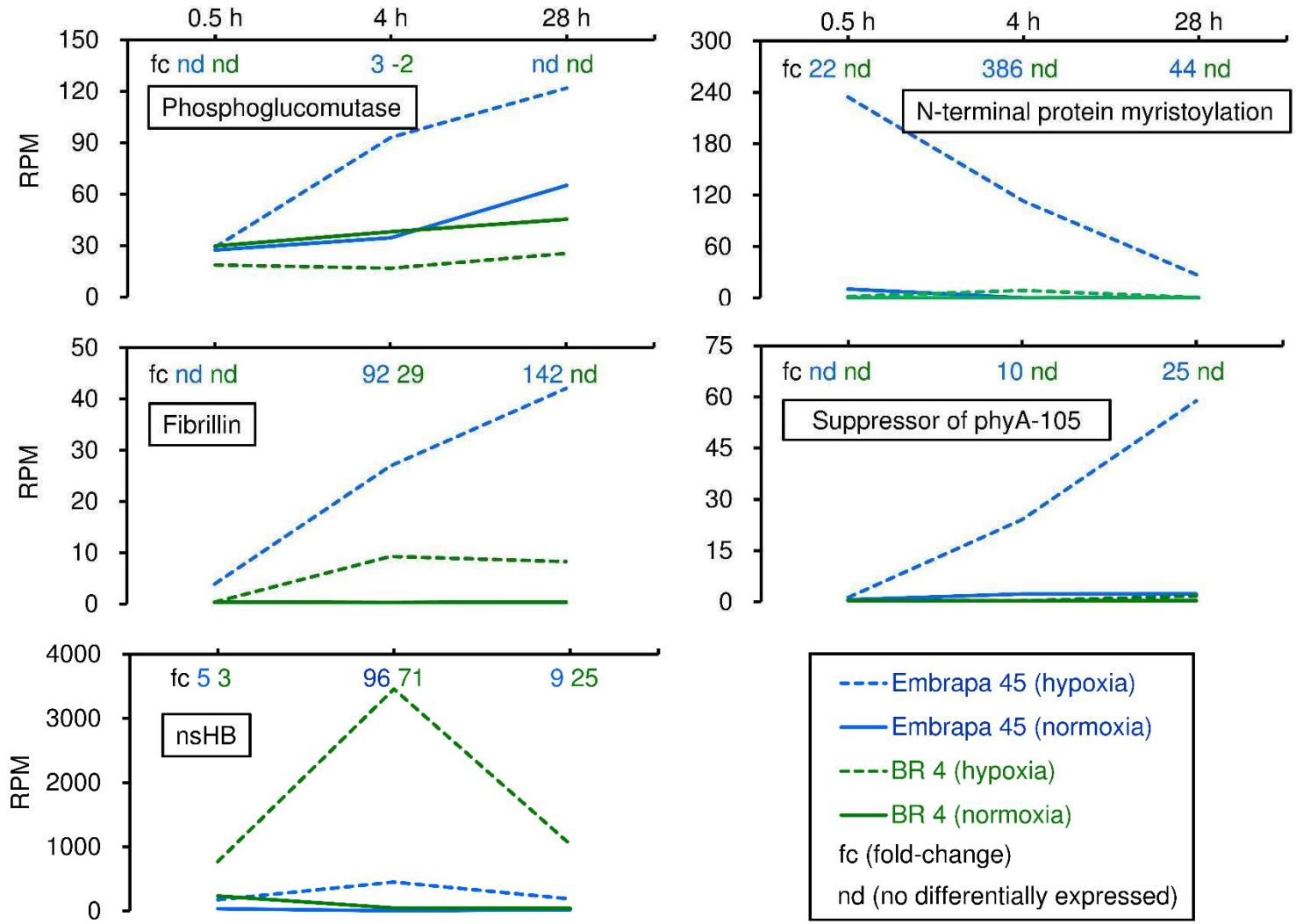

0.5 h
4 h
28 h
fc nd nd
3 -2
nd nd
Phosphoglucomutase
RPM
150
120
90
60
30
0
fc 22 nd
386 nd
44 nd
N-terminal protein myristoylation
300
240
180
120
60
0
fc nd nd
92 29
142 nd
Fibrillin
50
40
30
20
10
0
fc nd nd
10 nd
25 nd
Suppressor of phyA-105
75
60
45
30
15
0
fc 5 3
96 71
9 25
nsHB
4000
3000
2000
1000
0
Embrapa 45 (hypoxia)
Embrapa 45 (normoxia)
BR 4 (hypoxia)
BR 4 (normoxia)
fc (fold-change)
nd (no differentially expressed)

Figure 4

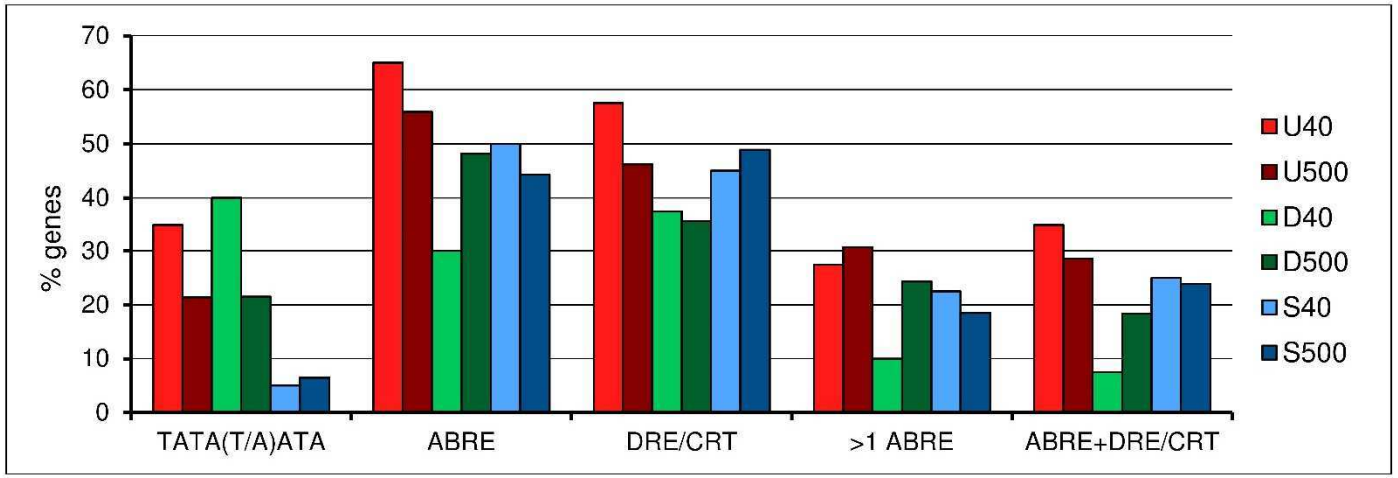
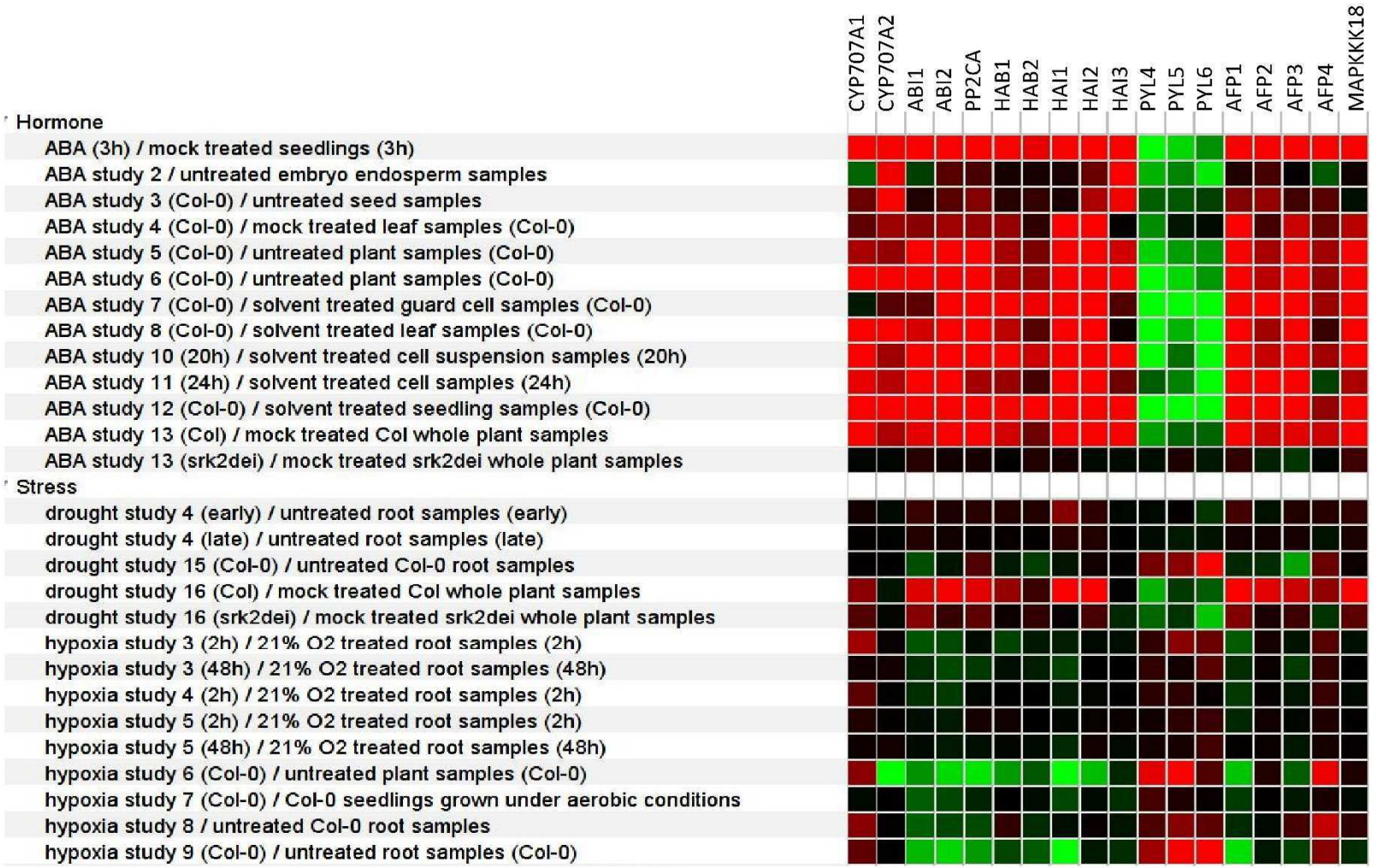

70
60
50
40
30
20
10
0
% genes
TATA(T/A)ATA
ABRE
DRE/CRT
>1 ABRE
ABRE+DRE/CRT
U40
U500
D40
D500
S40
S500

Figure 5


CYP707A1
CYP707A2
ABI1
ABI2
PP2CA
HAB1
HAB2
HAI1
HAI2
HAI3
PYL4
PYL5
PYL6
AFP1
AFP2
AFP3
AFP4
MAPKKK18
Hormone
ABA (3h) / mock treated seedlings (3h)
ABA study 2 / untreated embryo endosperm samples
ABA study 3 (Col-0) / untreated seed samples
ABA study 4 (Col-0) / mock treated leaf samples (Col-0)
ABA study 5 (Col-0) / untreated plant samples (Col-0)
ABA study 6 (Col-0) / untreated plant samples (Col-0)
ABA study 7 (Col-0) / solvent treated guard cell samples (Col-0)
ABA study 8 (Col-0) / solvent treated leaf samples (Col-0)
ABA study 10 (20h) / solvent treated cell suspension samples (20h)
ABA study 11 (24h) / solvent treated cell samples (24h)
ABA study 12 (Col-0) / solvent treated seedling samples (Col-0)
ABA study 13 (Col) / mock treated Col whole plant samples
ABA study 13 (srk2dei) / mock treated srk2dei whole plant samples
Stress
drought study 4 (early) / untreated root samples (early)
drought study 4 (late) / untreated root samples (late)
drought study 15 (Col-0) / untreated Col-0 root samples
drought study 16 (Col) / mock treated Col whole plant samples
drought study 16 (srk2dei) / mock treated srk2dei whole plant samples
hypoxia study 3 (2h) / 21% O2 treated root samples (2h)
hypoxia study 3 (48h) / 21% O2 treated root samples (48h)
hypoxia study 4 (2h) / 21% O2 treated root samples (2h)
hypoxia study 5 (2h) / 21% O2 treated root samples (2h)
hypoxia study 5 (48h) / 21% O2 treated root samples (48h)
hypoxia study 6 (Col-0) / untreated plant samples (Col-0)
hypoxia study 7 (Col-0) / Col-0 seedlings grown under aerobic conditions
hypoxia study 8 / untreated Col-0 root samples
hypoxia study 9 (Col-0) / untreated root samples (Col-0)

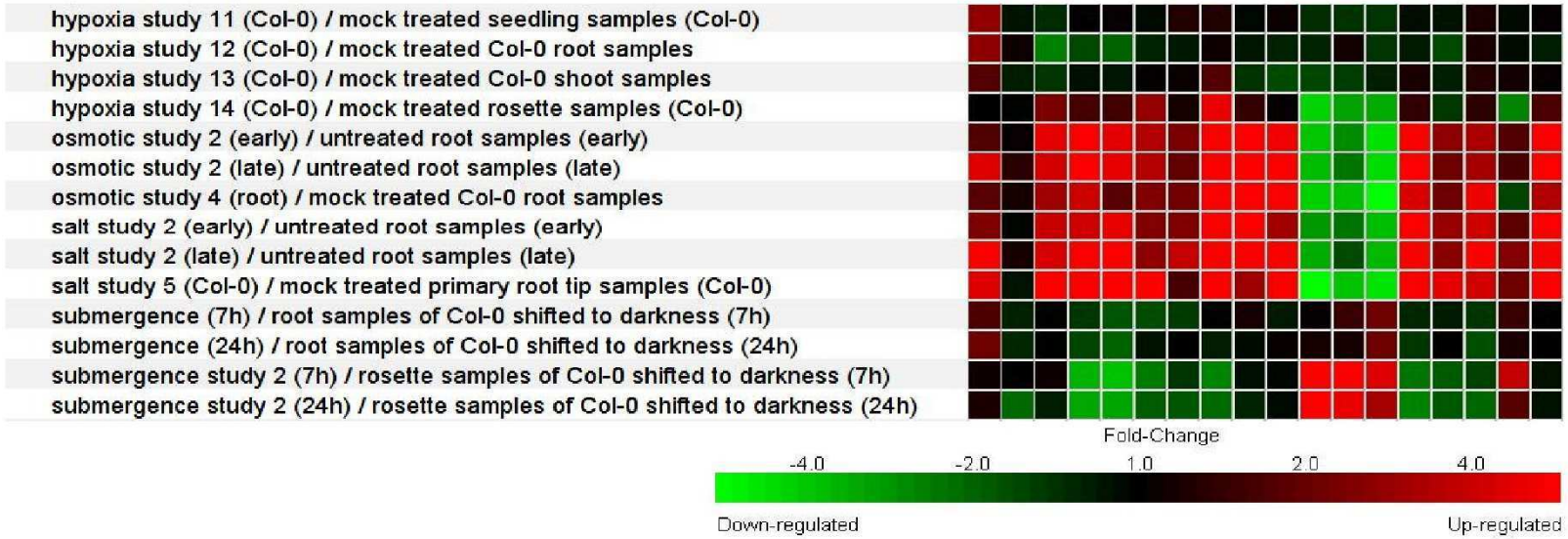

hypoxia study 11 (Col-0) / mock treated seedling samples (Col-0)
hypoxia study 12 (Col-0) / mock treated Col-0 root samples
hypoxia study 13 (Col-0) / mock treated Col-0 shoot samples
hypoxia study 14 (Col-0) / mock treated rosette samples (Col-0)
osmotic study 2 (early) / untreated root samples (early)
osmotic study 2 (late) / untreated root samples (late)
osmotic study 4 (root) / mock treated Col-0 root samples
salt study 2 (early) / untreated root samples (early)
salt study 2 (late) / untreated root samples (late)
salt study 5 (Col-0) / mock treated primary root tip samples (Col-0)
submergence (7h) / root samples of Col-0 shifted to darkness (7h)
submergence (24h) / root samples of Col-0 shifted to darkness (24h)
submergence study 2 (7h) / rosette samples of Col-0 shifted to darkness (7h)
submergence study 2 (24h) / rosette samples of Col-0 shifted to darkness (24h)
Fold-Change
-4.0
-2.0
1.0
2.0
4.0
Down-regulated
Up-regulated

Figure 6

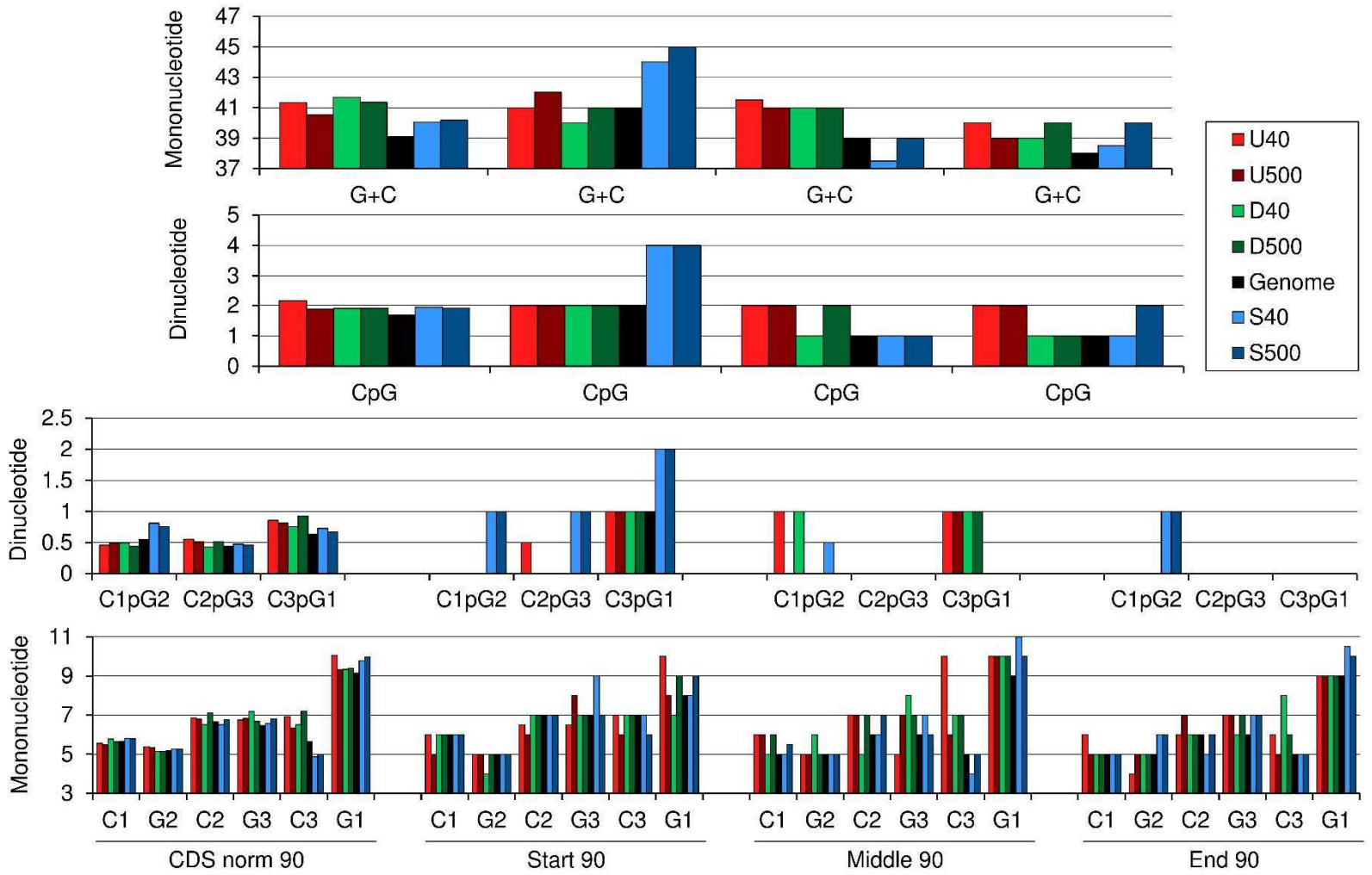

Mononucleotide
47
45
43
41
39
37
G+C
G+C
G+C
G+C
Dinucleotide
5
4
3
2
1
0
CpG
CpG
CpG
CpG
U40
U500
D40
D500
Genome
S40
S500
Dinucleotide
2.5
2
1.5
1
0.5
0
C1pG2 C2pG3 C3pG1
C1pG2 C2pG3 C3pG1
C1pG2 C2pG3 C3pG1
C1pG2 C2pG3 C3pG1
Mononucleotide
11
9
7
5
3
C1 G2 C2 G3 C3 G1
C1 G2 C2 G3 C3 G1
C1 G2 C2 G3 C3 G1
C1 G2 C2 G3 C3 G1
CDS norm 90
Start 90
Middle 90
End 90

Figure 7

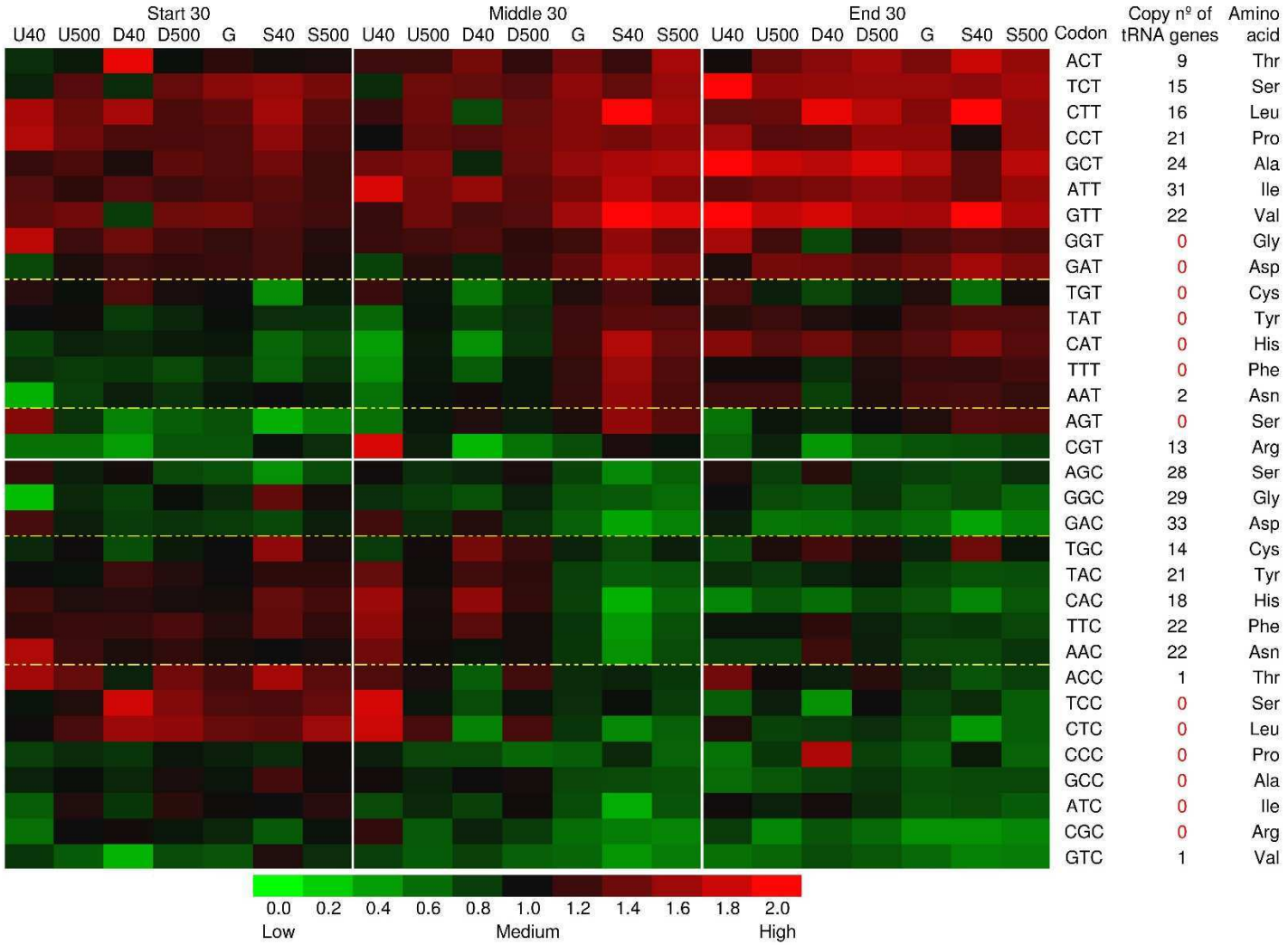

Start 30
Middle 30
End 30
U40 U500 D40 D500 G S40 S500
Codon
Copy nº of tRNA genes
Amino acid
ACT 9 Thr
TCT 15 Ser
CTT 16 Leu
CCT 21 Pro
GCT 24 Ala
ATT 31 Ile
GTT 22 Val
GGT 0 Gly
GAT 0 Asp
TGT 0 Cys
TAT 0 Tyr
CAT 0 His
TTT 0 Phe
AAT 2 Asn
AGT 0 Ser
CGT 13 Arg
AGC 28 Ser
GGC 29 Gly
GAC 33 Asp
TGC 14 Cys
TAC 21 Tyr
CAC 18 His
TTC 22 Phe
AAC 22 Asn
ACC 1 Thr
TCC 0 Ser
CTC 0 Leu
CCC 0 Pro
GCC 0 Ala
ATC 0 Ile
CGC 0 Arg
GTC 1 Val
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0
Low
Medium
High

Figure 8

Figure 1. Grain yields of two soybean cultivars under two moisture regimes in the soil. Control indicates well-watered conditions, 70% available water in the soil. n = 4, except for BR 4 under waterlogging (n = 3). Each repetition were consisted of 75 plants, from which grain yield were converted to 200,000 plants per hectare. Means values (± S.E.M.) followed by different capital letters between cultivars under same soil condition, and lowercase letters between soil conditions for same cultivar, significantly differ according to Tukey test 5%.

Figure 2. Validation of pairwise RNA-seq data through qRT-PCR. Six hypoxia-responsive genes (TPS: Trehalose-6-Phosphate Synthase, Glyma17g07530; APX: Ascorbate Peroxidase, Glyma12g07780; SUSY: Sucrose Synthase, Glyma13g17420; AOX: Alternative Oxidase, Glyma04g14800; nsHB: non-symbiotic Hemoglobin, Glyma11g12980; NR: Nitrate Reductase, Glyma06g11430) were analyzed in two soybean cultivars (Embrapa 45 = blue; BR 4 = green). The transcript abundance of the target genes from plants subjected to hypoxic conditions for different periods of time was compared with the respective controls (normoxic condition). Differential gene expression statistically significant: *P < 0.05, **P < 0.01, and ***P < 0.001. From qRT-PCR, raw data was normalized using the ELF1B and ACTB reference genes (Nakayama et al., 2014).

Figure 3. Number and GO enrichment analysis of up-, down-, and stable-regulated genes in Embrapa 45 (E), BR 4 (B), and in both soybean cultivars (C). Differentially expressed genes: fold-change ≥ 2 (up), ≤ -2 (down); adj. *p*-value ≤ 0.01; RPM ≥ 9 (control or stress datasets). Stable-regulated genes: fold-change ≤ 1.1 or ≥ -1.1; RPM ≥ 9 (control and stress datasets). RPM ≥ 9: at least 20 reads in all datasets. Red GO names are cited in the text.

Figure 4. Relative expression (fc) and expression level (in RPM) in two soybean cultivars along three time points. The genes analyzed were phosphoglucomutase (Glyma05g34790), N-terminal protein myristoylation (Glyma06g03430), fibrillin (Glyma10g32620), suppressor of phyA-105 (Glyma06g37080), and non-symbiotic hemoglobin (nsHB; Glyma11g12980). Differentially expressed genes: fold-change ≥ 2 (up), ≤ -2 (down); adj. *p*-value ≤

0.01; RPM ≥ 9 (control or stress datasets). RPM ≥ 9: at least 20 reads in all datasets.

Figure 5. Percentage of genes with which putative promoter region contain consensus TATA box, ABRE, and/or DRE/CRT motifs. Statistical significances from pairwise comparisons are provided in additional file 2.

Figure 6. Thale cress genes related with ABA metabolism and signaling dependent of SRK2D/E/I responding to diverse treatments (from Genevestigator microarray database; Hruz et al., 2008).

Figure 7. Median compositional features for full (normalized to 90 nucleotides), start, middle, and end 90 nucleotides length from different soybean genes groups. The groups are formed by top 40 and 500 ranked up- (U), down- (D), and stable-genes (S) to hypoxia, and genome. Phosphate bonds in dinucleotides are represented by “p”, and numbers are nucleotide position in codon. Statistical significance from pairwise comparisons are provided in additional file 4.

Figure 8. Heat map of RSCU (Relative Synonymous Codon Usage) for start, middle, and end 30 CDS (coding sequence) codons ending in pyrimidine (C and T) from different soybean genes groups. The groups are formed by top 40 and 500 ranked up- (U), down- (D), and stable-genes (S) to hypoxia, and genome (G). Asp, Cys, Tyr, His, Phe, and Asn are coding by 2-fold degenerate pyrimidine ending codons.

Additional file 1. Samples sizes used for structural and compositional analysis among groups of genes.

genes groups	promotor cis elements	gene structure	CDS			
			CpG		C1pG2, C2pG3, C3pG1	
			full	start 90, middle 90, end 90	full	start 90, middle 90, end 90
U40	40	40	40	34	40	34
U500	381	500	500	424	500	424
D40	40	40	40	37	40	37
D500	386	500	500	467	500	467
Genome	---	54174	54174	46644	10835	9329
S40	40	40	40	38	40	38
S500	352	500	500	490	500	490

genes groups	CDS				
	G+C		C1, C2, C3, G1, G2, G3		RSCU
	full	start 90, middle 90, end 90	full	start 90, middle 90, end 90	start 30, middle 30, end 30
U40	40	34	40	34	34
U500	500	424	500	424	424
D40	40	37	40	37	37
D500	500	467	500	467	467
Genome	54174	46644	10835	9329	46644
S40	40	38	40	38	38
S500	500	490	500	490	490

Additional file 2. Fisher's pairwise comparisons between genes groups.

	Promoters (%) with					Genes (%)	
	Consensus TATA-Box TATA(T/A)ATA	**P$ABRE**	**P$DREB**	**ABRE >1**	**P$ABRE+P$DREB**	**Intronless**	**Alternative splicing**
U40 vs D40	0.82	**0.00**	0.12	0.08	**0.01**	0.82	1.00
U40 vs S40	**0.00**	0.26	0.37	0.80	0.46	**0.00**	**0.00**
D40 vs S40	**0.00**	0.11	0.65	0.22	0.07	**0.00**	**0.00**
U500 vs D500	1.00	0.04	**0.00**	0.05	**0.00**	0.89	**0.01**
U500 vs S500	**0.00**	**0.00**	0.51	**0.00**	0.15	**0.00**	**0.00**
D500 vs S500	**0.00**	0.30	**0.00**	0.06	0.07	**0.00**	**0.00**

Additional file 3. Differentially expressed genes (DEGs) from hypoxic roots of Embrapa 45 (E) and BR 4 (B) soybean cultivars orthologs of thale cress genes related with ABA metabolism and signaling dependent of SRK2D/E/I (from Genevestigator microarray database; Hruz et al., 2008).

Glyma 1.1 ID	Top TAIR10 Blastp Hit	Ratio					
		E-0.5h	E-4h	E-28h	B-0.5h	B-4h	B-28h
Glyma09g07650	AT4G26080.1|Symbols:ABI1, AtABI1|Protein phosphatase 2C	0.0	-6.1	-2.5	0.0	-9.3	-3.7
Glyma13g16640	AT4G26080.1|Symbols:ABI1, AtABI1|Protein phosphatase 2C	0.0	2.2	2.1	0.0	0.0	3.0
Glyma12g10370	AT1G05100.1|Symbols:MAPKKK18|mitogen-activated protein	43.0	218.9	0.0	35.0	331.0	22.5
Glyma01g43460	AT2G29380.1|Symbols:HAI3|highly ABA-induced PP2C gene	62.8	485.1	42.2	41.1	930.0	42.0
Glyma02g41750	AT3G11410.1|Symbols:ATPP2CA, AHG3, PP2CA|protein pho	0.0	-7.5	0.0	0.0	-3.5	0.0
Glyma11g34410	AT3G11410.1|Symbols:ATPP2CA, AHG3, PP2CA|protein pho	0.0	0.0	2.3	2.8	0.0	2.3
Glyma14g07210	AT3G11410.1|Symbols:ATPP2CA, AHG3, PP2CA|protein pho	0.0	0.0	2.6	0.0	2.6	2.5
Glyma14g13020	AT1G72770.3|Symbols:HAB1|homology to ABI1|chr1:273909	0.0	-2.1	0.0	0.0	-3.4	0.0
Glyma04g05661	AT1G72770.3|Symbols:HAB1|homology to ABI1|chr1:273909	0.0	-2.0	0.0	-3.1	0.0	0.0
Glyma17g33410	AT1G72770.3|Symbols:HAB1|homology to ABI1|chr1:273909	0.0	0.0	0.0	0.0	-2.2	0.0
Glyma01g02290	AT5G46790.1|Symbols:PYL1, RCAR12|PYR1-like 1|chr5:189	0.0	-3.5	-3.4	0.0	0.0	0.0
Glyma08g27010	AT1G13740.1|Symbols:AFP2|ABI five binding protein 2|chr1:4	0.0	0.0	0.0	0.0	-3.6	0.0
Glyma02g13120	AT1G13740.1|Symbols:AFP2|ABI five binding protein 2|chr1:4	0.0	0.0	2.4	0.0	0.0	2.8
Glyma01g07390	AT1G13740.1|Symbols:AFP2|ABI five binding protein 2|chr1:4	0.0	2.5	3.4	0.0	2.4	4.4
Glyma09g35250	AT4G19230.1|Symbols:CYP707A1|cytochrome P450, family 7	3.1	0.0	2.2	4.6	0.0	5.2
Glyma01g35660	AT4G19230.1|Symbols:CYP707A1|cytochrome P450, family 7	6.5	4.1	5.4	8.3	4.6	7.5
Glyma16g20490	AT4G19230.1|Symbols:CYP707A1|cytochrome P450, family 7	0.0	19.6	142.1	0.0	0.0	19.9

Glyma11g35670	AT5G05440.1|Symbols:PYL5, RCAR8|Polyketide cyclase/deh	0.0	0.0	-2.0	-2.0	0.0	-2.1
Glyma18g43676	AT2G38310.1|Symbols:PYL4, RCAR10|PYR1-like 4|chr2:1605	0.0	0.0	-2.7	0.0	0.0	0.0
Glyma07g19120	AT2G38310.1|Symbols:PYL4, RCAR10|PYR1-like 4|chr2:1605	-2.3	0.0	0.0	0.0	0.0	0.0
Glyma02g42990	AT5G05440.1|Symbols:PYL5, RCAR8|Polyketide cyclase/deh	0.0	0.0	-2.0	-2.6	2.4	0.0
Glyma14g06100	AT5G05440.1|Symbols:PYL5, RCAR8|Polyketide cyclase/deh	0.0	3.0	0.0	0.0	3.9	0.0

DEGs: fold-change ≥ 2 (up), ≤ -2 (down); adj. *p* -value ≤ 0.01; RPM ≥ 9 (control or stress datasets). RPM ≥ 9: at least 20 reads in all datasets.

Additional file 4. Mann-Whitney pairwise comparisons.

	Introns number	**Gene length (pb)**	**CDS length (pb)**
U40 vs D40	0.04	0.74	0.26
S40 vs U40	**0.00**	**0.00**	**0.00**
S40 vs D40	**0.00**	**0.00**	**0.01**
U500 vs D500	0.46	0.02	**0.00**
S500 vs U500	**0.00**	**0.00**	**0.00**
S500 vs D500	**0.00**	**0.00**	**0.00**

	CDS G+C norm 90	**Start 90 G+C**	**Middle 90 G+C**	**End 90 G+C**
U40 vs D40	0.71	0.15	0.34	0.27
S40 vs U40	0.01	0.27	**0.00**	0.69
S40 vs D40	0.08	0.04	0.05	0.51
U500 vs D500	**0.00**	0.54	0.02	0.02
S500 vs U500	0.09	**0.00**	**0.00**	**0.00**
S500 vs D500	**0.00**	**0.00**	**0.00**	0.55

	CDS CpG norm 90	**Start 90 CpG**	**Middle 90 CpG**	**End 90 CpG**
U40 vs D40	0.53	0.65	0.82	0.10
S40 vs U40	0.98	**0.00**	0.16	0.45
S40 vs D40	0.53	**0.00**	0.62	0.29
U500 vs D500	0.58	**0.00**	0.51	0.13
S500 vs U500	0.09	**0.00**	**0.00**	0.93
S500 vs D500	0.35	**0.00**	**0.00**	0.10

	CDS norm 90			Start 90			Middle 90			End 90		
	C1pG2	**C2pG3**	**C3pG1**	**C1pG2**	**C2pG3**	**C3pG1**	**C1pG2**	**C2pG3**	**C3pG1**	**C1pG2**	**C2pG3**	**C3pG1**
U40 vs D40	0.89	0.59	0.54	0.76	0.31	0.50	0.66	0.14	0.43	0.38	0.26	0.75
S40 vs U40	**0.00**	0.23	0.18	**0.00**	0.10	0.11	0.52	0.96	0.03	0.19	0.03	0.53
S40 vs D40	**0.01**	0.93	0.50	**0.00**	0.02	0.21	0.88	0.11	0.17	0.03	0.20	0.82
U500 vs D500	0.02	0.99	0.06	0.06	0.04	0.20	**0.00**	0.23	0.99	0.10	0.11	0.50
S500 vs U500	**0.00**	0.06	0.02	**0.00**	**0.00**	**0.00**	0.86	**0.00**	**0.00**	**0.00**	0.01	0.04
S500 vs D500	**0.00**	0.08	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.34	**0.01**

	Start 90 G+C vs Middle 90 G+C	Start 90 G+C vs End 90 G+C	Middle 90 G+C vs End 90 G+C
S40	**0.00**	**0.01**	0.22
D40	0.77	0.22	0.11
U40	0.34	0.02	**0.01**
S500	**0.00**	**0.00**	**0.01**
D500	0.34	**0.00**	**0.00**
U500	0.04	**0.00**	**0.00**

	Start 90 CpG vs Middle 90 CpG	Start 90 CpG vs End 90 CpG	Middle 90 CpG vs End 90 CpG
S40	**0.00**	**0.00**	0.60
D40	0.61	0.19	0.12
U40	0.61	0.65	0.22
S500	**0.00**	**0.00**	0.03
D500	0.09	0.05	**0.00**
U500	0.59	**0.00**	**0.00**

	CDS norm 90						First 90					
	C1	**G2**	**C2**	**G3**	**C3**	**G1**	**C1**	**G2**	**C2**	**G3**	**C3**	**G1**
U40 vs D40	0.72	0.62	0.84	0.75	0.64	0.54	0.53	0.52	0.98	0.57	0.97	0.06
S40 vs U40	0.35	0.77	0.70	0.10	**0.00**	0.73	0.49	0.53	0.77	0.05	0.73	0.13
S40 vs D40	0.62	0.95	0.70	0.04	**0.00**	0.31	0.98	0.27	0.92	0.07	0.87	0.46
U500 vs D500	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	0.77	0.14	0.85	0.02	0.08	**0.00**	0.10
S500 vs U500	**0.00**	0.22	0.75	0.23	**0.00**	**0.00**	0.45	0.82	0.03	0.90	0.28	0.15
S500 vs D500	0.13	0.05	**0.00**	0.07	**0.00**	**0.00**	0.48	0.66	0.82	0.08	0.05	0.83

	Middle 90						End 90					
	C1	**G2**	**C2**	**G3**	**C3**	**G1**	**C1**	**G2**	**C2**	**G3**	**C3**	**G1**
U40 vs D40	**0.01**	0.16	0.04	**0.00**	**0.01**	0.68	0.38	0.39	0.84	0.14	0.07	0.06
S40 vs U40	0.08	0.63	0.05	0.14	**0.00**	0.20	0.18	0.01	0.49	0.37	0.01	0.04
S40 vs D40	0.35	0.26	0.95	0.02	**0.00**	0.35	0.61	0.23	0.33	0.58	**0.00**	**0.00**
U500 vs D500	0.84	0.29	0.05	0.50	0.10	0.87	0.62	0.68	0.77	0.76	0.02	0.89
S500 vs U500	0.13	0.77	0.42	0.56	**0.00**	**0.01**	0.42	**0.00**	0.06	1.00	**0.00**	**0.00**
S500 vs D500	0.18	0.44	0.25	0.22	**0.00**	0.01	0.73	**0.00**	0.03	0.74	**0.00**	**0.00**

Additional file 5. Differentially expressed genes (DEGs) related with modification of I34, Q34, yW37, pseudouridine, and methylation (tRNA and DNA methyltransferases), as well as aminoacylation of tRNAs (amino acid charging) in hypoxic roots of Embrapa 45 (E) and BR 4 (B) soybean cultivars.

Glyma 1.1 ID	Top TAIR10 Blastp Hit	Ratio					
		E-0.5h	E-4h	E-28h	B-0.5h	B-4h	B-28h
Glyma08g06900	AT1G58290.1|Symbols:HEMA1|Glutamyl-tRNA reductase family pr	3.0	0.0	0.0	5.0	0.0	0.0
Glyma17g34070	AT4G31180.2|Symbols:|Class II aminoacyl-tRNA and biotin synthet	2.6	4.3	0.0	2.7	3.6	0.0
Glyma14g11711	AT4G31180.2|Symbols:|Class II aminoacyl-tRNA and biotin synthet	9.1	10.2	5.2	6.7	16.0	5.7
Glyma11g20660	AT2G28450.1|Symbols:|zinc finger (CCCH-type) family protein|chr2	0.0	-2.5	0.0	0.0	-3.0	0.0
Glyma15g06160	AT3G57150.1|Symbols:NAP57, AtNAP57, CBF5, AtCBF5|homolog	0.0	-3.0	-2.2	0.0	-3.0	0.0
Glyma08g18860	AT5G14460.1|Symbols:|Pseudouridine synthase family protein|chr5	0.0	-3.0	0.0	0.0	-3.3	0.0
Glyma20g23060	AT1G56345.1|Symbols:|Pseudouridine synthase family protein|chr1	3.1	4.9	3.6	6.2	10.7	6.4
Glyma0448s00210	AT2G30320.1|Symbols:|Pseudouridine synthase family protein|chr2	0.0	14.8	3.8	0.0	22.3	9.0
Glyma11g36620	AT4G13940.1|Symbols:HOG1, EMB1395, SAHH1, MEE58, ATSAH	0.0	-13.5	-32.1	0.0	-9.9	-32.2
Glyma05g28480	AT4G13940.1|Symbols:HOG1, EMB1395, SAHH1, MEE58, ATSAH	0.0	-6.9	-4.1	0.0	-3.8	-2.3
Glyma08g11480	AT4G13940.1|Symbols:HOG1, EMB1395, SAHH1, MEE58, ATSAH	0.0	-5.6	-4.0	0.0	-4.4	-2.5
Glyma20g22170	AT5G26710.1|Symbols:|Glutamyl/glutaminyl-tRNA synthetase, clas	3.0	5.5	0.0	3.1	7.7	2.8
Glyma02g03750	AT5G26830.1|Symbols:|Threonyl-tRNA synthetase|chr5:9437351-9	0.0	0.0	-2.2	0.0	-2.9	0.0
Glyma17g02440	AT5G18860.1|Symbols:|inosine-uridine preferring nucleoside hydro	0.0	-2.1	-4.4	0.0	-3.0	-3.3
Glyma0048s00280	AT1G05620.2|Symbols:URH2|uridine-ribohydrolase 2|chr1:167960	0.0	-2.4	0.0	0.0	-2.1	0.0
Glyma01g00960	AT4G26300.1|Symbols:emb1027|Arginyl-tRNA synthetase, class Ic	0.0	0.0	-2.3	0.0	-2.6	0.0
Glyma01g00940	AT4G26300.1|Symbols:emb1027|Arginyl-tRNA synthetase, class Ic	0.0	10.3	4.9	0.0	7.8	3.8
Glyma01g01120	AT1G69770.1|Symbols:CMT3|chromomethylase 3|chr1:26248496-2	0.0	0.0	-6.0	0.0	0.0	-33.1
Glyma04g36150	AT5G49160.1|Symbols:MET1, MET2, METI, DDM2, DMT01, DMT1	0.0	0.0	-33.0	0.0	0.0	0.0
Glyma06g18790	AT5G49160.1|Symbols:MET1, MET2, METI, DDM2, DMT01, DMT1	0.0	-4.6	-4.6	0.0	0.0	-9.5
Glyma19g00250	AT5G14620.1|Symbols:DRM2, DMT7|domains rearranged methyltr	0.0	-2.4	0.0	0.0	-2.5	0.0
Glyma07g39030	AT4G04670.1|Symbols:|Met-10+ like family protein / kelch repeat-c	0.0	-4.3	-2.8	0.0	0.0	0.0

Glyma01g38150	AT5G66750.1|Symbols:DDM1, CHR01, CHR1, CHA1, SOM4, S	0.0	0.0	-11.8	0.0	0.0	-5.1
Glyma11g07220	AT5G66750.1|Symbols:DDM1, CHR01, CHR1, CHA1, SOM4, S	0.0	0.0	-3.0	0.0	0.0	0.0
Glyma10g41920	AT1G25350.1|Symbols:OVA9|glutamine-tRNA ligase, putative / g	0.0	0.0	2.6	0.0	0.0	2.3
Glyma20g25120	AT1G25350.1|Symbols:OVA9|glutamine-tRNA ligase, putative / g	0.0	3.1	2.2	0.0	5.6	3.1
Glyma05g35021	AT3G02760.1|Symbols:|Class II aaRS and biotin synthetases sup	0.0	-3.3	0.0	0.0	-3.1	0.0
Glyma01g41660	AT1G09620.1|Symbols:|ATP binding;leucine-tRNA ligases;amino	0.0	0.0	2.3	0.0	0.0	2.0
Glyma20g25160	AT5G38830.1|Symbols:|Cysteinyl-tRNA synthetase, class Ia famil	0.0	-2.1	0.0	0.0	-2.9	0.0
Glyma02g05280	AT5G66180.1|Symbols:|S-adenosyl-L-methionine-dependent me	0.0	-4.0	0.0	0.0	-3.5	0.0
Glyma11g05250	NA	0.0	0.0	-5.0	0.0	0.0	0.0
Glyma13g35870	AT4G27340.1|Symbols:|Met-10+ like family protein|chr4:136873(	0.0	0.0	2.6	0.0	0.0	2.2
Glyma04g10706	NA	0.0	4.8	0.0	0.0	3.0	0.0

DEGs: fold-change ≥ 2 (up), ≤ -2 (down); adj. *p* -value ≤ 0.01; RPM ≥ 9 (control or stress datasets). RPM ≥ 9: at least 20 reads in all datasets.

Additional file 6. Differentially expressed genes (DEGs) related with endoplasmic reticulum stress [unfolded protein response (UPR) and ER stress-induced plant-specific cell death signaling pathways (Silva et al., 2015)] analyzed in hypoxic soybean roots of Embrapa 45 (E) and BR 4 (B) soybean cultivars.

Glyma 1.1 ID	Top TAIR10 Blastp Hit	Ratio					
		E-0.5h	E-4h	E-28h	B-0.5h	B-4h	B-28h
UPR							
Glyma02g19754	AT1G42990.1|Symbols:ATBZIP60, BZIP60|basic region/leuc	0.0	0.0	2.7	0.0	0.0	3.2
Glyma15g05050	AT1G07960.1|Symbols:ATPDIL5-1, PDIL5-1|PDI-like 5-1|ch	0.0	-2.6	0.0	0.0	-2.5	0.0
Glyma10g36170	AT1G35620.1|Symbols:ATPDIL5-2, ATPDI8, PDI8, PDIL5-2	0.0	-2.3	-2.2	0.0	-2.1	0.0
Glyma20g02951	AT5G56360.1|Symbols:PSL4|calmodulin-binding protein|chr	0.0	-2.6	-2.1	0.0	0.0	0.0
Glyma09g04290	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	-7.1	-8.0	0.0	-21.3	-3.4
Glyma15g15330	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	-8.8	-4.1	0.0	-5.6	-4.7
Glyma17g03540	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	-2.4	0.0	0.0	-3.1	-2.2
Glyma17g03520	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	-2.7	-3.6	0.0	0.0	0.0
Glyma07g37070	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	-2.2	0.0	0.0	0.0	0.0
Glyma07g37080	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	0.0	0.0	0.0	0.0	2.4
Glyma03g39110	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	0.0	2.1	0.0	0.0	2.3	2.2
Glyma10g28910	AT4G02080.1|Symbols:ASAR1, ATSARA1C, ATSAR2, SAI	12.7	3.3	0.0	7.2	8.2	0.0
Glyma14g36840	AT4G35580.2|Symbols:NTL9|NAC transcription factor-like 9	0.0	0.0	0.0	0.0	0.0	2.3
Glyma04g40450	AT4G35580.1|Symbols:NTL9|NAC transcription factor-like 9	2.5	0.0	0.0	3.7	0.0	0.0
Glyma06g14290	AT4G35580.1|Symbols:NTL9|NAC transcription factor-like 9	3.0	3.5	2.6	3.1	5.3	4.0
Glyma04g38630	AT5G63840.1|Symbols:RSW3, PSL5|Glycosyl hydrolases fa	0.0	-2.4	0.0	0.0	-2.7	0.0
Glyma17g37820	AT4G24190.1|Symbols:SHD, HSP90.7, AtHsp90.7, AtHsp9	0.0	-4.3	-4.2	-2.1	-3.4	-2.8
Glyma14g40320	AT4G24190.1|Symbols:SHD, HSP90.7, AtHsp90.7, AtHsp9	0.0	-2.8	-3.2	0.0	-3.0	0.0
Glyma19g41690	AT2G47470.1|Symbols:ATPDIL2-1, UNE5, MEE30, PDI11,	0.0	0.0	-7.6	0.0	0.0	-14.3
Glyma03g39130	AT2G47470.1|Symbols:ATPDIL2-1, UNE5, MEE30, PDI11,	0.0	0.0	-6.8	0.0	0.0	0.0
Glyma02g01750	AT2G47470.1|Symbols:ATPDIL2-1, UNE5, MEE30, PDI11,	0.0	0.0	-2.1	0.0	0.0	0.0

Glyma10g42180	AT3G02540.1|Symbols:RAD23-3, RAD23C|Rad23 UV excis	0.0	0.0	0.0	0.0	0.0	2.2
Glyma17g03610	AT3G17000.1|Symbols:UBC32|ubiquitin-conjugating enzyme	0.0	0.0	0.0	0.0	2.4	0.0
Glyma06g15840	AT5G09330.4|Symbols:VNI1|NAC domain containing protein	0.0	0.0	0.0	0.0	-2.1	0.0
Glyma05g32470	AT5G09330.4|Symbols:VNI1|NAC domain containing protein	0.0	0.0	2.5	0.0	0.0	0.0
Glyma04g39140	AT5G09330.4|Symbols:VNI1|NAC domain containing protein	0.0	0.0	3.0	0.0	0.0	0.0
Glyma08g16630	AT5G64060.1|Symbols:anac103, NAC103|NAC domain con	0.0	0.0	3.3	0.0	0.0	2.8
Glyma19g30681	AT2G40950.1|Symbols:BZIP17|Basic-leucine zipper (bZIP) t	0.0	0.0	2.0	0.0	0.0	0.0
Glyma12g31620	AT5G22060.1|Symbols:ATJ2, J2|DNAJ homologue 2|chr5:7	0.0	-3.0	-2.0	0.0	-2.5	0.0
Glyma11g17930	AT5G22060.1|Symbols:ATJ2, J2|DNAJ homologue 2|chr5:7	0.0	0.0	0.0	0.0	0.0	2.1
Glyma03g27030	AT5G22060.1|Symbols:ATJ2, J2|DNAJ homologue 2|chr5:7	0.0	0.0	2.4	0.0	0.0	0.0
Glyma12g10150	AT5G22060.1|Symbols:ATJ2, J2|DNAJ homologue 2|chr5:7	0.0	0.0	0.0	0.0	0.0	2.7
Glyma07g14540	AT5G22060.1|Symbols:ATJ2, J2|DNAJ homologue 2|chr5:7	0.0	2.4	2.5	0.0	0.0	0.0
Glyma14g05520	AT1G04980.1|Symbols:ATPDIL2-2, ATPDI10, PDI10, PDIL2	0.0	0.0	-2.2	0.0	-2.0	0.0
Glyma02g43460	AT1G04980.1|Symbols:ATPDIL2-2, ATPDI10, PDI10, PDIL2	0.0	-2.0	0.0	0.0	0.0	0.0
Glyma04g42690	AT1G77510.1|Symbols:ATPDIL1-2, PDI6, ATPDI6, PDIL1-2	0.0	-3.0	-2.1	0.0	-3.4	0.0
Glyma06g12090	AT1G77510.1|Symbols:ATPDIL1-2, PDI6, ATPDI6, PDIL1-2	2.4	6.2	4.0	2.0	8.0	7.4
Glyma10g28890	AT1G09210.1|Symbols:CRT1b, AtCRT1b|calreticulin 1b|chr1	0.0	-2.7	-5.0	0.0	-2.6	-2.4
Glyma20g23080	AT1G09210.1|Symbols:CRT1b, AtCRT1b|calreticulin 1b|chr1	0.0	-2.8	-3.1	0.0	-2.3	0.0
Glyma12g05460	AT1G08450.1|Symbols:CRT3, PSL1, EBS2, AtCRT3|calreti	0.0	-4.3	-3.4	0.0	-4.8	0.0
Glyma11g13460	AT1G08450.1|Symbols:CRT3, PSL1, EBS2, AtCRT3|calreti	0.0	0.0	-2.1	0.0	-2.4	0.0
Glyma06g17060	AT5G61790.1|Symbols:CNX1, ATCNX1|calnexin 1|chr5:248	0.0	-4.1	-5.1	0.0	-4.4	-2.7
Glyma04g38000	AT5G61790.1|Symbols:CNX1, ATCNX1|calnexin 1|chr5:248	0.0	-4.5	-4.5	0.0	-3.3	-3.2
Glyma08g00920	AT5G61790.1|Symbols:CNX1, ATCNX1|calnexin 1|chr5:248	0.0	-2.8	-3.6	0.0	-2.6	0.0
Glyma05g33330	AT5G61790.1|Symbols:CNX1, ATCNX1|calnexin 1|chr5:248	0.0	-2.5	0.0	0.0	-2.2	0.0
Glyma05g36600	AT5G42020.1|Symbols:BIP, BIP2|Heat shock protein 70 (Hs	0.0	-4.9	-9.7	0.0	-5.4	-6.0
Glyma08g02940	AT5G42020.1|Symbols:BIP, BIP2|Heat shock protein 70 (Hs	0.0	-4.5	-4.1	-2.4	-3.8	-2.6
Glyma08g02960	AT5G42020.1|Symbols:BIP, BIP2|Heat shock protein 70 (Hs	0.0	-4.9	-2.4	0.0	-4.8	-2.3
Glyma05g36620	AT5G42020.1|Symbols:BIP, BIP2|Heat shock protein 70 (Hs	0.0	-3.5	-3.3	0.0	-4.1	-2.0

Glyma08g20020	AT1G71220.2|Symbols:EBS1|UDP-glucose:glycoprotein glu	0.0	-2.4	0.0	0.0	-2.5	0.0
Glyma20g27980	AT1G61790.1|Symbols:|Oligosaccharyltransferase complex/	0.0	-2.1	0.0	0.0	-2.1	0.0
Glyma10g39750	AT1G61790.1|Symbols:|Oligosaccharyltransferase complex/	0.0	-2.3	0.0	0.0	0.0	0.0
Glyma12g07260	AT5G60640.1|Symbols:ATPDIL1-4, PDI2, ATPDI2, PDIL1-4	0.0	-2.8	0.0	0.0	-3.2	0.0
Glyma11g20630	AT5G60640.1|Symbols:ATPDIL1-4, PDI2, ATPDI2, PDIL1-4	0.0	-2.3	0.0	0.0	-3.6	0.0
Glyma13g40130	AT5G60640.1|Symbols:ATPDIL1-4, PDI2, ATPDI2, PDIL1-4	0.0	-2.3	0.0	0.0	-2.3	0.0
Glyma13g27600	AT4G04860.1|Symbols:DER2.2|DERLIN-2.2|chr4:2460135-	0.0	-2.3	0.0	0.0	0.0	0.0
Glyma16g01680	AT1G17280.2|Symbols:UBC34|ubiquitin-conjugating enzyme	0.0	0.0	0.0	0.0	-2.0	0.0
Glyma08g05030	AT2G21270.2|Symbols:UFD1|ubiquitin fusion degradation 1	0.0	0.0	0.0	0.0	0.0	2.1
Glyma15g04140	AT2G29070.2|Symbols:|Ubiquitin fusion degradation UFD1 f	0.0	0.0	2.7	0.0	0.0	3.0
Glyma14g05050	AT4G15420.1|Symbols:|Ubiquitin fusion degradation UFD1 f	0.0	2.9	0.0	0.0	4.8	0.0
Glyma19g03370	AT5G15400.1|Symbols:|U-box domain-containing protein|chr	0.0	-2.5	0.0	0.0	-2.5	0.0
Glyma08g16780	AT2G31955.2|Symbols:CNX2|cofactor of nitrate reductase a	0.0	-3.7	0.0	0.0	-14.2	0.0
Glyma20g23590	AT1G01290.2|Symbols:CNX3|cofactor of nitrate reductase a	0.0	-2.3	-3.1	0.0	-3.8	0.0
Glyma15g42290	AT2G31955.2|Symbols:CNX2|cofactor of nitrate reductase a	0.0	0.0	-3.1	0.0	0.0	-4.3
Glyma06g47960	AT4G10100.3|Symbols:CNX7, SIR5|co-factor for nitrate, red	0.0	2.7	0.0	0.0	2.2	3.0
Glyma08g05930	AT1G18260.1|Symbols:|HCP-like superfamily protein|chr1:6	0.0	0.0	0.0	0.0	-2.2	0.0
Glyma05g33770	AT1G18260.1|Symbols:|HCP-like superfamily protein|chr1:6	0.0	0.0	0.0	0.0	0.0	2.2
Glyma03g36600	AT3G63000.1|Symbols:NPL41|NPL4-like protein 1|chr3:232	0.0	-2.3	-2.7	0.0	0.0	0.0
Glyma02g43770	AT4G15410.1|Symbols:PUX5|serine/threonine protein phosp	0.0	-2.4	0.0	0.0	0.0	0.0
Glyma15g18220	AT4G29870.1|Symbols:|Oligosaccharyltransferase complex/	-2.6	-2.6	-2.3	0.0	-3.8	-2.8
Glyma09g06930	AT4G29870.1|Symbols:|Oligosaccharyltransferase complex/	0.0	-2.6	-2.4	0.0	0.0	-2.1
Glyma08g36061	AT1G14570.1|Symbols:|UBX domain-containing protein|chr1	0.0	0.0	0.0	0.0	0.0	2.2
Glyma11g01330	AT3G16090.1|Symbols:|RING/U-box superfamily protein|chr3	0.0	-2.3	-2.0	0.0	-2.3	0.0
Glyma17g14020	AT4G11740.1|Symbols:SAY1|Ubiquitin-like superfamily prot	0.0	2.1	0.0	0.0	0.0	0.0
Glyma05g03480	AT4G11740.1|Symbols:SAY1|Ubiquitin-like superfamily prot	0.0	0.0	0.0	0.0	0.0	2.1
Glyma19g24300	AT3G27310.1|Symbols:PUX1|plant UBX domain-containing	0.0	-2.2	0.0	0.0	0.0	0.0
Glyma15g01880	AT3G16110.1|Symbols:ATPDIL1-6, ATPDI4, PDI4, PDIL1-6	0.0	0.0	0.0	0.0	-5.6	0.0
Glyma10g06480	AT5G03340.1|Symbols:|ATPase, AAA-type, CDC48 protein	0.0	-2.9	-3.2	0.0	-2.2	-2.0

Glyma13g20680	AT5G03340.1|Symbols:|ATPase, AAA-type, CDC48 protein	0.0	-2.8	0.0	0.0	-2.8	-2.5
Glyma12g30060	AT5G03340.1|Symbols:|ATPase, AAA-type, CDC48 protein	0.0	-2.0	0.0	0.0	0.0	0.0
Glyma16g22530	AT2G17200.1|Symbols:DSK2|ubiquitin family protein|chr2:74	0.0	0.0	0.0	0.0	0.0	2.6
Cell Death							
Glyma02g19754	AT1G42990.1|Symbols:ATBZIP60, BZIP60|basic region/leuc	0.0	0.0	2.7	0.0	0.0	3.2
Glyma04g01460	AT4G34460.1|Symbols:AGB1, ELK4, ATAGB1|GTP binding	0.0	-4.4	0.0	0.0	-5.5	0.0
Glyma12g04810	AT4G34460.1|Symbols:AGB1, ELK4, ATAGB1|GTP binding	0.0	-3.1	0.0	0.0	-3.0	0.0
Glyma11g12600	AT4G34460.1|Symbols:AGB1, ELK4, ATAGB1|GTP binding	0.0	0.0	0.0	0.0	-2.3	0.0
Glyma12g02540	AT2G17040.1|Symbols:anac036, NAC036|NAC domain cor	0.0	0.0	0.0	6.2	0.0	0.0
Glyma13g10640	AT5G42050.1|Symbols:|DCD (Development and Cell Death)	0.0	-2.3	0.0	0.0	-2.6	0.0
Glyma20g16100	AT5G42050.1|Symbols:|DCD (Development and Cell Death)	0.0	-2.3	0.0	0.0	-2.0	0.0
Glyma05g36700	AT5G42050.1|Symbols:|DCD (Development and Cell Death)	0.0	0.0	0.0	3.0	0.0	2.1
Glyma06g16440	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	2.3	0.0	-3.1	2.9	0.0	0.0
Glyma02g26480	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	0.0	8.3	3.3	0.0	10.0	6.7
Glyma06g11970	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	2.9	6.7	4.2	3.0	7.5	5.3
Glyma04g38560	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	6.8	5.9	0.0	6.1	17.3	0.0
Glyma14g24220	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	2.6	9.8	3.2	3.0	16.2	7.2
Glyma05g32850	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	12.9	16.4	7.0	8.3	9.9	4.4
Glyma04g42800	AT1G01720.1|Symbols:ATAF1, ANAC002|NAC (No Apical	9.8	21.3	7.7	7.7	25.7	10.9
Glyma19g30681	AT2G40950.1|Symbols:BZIP17|Basic-leucine zipper (bZIP) t	0.0	0.0	2.0	0.0	0.0	0.0
Glyma17g34900	AT4G32940.1|Symbols:GAMMA-VPE, GAMMAVPE|gamma	0.0	-10.0	-3.5	0.0	-6.9	-2.0
Glyma14g10620	AT4G32940.1|Symbols:GAMMA-VPE, GAMMAVPE|gamma	0.0	0.0	0.0	0.0	0.0	-2.3
Glyma14g05980	AT4G14270.1|Symbols:|Protein containing PAM2 motif whic	0.0	3.0	-4.2	0.0	4.6	-2.5
Glyma19g25870	AT3G27090.1|Symbols:|DCD (Development and Cell Death)	0.0	0.0	2.2	0.0	0.0	0.0
Glyma02g42860	AT4G14270.1|Symbols:|Protein containing PAM2 motif whic	2.2	8.7	2.5	2.4	11.5	3.8

DEGs: fold-change ≥ 2 (up), ≤ -2 (down); adj. *p* -value ≤ 0.01; RPM ≥ 9 (control or stress datasets). RPM ≥ 9: at least 20 reads in all datasets.

Additional file 7. Primer pairs used in the study.

Name	**Gene locus**	**Forward primer sequence**	**Reverse primer sequence**
SUSY	Glyma13g17420	TGTTGTTGCATGATTTGGATCTTG	CACGGCTTAAAATTGAATTGATGG
NR	Glyma06g11430	AATCCCATGCAAGCTCATCTCC	CAAACCCATGAGTAGGTCGTCG
nsHB	Glyma11g12980	TGAAACCACCCTCCTCTTAGACTCC	GAATGCGAAACACTTCCCAACC
AOX	Glyma04g14800	CTTCTTGGGGTATTTGTTGTCACCT	ATTGCCCTTGTCCAGCTCCTTA
APX	Glyma12g07780	CACGGTGCCCATAATATTTCTCTC	CAACCCAACTCCAATCATCATCAC
TPS	Glyma17g07530	TGCGTAACAGCAATCTTTCAACAA	CCCAAGACCTAATCCACCAACC
ELF1B[A]	Glyma02g44460	GTTGAAAAGCCAGGGGACA	TCTTACCCCTTGAGCGTGG
ACTB[B]	Glyma15g05570	GAGCTATGAATTGCCTGATGG	CGTTTCATGAATTCCAGTAGC

Primers obtained from: A (Jian et al. 2008), B (Byfield et al. 2006).

CAPÍTULO III (ARTIGO)

Setaria viridis floral-dip: A simple and rapid *Agrobacterium*-mediated transformation method

Keywords Model plant; Green millet; Red fluorescent protein

Abstract *Setaria viridis* was recently described as a new monocotyledonous model species for C4 photosynthesis research and genetic transformation. It has biological attributes (rapid life cycle, small genome, diploid, short stature, and simple growth requirements) that make it suitable for use as a model plant. We report an alternative method of *S. viridis* transformation using floral dip to circumvent the necessity of tissue culture phase for transgenic plant regeneration. *S. viridis* spikes at boot stage were selected to be immersed in *Agrobacterium* suspension. T1 seeds could be identified in 1.5–2 months after floral dipping. We demonstrated through molecular analysis and RFP expression that seeds and resulting plants from dipped inflorescences were transformed. Our results suggest the feasibility of *S. viridis* floral dip transformation as a time-saving and cost-effective compared with traditional methods. To our knowledge, this is the first report using floral dip in *S. viridis* as an *Agrobacterium*-mediated transformation method.

Grass crop plants are the major sources of human food–feed–fiber–fuel in the world. The most agriculturally important grasses are rice, wheat, maize, sorghum, and sugarcane. *Setaria viridis* belongs to the Poaceae family, subfamily Panicoideae that includes the most agronomically important grass crops. *S. viridis* has a number of characteristics that makes it interesting as a model plant for genetics and functional genomics studies (Diao et al., 2014). Plant transformation typically requires refined regeneration methods that involve multiples steps from *in vitro* culture to greenhouse conditions. *Agrobacterium*-mediated transformation by vacuum infiltration of *Arabidopsis* inflorescences was first reported by Bechtold et al. (1993) and further modified as floral dip method by Clough and Bent (1998). Floral dip method has been used in genetic

engineering strategies as it directly produces genetically modified seeds bypassing the laborious tissue culturing procedures. In addition, tissue culture could generate transformed plants harboring undesirable mutations from somaclonal variation (Bent, 2000). This method, indubitably, accelerated *Arabidopsis thaliana* studies as a model plant, quickly advancing our understanding of several other dicotyledonous crops. For that reason, the development of a simple and rapid transformation method for *S. viridis* as a C4 model for monocotyledonous plants could help to achieve results comparable with those in *A. thaliana*.

Three seeds of *S. viridis* (accession A10.1) were cultivated in ten square plastic pots (230 ml) with Supersoil® (Scotts Miracle-Gro Company) and grown in an air-conditioned greenhouse with a 26 ºC/22 ºC day/night cycle with a 14 h photoperiod. About 26–30 days after planting, *S. viridis* spikes at the boot stage (Blaise et al., 1992) were used for floral dipping (Fig. 1a).

In order to transform *Setaria* plants, *Agrobacterium tumefaciens* (AGL1 strain) cultures containing the vector pANIC 6A (Mann et al., 2012) was prepared. The bacteria were cultured in 500 ml of LB medium containing 50 mg/l kanamycin at 30 ºC to OD_{600} = 1.0 and centrifuged at 5000 × g for 15 min. The pellet was washed with 10 ml of MM buffer (10 mM MES, 10 mM MgCl2, pH 5.6) and resuspended in 50 ml of fresh MM buffer. The addition of tobacco (*Nicotiana tabacum*) extract is known to improve the gene transfer by agrobacteria in monocots (Fursova et al., 2012). Therefore, tobacco leaf extract was prepared as following: 40 g of fresh leaves were cut into 2 cm^2 pieces in 1l of MM buffer and stirred gently for 1h. The solution was sieved and the liquid phase was collected. Prior to inoculation, the solution (1l) was supplemented with 1ml of 50mM acetosyringone, 200 µl of Silwet-L77 and 50 g of sucrose. The final solution was mixed with the bacterial suspension used for the floral dip procedure. Spikes of one month-old *S. viridis* plants at boot stage (Fig. 1a) were infiltrated with bacterial suspension carrying the pANIC 6 A. Briefly, the aerial part of plants containing the spikes were immersed in *Agrobacterium* suspension inside of a Secador® Techni-dome® 360 vacuum desiccator and subjected to vacuum infiltration (80 kPa) for 10 min (Fig. 1b). The spikes were covered with plastic bags for 24h (Fig. 1c). The co-cultivation was performed for 3 days under greenhouse conditions without irrigation. After 1.5–2 months seeds were

screened for red fluorescence using a Leica MZ16F stereo microscope equipped with an XCite EXFO fluorescence illumination source and a filter combination for fluorescence in the red region of the spectrum (excitation: 558 nm, emission 580 nm). Digital images were collected using a digital camera (Retiga 2000 Fast Cooled 12-bit) (Fig. 1d and e). Mature seeds were collected and germinated in half-strength semi-solid MS medium containing 30 mg/l hygromycin B for selection of transformed seedlings. DNA was extracted from plants surviving on hygromycin-containing medium and processed according to Doyle and Doyle (Doyle and Doyle, 1990). The molecular analysis of the transformed plants was performed by PCR. Each reaction (20 µl) contained 20 mM Tris–HCl (pH 8.4), 50 mM KCl, 2 mM MgCl2, 200 µM of each dNTP, 1.0 U of Taq polymerase, 50 ng DNA, and 5 µM of each oligonucleotide specific to *pporRFP* gene. The *pporRFP* primer sequences were 5′-GGCCATTATACGTGCGACTT-3′ and 5′-CAACGAGAAATTGCACATGC-3′ to generate 159 bp PCR product (Fig. 1f). The samples were submitted to the following amplification cycles in a Bio-Rad iCycler at 95 C for 30 s, followed by 40 cycles of 95 ºC for 30 s, 50 ºC for 45 s, 68 ºC for 40 s, and 68 ºC for 5 min. The amplified fragments were subjected to electrophoresis in 1% agarose gel containing 0.5 µg/ml ethidium bromide and visualized in UV transilluminator.

The appropriate stage for dipping appears to be from the mid to late uninucleate microspore stage when the spike has not emerged from the sheath (boot stage) (Zale et al., 2009). This stage coincides with the highest frequency uninucleate microspore in the anther (Liu et al., 2002). In addition, in rice, the anther/microspores were the preferential targets of the floral-dip transformation (Rod-in et al., 2014). To date, the target tissue for the T-DNA transfer to *S. viridis* has not been determined. The visual screening in RFP-expressing seeds could not discriminate which part of the seed was transformed (i.e., endosperm, embryo etc.). Therefore, mature seeds were germinated on selective medium and the surviving seedlings were analyzed by PCR to confirm that transgenic plants were generated. From one thousand seeds, 190 survived on hygromycin containing medium. The PCR analysis showed that among them six plants were confirmed to be transformed (transformation efficiency of 0.6% – number of transformants/number of seed set). To our knowledge, this is the first report using floral dip in *S. viridis* as an *Agrobacterium*-mediated transformation method.

Therefore, we confirmed the feasibility of floral dip for *S. viridis*. This method could accelerate genetic and genomics studies of important food–feed–fiber–fuel crops such as sugarcane, *Miscanthus*, elephant grass, *Brachiaria*, *Sorghum*, and corn. Currently, we are working on the optimization of different steps involved in this method to improve the transformation efficiency.

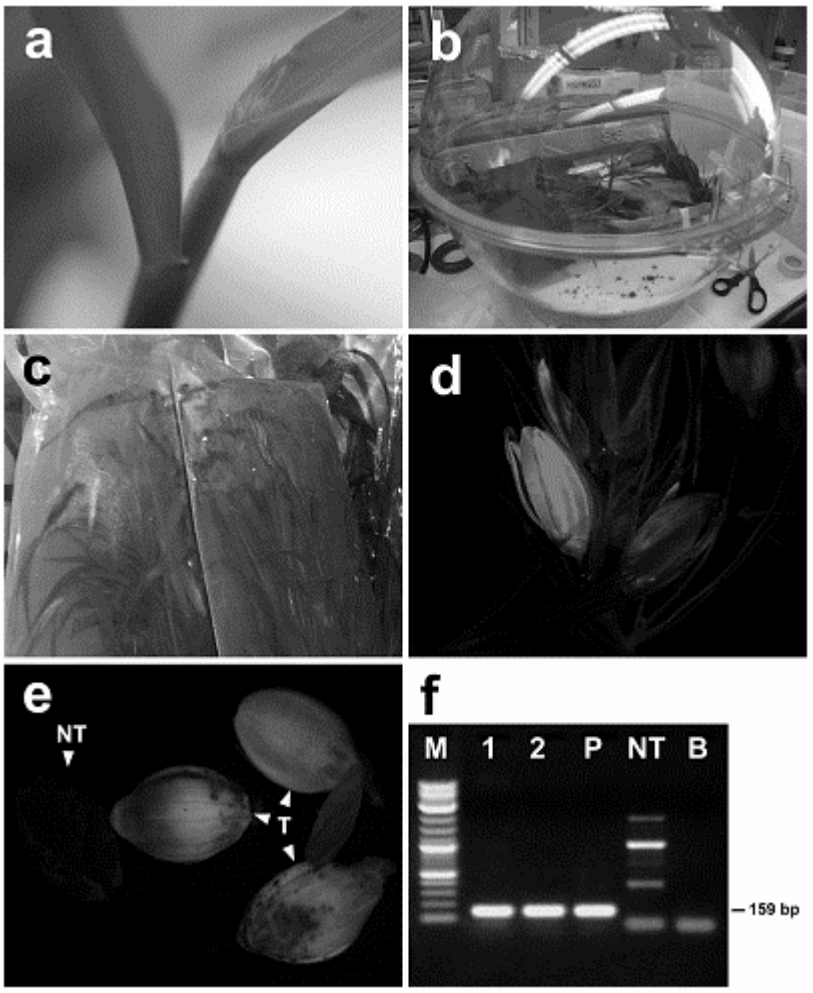


Fig. 1.*Setaria viridis* floral dip transformation. (a) Inflorescence developmental stage (boot stage) selected for transformation. (b) Spikes dipped in *Agrobacterium* suspension inside a desiccator for vacuum-assisted *Agrobacterium* infiltration. (c) Spikes covered with plastic bags after infiltration to keep moisten for 24 h. (d) Spike showing RFP-expressing immature seed. (e) RFP-expressing mature seeds (NT – non-transgenic, T – transgenic seeds, indicated by arrow heads). (f) PCR analysis of surviving plants on hygromycin-containing medium (M: molecular weight marker-2-log DNA ladder, lanes 1 and 2: transgenic plants showing expected 159 bp *porpRFP*-specific band, NT: non-transgenic plant, P: positive control–plasmid pANIC 6 A, B: polymerase chain reaction amplification without DNA template).

References

Bechtold N, Ellis J, Pelletier G (1993) In planta *Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C R Acad Sci **316**: 1194–1199

Bent AF (2000) Arabidopsis in Planta Transformation. Uses, Mechanisms, and Prospects for Transformation of Other Species. Plant Physiol **124**: 1540–1547

Blaise M, Girardin P, Millet B (1992) Developmental stages and floral ontogenesis of foxtail millet Setaria italica (L) P Beauv. Agronomie **12**: 141–156

Clough SJ, Bent AF (1998) Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735–743

Diao X, Schnable J, Bennetzen JL, Li J (2014) Initiation of Setaria as a model plant. Front Agric Sci Eng **1**: 16

Doyle J, Doyle J (1990) Isolation of plant DNA from fresh tissue. Focus (Madison) **12**: 13–15

Fursova O, Pogorelko G, Zabotina OA (2012) An efficient method for transient gene expression in monocots applied to modify the *Brachypodium distachyon* cell wall. Ann Bot **110**: 47–56

Liu W, Zheng MY, Polle EA, Konzak CF (2002) Highly efficient doubled-haploid production in wheat (*Triticum aestivum* L.) via Induced microspore embryogenesis. Crop Sci **42**: 686–692

Mann DGJ, LaFayette PR, Abercrombie LL, King ZR, Mazarei M, Halter MC, Poovaiah CR, Baxter H, Shen H, Dixon RA, et al (2012) Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. Plant Biotechnol J **10**: 226–236

Rod-in W, Sujipuli K, Ratanasut K (2014) The floral-dip method for rice (*Oryza sativa*) transformation. J Agric Technol **10**: 467–474

Zale JM, Agarwal S, Loar S, Steber CM (2009) Evidence for stable transformation of wheat by floral dip in *Agrobacterium tumefaciens*. Plant Cell Rep **28**: 903–913

Conclusões Gerais

No presente estudo, das diferenças entre as duas cultivares, Embrapa 45 apresentou menor número de genes *up*-, maior quantidade de genes *down*-regulados e maior expressão dos genes que codificam *phosphoglucomutase* (Glyma05g34790), *unknown protein related to N-terminal protein myristoylation* (Glyma06g03430), *protein suppressor of phyA-105* (Glyma06g37080) e *fibrillin* (Glyma10g32620). Dados de RNA-seq e qRT-PCR da hemoglobina não simbionte Glyma11g12980 indicaram divergência estrutural desse gene entre as cultivares. Em comum entre as duas cultivares, mudanças na abundância de transcritos envolvidos no metabolismo de aminoácidos e derivados sugerem perturbação destes em modificações do tRNA, acurácia/eficiência traducional e estresse do retículo endoplasmático (RE) sob hipóxia. Também observamos que grupos de genes (estáveis, *up*- e *down*-regulados) diferem quanto à frequência de elementos *cis* TATA box, ABREs (*ABA-responsive elements*) e CRT/DREs (*C-repeat/dehydration-responsive elements*), assim como em estrutura, composição e uso de códons sinônimos. Especulamos que *cis* elementos ABREs podem mediar expressão gênica independentemente de ABA em raízes de soja sob hipóxia. Por fim, a nosso conhecimento, somos os primeiros a demonstrar viabilidade de se transformar *Setaria viridis* (espécie modelo de monocotiledôneas C4) mediada por *Agrobacterium* por meio da técnica *floral dip*, que não demanda cultura de tecidos nem regeneração de plantas transgênicas. Assim como em arabidopsis, espécie modelo de dicotiledôneas C3, *floral dip* em *S. viridis* pode acelerar estudos genéticos e de genômica funcional (inclusive ao estresse de alagamento) de gramíneas importantes para alimentação humana e animal, fibras e biocombustíveis, tais como cana-de-açúcar, Miscanthus, capim-elefante, Brachiaria, sorgo e milho.

Contribuições Complementares

Durante meu doutorado sanduíche, clonei 33 genes e transformei 30 deles em plantas de arabidopsis para posterior análise funcional. Também adaptei a tecnologia de engenharia genética de precisão CRISPR (*Clustered Regularly*

Interspaced Short Palindromic Repeats) para plantas. A obtenção de linhagens homozigotas dos transgenes para posterior fenotipagem e testes do funcionamento do CRISPR está em andamento.