

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

O Gene Codificador da "Proteína-cinase Ativada por Mitógenos" 5 (MPK5) de
Eucalyptus grandis

JULIANA DAMETTO KRÁS BORGES

Dissertação submetida ao Programa de Pós-
graduação em Biologia Celular e Molecular
do Centro de Biotecnologia da UFRGS
como requisito parcial à obtenção do grau
de Mestre

Orientador: Giancarlo Pasquali

Porto Alegre, julho de 2014

Instituição e Fonte Financiadora

Este trabalho foi realizado no Laboratório de Biologia Molecular Vegetal do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul.

Como fonte financiadora, este trabalho contou com bolsa e recursos financeiros do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Ministério da Ciência, Tecnologia e Inovação (MCTI).

Índice

Lista de abreviaturas	5
Resumo.....	7
<i>Abstract</i>	8
1. Introdução	9
1.1. Proteínas-cinases	9
1.1.1 Proteínas-cinases ativadas por mitógenos	9
1.1.2 Classificação das PKs	12
1.1.3 MAPKs em plantas	13
1.1.4 Regulação de MAPKs	14
1.1.4.1 Regulação pós-traducional	14
1.1.4.2 Regulação transcricional e pós-transcricional.....	15
1.2. Estresse vegetal	16
1.2.1 Sinalização de estresse em plantas	17
1.2.2 ABA e estresses.....	21
1.2.3 Estresse por salinidade	23
1.2.4 Estresse por seca.....	24
1.3. Plantas modificadas geneticamente com componentes das cascatas de MAPKs.....	25
1.4. Importância comercial e melhoramento de <i>Eucalyptus</i>	26
2. Objetivos	28
3. Materiais e métodos	29
3.1. Análises <i>in silico</i>	29
3.1.1 Avaliação do número de cópias e da estrutura gênica de <i>Egmpk5</i>	29
3.1.2 Construção de mapa cromossômico.....	30
3.1.3 Análise da região promotora do gene <i>Egmpk5</i>	30
3.1.4 Sequência peptídica, árvore filogenética e análise de domínios	30
3.2. Análises da expressão de <i>mpk5</i> em <i>E. grandis</i>	31
3.2.1 Material vegetal e tratamentos	31
3.2.2 Extração de RNA total e síntese de cDNA.....	31
3.2.3 Avaliação da expressão relativa de <i>Egmpk5</i> por RT-qPCR	31
3.3. Caracterização de plantas transformadas de <i>N. tabacum</i>	33
3.3.1 Assepsia de sementes, meios de cultivo, germinação e cultura <i>in vitro</i> de <i>N. tabacum</i> SR1.....	33
3.3.2 Cultura hidropônica.....	34

3.3.3 Detecção de transgenes em plantas transformadas por PCR.....	34
4. Resultados	35
4.1. Análises <i>in silico</i> do gene <i>mpk5</i> de <i>E. grandis</i>	35
4.1.1 Definição do número de cópias e da estrutura gênica de <i>Egmpk5</i>	35
4.1.2 Análise da região promotora	46
4.1.3 Análise do peptídeo deduzido de MPK5 de <i>E. grandis</i>	48
4.2. Caracterização do padrão de expressão de <i>mpk5</i> em <i>E. grandis</i>	52
4.3. Identificação e caracterização de plantas de tabaco superexpressando <i>Egmpk5</i>	57
4.3.1 Análise por PCR.....	57
4.3.2 Germinação e classificação de genótipos de plantas transgênicas quanto à homo/hemizigose	59
4.3.3 Teste das plantas transgênicas quanto à tolerância ao estresse salino.....	62
4.3.4 Teste das plantas transgênicas quanto à tolerância à seca.....	63
5. Discussão	66
5.1. Análise <i>in silico</i> do gene <i>mpk5</i> de <i>E. grandis</i> e proteína deduzida.....	66
5.2. Análises da expressão de <i>mpk5</i>	69
5.3. Análise de plantas de tabaco expressando <i>Egmpk5</i>	71
6. Conclusão e perspectivas	73
7. Referências bibliográficas	74
8. Anexo I.....	83
9. Anexo II	86

Lista de abreviaturas

°C – graus Celsius

μL – microlitros

ABA – ácido abscísico

ABRE – *ABA-responsive element*

ATP – adenosina trifosfato

BLAST – *Basic Local Alignment Search Tool*

CaMV 35S – promotor do vírus do mosaico da couve-flor

cDNA – DNA complementar

CDS – Sequência de DNA codificante (do inglês, *coding DNA sequence*)

DNA – ácido desoxirribonucleico (do inglês *desoxiribonucleic acid*)

dNTPs – desoxirribonucleosídeos trifosfatados

EST – *Expressed Sequence Tag*

GTP – guanosina trifosfato

IP3 – inositol 1,4,5-trifosfato

Kin – cinetina

MAPK – proteína-quinase ativada por mitógenos

MAPKK ou MAP2K – MAP quinase-quinase

MAPKKK ou MAP3K – MAP quinase-quinase-quinase

mg – miligramas

mL – mililitros

mM – milimolar

NAA – ácido naftalenoacético

ng – nanogramas

Nos – região terminadora da nopalina sintase

PA – ácido fosfatídico

PCR – reação em cadeia da polimerase (do inglês *polymerase chain reaction*)

PK – proteína-cinase (do inglês *protein kinase*)

PL – fosfolipase

RNA – ácido ribonucleico (do inglês *ribonucleic acid*)

RT-qPCR – reação em cadeia da polimerase quantitativa, ou em tempo real, precedida de transcrição reversa

STK – serina/treonina cinase (do inglês *serine/threonine kinase*)

TK – tirosina cinase (do inglês *tyrosine kinase*)

Resumo

As proteínas-cinases ativadas por mitógenos (MAPKs) participam de rotas de transdução de sinais universais em eucariotos e compõem cascatas de fosforilação que levam as células a responder a estímulos extracelulares. Em plantas, MAPKs estão associadas a processos de desenvolvimento e na resposta a hormônios e a estresses bióticos e abióticos. Em trabalhos anteriores de nosso grupo, o gene *Egmpk5* de *Eucalyptus grandis* foi clonado e seu padrão de expressão foi caracterizado em plantas submetidas a diferentes tratamentos. Ainda, plantas de *Nicotiana tabacum* SR1 foram transformadas com *Egmpk5* sob regulação do promotor CaMV 35S para futuro estudo de função do gene. No presente trabalho, foi avaliada mais amplamente a estrutura, a expressão e o produto deste gene. Análises *in silico* revelaram que *Egmpk5* apresenta-se como cópia única no genoma, está organizado em seis éxons e seus cinco íntrons possuem sítios canônicos de *splicing* de RNA. O produto deduzido deste gene apresenta em sua sequência a assinatura do domínio proteína-cinase e os motivos de ligação a ATP e de ativação/dupla fosforilação, indicando a possibilidade de ser uma proteína funcional. Árvore filogenética foi construída pelo método de Máxima Verossimilhança utilizando o algoritmo MUSCLE para alinhamento das sequências e um valor de *bootstrap* de 500 repetições com sequências peptídicas similares à proteína deduzida EgMPK5 e permitiu a identificação de MAPKs de outras plantas com função já comprovada na sinalização da divisão celular e na resposta a estresses abióticos, ferimento e patógenos. Análise de RT-qPCR mostrou a expressão de *Egmpk5* em níveis comparáveis em raízes, caules e folhas de plantas de *E. grandis* tratadas com água, e foi detectado um aumento de sua expressão em folhas tratadas com ácido abscísico (ABA) e em todos os órgãos tratados com solução de cloreto de sódio a 200 mM. O estudo da região promotora de *Egmpk5* revelou um grupo de possíveis elementos de ação *cis* relacionados a respostas a estresses. Seis linhagens transformadas de tabaco que tiveram seu estado transgênico confirmado por PCR foram desafiadas com tratamentos de seca e cloreto de sódio, e linhagens mais tolerantes foram identificadas. Ao final destas diversas análises, os resultados obtidos sugerem a participação de *Egmpk5* em respostas a estresses abióticos.

Abstract

Mitogen-activated protein kinases (MAPKs) are part of phosphorylation cascades found in all eukaryotes and are responsible for transducing extracellular stimuli into intracellular responses. In plants, MAPK cascades are involved in growth and development processes and have roles in the response to hormones and biotic and abiotic stresses. In our previous studies, the *Egmpk5* gene from *Eucalyptus grandis* was cloned and its expression profile was characterized in plants subjected to different treatments. Also, *Nicotiana tabacum* SR1 plants were transformed with *Egmpk5* under the control of the CaMV 35S promoter for further function studies. In the present study the structure, expression profiles and product of this gene were further characterized. *In silico* analyses revealed that the *Egmpk5* gene is present as a single copy in the *E. grandis* genome, and that its structure comprises six exons and five introns with conserved sites for transcript splicing. The deduced product of this gene presents the protein kinase domain signature, ATP-binding site and activation/dual phosphorylation motifs in its structure. A phylogenetic tree was built with sequences similar to the deduced EgMPK5 peptide and lead to the identification of plant MAPKs that have had proven activity in cell division signaling and in the response to abiotic stresses, wounding and pathogens. RT-qPCR analysis showed equal expression of *Egmpk5* in roots, stems and leaves of plants treated with water and we detected an increase of expression in leaves treated with abscisic acid (ABA) and also increase in expression in all three organs upon treatment with 200 mM sodium chloride. Analysis of the promoter region of *Egmpk5* revealed a group of putative *cis*-acting elements related to stress responses. Six transformed tobacco lines were confirmed by PCR and were assayed with drought and sodium chloride treatments for the identification of more tolerant individuals. After all analyses performed, results allowed us to suggest that *Egmpk5* is involved in abiotic stresses responses.

1. Introdução

1.1 Proteínas-cinases

Proteínas-cinases (PKs, do inglês, *Protein Kinases*) realizam a fosforilação de substratos celulares, um importante mecanismo de regulação nas células eucarióticas que participa em processos como proliferação, metabolismo, motilidade, transporte de membranas e expressão gênica (HANKS, 2003). A reação de fosforilação é uma modificação covalente e reversível de proteínas e se dá pela transferência de um grupamento fosfato terminal (γ) de um nucleotídeo purínico trifosfatado (ATP ou GTP) para hidroxilas de resíduos de aminoácidos específicos, que podem ser tanto serina e treonina ou unicamente tirosina, variando de acordo com a classe da enzima que cataliza a reação (UBERSAX & FERREL, 2007).

A adição de grupamento fosfato, com duas cargas negativas, causa mudanças eletrostáticas na superfície das proteínas, o que leva a mudanças conformacionais que alteram a atividade catalítica e a ligação a substratos (BERG *et al.*, 2002). Proteínas-cinases têm um domínio catalítico que liga e fosforila proteínas-alvo, e uma região de regulação que é também controlada por fosforilação (MANNING *et al.*, 2002). A atividade de fosforilação por cinases tem grande capacidade de amplificação de sinal, já que as moléculas-alvo podem ser também enzimas e, portanto, também capazes de catalizar reações em outras moléculas (BERG *et al.*, 2002).

Estas proteínas integram a superfamília de proteínas-cinases eucarióticas (ePK, do inglês, *eukariotic Protein Kinases*), cujos membros apresentam um domínio bastante conservado (HANKS, 2003). A família ePK é dividida de acordo com os substratos reconhecidos por seus membros, o que dá origem a duas grandes classes: serina/treonina cinases (STKs, do inglês, *Serine/Threonine Kinases*) e tirosina cinases (TKs, do inglês, *Tyrosine Kinases*) (SRINIVASAN *et al.*, 2010).

1.1.1 Proteínas-cinases ativadas por mitógenos

Em meados dos anos 1980, estudos em adipócitos sobre o efeito da insulina e de fatores de crescimento sobre a fosforilação intracelular de proteínas levaram à detecção

de STKs associadas a microtúbulos e à proposição de um mecanismo de sinalização possivelmente relacionado com a classe dos receptores TKs. Estudos posteriores passaram a denominá-las proteínas-cinases ativadas por mitógenos (MAPKs do inglês, *Mitogen-activated Protein Kinases*), e a similaridade entre MAPKs de mamíferos e leveduras serviu para a proposta de um mecanismo conservado com ação na proliferação e na diferenciação celular (AVRUCH, 2007).

As MAPKs pertencem à classe das STKs e se organizam hierarquicamente em três módulos capazes de transmitir e amplificar sinais na célula. A cascata que elas formam consiste na ação consecutiva de três componentes enzimáticos: MAPKKKs (MAP3Ks), MAPKKs (MAP2Ks) e MAPKs. As MAPKs estão ativadas quando resíduos de treonina e tirosina de seu sítio de regulação estão fosforilados. Esta fosforilação das MAPKs é feita por MAP2Ks que são, por sua vez, ativadas por fosforilação de seus resíduos de serina ou treonina em uma reação catalizada por MAP3Ks (HAMEL *et al.*, 2006; RODRIGUEZ *et al.*, 2010) (**Figura 1**).

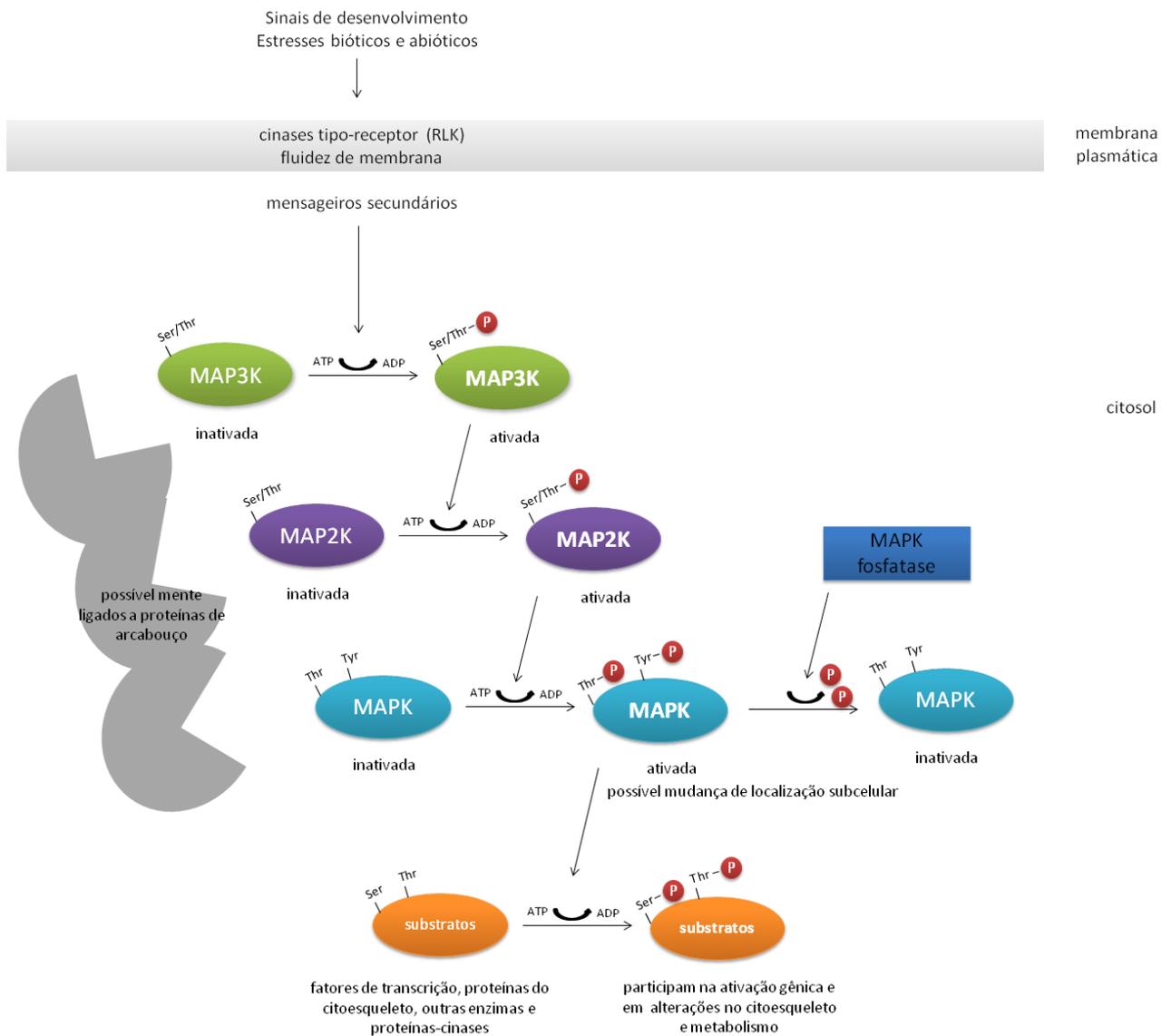


Figura 1: No esquema adaptado das revisões de RODRIGUEZ *et al.*, 2010 e ŠAMAJOVÁ *et al.*, 2013a e 2013b, é mostrado o modelo geral de uma cascata de transdução de sinais por MAPKs em plantas. Mudanças na fluidez da membrana e estimulação de cinases tipo-receptor levam à geração de segundos mensageiros que ativam MAP3Ks no citosol por fosforilação de seus resíduos de serina (Ser) ou treonina (Thr). MAP3Ks, quando ativadas, fosforilam resíduos de Ser ou Thr do motivo conservado S/T-X₃₋₅-S/T de MAP2Ks. Em sequência, MAPKs são ativadas mediante fosforilação de resíduos de Thr e tirosina (Tyr) de seu motivo conservado T-X-Y. A regulação da atividade de MAPKs pode ser controlada mediante reações de desfosforilação, catalisadas por tirosina e serina/treonina fosfatases. Possivelmente, a cascata das MAPKs é mantida unida mediante interação com proteínas de arcabouço (*scaffold proteins*), o que pode contribuir para a especificidade na sinalização. MAPKs ativadas podem tanto permanecer no citoplasma como realocar para o núcleo, locais onde agem fosforilando outras proteínas-cinases e enzimas, fatores de transcrição, proteínas citoplasmáticas e do citoesqueleto. Como consequência, há modulação na expressão gênica, reprogramação no desenvolvimento e resposta a estresses.

Entre os mamíferos, foram identificadas quatro principais cascatas de MAPKs que são conhecidas como ERK1/2, ERK5, p38 e JNK de acordo com as enzimas que as compõem. Entre eucariotos, foram encontrados ortólogos como, por exemplo, a enzima Hog1MAPK de *Saccharomyces cerevisiae*, que é ortóloga de p38 (BARDWELL & SHAH, 2006) e também diversas MAPKs de plantas, que são mais próximas da classe ERK de mamíferos (LEE *et al.*, 2008). Durante os 1,6 bilhões de anos em que plantas e animais divergiram, as funções dos componentes das MAPKs diferenciaram-se bastante (JONAK *et al.*, 2002). Os genomas vegetais codificam um número elevado de componentes da cascata das MAPKs comparado a humanos – em *Arabidopsis thaliana*, por exemplo, são descritas 20 MAPKs e 10 MAP2Ks, e em humanos há 10 MAPKs e 7 MAP2Ks (HAMEL *et al.*, 2006) – o que forma uma rede de transdução de sinais especializada e complexa (LEE *et al.*, 2008). Em *A. thaliana*, esta falta de proporcionalidade entre o número de MAPKs e MAP2Ks resulta em “conversa-cruzada” entre os módulos que fazem a sinalização (JONAK *et al.*, 2002).

1.1.2 Classificação das PKs

O domínio catalítico das ePKs foi descrito por Hanks & Hunter em 1995. Este foi definido após estudos de homologia de diversas sequências, e se caracteriza por uma região de 250-300 resíduos de aminoácidos que pode ser subdividida em 12 regiões menores, ou subdomínios. Análise da estrutura tridimensional de MAPKs permitiu determinar que estes resíduos conservados são de grande importância para a função enzimática, já que participam na ligação e na orientação da molécula doadora de fosfato e do substrato à enzima, além de realizarem a transferência do grupamento fosfato durante a catálise (HANKS & HUNTER, 1995; revisado em HANKS, 2003).

O domínio catalítico das ePKs possui ainda uma importante região reguladora que se localiza entre os domínios VII e VIII e se caracteriza pelo consenso treonina-aminoácido-tirosina (T-X-Y). Como estes resíduos encontram-se em uma região de estrutura em alça (ou *T-loop*, em inglês), a fosforilação destes por MAP2Ks causa mudanças conformacionais capazes de ativar as MAPKs (ICHIMURA *et al.*, 2002). Outro motivo importante em MAPKs é o sítio C-terminal de ancoragem denominado CD (do inglês, *C-terminal Docking site*), localizado fora do domínio catalítico. Este é

responsável pela interação com MAP2Ks, fosfatases e proteínas de arcabouço, tendo também papel na regulação da atividade enzimática (DÓCZI *et al.*, 2012).

1.1.3 MAPKs em plantas

A pesquisa de MAPKs vegetais teve início com a identificação de MsERK1 de *Medicago sativa* e a cinase D5 de *Pisum sativum* em 1993 (revisado em ZHANG, 2008). Ainda no início dos anos 1990, foram clonados genes codificadores de MAPKs de *A. thaliana* e *Nicotiana tabacum* (MIZOGUCHI *et al.*, 1993; WILSON *et al.*, 1993). Desde então, foram descritos diversos genes codificadores de MAPKs e seus perfis de transcrição foram caracterizados em organismos como *Avena sativa* (HUTTLY & PILLIPS, 1995), *Brachypodium distachyon* (CHEN *et al.* 2012), *Chorispora bungeana* (ZHANG *et al.* 2006), *Hordeum vulgare* (KNETSCH *et al.*, 1996), *Malus domestica* (ZHANG *et al.*, 2013), *Oryza sativa* (FU *et al.*, 2002), *Petunia juss* (DECROOCQ-FERRANT *et al.*, 1995), *Populus trichocarpa* (NICOLE *et al.*, 2006), *Solanum lycopersicum* (KONG *et al.*, 2012), *Triticum aestivum* (LIAN *et al.*, 2012) e *Zea mays* (LALLE *et al.*, 2005).

Em 2002, Ichimura *et al.* descreveram a possibilidade de agrupar MAPKs e MAP2Ks a partir de análises filogenéticas com proteínas de outras espécies e chegaram a propor uma nomenclatura baseada na homologia das sequências peptídicas. Foi proposta a divisão destas proteínas em quatro grupos, nomeados de A a D, de acordo com o consenso de sequência na região reguladora: as proteínas que contêm o motivo treonina-ácido glutâmico-tirosina (subtipo TEY) e aquelas que contêm o motivo treonina-ácido aspártico-tirosina (subtipo TDY). O subtipo TEY está presente nas proteínas dos grupos A, B e C, enquanto as MAPKs do grupo D apresentam o motivo TDY.

As MAPKs do tipo TEY são as mais amplamente caracterizadas e há evidências de que elas tenham um papel importante na resposta a estresses bióticos (ROMEIS, 2001; ZHANG & KLESSIG 2001), metais pesados (JONAK *et al.*, 2004), seca (LI *et al.*, 2012), salinidade (TENA *et al.*, 2001; NAKAGAMI *et al.*, 2005), estresse oxidativo (SAMUEL *et al.*, 2000), fermento (MAYROSE & SESSA, 2004), hormônios vegetais (KNETSCH *et al.*, 1996; MOCKAITIS *et al.*, 2000; BURNETT *et al.*, 2000), na

citocinese da célula vegetal (SOYANO *et al.*, 2003) e no desenvolvimento do pólen (VORONIN *et al.*, 2004).

À medida que novas MAPKs foram identificadas, diversos trabalhos também conseguiram agrupá-las de acordo com o proposto por Ichimura *et al.* (2002). Assim, MAPKs de espécies eudicotiledôneas e monocotiledôneas permitiram validar o sistema de grupos proposto e levaram a uma ideia inicial do grau de conservação de sequências e atividades desta classe de proteínas em plantas (HAMEL *et al.*, 2006).

1.1.4 Regulação de MAPKs

1.1.4.1 Regulação pós-traducional

As cascatas de MAPKs, por operarem no núcleo central de redes de transdução de sinal, possuem atividades importantes para a planta adaptar seu metabolismo a condições adversas do ambiente e, assim, o entendimento de como os sinais percebidos pelas células são convertidos em respostas intracelulares é importante para o estudo da adaptação ao estresse (SMÉKALOVÁ *et al.*, 2013).

Em mamíferos, a duração da ativação de algumas MAPKs depende do tipo de estímulo e isso gera respostas com diferença temporal e diferentes alterações fisiológicas (MARSHALL, 1995; MURPHY *et al.*, 2002). De forma similar em tabaco, a ativação transiente de uma MAP2K e da MAPK sobre a qual ela atua (NtSIPK) induz a expressão de genes de defesa, mas a ativação contínua desta mesma MAP2K fosforila outra MAPK (NtWIPK), que causa morte celular (LIU *et al.*, 2013).

Em plantas, a interação entre uma MAP2K e MAPKs diferentes leva a respostas distintas. Por exemplo, em *A. thaliana*, a MAP2K chamada MEKK1 ativa as MAPKs MPK3 e MPK6 na sinalização em resposta a flagelina, mas também ativa MPK4 e MPK6 em estresses por frio e salinidade (ASAI *et al.*, 2002; TEIGE *et al.*, 2004).

Há evidência de redundância de sinalização na parte final da cascata mediada por MAPKs, ou seja, nos substratos-alvo de MAPKs ativadas. Um único substrato pode ser fosforilado por diferentes MAPKs como, por exemplo, a MAP65-1 (uma proteína

associada a microtúbulos) que é fosforilada por três MAPKs, MPK3, MPK4 e MPK6 (HOEHENWARTER *et al.*, 2013) em *A. thaliana*. Nesta mesma planta, foi descrito um conjunto de 9 MAP2Ks e 10 MAPKs que é capaz de fosforilar 570 substratos diferentes, entre eles fatores de transcrição envolvidos na regulação de desenvolvimento, defesa e respostas a estresses (POPESCU *et al.*, 2009).

Além da regulação positiva na ativação, MAPKs são também inativadas mediante desfosforilação. MAPK-fosfatases são enzimas de dupla especificidade, capazes de desfosforilar resíduos de serina, treonina e tirosina e agem, dessa forma, como reguladores negativos das MAPKs (LI *et al.*, 2012). Ao desativar MAPKs, as MAPK-fosfatases são capazes de interferir na sinalização hormonal, na resposta a estresse abiótico e em respostas de crescimento (LUMBRERAS *et al.*, 2010).

Em mamíferos e leveduras, os componentes da cascata MAPK são mantidos unidos por proteínas chamadas “arcabouço” (do inglês, *scaffold proteins*), que são capazes de organizar os componentes e direcioná-los para compartimentos subcelulares e, assim, contribuem para a especificidade da sinalização (revisado em PLOTNIKOV, 2011). Em arroz foi descrito um complexo proteico envolvendo uma GTPase e a forma ativa de uma MAPK responsiva a estresse, o que foi considerado uma das primeiras evidências da presença de *scaffold proteins* em plantas (LIEBERHERR *et al.*, 2005). Outro exemplo da função da compartimentalização subcelular se dá em *P. crispum*, em que células tratadas com elicitador levaram três MAPKs citosólicas a se localizarem no núcleo, onde provavelmente interagem com fatores de transcrição (revisado em NICOLE *et al.*, 2006).

1.1.4.2 Regulação transcricional e pós-transcricional

As MAPKs são também reguladas de forma transcricional e pós-transcricional. A ocorrência de *splicing* alternativo já foi observada nos transcritos codificadores da ERK1 de mamíferos (YUNG *et al.*, 2000). Em plantas, *splicing* foi também descrito em uma *map3k* de *A. thaliana* (NISHIHAMA *et al.*, 1997), em duas *map2ks* de *B. distachyon* (CHEN *et al.*, 2012) e em quatro *mapks* de arroz (LIU & XUE, 2007).

A expressão de genes *mapks* de plantas pode ser restrita a alguns órgãos e, assim, participar de um controle espacial ou temporal. Como exemplo, o gene *ntf4*, que

codifica uma MAPK de *N. tabacum*, é expresso somente em grãos de pólen e embriões (VORONIN *et al.*, 2001). Já em *A. thaliana*, os genes parálogos *mpk3* e *mpk6* têm padrões de expressão diferentes em diversos órgãos, sendo que o primeiro tem expressão diferencial e, o segundo, constitutiva (LEE *et al.*, 2008). Além disso, genes *mapks* vegetais têm a expressão induzida em resposta a estresses e estudos já descreveram o efeito de tratamentos como ozônio, peróxido de hidrogênio, etileno, ácido abscísico (ABA), ácido jasmônico, frio, seca, calor e auxinas. Muitos dos genes descritos são inclusive ativados por mais de um tratamento (conforme revisão por SINHA *et al.*, 2011). A regulação transcricional destes genes é uma forma importante de controle e o estudo de seu padrão de expressão pode ajudar a melhor entender a sinalização de processos celulares e resposta a estresses.

1.2 Estresse vegetal

O estresse vegetal é definido como um fator externo que exerce uma influência desvantajosa sobre a planta e inibe o funcionamento normal do organismo (TAIZ & ZEIGER, 2006). A sobrevivência do indivíduo depende da percepção dos sinais e adaptação a estas condições não ideais. Plantas ajustam-se a mudanças no ambiente pela ativação de cascatas de transdução de sinal que controlam e coordenam respostas bioquímicas e fisiológicas necessárias à adaptação (HUANG *et al.*, 2012).

Diversos fatores podem causar estresse em plantas e eles podem ser divididos em abióticos e bióticos. Exemplos de estresses abióticos são frio, calor, salinidade, seca, excesso de água, radiação, poluição, estresse oxidativo e privação de nutrientes. Já estresses bióticos são causados por patógenos como vírus, bactéria, fungos, insetos, nematoides e herbívoros vertebrados. O estresse pode ser medido com relação à sobrevivência da planta, à produtividade agrícola, ao crescimento e aos processos primários de assimilação (TUTEJA & MAHAJAN, 2005).

Algumas plantas podem ser mais aptas do que outras para enfrentar ambientes desfavoráveis, o que é chamado de tolerância ao estresse. Dessa forma, um ambiente estressante para uma planta pode não o ser para outra (TAIZ & ZEIGER, 2006). Frio, seca e alta salinidade são as condições de estresse que mais comumente afetam o crescimento de plantas de interesse econômico. O entendimento de como as plantas

transmitem os sinais de estresse e conseguem ativar respostas adaptativas é importante para o desenvolvimento de espécies mais tolerantes (HUANG *et al.*, 2012).

1.2.1 Sinalização de estresse em plantas

Uma rota de transdução de sinal pode ser iniciada com a percepção do sinal por receptores de membrana e pela membrana celular. A membrana plasmática tem um papel na percepção e transmissão de sinais externos durante estresses abióticos, principalmente por sofrer alterações em suas propriedades físicas e composição. (HUANG *et al.*, 2012)

Em plantas, há diferentes mecanismos de sinalização que envolvem fosfolipídeos. Enzimas como as fosfolipases C, D e A2 (PLC, PLD e PLA2) e diacilglicerol pirofosfatase (DGPP) são ativadas diferencialmente mediante estresses com diferentes graus de severidade. (MUNNIK & MEIJER, 2001)

Estudos em alfafa, tomateiro e *Chlamydomonas* demonstraram que o estresse osmótico ativa PLD. Em *A. thaliana*, esta enzima cliva fosfolipídeos de membrana e produz ácido fosfatídico (PA) em resposta a estresse por seca (TUTEJA & MAHAJAN, 2005). Também em *A. thaliana*, foi demonstrado que o aumento transiente nos níveis de PA ativa uma MAPK que participa na detoxificação de íons Na⁺ da célula (YU *et al.*, 2010).

O estresse osmótico leva também à sinalização por fosfolipídeos mediante ativação de PLC. Esta enzima cataliza a hidrólise de fosfolipídios de membrana como o fosfatidilinositol-4,5-bifosfato, que gera as moléculas sinalizadoras inositol 1,4,5-trifosfato (IP3), diacilglicerol e PA. IP3 é capaz de alterar a os níveis de Ca⁺² nos compartimentos de armazenamento, e o aumento nos seus níveis já foi relacionado à resposta a estresses hiperosmóticos em *A. thaliana* (DEWALD *et al.*, 2001; revisado em TUTEJA & MAHAJAN, 2005) (**Figura 2**).

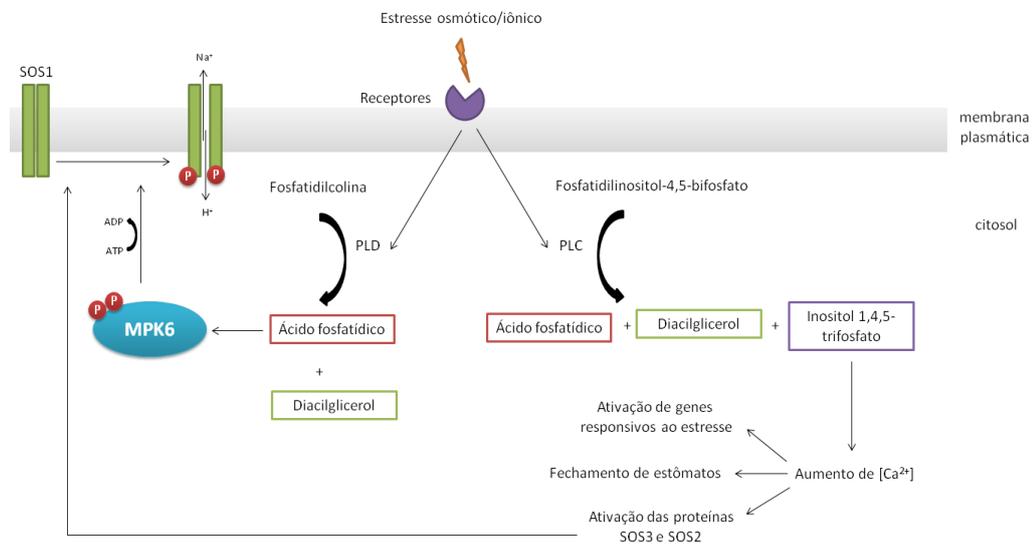


Figura 2: No esquema adaptado de TUTEJA & MAHAJAN, 2005 e YU *et al.*, 2010 são mostrados dois mecanismos de sinalização celular que envolvem fosfolipídeos. Estresse osmótico é sentido por receptores na membrana plasmática, e dependendo da intensidade e duração, há ativação das fosfolipases C (PLC) ou fosfolipase D (PLD). A rota da PLD envolve hidrólise de fosfatidilcolina em ácido fosfatídico e diacilglicerol. Em *A. thaliana*, tratamento com alta salinidade induz o acúmulo de moléculas de ácido fosfatídico por esta via, e este fosfolipídeo é capaz de ligação a formas ativadas de uma MAPK, MPK6. Esta cinase é então estimulada a fosforilar a proteína SOS1, responsável pelo antiporte Na^+/H^+ que participa na detoxificação da célula. A rota da PLC envolve a hidrólise de fosfatidilinositol-4,5-bisfosfato, que gera as moléculas sinalizadoras inositol 1,4,5-trifosfato, diacilglicerol e ácido fosfatídico. Estresse por seca em *A. thaliana* e *Craterostigma plantagineum* causam ativação da PLC e aumento nas concentrações citosólicas de Ca^{2+} em decorrência do acúmulo de inositol 1,4,5-trifosfato. Proteínas ligadoras de Ca^{2+} são capazes de sentir estas mudanças e levam à ativação de genes responsivos ao estresse, ao fechamento de estômatos e à ativação das proteínas SOS3 e SOS2, que por sua vez ativam SOS1. Neste ponto, há uma possível convergência no substrato-alvo das rotas de sinalização via PLC e PLD, já que ambas são capazes de modular a proteína SOS1 e o antiporte Na^+/H^+ (para mais detalhes da via SOS, ver **Figura 3**).

A alteração nos níveis citosólicos de íons cálcio (Ca^{2+}) é sentida por proteínas ligadoras de cálcio que sofrem mudanças conformacionais dependentes do ligante. Estas proteínas interagem com outros substratos e dão origem a cascatas de fosforilação que têm como alvo tanto genes responsivos ao estresse como fatores de transcrição que controlam tais genes (HUANG *et al.*, 2012). Os produtos dos genes responsivos a estresse levam à adaptação da planta e a ajudam a sobreviver em condições desfavoráveis. Um exemplo deste mecanismo são os genes *sos* (do inglês, *salt overly sensitive*) de *A. thaliana*, que formam uma rota capaz de conferir tolerância à salinidade (revisado em JI *et al.*, 2013). Esta rota, que inclui a interação das proteínas SOS1, 2 e 3, leva à exclusão do excesso de íons Na^+ pela atividade de um antiporto Na^+/H^+ na membrana plasmática e pelo sequestro destes íons em excesso para compartimentos intracelulares (HUANG *et al.*, 2012) (**Figura 3**).

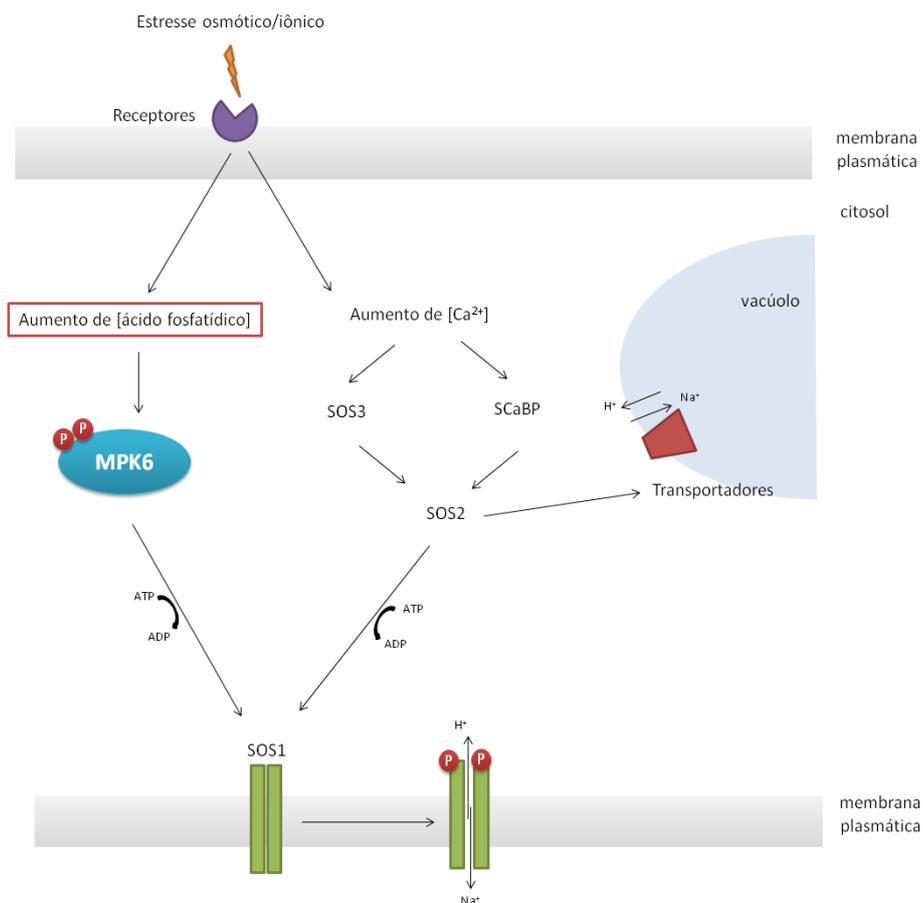


Figura 3: No esquema adaptado de HUANG *et al.*, 2012 e JI *et al.*, 2013 é mostrado o modelo de sinalização por proteínas SOS. Estresse osmótico é transmitido por componentes ainda não identificados para o citosol, onde ocorre aumento nas concentrações de Ca^{2+} . A proteína SOS3, que possui sítios de ligação a Ca^{2+} , sofre mudanças conformacionais que permitem sua interação com a proteína SOS2, uma serina/treonina cinase. Em células de raízes de *A. thaliana*, esta interação libera a proteína SOS2 de um estado de auto-inibição e promove a realocação do complexo SOS3-SOS2 para a membrana plasmática. Já em células de partes aéreas de *A. thaliana*, SOS2 é alternativamente ativada pela proteína SCABP8, e o complexo SCABP8-SOS2 é também translocado para a membrana plasmática. Na membrana plasmática, a proteína SOS2 fosforila e ativa a proteína de membrana SOS1, responsável pelo antiporto Na^+/H^+ que contribui para a exclusão de Na^+ da célula. Outro evento desencadeado por esta via é a fosforilação de transportadores do tonoplasto por SOS2, que causa sequestro de Na^+ para o vacúolo. Conforme já mencionado, em *A. thaliana* foi descrita a participação de uma proteína MAPK (MPK6) na ativação de SOS1, e, desta forma, a rota SOS não ocorre por interação exclusiva entre as proteínas SOS e não é a única via responsável pela detoxificação de Na^+ das células.

Os genes responsivos a estresse podem ser classificados de acordo com a rapidez de sua ativação. Alguns genes de ativação imediata (do inglês, *early induced genes*) levam a uma resposta rápida e são induzidos em de minutos após a percepção do sinal, sendo comumente expressos de forma transiente. São exemplos genes codificadores de fatores de transcrição cuja indução não requer síntese de novas proteínas. Outros genes de ativação tardia (do inglês, *late induced genes*), que levam mais tempo para serem induzidos, são ativados horas após a percepção do sinal e são comumente mantidos na forma ativa por mais tempo. São exemplos alguns dos principais genes responsivos ao estresse, como genes RD (do inglês, *responsive to dehydration*), KIN (*cold induced*) e COR (*cold responsive*), que codificam proteínas necessárias para a síntese de outros compostos, como as proteínas do tipo LEA (do inglês, *late embryogenesis abundant*), proteínas estabilizadoras de membrana, antioxidantes e proteínas para síntese de osmólitos (TUTEJA & MAHAJAN, 2005; TAIZ & ZEIGER, 2006).

As mudanças na expressão gênica causadas por estresse podem induzir a produção de hormônios como ABA, ácido salicílico e etileno. Estas moléculas amplificam o sinal inicial e dão origem a uma segunda cascata de sinalização que pode seguir tanto a mesma rota como caminhos diferentes (TUTEJA & MAHAJAN, 2005).

1.2.2 ABA e estresses

ABA é um fitormônio com papel importante na resposta a vários sinais de estresse, sendo que sua aplicação exógena é capaz de simular o efeito de uma situação de estresse. Como muitos estresses abióticos causam um desbalanço osmótico na célula, há uma sobreposição no padrão de expressão de genes de resposta ao estresse por frio, seca, alta salinidade e aplicação de ABA. Assim, ABA e sinais gerados por estresses compartilham elementos comuns nas suas rotas de sinalização (HUANG *et al.*, 2012) (**Figura 4**).

O estresse osmótico causado por sal ou seca é transmitido por pelo menos duas rotas, uma ABA-dependente e outra ABA-independente. O frio, por exemplo, exerce efeitos na expressão gênica principalmente em uma rota que independe de ABA. A expressão ABA-dependente envolve elementos de ação *cis* da família ABRE (do inglês, *ABA-responsive elements*). Os componentes das rotas que dependem ou não de ABA

comunicam-se ou, mesmo, convergem para um mesmo ponto na rota de sinalização (TUTEJA & MAHAJAN, 2005) (Figura 4).

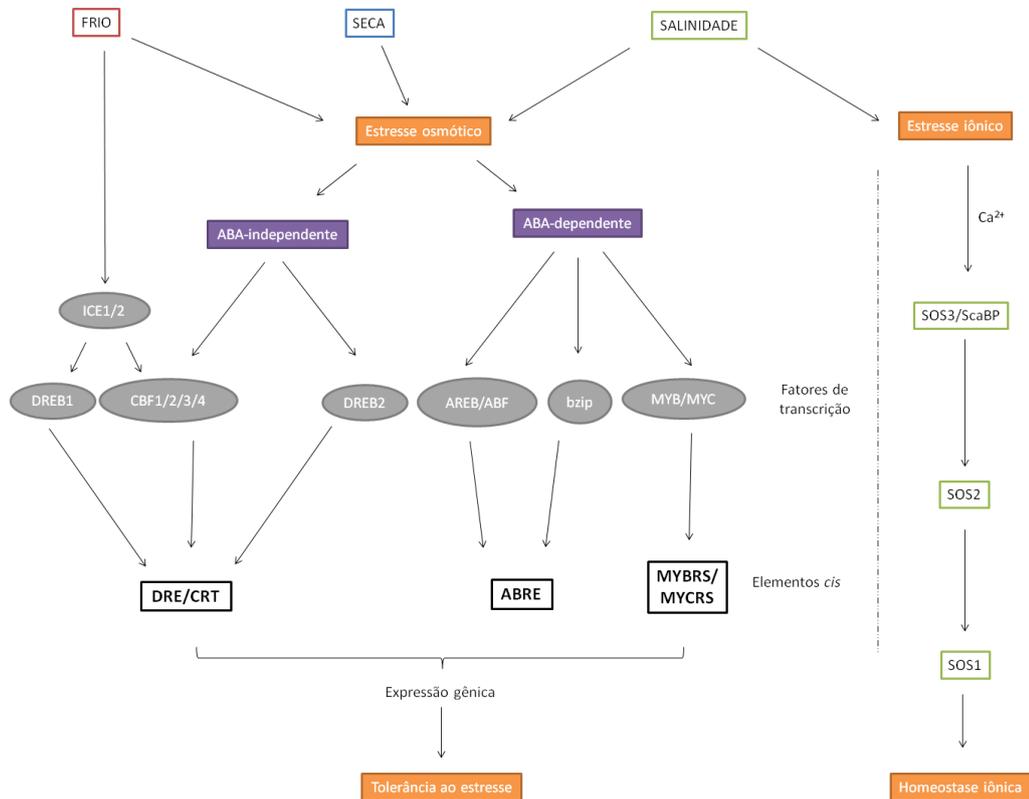


Figura 4: Esquema adaptado de HUANG *et al.*, 2012 mostrando os principais elementos envolvidos na resposta vegetal a ABA e a sinais de estresse. A rota de sinalização leva à ativação de genes e proteínas e tem papel na manutenção da homeostase celular. A resposta a seca, sal e frio é mediada por diferentes fatores de transcrição que ativam genes responsivos ao estresse, e resulta muitas vezes em uma “conversa cruzada” entre os componentes destas rotas e daquelas que envolvem ABA. Tratamentos com frio, salinidade e seca geram estresse osmótico nas células, e este é mediado por fatores de transcrição como DREB2A, DREB2B, bZip, MYC e MYB, os quais interagem com elementos *cis* como CRT/DRE, ABRE e MYCRE/MYBRE e atuam na ativação da transcrição de genes de resposta a estresse. Tais fatores de transcrição são produtos de genes de ativação imediata, e sua expressão é induzida nos primeiros minutos do estresse. Tratamentos com salinidade causam, adicionalmente ao estresse osmótico, um estresse iônico que é caracterizado principalmente pela ativação de proteínas da rota SOS para detoxificação de íons Na⁺ da célula.

A cascata de sinalização por MAPKs está envolvida na sinalização por ABA. Contudo, a sinalização por este hormônio envolve muitos componentes e é bastante complexa. MAPKs já foram implicadas na resposta a ABA em células-guarda e na defesa antioxidante e diversas espécies vegetais possuem genes que são geralmente induzidos por ABA (LIU *et al.*, 2012).

1.2.3 Estresse por salinidade

A salinidade é um dos principais estresses ambientais e representa problemas para a produção comercial de plantas. Este estresse interfere no crescimento da planta por exercer condições semelhantes à seca e pela toxicidade iônica (HUANG *et al.*, 2012).

A exposição à salinidade e a meios hiperosmóticos leva a eventos de sinalização que incluem a formação de espécies reativas de oxigênio, a produção de ácido fosfatídico, aumento transiente dos níveis citosólicos de Ca^{2+} e acumulação de óxido nítrico (SMÉKALOVÁ *et al.*, 2013). Os eventos fisiológicos desencadeados por estresse salino e por seca sobrepõem-se. A deposição de sal no solo gera zonas de baixo potencial hídrico, fazendo com que as plantas tenham mais dificuldade na aquisição de água e nutrientes. Os principais íons envolvidos na sinalização do estresse salino são Na^+ , K^+ , H^+ e Ca^{2+} (TUTEJA & MAHAJAN, 2005) e o estresse salino causa um desbalanço na homeostase iônica, resultando em um excesso tóxico de íons Na^+ no citoplasma e em redução de K^+ (HUANG *et al.*, 2012).

Como o estresse por cloreto de sódio (NaCl) leva à redução de água intracelular, plantas são capazes de acumular metabólitos chamados solutos compatíveis para proteger sua proteínas. Estes consistem em osmólitos como açúcares e poliálcoois, além de metabólitos com carga residual como glicina-betaína e prolina. O acúmulo destas substâncias favorece o ajuste osmótico permitindo que a célula não perca água. (TUTEJA & MAHAJAN, 2005; TAIZ & ZEIGER, 2006)

Tratamentos com sal e meios hiperosmóticos ativam proteínas da família das MAPKs. Em *A. thaliana* e em arroz, foram descritas MAP3Ks que atuam na regulação positiva e negativa da resposta à salinidade e a meios hiperosmóticos. Estas repostas são também desencadeadas por seca e ABA (KIM *et al.*, 2012). Além disso, proteínas

MAP2Ks nestas duas espécies vegetais também estão implicadas na resposta ao estresse salino e têm ação sobre as MAPKs MPK4, MPK6 e MPK3 (TEIGE *et al.*, 2004).

Em *A. thaliana*, estudos comprovaram que o estresse hiperosmótico e o estresse salino são capazes de causar expressão diferencial dos genes *mpk3*, *mpk4* e *mpk6* (DROILLARD *et al.*, 2004; KIM *et al.*, 2011). A proteína MPK6 age na fosforilação da proteína SOS1 e, assim, participa na resposta de detoxificação de sódio descrita anteriormente (YU *et al.*, 2010). Em longo prazo, esta proteína é capaz de ativar genes responsivos a sal mediante a fosforilação de ZAT6, um fator de transcrição do tipo dedo-de-zinco (LIU *et al.*, 2013).

1.2.4. Estresse por seca

O estresse vegetal causado pela deficiência de água (seca) pode causar desidratação das membranas celulares, formação de poros e desprendimento de proteínas de membrana. Esta perda de integridade afeta a compartimentalização celular e compromete a atividade enzimática. A seca limita o crescimento da planta por causar diminuição na fotossíntese, além de o solo seco interferir na disponibilidade de nutrientes para o organismo (HUANG *et al.*, 2012). Outro efeito da seca em plantas é a redução no crescimento vegetativo, que ocorre principalmente em caules e partes aéreas pela diminuição da atividade de ciclinas dependentes de cinases, o que resulta em divisão celular mais lenta. A redução na expansão celular é benéfica para plantas sob estresse hídrico por diminuir a área de folha que é exposta, resultando em menor transpiração (TUTEJA & MAHAJAN, 2005).

O estresse osmótico causado pela seca também faz com que folhas minimizem a perda de água por transpiração pelos estômatos (HUANG *et al.*, 2012). Este processo, que é conhecido como fechamento hidroativo, pode ser regulado pelo hormônio ABA e promove o efluxo de íons K^+ das células-guarda, o que resulta na perda da pressão de turgor e leva ao fechamento dos estômatos (TUTEJA & MAHAJAN, 2005).

Em plantas, a sinalização do estresse por seca induz a produção de fosfolipídeos e a formação de espécies reativas de oxigênio. Como condições de seca são persistentes por períodos extensos, as respostas em longo prazo envolvem mudanças na expressão

gênica. Estes processos são mediados por fatores de transcrição como os WRKYs (SHEN *et al.*, 2012) e do tipo zíper-de-leucinas (TANG *et al.*, 2012).

A sinalização por MAPKs regula a ativação transcricional e as respostas de curto e médio prazo em plantas expostas a condições de pouca água (SMÉKALOVÁ *et al.*, 2013). Em *A. thaliana*, os genes *mpk9* e *mpk12* são induzidos pelo estresse de seca e sua expressão se dá em células-guarda em resposta a ABA (JAMMES *et al.*, 2009). Já a proteína MPK6, que também responde ao estresse de seca, tem sua atividade de fosforilação aumentada mediante acumulação de radicais livres e fosfolipídeos (YU *et al.*, 2010), mas não responde a ABA (TSUGAMA *et al.*, 2012). Como consequência do estresse por seca, as alterações metabólicas da célula incluem também a mudança nos níveis de expressão de genes codificadores de proteínas do tipo LEA/desidrinas, na síntese de chaperonas que protegem outras proteínas e de proteinases para degradação de proteínas desnaturadas (TUTEJA & MAHAJAN, 2005).

1.3 Plantas modificadas geneticamente com componentes das cascatas de MAPKs

Como exposto, os componentes da cascata de sinalização das MAPKs participam em importantes respostas celulares, sendo reguladores da expressão gênica, da citocinese e do desenvolvimento. A atividade destes módulos em plantas sob estresse torna as MAPKs um possível alvo para aplicações biotecnológicas (ŠAMAJOVÁ *et al.*, 2013b).

A expressão heteróloga do gene *map3k* de *N. tabacum* em milho aumentou a tolerância das plantas ao frio, ao calor e à salinidade (SHOU *et al.*, 2004). Adicionalmente, a superexpressão de um gene *map2k* de *A. thaliana* foi capaz de conferir tolerância ao frio e à salinidade. Mutantes de *A. thaliana* que não expressam este mesmo gene *map2k* mostraram maior sensibilidade a estes estresses (TEIGE *et al.*, 2004).

Os efeitos da expressão heteróloga de genes codificadores de MAPKs já foram demonstrados frente a estresses bióticos e abióticos. Um gene de pepino (*Csnmapk*) conferiu tolerância ao estresse osmótico em tabaco (XU *et al.*, 2009). Um gene *mapk* de milho (*Zmmpk4*) expresso em tabaco aumentou a tolerância ao frio (ZHOU *et al.*, 2012), e a expressão do gene *ZmSIMK1* de milho em *A. thaliana* permitiu a geração de

plantas tolerantes a alta salinidade (GU *et al.*, 2010). Foi demonstrado que o gene *Ghmpk2* de algodoeiro está envolvido na regulação da tolerância ao sal e à seca em tabaco (ZHANG *et al.*, 2011), e que *Ghmpk16* foi capaz de conferir resistência a fungos e bactérias em plantas de *A. thaliana* transformadas (SHI *et al.*, 2011).

Assim, estudos sobre a sinalização de estresses vegetais e o impacto da superexpressão de genes capazes de sinalizar tais estresses têm sido utilizados para o desenvolvimento de plantas transgênicas com características de tolerância bastante promissoras. O entendimento das interações nas rotas de sinalização, da regulação e dos produtos dos genes vegetais codificadores de MAPKs pode ser aplicado na produção de plantas transgênicas capazes de tolerar condições de estresse com minimização de perdas de crescimento e produção (TUTEJA & MAHAJAN, 2005).

1.4 Importância comercial e melhoramento de *Eucalyptus*

O gênero *Eucalyptus* possui 900 espécies nativas da Austrália, local onde o clima varia desde condições úmidas e temperadas até zonas quentes e áridas (BEDON *et al.*, 2012). A boa adaptação destas espécies a diferentes condições de clima e solo se dá pela grande diversidade genética dos indivíduos, que inclui a expressão de genes possivelmente envolvidos em processos adaptativos (conforme revisado em VILLAR *et al.*, 2011). Algumas espécies de *Eucalyptus* estão entre as árvores que apresentam mais rápido crescimento em áreas tropicais e sub-tropicais, e esta característica, aliada ao seu bom conteúdo de fibras, torna estas espécies importantes fontes de matéria-prima para as indústrias de polpa de celulose e papel (VALDÉS *et al.*, 2013).

Dados de 2008 estimaram que plantações de *Eucalyptus* ocupavam mundialmente cerca de 20 milhões de hectares distribuídos entre Austrália, África do Sul e países da América do Sul, da Ásia, e do sudoeste da Europa (BEDON *et al.*, 2012). Durante o ano 2012, a área plantada com *Eucalyptus* chegou a 5,1 milhões de hectares distribuídos em 17 estados brasileiros, o que representa cerca de 70% da área de plantio florestal do país (ABRAF, 2013). No Brasil, a produção de *Eucalyptus* é majoritariamente empregada na indústria de papel e celulose (72,5% da área plantada), tendo também aplicação na siderurgia a carvão vegetal e fabricação de painéis de madeira. Dados de 2012 colocaram o Brasil entre os maiores produtores mundiais de

celulose (4º lugar) e papel (9º lugar) e estes níveis de produção, que contam com outras fontes, têm também a contribuição das florestas de *Eucalyptus* (BRACELPA, 2013).

Nos últimos 40 anos, programas de melhoramento de espécies de *Eucalyptus* focaram-se no aumento da produtividade pela seleção e propagação clonal de plantas com maior produção de biomassa e conteúdo de polpa. No entanto, a expansão das plantações para ambientes que imponham desafios ao desenvolvimento das plantas, e mesmo mudanças climáticas em áreas que já abrigam plantações, são importantes fatores a serem levados em conta no melhoramento (NAVARRETE-CAMPOS *et al.*, 2013; VALDÉS *et al.*, 2013). Fatores como frio, seca e salinização do solo podem prejudicar o crescimento e a produção destas árvores, e, portanto o entendimento dos mecanismos de tolerância aos estresses é fundamental para o desenvolvimento de indivíduos tolerantes (BEDON *et al.*, 2011).

Ao contrário dos muitos estudos em plantas modelo, pouco se sabe ainda sobre o mecanismo de adaptação ao estresse em *Eucalyptus*. Trabalhos elucidaram respostas morfológicas e fisiológicas (PITA & PARDOS, 2001; COSTA E SILVA *et al.*, 2008), e, mais recentemente, estudos de proteômica e transcriptômica buscaram possíveis elementos que participam na resposta ao estresse hídrico (VILLAR *et al.*, 2011; BEDON *et al.*, 2012; THUMMA *et al.*, 2012).

A espécie *Eucalyptus grandis* é uma das mais difundidas em plantações florestais no Brasil, sendo inclusive importante para a geração de híbridos com alta produtividade e boa qualidade de fibra (SOUZA *et al.*, 2004). Assim, considerando-se a disponibilidade do sequenciamento do genoma desta espécie, é possível a busca por genes de interesse ao melhoramento. Dada a importância de MAPKs na sinalização de estresse em vegetais, o estudo mais aprofundado do gene *Egmpk5* de *E. grandis* poderá contribuir para o entendimento do mecanismo de adaptação desta planta a condições adversas do ambiente.

2. Objetivos

2.1 Objetivo Geral

Caracterizar o gene *mpk5* de *Eucalyptus grandis* e o efeito de sua expressão heteróloga em plantas

2.2 Objetivos específicos:

- Com base na sequência de *mpk5* de *E. grandis*, obter sequências de outras *mapks* em bancos de dados e realizar análises filogenéticas;
- Caracterizar a estrutura e determinar o número de cópias do gene *Egmpk5* no genoma de *E. grandis*, identificar a localização cromossômica de membros de uma possível família gênica *mpk5* e analisar potenciais regiões promotoras;
- Caracterizar o padrão de expressão gênica de *Egmpk5* em raízes, caules e folhas de *E. grandis* em condições normais e em respostas a diferentes sinais por RT-qPCR;
- Caracterizar o estado transgênico de plantas de tabaco transformadas com o gene *Egmpk5* por PCR;
- Caracterizar o padrão de herança de *Egmpk5* em progênies de plantas transgênicas de tabaco selecionadas;
- Desafiar as plantas transgênicas de tabaco frente à salinidade e à seca e caracterizar as respostas.

3. Materiais e métodos

3.1 Análises *in silico*

3.1.1 Avaliação do número de cópias e da estrutura gênica de *Egmpk5*

A sequência gênica estudada neste trabalho partiu de um estudo anterior de nosso grupo, que isolou o cDNA de *Egmpk5* a partir de ESTs (do inglês, *Expressed Sequence Tags* ou *ESTs*) do banco de dados do projeto GENOLYPTUS – Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*. O clone de cDNA selecionado apresentou 5.068 pb, foi completamente sequenciado e teve sua expressão parcialmente caracterizada frente a diversos estímulos (KÖRBES, 2006).

Neste trabalho, a sequência do cDNA de *Egmpk5* foi utilizada para BLASTn (*nucleotide BLAST*; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) no banco de dados *Phytozome* (*Eucalyptus grandis* *Genome Project*, <http://www.phytozome.net/eucalyptus>) a fim de se obter informações sobre a localização cromossômica deste gene. Os resultados desta busca foram dados em termos de escores de alinhamento com sequências genômicas e valores estimados ou *E-values*.

Para determinação da estrutura gênica de *Egmpk5*, as sequências obtidas a partir do genoma foram alinhadas com ClustalW (THOMPSON *et al.*, 1994). Alinhamentos manuais e comparação com sequências similares de outras espécies descritas na literatura também foram empregados nesta etapa e, assim, foi determinada a organização de íntrons e éxons deste gene.

Tradução e manipulação de sequências nucleotídicas foram feitas com as ferramentas do *Molecular Toolkit* (disponível em <http://www.vivo.colostate.edu/molkit/index.html>) e representação gráfica da estrutura gênica foi feita com criação manual de escala e desenho no *software Microsoft Office PowerPoint 2007*.

3.1.2 Construção de mapa cromossômico

Dados relativos ao tamanho de *scaffolds* montados a partir do sequenciamento do genoma de *E. grandis* foram obtidos no servidor do banco *Phytozome*. A localização no genoma da possível sequência de *Egmpk5* obtida por BLAST foi também pesquisada. Estes dados foram inseridos no programa *MapChart* versão 2.2 (VOORRIPS, 2002) para criação do mapa.

3.1.3 Análise da região promotora do gene *Egmpk5*

Da sequência genômica selecionada a partir do resultado obtido por BLAST com cDNA de *Egmpk5* no *Phytozome*, 2 kb anteriores ao códon de início de tradução foram selecionados para análise de elementos *cis* promotores. Estas análises foram realizadas pro comparação aos dados disponíveis no *PlantCARE* (LESCOT *et al.*, 2002), que reúne motivos conhecidos de elementos regulatórios de ação *cis* em plantas.

3.1.4 Sequência peptídica deduzida, análise filogenética e de domínios de EgMPK5

A sequência peptídica deduzida (EgMPK5) foi obtida pela tradução *in silico* da sequência do clone de cDNA de *Egmpk5* após remoção manual das sequências de íntrons. Esta sequência resultante foi utilizada como isca na busca por proteínas similares em outras plantas por BLASTp no GenBank. Sequências peptídicas de MAPKs de plantas com no mínimo 80% de similaridade à EgMPK5 foram utilizadas para construção de uma árvore filogenética no programa *MEGA5* (TAMURA *et al.*, 2011) pelo método de Máxima Verossimilhança (*Maximum Likelihood*), utilizando o algoritmo *MUSCLE* (EDGAR, 2004) para alinhamento das sequências e um valor de *bootstrap* de 500 repetições.

Domínios conservados na sequência peptídica foram encontrados pelo emprego de BLASTp no GenBank e a análise dos subdomínios de STKs foi feito por comparação com a literatura e alinhamentos de sequências proteicas com ClustalW. Representação gráfica das sequências consenso obtidas foi feita com o programa *WebLogo* (CROOKS *et al.*, 2004).

3.2 Análises da expressão de *mpk5* em *E. grandis*

3.2.1 Material vegetal e tratamentos

Sementes de *E. grandis* foram plantadas em copos plásticos contendo a mistura solo/vermiculita na proporção de 1:1. As plantas foram mantidas em câmara de cultivo sob fotoperíodo de 16 h de luz e 8 h de escuro a 28 °C, sendo irrigadas duas vezes por semana. Para avaliação da resposta a tratamentos, plantas de 90 dias, em triplicata, foram irrigadas com 20 mL de soluções de ácido naftalenoacético (NAA) a 50 µM, ácido abscísico (ABA) a 100 µM, cinetina (kin) a 10 µM e cloreto de sódio (NaCl) a 200 mM. Para o tratamento de seca, três plantas foram mantidas sem irrigação. Após 3 dias dos tratamentos (ou 7 dias no caso da seca), raízes, caules e folhas foram coletados separadamente, congelados em nitrogênio líquido e mantidos a -80 °C até extração dos ácidos nucleicos.

3.2.2 Extração de RNA total e síntese de cDNA

A extração de RNA total foi feita pelo método CTAB (DOYLE & DOYLE, 1987) e a avaliação da qualidade das extrações foi realizada por inspeção visual de bandas características de preparações de RNA total íntegro em gel de agarose a 1% corado com brometo de etídeo após eletroforese e quantificação por método espectrofotométrico.

Amostras de RNA foram tratadas com *RQ1 RNase-Free DNaseI* (Promega) para eliminação de DNA contaminante segundo as instruções do fabricante. A síntese de cDNA foi realizada com a enzima transcriptase reversa M-MuLV (New England Biolabs) conforme instruções do fabricante.

3.2.3 Avaliação da expressão relativa de *Egmpk5* por RT-qPCR

Reações em cadeia da DNA polimerase quantitativas ou em tempo real (qPCRs) precedidas de transcrição reversa (RT) foram realizadas no equipamento *ABI 7500 Real-Time PCR System* (Applied Biotechnology) utilizando-se o agente fluoróforo

SYBR Green (Invitrogen) como corante. As reações foram conduzidas em placas de 96 poços contendo 20 µL de volume total, sendo metade composta pelas preparações de cDNA na diluição 1:25 e a outra metade composta por uma mistura de reação. A mistura de reação foi composta pelos seguintes reagentes: *primers* direto (*forward*) e reverso (*reverse*) a 0,2 µM cada, desoxirribonucleosídeos trifosfatados ou dNTPs (dATP, dGTP, dCTP e dTTP) a 25 µM cada, 0,25 U da enzima *GoTaq DNA polimerase* (Promega) e tampão de reação fornecido (Promega) na concentração final de 1x, cloreto de magnésio a 3 mM, solução de *SYBR Green* (Invitrogen) na concentração final de 0,1x, e água purificada por osmose reversa (MiliQ) e autoclavada.

Os pares de *primers* utilizados nas RT-qPCRs foram específicos para os seguintes *amplicons*: (i) região de 66 pb do gene *Egmpk5* (direto 5'-GATAACAAGATCGACGCGAA-3' e reverso 5'-TCTGTGGCGGTGGAATAATA-3'); (ii) fragmento de 152 pb do gene *Ec08* (fator de alongação de transcrição s-II; direto 5'-TCCAATCCGAGTCGCTGTCATTG-3' e reverso 5'-TGATGAGCCTCTCTGGTTTGACCT-3'); e (iii) fragmento de 141 pb do gene *Ec04* (cdk; direto 5'-TACAAGCGCTGTTGATATGTGGGC-3' e reverso 5'-TTGCCAATGAGGCGGATTCACAAG-3') (OLIVEIRA *et al.*, 2012).

As condições de reação foram as seguintes: ativação da enzima por calor (*hot start*) a 94 °C por 300 s, seguida por 40 ciclos de desnaturação (94 °C por 15 s), anelamento (60 °C por 10 s) e síntese (72 °C por 35 s). A leitura da fluorescência foi realizada a 60 °C por 35 s em cada ciclo. A curva de dissociação (T_m) foi obtida por aquecimento de 50 a 94 °C (variação ou Δ de 0,4 °C/s).

As RT-qPCRs para cada gene foram feitas em quadruplicatas, sendo sempre considerada uma reação em branco, composta apenas pela mistura de reação e água ultrapura.

Para análise da expressão relativa de *mpk5* foi utilizado o método $2^{-\Delta\Delta Ct}$ (LIVAK & SCHMITTGEN, 2001) e a análise estatística foi realizada pelo Teste T de Student assumindo-se uma distribuição de dados bicaudal e variância independente entre as amostras comparadas, utilizando-se o programa *Microsoft Office Excel 2007*. As diferenças foram consideradas estatisticamente significativas quando o valor de probabilidade foi $P < 0,05$.

3.3 Caracterização de plantas transformadas de *N. tabacum*

As plantas de *N. tabacum* SR1 transformadas com o gene *Egmpk5* estudadas nesta parte do trabalho foram oriundas de experimentos anteriores de nosso grupo (KÖRBES, 2006; KIRCH, 2008). Para tanto, a sequência de cDNA de *Egmpk5* foi adaptada ao promotor 35S do vírus-do-mosaico-da-couve-flor (CaMV 35S) e ao terminador da nopalina sintase (*nos*) no plasmídeo pMOG463 (Mogen). O cassete de expressão assim constituído foi transferido para o vetor binário pCAMBIA2300 (CAMBIA), utilizado para a transformação de *Agrobacterium tumefaciens* LBA4404. Plantas de *N. tabacum* SR1 foram transformadas com o método de cocultura de discos foliares (DEBLOCK *et al.*, 1984; HORSCH *et al.*, 1984) com a linhagem resultante *A. tumefaciens* LBA4404::pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos}.

No presente trabalho, a caracterização de seis plantas de tabaco regeneradas foi realizada, tanto molecularmente como em bioensaios.

3.3.1 Assepsia de sementes, meios de cultivo, germinação e cultura *in vitro* de *N. tabacum* SR1

Todas as manipulações de meios de cultivo e materiais vegetais para cultura *in vitro* foram realizadas em cabine de fluxo laminar horizontal. Sementes transgênicas e controles de tabaco foram submetidas à assepsia superficial mergulhando-as em cerca de 50 mL de etanol a 70% por 2 min seguido de 50 mL de hipoclorito de sódio a 1,5% por 20 min e 4 lavagens em 50 mL de água destilada estéril.

As sementes foram dispostas na superfície de meio MS (MURASHIGE & SKOOG, 1963) contendo metade da quantidade de sais (½MS; Sigma) acrescido de ágar a 1,5 % em placas de Petri ou frascos de cultura. As placas contendo sementes foram incubadas em câmara de cultivo tipo B.O.D a 25 °C no escuro por 2 dias seguido de incubação sob fotoperíodo de 16 h luz e 8 h escuro na mesma B.O.D. Para testar a viabilidade das sementes após assepsia e para controles, o meio ½MS foi utilizado sem adição de outras substâncias. Nos testes de tolerância à salinidade e de seleção em canamicina, foram adicionados, respectivamente, NaCl para concentrações finais de 100, 200 e 300 mM e o antibiótico canamicina para concentração final de 100 mg/L (Kan¹⁰⁰).

A taxa de sobrevivência das plantas germinadas foi acompanhada e registrada fotograficamente a cada 5 dias por até 30 dias. Para definição das plantas resistentes à canamicina, foi avaliada a capacidade de germinação das sementes e, em caso de germinação, a capacidade das plântulas manterem seu desenvolvimento.

3.3.2 Cultura hidropônica

Para ensaios de tolerância à seca, plantas transgênicas e controles de tabaco mantidas *in vitro* sob seleção (ou não) com Kan¹⁰⁰ foram fixadas a placas de isopor suspensas em bacias plásticas contendo meio ½MS e aclimatadas durante uma semana. As plantas foram mantidas em sala de cultivo a 25 °C sob fotoperíodo de 16 h luz e 8 h escuro. Para teste de tolerância à seca, as plantas foram retiradas do meio hidropônico e suas raízes deixadas expostas ao ar ambiente. Após 4 dias, as plantas foram recolocadas em meio nutritivo e acompanhadas com registros fotográficos até o 36º dia de ensaio.

3.3.3 Detecção de transgenes em plantas transformadas por PCR

Extração de DNA de tabaco foi realizada pelo método CTAB (DOYLE & DOYLE, 1987). PCRs foram usadas para confirmação do estado transgênico das plantas de tabaco. Foi utilizado o par de *primers 35S forward* (5'-CGACACTCTCGTCTA CTCCAAG-3') e *MAPK-RT reverse* (5'-TCTGTGGCGG TGGAATAATA-3') para amplificação de um fragmento de 864 pb incluindo parte da região promotora e parte do transgene do cassete P_{35S}-*Egmpk5*-T_{nos}. Nas PCRs, foram usados 75 ng de DNA de plantas de tabaco e 10 ng de DNA plasmidial da construção pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos} como controle positivo. As reações foram constituídas por 10 µL de Tampão PCR 5x (Promega); 2 µL de MgCl₂ a 25 mM (Promega); 2 µL de solução de dNTP's (dATP, dGTP, dCTP e dTTP a 5 mM cada); 0,25 µL de *GoTaq DNA* polimerase a 5 U/µL (Promega); 2,5 µL de cada *primer* a 10 µM; amostra de DNA e água ultrapura para completar um volume final de 50 µL. O programa básico de termociclagem incluiu uma etapa a 96 °C por 3 min, seguida de 35 ciclos em temperaturas de 96 °C, 50 °C e 72 °C por 30 s cada, extensão final a 72 °C por 5 min e armazenagem a 4 °C até retirada do termociclador para gelo ou armazenagem em congelador (-20 °C).

4. Resultados

Os dados a seguir apresentados são oriundos de atividades para caracterização do gene *mpk5* de *E. grandis* nesta planta e também do efeito de sua expressão em plantas de tabaco. O presente trabalho baseia-se em trabalhos anteriores de nosso grupo e utiliza a sequência de cDNA de *Egmpk5* clonada de *E. grandis* por KÖRBES (2006) para estudos mais detalhados de estrutura, padrão de expressão e produto deste gene. Ainda, as análises de expressão heteróloga são baseadas na caracterização de seis plantas de tabaco transformadas previamente (KÖRBES, 2006; KIRCH, 2008). Desta forma, os resultados obtidos estão apresentados em três grandes tópicos que abrangem análises *in silico*, caracterização do padrão de expressão gênica de *Egmpk5* em *E. grandis* e caracterização de plantas de tabaco com expressão heteróloga de *Egmpk5* frente a estresses.

4.1 Análises *in silico* do gene *mpk5* de *E. grandis*

4.1.1 Definição do número de cópias e da estrutura gênica de *Egmpk5*

Segundo o banco de dados *Phytozome* (<http://www.phytozome.net/>), o genoma de *E. grandis* possui 691 milhões de pares de bases. Neste banco, as sequências foram agrupadas primariamente em *contigs* ou regiões contíguas que representam sobreposição, ou consenso, de sequências de DNA. As regiões contíguas foram organizadas, finalmente, em grupos de ligação, levando à montagem de *scaffolds* e cromossomos. A montagem do genoma de *E. grandis* traz, em sua versão 1.1, onze principais *scaffolds*/cromossomos, que representam 93% do total de *loci* com transcritos codificadores de proteínas.

A busca por BLASTn neste banco de dados revelou duas sequências genômicas com grande similaridade ao clone de cDNA de *Egmpk5* (KÖRBES 2006): uma delas encontra-se no *scaffold 7* (posições de 28.429.209 a 28.436.392) e tem máximo *score* de alinhamento de 3.152 *bits*, e a outra sequência encontra-se no *scaffold 12* (posições de 345.288 a 352.092) e apresenta um *score* máximo de 2.008 *bits*. Para ambos os resultados, o *E-value* foi de zero. Outras sequências obtidas por esta busca não foram

analisadas mais detalhadamente porque não tiveram boa cobertura de alinhamento, o que foi verificado por valores de *score* muito menores, bem como maiores valores de *E-value*.

As duas sequências selecionadas do banco de dados *Phytozome* foram mais amplamente analisadas para determinação do número de cópias do gene *Egmpk5* no genoma. Primeiramente, estas foram alinhadas à sequência do clone de cDNA de *Egmpk5* e, por comparação com sequências nucleotídicas de *mapks* já identificadas em outras espécies, foi possível designar possíveis éxons e íntrons nas estruturas dos genes (Figura 5).

```
a)
>mpk5_scaffold7
TTTGCCAATTGAGGAGCATGATAATGATGTCGTAAGCTGTATCCCCAAGGAAGATTCTCTCCGTCGGCAGATCCAAGGGCGCGAATGC
AGTCCGAGTCCAATCGGCCCTTATTGCGTGTGGAATGCATGCCTGGATTAGGTTTGTAGAAGTTGTTGATATATCTCGGGTAGCGTGT
TTAGAGTCCAATAATAATCTTGTCCATCATATATGTTTGCCTTCATTCTTAGTTCCTCATGGGGTTAAGTACCAAAAAAAGTCAACT
TGTATATATTGATACTAATTTAGTTATAAACTTTTAAATGTGTTAATTTAGTTCCTAATATTTTATATTAATATCAATTCAGTT
TATTCGATCAATTTAATTTGAAATTTATTGATATGAATACGACCATGTATTAATTTTAGATTTTTTTATCATTTTTTATTAACAT
TTTTAACTTTTTCCCTTTTTCTTGTGTTTCCATTTTTTCCTTTTCGTTTTTCTCTTCTTCCCTTGGCTTGTGGCCGATACTTGC
TAGATTTGGTGAGGGCAACTCTCGCCTATAGCCATTTGCTGGCCATGTTGAGTACCAGCCACAGCCACAGGCCAATGAAAAAGAATG
AAAAGAAATGAAAAAACAACAAAAAAGGAAATTTAAAAATTAATAATAAAAAATTTGAAAAAAATAGTAAAAATTTATAAAAAA
AGTTGTCTACATTAGTGTACGTAAGATGTTTAGCGTTCACGTCAGCAGTTCCCGTCTAAATTAAC TAGATGGACTCAATTACATTAA
TGTA AAAAGATTTATTATTGAATGATTTAGTTGAAATGTTAAGATTGATTTAGTATCAATGCATTGATTA AAAATTTTTTTTGATA
CTTTCTCTTCTTTTCATGATATCATTTATTAGGAGACGACTTATAAAAAGGTTTCATCAATTAGAAAGTATTTGGAAGGGGATTGA
TTGGAGTATAAATCTGATTTACCTAGAGACCAATCCAACCACTCTATCTTCTTCTCGTCTGGTGCAGTTGACATGATTG
GAGCCATTGACCCAATGACCCATCTCTCTTACTATAAATTAGATTTCATAAATGCAAAATGAGCTTTCACATATGTAAC TAGAACA
ATTTACACAGTGATTAAGTTATAACGAAGTGACACAGAAATATATCATAAATCAACGACGACAACCTTCGAGAAGTCATTATAAAT
CTCCTGAACTCTGGCACAATACAAAATTAGCCCTCAAGCTTATTTCTATCGCGGGCCACGATCTTTCTCAGTATGGATGGCAAT
TCAAGTGCATCCAGAAAAACAAACCTAAACCTCATTCTTCTCGACCGAGTTCTTCTCACCGGAGATGATCAGCCGAGACGGCCGAG
AGGAATATTTGACAAAGTACGAGCGAATAAGATCGCTTCGGCCTGAACAGTCAATTCGAGAGGTGGATATAGAGCCGCCCTAATCCT
TAAATTAGCATCTTTTTTCGGATGACATTCTTTAATCAATTA AAAAAGTCAAAAAGAAAGAAAGAGAAGGAAGAAAGAAAGAA
AAGGGTCAAAAAAAGAAAGAGGGGAGAAATACAGGTGGTGAATCATCACAGGGCAGATGAACAGCAGCAGCAGCGGGCGCAGCA
GAAAGCAACAAAAACAGGGCGAAAGGAAAACGACCTTTCCCACTCAAAAGCTCCTCCCTTTTTCATTTGATTTCTGATTTCCACACA
GGCACAGGGAGACAAGGGGACCGAGTCAGTGAGTCGGAACGGTTGACCGGGAATCTCCCCCCCCAACAAAAGCCATCGCCCACTC
AGGCCAACAGTCAAACCCACCTTAACCGAATTCGCCAGATCAATCCCTCTCTCTCTCCATGTAACGGAATCGCGACCCCA
AATCTAGGGCTTTCTTCTCTCTTTCTCTCTATTTCTACCACCACCATCACACCACCGATGACGGCGGGGCTCTCAGCCG
CCGATACCGTCAATGTCGGAGGCGCGCCGGCGCAGCAGCAGCAGCAGCGCAGCGCAGCCGAGCCGAGCCGAGCAACATCCCGG
CGACGCTCAGCCACGGGGCGCTTCATCCAGTACAACATCTTCGGCAACATCTTCGAGGTCACCGCCAAGTACAAGCCCCCATCAT
GCCATCGGCAAGGGCGCCTACGGCATCGTCTGgttaggtgtcggagaagctcgggcttgctgctgataatgcatcatcgacttgata
tagctgcttgctggttaatttctctcatgagcgtatgctaaagtcttcttttttatttgcttttgattgggttgctctctttaa
tccatctctctctggtttcagCTCAGCTTTGAATTCGGAGACGACAGCAGCAGCTGGCCATAAAGAAGATTGCTAATGCTTTCCGATAACA
AGATCGATGCGAAGAGGACTCTCCGTGAGATCAAGCTTCTCCGGCACATGGACCATGAAAACgtcagtgatccccatgctgcaagaat
tttgctgcttacgtgattgagcatggctttctgggttggttgaagtgtgtgtttctcttgatgggtggctatgggtgtggcggtg
ccatcgatattgtgcaatggtgcatttacgttgctctctgctcagttctaaagtcaacgaatgtgcagtaacaggtgaattagcatt
cctgctgtggagaacttaagagttagtgacgaaatagtagagctgtgaatcttctatttcaactagaagggtattgattgctt
aagagaatcagaacctcccatttattataggtaccatgctcagccagttgttaatgagacagagaacaactgaaacgcccaggtta
atattattgcatcatgcatcaagatatagaaggtcttattgtcactgtgctacttacagagttgaacgtcccttctttgaaactttca
gttcatatgtcttttaagtgctggtgctaggaacacgggaagatctctgaaagatctagtgagaagacttttgctatctggttcaag
tagttaaacttaacagatcgactgtcaagtatagctggctaaatagctgtgtttgctatcttttccacatcacaacac
tttcgtaggtgaaatattagaagtcctcaagcagatgattgttttggtaggttttaattgaaattgcattttccatgaaattttta
aatctttatagtttgagctgagtaaatgttgataggtgacttggttgaattatagcttttggtggttagaaagtaacagcaaatg
aaaaggaagtggtacgtatgggtaagtgtgtattttattttgtttgacagtttttaacatttctagaaggtcaccctcgtgtgtt
ttaaatacatgatacagaattttacctcattctctctctctctctctctctctctctctctctctctctctctctctctctctctc
tctctcactcttattttctagataattcacatttctctctctctctctctctctctctctctctctctctctctctctctctctctct
gtacgatcccatttcagtcgaaataaatgatagacatggtatttgggtatcattgatgactattctatggtggaggcaaaaattctc
aaaagtttaattctattatactgatatagaagggcactgatactcatcattcttagagatggaaataggtggttatatcctgactgga
tgcaatttgtaattggtatgaaagtagtcttatttggtataaagaacaagtttgaaactatgattggtgagaatccacatgacagtgga
ccttaaatgactacatttattaaaggcgttggtctcctgcagatacaaaaataagctctttaaatcaatgcagaactggtattgcatg
gcaattttaaagaaacttgctttcagttattatgattcttctgtatgcttattgtgaggttctggtattgtaattgcagTTCGTGGCAAT
AGGGATATTTCCACCGCCACAGAGAGAGGTGTTCAATGATGTTTATATTTGCATATGAGCTTATGGACACTGATCTGCATCAAAATTA
TTCGTTCCAACCAAGCATTTGCTGAGGAGCATTGTCAGgtttcattttgccaatttactgtacgatattggcgcttctctctagatggg
atattgcttctctgctagctagctggttttaattatgctagcgggagttgcatccgctgcttctctgctattgctggtgcttatt
```

aggttattgggaattgacccaaaaaaagttattgggaactttcaaataaaactctgtatcatagtaagaagccaagagagcaagat
catacaactcttagctgtcaaatatcgaaacgggatattacaataattttttttttttctgcatgtacagaaaataaagggggcacta
aatagctggggatcatatctaaactagttattccatatagttaactgttggctggaatttcatccagccaactcctgggataaaacatt
gaaagtggtcttctatggcaattatataatcatcaattaaaagataattagagctacggagtaataagttggagacttctctacggg
aatatgaaccgagactttatgggtcttatacaaatgattatgttagtgatagtttagctgtaaaatagtagcgggagatcattattttaga
cacaatttaaggatctttatggaaattattaggccagaataaatttgaaccggtggggtgatgaagaatagtaaaagaatacttct
aggtttgcttatttatgggaacttgatgatagagactgcttttccattctgctacatggcattgtagcagtatccatttctgttctt
tgcttgatagaagatgcatgaaggtatcatcataaattgatccctctttttctacaacactagaattgaaatgctgacgggtgctg
atcatttagcagatatactgctcatataaattataactcaagtcactgcatgtggtgctgatgactatccatttctgtctttttg**ca**gTAT
TTTCTATATCAGATCTTGGCAGGATTAATAACATACATCTGCAAAATGTTCTGCATAGAGACTTGAAGCCAGCAATCTTCTCTTAA
ATGCCAAATTGCGATTTGAAAATATGTGATTTGGACTAGCTCGTGCACCTTCTGAAACTGATTTTATGACAGAATATGTTGTCACAAG
ATGGTACCGTGCACCAGAGCTATTGTTAAATCTTTCAGACTATACGGCGGCAATAGATGTATGGTCTGTAGGCTGTATCTTTATGGAA
CTAATGGATCGGAAACCCTTGTTCCTGGCAGAGACCATGTGCAACAGCTGCGTTTGTGATGGAG**G**taacttttaagaagttgatta
ggtgtgaaacagttctcttggagaagagtttcttccatgtcgctcttgcattttttgttttctcgaaattttgcccgttgggag
ctcttctgttatttggtaacttaaaccaaaaaactgcaactcaagatgcttctagataggcgttctcaactttcaactttctct
ttgcaccatttaactatttattgtgtcttctcactggagacggatttgatctagctctgttagaggttatgatgtgacatctggtagttatgc
ttgtgtgacctt**ca**gCTGATTTGGCACCCCATCAGAGGCAGAGTTGGGGTCTTAAATGAAAATGCTAAGAGATATATCAGACAGCTT
CCTCTGTACCGTCCGCAATCTTCACTGAAAAGTTTCCCCATGTCCACCCGCTTGAATCGATCTCGTTGAGAAGATGTTAACGTTTCG
ATCCCAGGCTGAGCTCACAG**G**tgcgaaaatgttctctgactttctgaagcatttctgcatgggctcaatttctgttcttctt
ctatccagctgtctgtgaaattgactgggtgctgctgactgcaactgctattgatattgaagagtagatattgacacagattt
gctcaatttctgcttggactgtgatgttcttccatctgcatctgcttcttctctgtgtggtattcatttctgttttgaagt
tgtaaatgcttttaggctgtaataatcttctgctctatctcagggatcaggttatgctaagggtgctaggatgtagcaaaaatagtg
gctgttattgtgagatcgataatgatctgtttagtttagttagaactcttctgctgggttaaagatgagttgggagtatttctac
ttaacttctgctcttctgctgctgtatgaagttatttcttctgctgcttctcaatttctgctcagtgataattcatttagggtcctat
cattcagaagaagtgaaacttctgaagatgtcaatgctcctgcccgtcaggttagtgcttggcatgtgatttggcaagtgccgaa
ccgactgcccctatttagttctgaaactctggcaatgaacaataaattagtagatagaacagggttcaaaatggaggttcatgca
gatcatttattctcatttcttactggtagaatgcttagagagatcttctttaaataaaactgagacacaccatt
tgcatctataatgaagaacccatcttctatgcttgaagcttgggtcatgtgctgagcaagagagataccatgctgattggtctt
gccaagtcgcttctgggtagatggggaggtgactgtctctgtgctcatcttggccagtttctgtcattaaatagttctcaa
gtcagaattgattttctgttggcctacagctctatttgaacgcatgaaaaccgataaaacaaccagtaattgctccttattctgt
tcaaaattcccccaaaaaatgtaataccaaaacttttcttcttcttgaacttgaagatcaataattcagtgataacttgcataaa
aagagtatacaactccaaagaattgcaagcttctattatgcaatcgataccaactctgttgataggattgtagcagtgctgcaa
gcaactgtgaaatgctcagctcaacagcagtgtaagcagctctcaactagtaagtgatctacaactcctcaatggtgggctggaact
caaaaatacataaataatagacatcactatagtgaaatttctatgctataaaatttctgacttcccttaagctagtgactgcaaa
ataggtataatagacataaattggttagatagatagatgcaagcttggactgctagggagccaacatagcagatagatgact
agtatacagacggcaaatggggccttataatgaaactaaattagtcatttattgatagtcactagcatagtaaacactacagaatgc
ataagttattctctgctgagagaatgctatgctgcaaatataatgctgctatgctagaagaggttagagaatacagcctccaatcc
acaacttttagctctgcttcttcttgaataggagcacgaaagcagtagtatttacttgatcttccaaaataaagcatgatttata
ctgcccacaaaatccaaagtgaatattttaaatttaatttgaacttcttctcctctctccaaacttttggatgaacttcttcca
tcttgagaccacaccgtcagaactgcaggagccaggagcctgtctccacgaatttccacacttgcacctcttagtctcctcaaaaga
aaaaagtggttagcaccatttgcagacatagatggccaaaatctctagtgaaagagataatgaactctggagtttgaagtttaagctc
atctagtcagtggttataagaagaagtagtcttgcaggggttggctcaggggtgtagtatctatcttagtaaggagaagccaattt
ttgaccatctgctgctgtacattgctctgctgctgagggcaagcttggacttggacttggtaacttggtaataatcgccatgcttctctgatt
cagaaaaccgatgcttgaaccgcagaaagaaagaggagaagttacctgtagtggacttggctcattatggttaattattggtacta
gtttccccagtttgcataaaaatgattgatatttggtaattggccttgttatatgctgaaaataaaaaaccacttggcttctccta
acacatgggagatgtagtagcatttcaatagctttaaagcgtgagctatgattcacagccttagaacatgtaaatattgctgtgag
ttcagcaagctgcatcagctatctcatgtagtatacaggctcaataggaatagtcctgcatatagtttcaattttagctgattg
tggcgtgtagtcttctgctattggccaaggtaatgtgacttggctgctctttatacagggccattgtagggtatagatttcttct
ggtgatagcaaaagggatctgtaaatcttctctatgttccccataaattcacttagtggtgctgctgctgtagcgggagacat
ctctgtttcaatgagatttatactcttaacatttagtttcttggta**ca**gTTGAAGAGGCAATTGGCTCATCCCTACCTAAACTCACT
GCACGACATCAGCGATGAGCCAACTTGCATGAATCCATTCAACTTTCAGCTTTGAGCAGCATGCACCTCACGGAGGAACAGATGAGGGAG
TTAATTTATAGGGAAGCGCTGCATTTAATCCCGAGTATCTACAG**TAA**

b)
>mpk5_scaffold12
TATGTAGCTTTTGTACTGGGTTTTAGAGCTCCATTGGCATTGCGGTGTTTCAATTTGGCATAACTCTTCTGGGCTCTTCTCTGGCTTTT
TTGCCAATTGAGGAGCATGATAATGATTTTCATAAGCTGTATCCTCAAGGAAGATTCTGAAGTTGGATTAGGCTCTGTAGAAGTTGCTGA
TGATCTCGGGCAGCGTGTATTAGCTCCAAATATAATCTTGTCCATCATATATGTTTGGCTTCACTCTTAGTTCTTCATGGGGATAA
GTACCAAAAAAAGTCAACTTGTATGTGCGATACAAAATTTAGTTATAAATTTTTAATTTGGTTAATTTAGTTTTCAACATTTTTATA
TTAATATCAATTCAGTTTTATCCGATTAATTTAATTTGAAAATTTATGATATGGACACCGACTATGCATAGATTTTTAATTTTTTAGAT
ATTTTTTCATTTTTTATTAACATTTTAAAAATTTTTCCCTTTTTCTTTGTTTTCCATTTTTTCTTTTCTTCTCTCTC
CCTTGGCTTATGGCCATTACTCGCCAAATTTGGTGAAGGCATCTCTCATCTATAGCCGTTTGTGGCCATGTTGAGTACCGGCCATAG
GCCAATGGAAGAAAAGGTAAGAAAAGAAAAGAAAAGAAAAGAAAATTTTTAAAAATTAATAAAAAATTTGGAAAAAAAATTTAG
TTAAAAATTTATGAAAAAGCTGTCTACATCAATGTTGCAATAGGACGCTCAACGTTACCGTTACCAATTTCCCGCTAAATTAACATG
ATGGACTCAATTGTATTAATAAAAAAGATTTATTTAATGAAATGATTTAATTAATAATGTTTAAGATTGATTTAGTATCAATGCAATG
ATTAATTTTATAGGAGACAGTACTTATAAAAAGTTCAATAATAGAAAATTTTGAAGGGGATTGATTGAAGTATATAACTCTGTA
TTTACCTAGAGACTAATCCGAACCATCAACTCTATCTTCTTTCGTCGGGTGTCGGATTGACATGATTGGAGCCGTCGACCCACAGC
CCATCTCTCTCTTACTATAAATTAGATTACATCAATTTAGTAAATGCAAAATTTAGCTTCCACTATAAGTAACTAGAACAACTTAACTG
TTATAACGAATGACAACCGGATATATTATAATTCGACAATTTGACAACCTTTCGAGAAGTTATTTACTACTCTCCGAAACTCCGGCAC
CATCAAAAGTTAGCCCTCAAACCTTATGTCTTTCGCGGGGCAGTCTTTCTCAGTATGGATGGCAATTTCAAGTGAATCCGAAAA
CAAACTGAACTCCATCTTCTCGGCCGAGTTCTTCTCATCTGGGAGACGATCAGGGGAGACGGCCCGAGAGGAATATCGAAGACAA
AGTAGGAGCGAACAAGATGACGCTGTCGGCTGCGCCGTAACAGCTCATTTTCGAGAGGTTGGATATAGAGCCGCGCCCTAACTGAAATTA
GCATCTTTTTTGGATGACATCTTTTAAATCAATTAATAAAGTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGGTA
AAAAAAGAAAAGAAAAGGGGGAGAAAATACAGGTGGTGAATCATCACAGGGCAGGTGAACAGAAGCAGCAGGCGCAGCAGAAAAGG
AAACAAAACAGGGCGAAAGAAAACGACCTTCCCACTCAAAGCTCTCCCTTTTTTCAATTTGCATTTCTGCATCCACACAGGCACA
GGACACAGGGGGACAGCTCAGTGAGTGGAAAGGTTGAGCCGGAATCTCCCCAACAAAAGCCGCTCAGCCCACTCAGCCCAA
CAGTCAAAACCCACCTTAACCGAATTTCCCCAGATCAATCCCTCTCTCTCTCCACTGTAACGGAAATCGCGACCCCAAATCTTA

Figura 5: Sequências genômicas de *E. grandis* com maior homologia ao cDNA de *Egmpk5* obtidas por BLASTn junto ao genoma disponível no *Phytozome*. Éxons estão representados por letras maiúsculas e fundo cinza claro; íntrons estão representados em letras minúsculas com seus sítios canônicos de *splicing* marcados em negrito. Os códons de iniciação e de terminação de transcrição também estão destacados em fundo cinza escuro. Estão apresentados, em letras maiúsculas e fundo branco, 2 kb anteriores ao códon de iniciação de tradução. (a) Gene *mpk5* presente no *scaffold* 7 entre as posições de 28.429.209 e 28.436.392 (b) Gene *mpk5* presente no *scaffold* 12 entre as posições 345.288 e 352.092; a jusante do quarto éxon estão marcados em itálico 102 nucleotídeos do íntron cuja identidade não foi confirmada pelo sequenciamento que gerou o genoma de *E. grandis*. Novas versões do sequenciamento poderão elucidar a sequência desta região.

A sequência de *Egmpk5* no *scaffold* 7, com 6 éxons, é a maior entre as duas analisadas, o que se dá pela ausência do éxon 5 de 181 pb e do íntron anterior no *scaffold* 12. Todas as regiões de íntrons identificadas possuem em suas extremidades as sequências conservadas GT e AG, que são sítios canônicos de *splicing* do transcrito (SZCZESNIAK *et al.*, 2013).

Para validar a definição de éxons destas sequências, os íntrons foram editados e duas possíveis sequências de DNA codificante (CDS) para *Egmpk5* foram obtidas. Elas apresentaram tamanhos distintos – 1.188 pb para *mpk5-scaffold* 7 e 1.011 pb segundo o *mpk5-scaffold* 12. Esta diferença numérica pode ser explicada em parte pela ausência do éxon 5 (181 pb) na sequência de menor tamanho.

A sequência codificante do clone de cDNA de *mpk5* foi alinhada com cada uma das duas CDS obtidas do genoma (dados apresentados no Anexo I, página 83). Desta forma, ficou evidente sua maior semelhança com a sequência do *scaffold* 7 (de 1.179 pb e 98,25 % de similaridade) do que com o *scaffold* 12 (de 993 pb e 82,34 %).

Como próximo passo na validação dos genes *mpk5* de *E. grandis*, as CDS foram traduzidas *in silico* e os peptídeos resultantes foram comparados com aquele deduzido do clone de cDNA de *mpk5*, conforme a **Figura 6** e a **Tabela 1**.

a)						
	10	20	30	40	50	60
AA_scf07						
AA__cDNA	MDGGAPQPADTVMSEAAPAQ	QQQQPQQQAQPQ	GIENIPATLSHGGRFIQ	YNI	FGNIFEVTA	
Prim.cons.	MDGGAPQPADTVMSEAAPAQ	QQQQPQQQAQPQ	GIENIPATLSHGGRFIQ	YNI	FGNIFEVTA	
	70	80	90	100	110	120
AA_scf07						
AA__cDNA	KYKPPIMP	IGKGAYGIVCSALN	SETNEHVAIKKIANAF	DNKIDAKRTLREIK	LLRHMDHE	
Prim.cons.	KYKPPIMP	IGKGAYGIVCSALN	SETNEHVAIKKIANAF	DNKIDAKRTLREIK	LLRHMDHE	
	130	140	150	160	170	180
AA_scf07						
AA__cDNA	NVVAIRDII	PPPPQREVNDVYI	AYELMDTDLHQI	IRSNQALSEEHCQ	YFLYQILRGLKYI	
Prim.cons.	NVVAIRDII	PPPPQREVNDVYI	AYELMDTDLHQI	IRSNQALSEEHCQ	YFLYQILRGLKYI	
	190	200	210	220	230	240
AA_scf07						
AA__cDNA	HSANVLRDLKPSN	LLLNANCDLKICDF	GLARVTSETDFMTEY	VVTRWYRAPELLN	SSD	
Prim.cons.	HSANVLRDLKPSN	LLLNANCDLKICDF	GLARVTSETDFMTEY	VVTRWYRAPELLN	SSD	
	250	260	270	280	290	300
AA_scf07						
AA__cDNA	YTAAIDVWSVGC	IFMELMDRKPLFP	GRDHVQQLRLLMEL	IGTPSEAE	LGFLEN	AKKYIR
Prim.cons.	YTAAIDVWSVGC	IFMELMDRKPLFP	GRDHVQQLRLLMEL	IGTPSEAE	LGFLEN	AKKYIR
	310	320	330	340	350	360
AA_scf07						
AA__cDNA	QLPLYRRQS	FTEKFPVHPLAID	LVEKMLTFDPRRL	LTVEEALAHPYL	NSLHDISDEPTC	
Prim.cons.	QLPLYRRQS	FTEKFPVHPLAID	LVEKMLTFDPRRL	LTVEEALAHPYL	NSLHDISDEPTC	
	370	380	390			
AA_scf07						
AA__cDNA	MNPFNFDFEQ	HALTEEQMRELI	YREALAFNPEYLQ			
Prim.cons.	MNPFNFDFEQ	HALTEEQMRELI	YREALAFNPEYLQ			
b)						
	10	20	30	40	50	60
AA_scf12						
AA__cDNA	MDGGGSSSRADTV	MSEVAPGAAAAA	ARRRSQ-----	QG-----	IEISG----	D
Prim.cons.	MDGG2222RADTV	MSE2AP222222	222222222222	QG	IEISG	222222222222
	70	80	90	100	110	120
AA_scf12						
AA__cDNA	AQPRGPLHPSTT	SS----ATSLRSP	-----			
Prim.cons.	A2222P22P2222	2222YGIV222L	2S22NEHVAIKKIANAF	DNKIDAKRTLREIK	LLRHMDH	
	130	140	150	160	170	180

AA_scf12	-----STSPSPCSARALRHRLLSFEFGD----ERARG----HKEDC-----
AA_cDNA	ENVVAIRDIIIPP--PQREVFNDVYIAYELMDTDLHQIIRSNQALSEEHCCQYFLYQILRGL
Prim.cons.	ENVVAIR222PPSCP222222222222E22DTDLH222R2NQAL22E2CQYFLYQILRGL
	190 200 210 220 230 240
AA_scf12	-----
AA_cDNA	KYIHSANVLRDLKPSNLLLNNANCDLKICDFGLARVTSETDFMTEYVWTRWYRAPELLLN
Prim.cons.	KYIHSANVLRDLKPSNLLLNNANCDLKICDFGLARVTSETDFMTEYVWTRWYRAPELLLN
	250 260 270 280 290 300
AA_scf12	-----
AA_cDNA	SSDYTAAIDVWSVGCIFMELMDRKPLFPGRDHVQQLRLLMELIGTPSEAEELGFLNENAKK
Prim.cons.	SSDYTAAIDVWSVGCIFMELMDRKPLFPGRDHVQQLRLLMELIGTPSEAEELGFLNENAKK
	310 320 330 340 350 360
AA_scf12	-----
AA_cDNA	YIRQLPLYRRQSFTEKFPVHPLAIDLVEKMLTFDPRLRLTVEEALAHPYLNSLHDISDE
Prim.cons.	YIRQLPLYRRQSFTEKFPVHPLAIDLVEKMLTFDPRLRLTVEEALAHPYLNSLHDISDE
	370 380 390
AA_scf12	-----
AA_cDNA	PTCMNPFNFDFEQHALTEEQMRELIYREALAFNPEYLQ
Prim.cons.	PTCMNPFNFDFEQHALTEEQMRELIYREALAFNPEYLQ

Figura 6: Alinhamento das sequências peptídicas deduzidas de MPK5 a partir do clone de cDNA e de duas sequências genômicas. O consenso entre as sequências está representado na linha “prim. cons”. (a) alinhamento da sequência obtida pelo clone de cDNA (representada por AA_cDNA; 395 resíduos) e pelo *scaffold* 7 (representada por AA_scf07, 395 resíduos). O tamanho total do alinhamento é de 395 resíduos de aminoácidos, sendo estes 100% idênticos (representados por asteriscos) entre as duas sequências. (b) alinhamento da sequência obtida pelo clone de cDNA (representada por AA_cDNA; 395 resíduos) e pelo *scaffold* 12 (representada por AA_scf12; 96 resíduos). O tamanho total do alinhamento é de 398 resíduos, sendo 30 (7%; representados por asteriscos) idênticos, 17 (4%; representados por dois pontos) fortemente similares, 22 (5%; representados por um ponto) fracamente similares e 329 (84%) diferentes entre as duas sequências.

Tabela 1: Comparação de peptídeos MPK5 deduzidos de duas sequências genômicas de *E. grandis* e pelo clone de cDNA *Egmpk5*.

	<i>Scaffold 7</i>	<i>Scaffold 12</i>	Clone cDNA
Tamanho do peptídeo deduzido (AA)	395	96	395
Similaridade dos peptídeos	<p>Idênticos à sequência deduzida do cDNA: 395 AA / 100 %</p> <p>Idênticos à sequência deduzida do <i>scaffold 12</i>: 30 AA / 7%</p> <p>Fortemente similares: 17 AA / 4%</p> <p>Fracamente similares: 22 AA / 5 %</p>	<p>Idênticos à sequência deduzida do cDNA e do <i>scaffold 7</i>: 30 AA / 7%</p> <p>Fortemente similares: 17 AA / 4%</p> <p>Fracamente similares: 22 AA / 5 %</p>	<p>Idênticos à sequência deduzida do <i>scaffold 7</i>: 395 AA / 100 %</p> <p>Idênticos à sequência deduzida do <i>scaffold 12</i>: 30 AA / 7%</p> <p>Fortemente similares: 17 AA / 4%</p> <p>Fracamente similares: 22 AA / 5 %</p>
Domínios conservados	<ul style="list-style-type: none"> • <i>STKc_Tey_MAPK_plant</i>¹ • <i>PKc_like superfamily</i>² 	Não foram encontrados prováveis domínios conservados	<ul style="list-style-type: none"> • <i>STKc_Tey_MAPK_plant</i>¹ • <i>PKc_like superfamily</i>²

¹ Domínio catalítico TEY de MAPKs em plantas; código de acesso [cd07858] no banco Conserved Domain Database disponível em <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>

² Domínio catalítico de PKs; código de acesso [cl09925] no banco Conserved Domain Database disponível em <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>

A sequência peptídica codificada pelo *scaffold 7* possui o tamanho esperado (395 aminoácidos) e não há códons de parada capazes de gerar interrupções na tradução deste peptídeo. A análise desta sequência por BLASTp no banco de dados GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) permitiu a identificação da presença dos domínios conservados do sítio de ativação de MAPKs em plantas (*STKc_Tey_MAPK_plant*; código de acesso [cd07858] no banco Conserved Domain Database disponível em <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) e do sítio catalítico de PKs (*PKc_like superfamily*; código de acesso [cl09925] no banco Conserved Domain Database disponível em <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>). Este peptídeo apresentou 100% de similaridade com a sequência do peptídeo deduzido do cDNA de *E. grandis* e, também, os mesmos domínios conservados.

A tradução *in silico* da CDS obtida de *mpk5-scaffold* 12 (1011 pb) geraria um peptídeo de 336 aminoácidos, mas foi verificada a presença de muitos códons de terminação no meio da sequência codificadora, o que torna seu tamanho real menor do que o esperado (com 96 aminoácidos). Este término prematuro da tradução faz com que o pequeno peptídeo codificado não apresente quaisquer motivos conservados de PKs conhecidos.

Como principal conclusão destas análises, foi possível estabelecer que provavelmente um único gene no genoma de *E. grandis* é responsável pelo transcrito correspondente ao clone de cDNA *Egmpk5* anteriormente isolado (KÖRBES, 2006), e este se localiza no *scaffold* 7. Na **Figura 7**, está representado o mapa cromossômico de *E. grandis* e a localização do gene *Egmpk5* conforme montagem disponível no banco de dados *Phytozome*. Os *scaffolds* 1-11 são denominados cromossomos A-K, sendo que *scaffolds* com numeração a partir de 12 não estão representados por ainda não fazerem parte da montagem cromossômica principal.

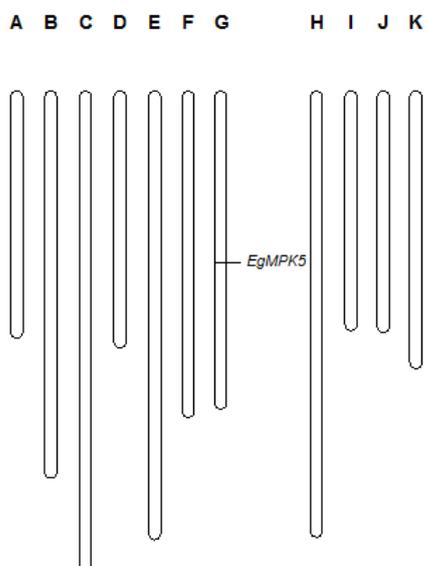


Figura 7: Localização e posição do gene *Egmpk5* no cromossomo G de *E. grandis*. A posição indicada refere-se à região limitada pelas posições 28.429.209 e 28.436.392 do *scaffold* 7 (ou cromossomo G).

Notavelmente, o clone de cDNA de *Egmpk5*, que foi completamente sequenciado (KÖRBES, 2006), possui 5.068 pb e é, portanto, maior que quaisquer das duas CDS encontradas no genoma de *E. grandis*. Novos alinhamentos permitiram elucidar que o cDNA possui, em verdade, 1.188 pb de sequência codificadora, sendo o restante correspondente a regiões de íntrons. O clone de cDNA trata-se, portanto, de um transcrito imaturo do gene *Egmpk5*, em que os quatro primeiros éxons foram unidos após remoção de regiões não codificadoras, mas na sequência ainda permaneceram os dois íntrons entre os dois últimos éxons (**Figura 8**). KÖRBES (2006) havia anteriormente concluído que o cDNA isolado representava um transcrito imaturo. Porém, a autora deduzira que o gene *mpk5* de *E. grandis* possuía três éxons e não cinco, como demonstrado no presente trabalho.

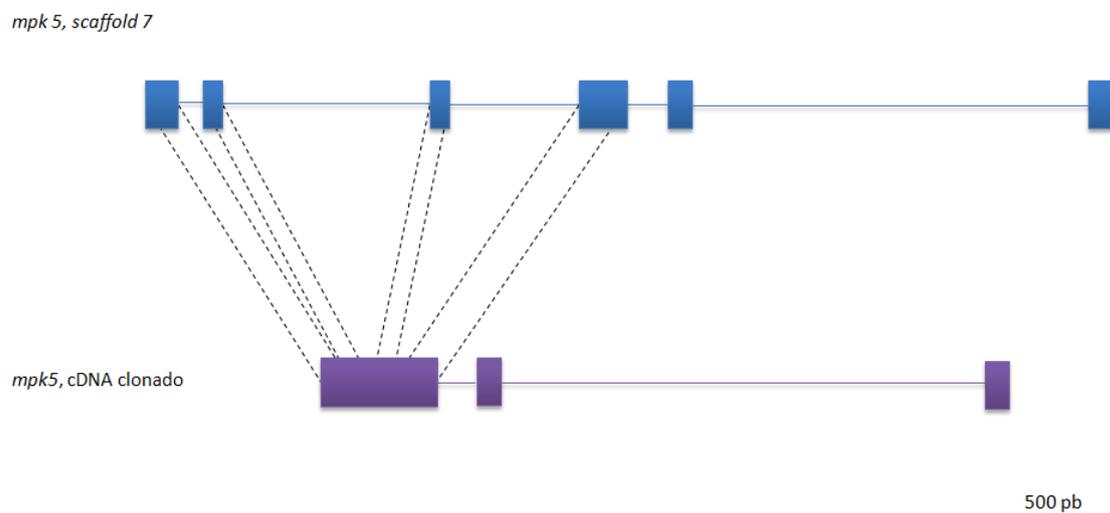


Figura 8: Comparação entre as estruturas do gene *mpk5* de *E. grandis* a partir de sua provável origem no cromossomo G de *E. grandis* e a partir do clone de cDNA isolado e caracterizado por KÖRBES (2006).

A busca por *mapks* de outras plantas com homologia a *mpk5* de *E. grandis* no GenBank por BLASTn resultou em 4 genes de diferentes organismos com estrutura descrita. Dessa forma, foram analisadas as sequências nucleotídicas de *mapks* de *Arabidopsis thaliana* (*mpk6*; AT2G43790.1; <http://www.arabidopsis.org/servlets/TairObject?accession=locus:2043904>), *Brachypodium distachyon* (*mpk6*; Bradi1g49100; CHEN *et al.*, 2012), *Populus trichocarpa* (*mpk6*; estExt_fgenes4_pm.C_LG_VII0025; NICOLE *et al.*, 2006) e *Solanum lycopersicum* (*mpk2*; AY261513.1; KONG *et al.*, 2012).

Na **Figura 9** está apresentada uma representação gráfica da organização de éxons e íntrons das sequências obtidas. Todas as sequências analisadas apresentaram 6 éxons interrompidos por íntrons de tamanhos variados. O tamanho dos íntrons não apresentou um padrão entre tamanho e localização nos genes das diferentes espécies, porém foi evidente que *A. thaliana* possui o gene com menores íntrons e que a sequência de *E. grandis* possui um dos maiores íntrons na extremidade 3' do gene. Editando-se os íntrons, as sequências resultantes apresentaram tamanhos bastante parecidos (resultado não apresentado), o que significa que todos os genes possivelmente codificam proteínas de tamanhos semelhantes. Os tamanhos dos éxons seguiram um padrão uniforme de acordo com a sua ordem nos genes e destaca-se (em **negrito** na **Figura 9**) a grande conservação de tamanho dos éxons internos (2 a 4) entre os genes das espécies analisadas.

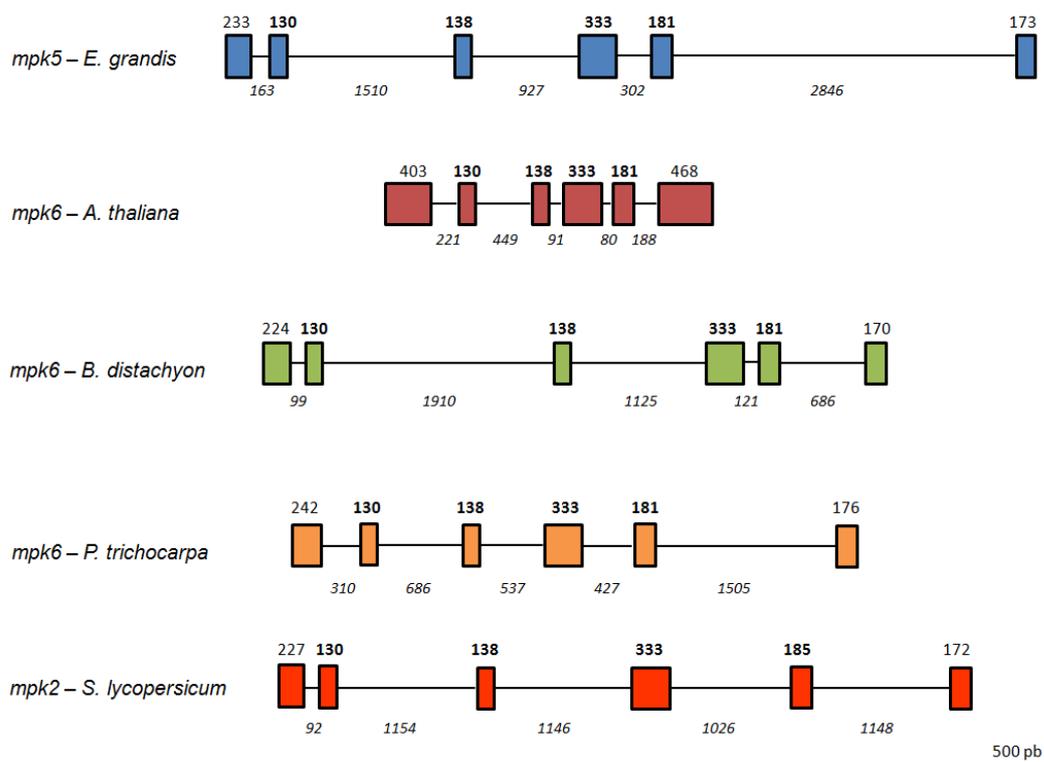


Figura 9: Comparação da estrutura de éxons (caixas) e íntrons (linhas) entre *mpk5* de *E. grandis* e genes *mapk* de outras espécies vegetais. Na parte superior de cada representação estão indicados os tamanhos dos éxons em pb, sendo os 4 éxons internos de cada sequência destacados em **negrito** e, na parte inferior, estão os tamanhos dos íntrons, em valores em *itálico*.

Os estudos *in silico* conduzidos até esta etapa permitiram-nos concluir que a sequência *mpk5* presente no cromossomo G de *E. grandis* possui uma região estrutural funcional, isto é, constituída por éxons e íntrons com extremidades canônicas de *splicing* e com elevado grau de conservação com outros genes codificadores de MAPKs oriundos de diferentes espécies vegetais. A seguir, um estudo da potencial região promotora, a montante do gene *Egmpk5*, está apresentado com vistas a fortalecer ainda mais a hipótese de que se trata, de fato, de um gene funcional.

4.1.2 Análise da região promotora do gene *Egmpk5*

A análise da região 5' de *Egmpk5*, correspondente à uma típica sequência promotora a montante da região transcrita, revelou possíveis elementos regulatórios de ação *cis* da transcrição deste gene. A análise desta região de 2000 pb e a busca junto ao banco *PlantCARE* (LESCOT *et al.*, 2002; <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) revelou 25 principais elementos, os quais estão representados na **Tabela 2** divididos em três grupos: elementos característicos de promotores eucarióticos, aqueles com envolvimento no crescimento e desenvolvimento de plantas, e elementos relacionados à resposta vegetal a reguladores do crescimento (hormônios) e a estresses.

Tabela 2: Possíveis elementos de ação *cis* presentes na potencial região promotora de *Egmpk5* segundo buscas no banco *PlantCARE*.

Processo Geral	Processo Específico	Nome do elemento	Motivo
Elementos <i>cis</i> envolvidos em crescimento e desenvolvimento de plantas	Envolvidos na resposta à luz	<i>Box 4</i>	ATTAAT
		<i>ACE</i>	AAAACGTTTA
		<i>ATCT-motif</i>	AATCTAATCC
		<i>CAT-motif</i>	GCATTC
		<i>G-Box</i>	CACGTA
		<i>G-box</i>	TACGTG
		<i>GA-motif</i>	AAGGAAGA
		<i>GT1-motif</i>	GGTTAA
		<i>MRE</i>	AACCTAA
	<i>TCT-motif</i>	TCTTAC	
	Envolvidos em expressão no endosperma	<i>GCN4_motif</i>	CAAGCCA
	Envolvido em regulação de rota de síntese de flavonoides	<i>MBSII</i>	AAAAGTTAGTTA
Elementos <i>cis</i> relacionados com resposta a hormônios e a estresse em plantas	Resposta a ácido abscísico	<i>ABRE</i>	TACGTG
	Resposta a ácido metiljasmônico	<i>CGTCA-motif</i>	CGTCA
		<i>TGACG-motif</i>	TGACG
	Resposta a ácido salicílico	<i>TCA-element</i>	GAGAAGAATA
	Resposta a auxina	<i>TGA-element</i>	AACGAC
	Resposta a etileno	<i>ERE</i>	ATTTCAAA
	Resposta a estresse por calor	<i>HSE</i>	AAAAAATTTTC
	Resposta a baixa temperatura	<i>LTR</i>	CCGAAA
	Resposta a giberelina	<i>GARE</i>	TCTGTTG
Resposta a elicitor fúngico	<i>Box-WI</i>	TTGACC	
Elementos <i>cis</i> comuns a promotores eucarióticos	Elemento que confere altos níveis de transcrição	<i>5UTR Py-rich stretch</i>	TTTCTTCTCT
	Elemento comum em regiões de promotores e <i>enhancers</i>	<i>CAAT-box</i>	CCAAT
	Principal elemento do promotor localizado próximo ao sítio -30 do início de transcrição	<i>TATA-box</i>	ATATAT

Com estas análises, foi possível verificar que a região de 2 kb anteriores ao gene *Egmpk5* apresenta prováveis elementos característicos de promotores funcionais de eucariotos. Uma vez que a sequência deste gene foi inicialmente isolada de *ESTs* (KÖRBES, 2006), estas análises *in silico* servem para sugerir a possível origem deste transcrito no genoma. A presença de possíveis elementos característicos de genes vegetais que estão implicados no metabolismo e em respostas a estresses e a hormônios indica provável função de *Egmpk5* também nestas respostas, e há correlação entre algumas destas e dados de expressão obtidos por RT-qPCR (ver item 5.2).

Para uma caracterização mais aprofundada de função, foi feita comparação *in silico* entre a sequência peptídica deduzida de MPK5 de *E. grandis* e de outras plantas que tiveram MAPKs implicadas em respostas a estresses e a tratamentos.

4.1.3 Análise do peptídeo deduzido MPK5 de *E. grandis*

Com a sequência peptídica deduzida de EgMPK5 (**Figura 6**) foram feitas buscas por sequências homólogas por BLASTp no GenBank. Foram analisadas 13 sequências de proteínas com no mínimo 80% de homologia à EgMPK5. As sequências obtidas com seus respectivos códigos de acesso foram as seguintes: *Glycine max* MPK2 (NP_001235426), *Medicago sativa* MSK7 (Q07176), *Pisum sativum* MAPK D5 (Q06060), *Arachis hypogaea* MAPK1 (AAZ23128.1); *Gossypium hirsutum* MPK6 (ADT65137.1), *Solanum lycopersicum* MPK2 (NP_001234355.1), *Solanum tuberosum* MPK2 (BAB93530.1), *Petroselinum crispum* MPK6 (AAN65179.1), *Oryza sativa* MPK6 (ACD76439.1), *Populus trichocarpa* MAPK6-2 (XP_002310398.2), *Brachypodium distachyon* MPK6 (NP_001266884.1), *Nicotiana tabacum* MPK4 (Q40532.1) e *Arabidopsis thaliana* MPK6 (NP_181907.1).

Uma árvore filogenética foi construída utilizando-se o programa MEGA 5 (TAMURA *et al.*, 2011) o que permitiu realizar o agrupamento de MAPKs semelhantes, com ramificações sustentadas por altos valores de *bootstrap* (**Figura 10**). Conforme demonstrado na árvore, sequências de MAPKs de plantas monocotiledôneas (*B. distachyon* e *O. sativa*) foram agrupadas à parte e, dentro do grande grupo de eudicotiledôneas, houve o agrupamento em famílias. Dois clados de MAPKs representaram famílias com mais de uma espécie, como é o caso da família Fabaceae, que contém as espécies *G. max*, *M. sativa*, *P. sativum* e *A. hypogaea* e, também, a família Solanaceae, com as espécies *S. lycopersicum* e *S. tuberosum*. As demais sequências não foram localizadas em ramificações com outras espécies.

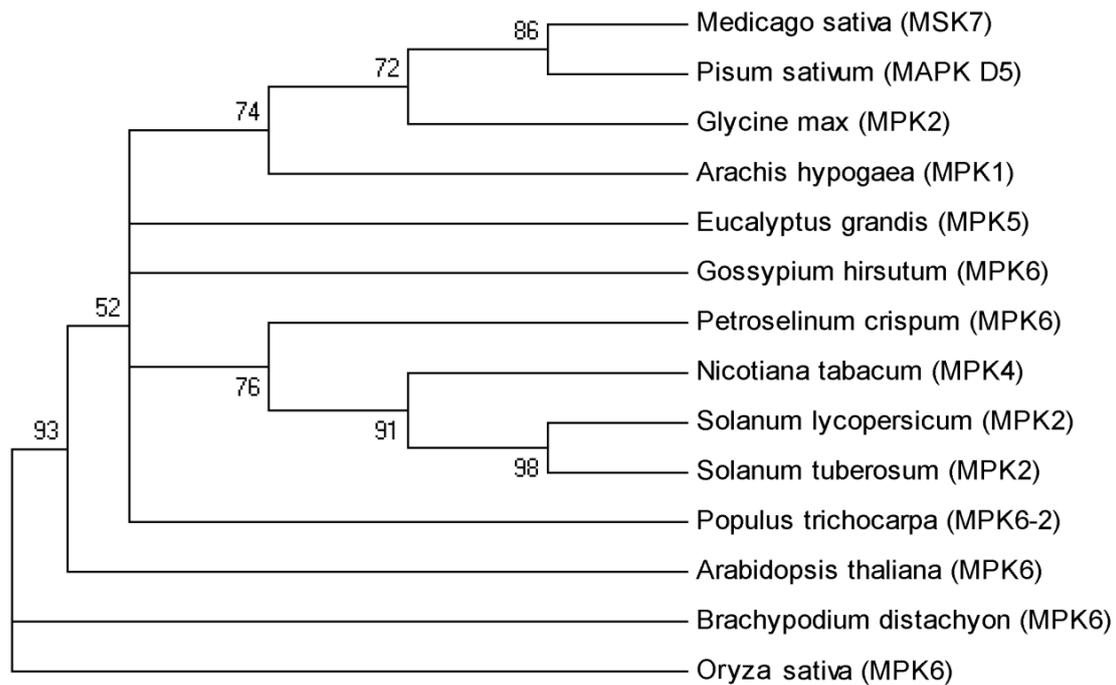
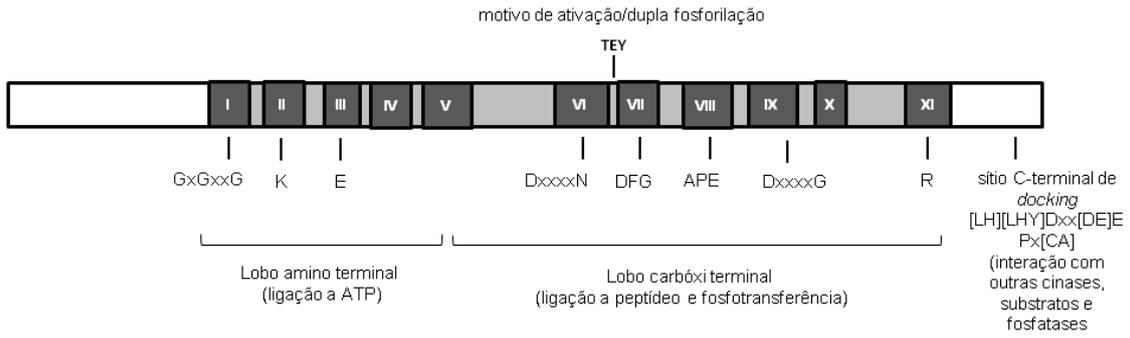


Figura 10: Árvore filogenética construída com MEGA5 utilizando 14 sequências peptídicas de MAPKs similares à MPK5 de *E. grandis*. Valores de *bootstrap* estão representados por números de 0 a 100 nas ramificações da árvore.

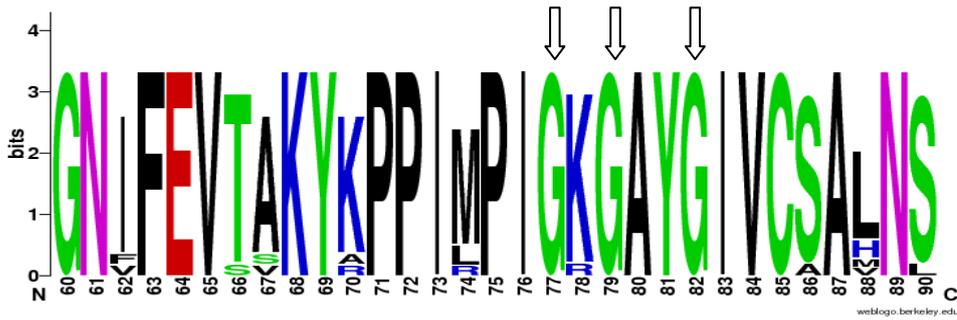
As sequências peptídicas utilizadas para a construção da árvore filogenética foram também avaliadas quanto à presença de domínios conservados. Análise por BLASTp no GenBank demonstrou a presença dos domínios característicos de PKs *STKc_Tey_MAPK_plant* e *PKc_like superfamily* em todas as sequências. Como o domínio de PKs é caracterizado por subdomínios com sequências-consenso e por sítios característicos, uma análise mais aprofundada foi conduzida com estes domínios encontrados.

As sequências completas das MAPKs estudadas foram alinhadas e o arquivo obtido foi analisado pelo programa disponível *on-line WebLogo*, que permite criar representações gráficas dos alinhamentos. As letras que representam aminoácidos são empilhadas em cada posição do alinhamento, sendo que seu tamanho é definido pela frequência relativa do resíduo em uma dada posição. Na **Figura 11** estão representadas as principais características encontradas no domínio PK das 14 sequências avaliadas, com base no esquema descrito por Hanks (2003).

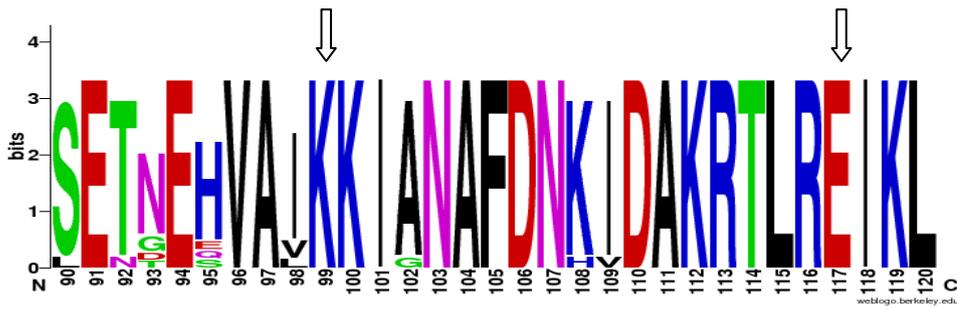
a)



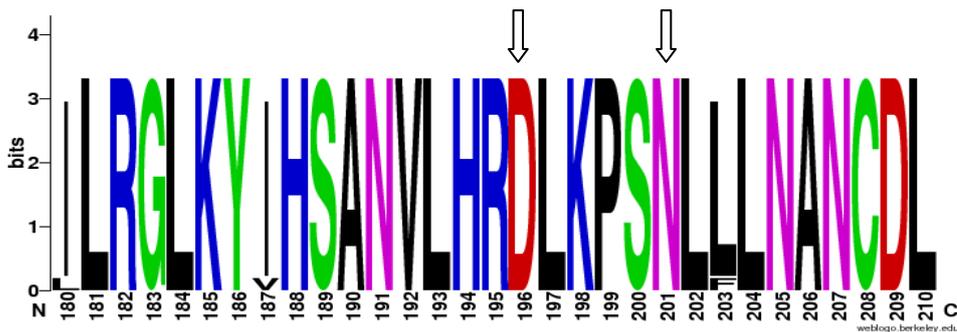
b)



c)



d)



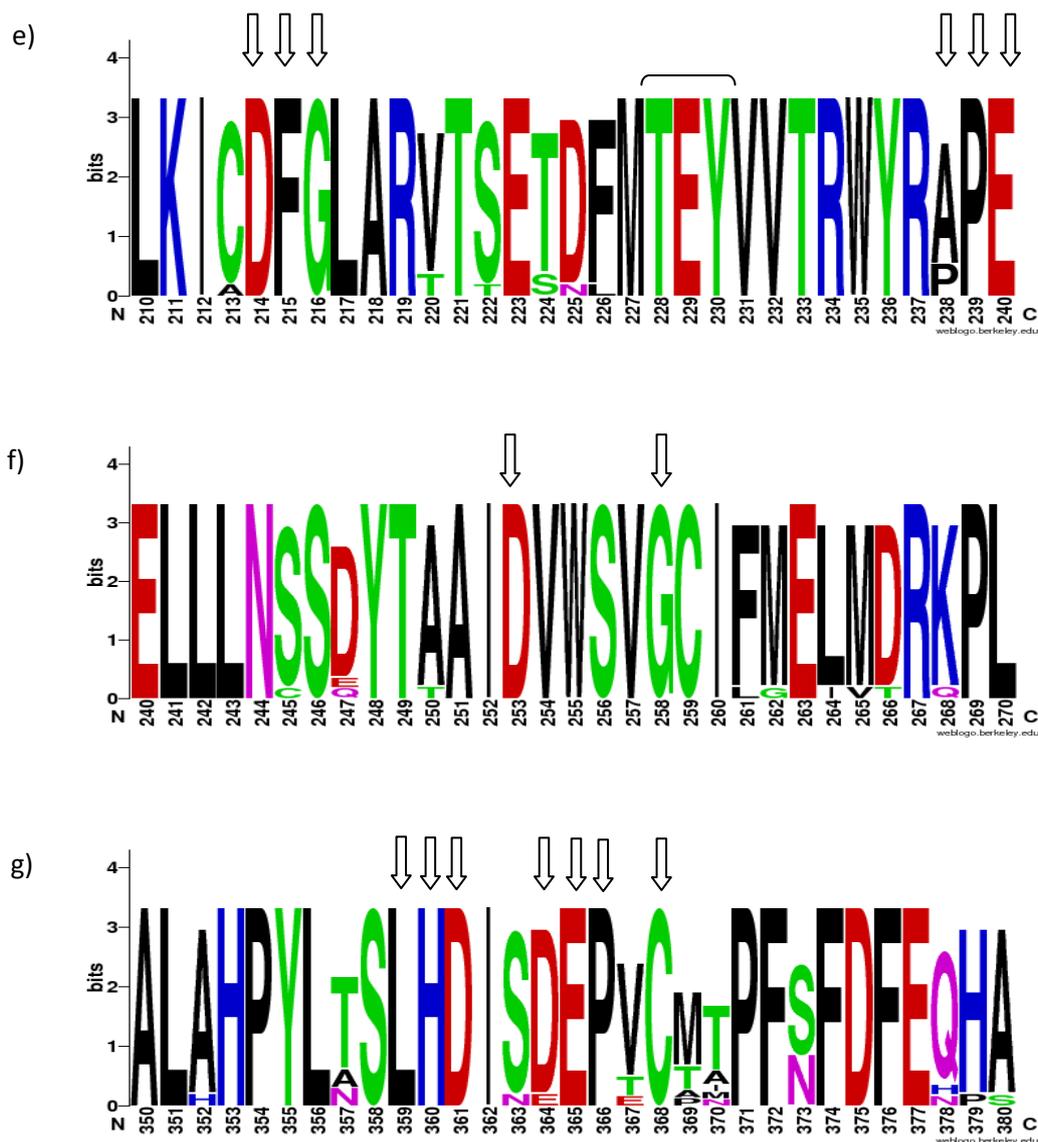


Figura 11: Representação dos subdomínios do domínio PK. Setas indicam consensos de sequência de importância (a) Esquema adaptado de Hanks (2003) onde estão mostrados os principais consensos de sequências dos subdomínios (numerais romanos) do domínio PK (fundo cinza); (b) Consenso GxGxxG do subdomínio I (onde “x” representa qualquer aminoácido); (c) Consensos K (subdomínio II) e E (subdomínio III); (d) Consenso DxxxxN (subdomínio VI); (e) Consensos DFG (subdomínio VII) e APE (subdomínio VIII) e motivo de ativação/dupla fosforilação TEY (entre os subdomínios VI e VII); (f) Consenso DxxxxG (subdomínio IX); (g) Consenso do sítio C-terminal de *docking* (*CD site*) de MAPKs [LH][LHY]Dxx[DE]EPx[CA] (colchetes representam os possíveis aminoácidos para uma dada posição).

Com as análises *in silico* do peptídeo deduzido MPK5 foram obtidas mais evidências de que a sequência gênica *mpk5* poderia ser corretamente editada e dar origem a um produto funcional. A árvore filogenética e a análise de motivos conservados entre as 14 MAPKs vegetais demonstrou sua similaridade, o que é bastante útil para inferirmos possíveis funções de MPK5 de *E. grandis*. A próxima etapa na

avaliação da possível função de *mpk5* foi o estudo de sua expressão frente a tratamentos que estimulam respostas celulares em plantas sob estresse.

4.2 Caracterização do padrão de expressão de *Egmpk5* em *E. grandis*

A expressão do gene *mpk5* em diferentes órgãos de *E. grandis* foi avaliada por RT-qPCR, medindo-se a quantidade relativa e momentânea de transcritos (*steady-state mRNA levels*), utilizando-se, para normalização dos dados, dois genes-referência desta planta recentemente descritos (genes *Ec08* e *Ec04*; OLIVEIRA *et al.*, 2012). Para estudos da expressão gênica espacial, foram utilizadas amostras de RNA total de raízes, caules e folhas de plantas germinadas e crescidas em copos plásticos contendo a mistura solo/vermiculita na proporção de 1:1 por 90 dias e tratadas com água. A análise da expressão gênica modulada por sinais incluiu estes mesmos órgãos e condições de germinação e crescimento, mas coletados a partir de plantas submetidas a diferentes tratamentos.

Os níveis relativos de transcritos de *Egmpk5* normalizados com os genes de referência estão representados na **Figura 12**. A raiz foi o órgão que apresentou os menores valores de expressão de *Egmpk5* no conjunto de análises realizadas e, por esta razão, os valores relativos de expressão obtidos para este órgão foram definidos como “um” (1), para efeitos de comparação com os valores de expressão do gene nos demais órgãos. Apesar da aparente expressão aumentada em caule e folha, a aplicação de análise estatística pelo teste T de Student permitiu revelar que as médias não foram estatisticamente diferentes, levando-se em consideração um intervalo de confiança de 95%. Assim, a análise espacial demonstrou que *Egmpk5* é expresso igualmente nos três órgãos avaliados de *E. grandis*.

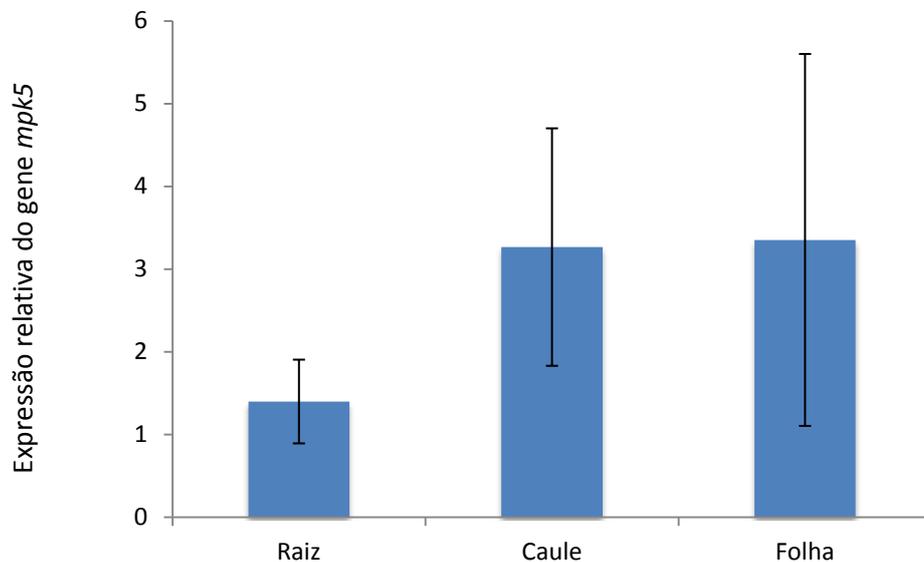


Figura 12: Dados de RT-qPCR da expressão relativa do gene *Egmpk5* em órgãos de *E. grandis* tratados com água após 90 dias de crescimento. Os valores estão normalizados com os genes constitutivos *Ec04* e *Ec08* de *E. grandis* e estão expressos em relação à raiz. O eixo “y” representa valores de expressão relativa calculados pelo método $2^{-\Delta\Delta Ct}$. O eixo “x” representa os três diferentes órgãos testados. Barras de erro representam valores de erro padrão para triplicatas biológicas.

Na análise de expressão de *Egmpk5* modulada por sinais, foram utilizados diferentes tratamentos em plantas de *E. grandis* crescidas por 90 dias: água, ácido abscísico (ABA), ácido naftalenoacético (NAA), cinetina (Kin), cloreto de sódio (NaCl) e seca. Nas **Figuras 13, 14 e 15** estão representados os níveis de expressão de *Egmpk5* em folhas, caules e raízes submetidas a estes tratamentos, sendo que os dados foram normalizados com os genes constitutivos *Ec04* e *Ec08* e foram representados em relação ao controle não tratado (plantas irrigadas com água).

Em folhas (**Figura 13**), a expressão de *Egmpk5* foi induzida por ABA com um aumento relativo à água de $17,7 \pm 6,57$ vezes, e por NaCl com um aumento de $98,6 \pm 35,03$ vezes. A exposição à seca e o tratamento com NAA não implicaram em valores de expressão diferentes dos observados em plantas tratadas com água. O tratamento com Kin pareceu aumentar a expressão de *Egmpk5*, mas a análise de probabilidade estatística revelou que estes valores não diferiram dos observados com o tratamento com água.

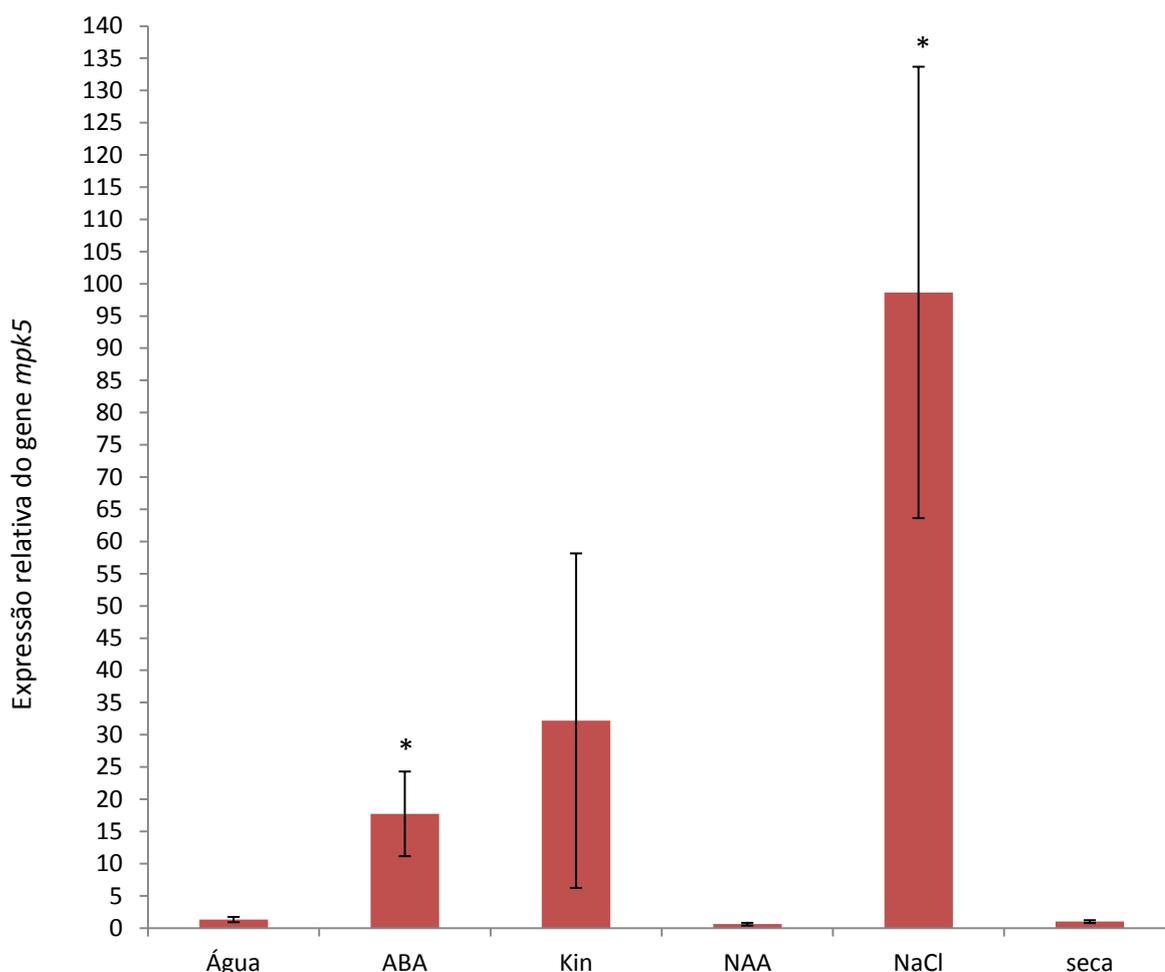


Figura 13: Dados de RT-qPCR da expressão relativa do gene *Egmpk5* em folhas de *E. grandis* sob diferentes tratamentos. Plantas de 90 dias foram irrigadas com 20 mL das soluções e foram coletadas após 3 dias. Para o teste de seca, plantas não foram irrigadas por 7 dias. Os valores estão normalizados com os genes constitutivos *Ec04* e *Ec08* de *E. grandis* e estão expressos em relação ao tratamento com água. O eixo “y” representa valores de expressão relativa calculados pelo método $2^{-\Delta\Delta Ct}$. O eixo “x” representa os tratamentos empregados. Barras de erro representam valores de erro padrão para triplicatas biológicas. Diferenças estatisticamente significativas estão indicadas por (*) $P < 0,05$ segundo teste T de Student.

A expressão de *Egmpk5* foi detectada em caules de plantas submetidas à seca e ao tratamento com ABA em valores equivalentes aos observados no tratamento com água (**Figura 14**). Diferença estatisticamente significativa foi verificada com NAA, que causou diminuição na expressão de *mapk5* ($0,37\pm 0,13$), e com NaCl, que causou aumento da expressão em relação ao controle em água ($9,59\pm 3,603$).

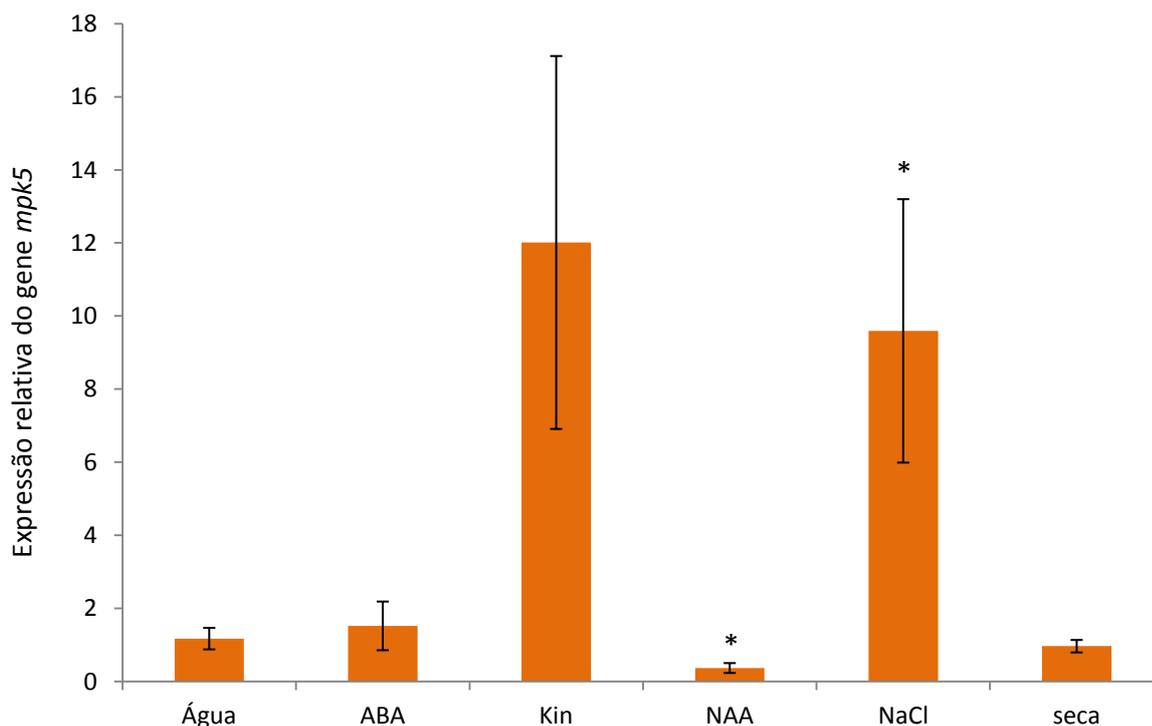


Figura 14: Dados de RT-qPCR da expressão relativa do gene *Egmpk5* em caules de *E. grandis* sob diferentes tratamentos. Plantas de 90 dias foram irrigadas com 20 mL das soluções e foram coletadas após 3 dias. Para o teste de seca, plantas não foram irrigadas por 7 dias. Os valores estão normalizados com os genes constitutivos *Ec04* e *Ec08* de *E. grandis* e estão expressos em relação ao tratamento com água. O eixo “y” representa valores de expressão relativa calculados pelo método $2^{-\Delta\Delta C_t}$. O eixo “x” representa os tratamentos empregados. Barras de erro representam valores de erro padrão para triplicatas biológicas. Diferenças estatisticamente significativas estão indicadas por (*) $P < 0,05$ segundo teste T de Student.

Em raízes (**Figura 15**), o tratamento com NaCl foi o único a causar aumento na expressão de *Egmpk5* ($149 \pm 76,7$). Os demais tratamentos não causaram expressão estatisticamente significativa em relação ao controle em água.

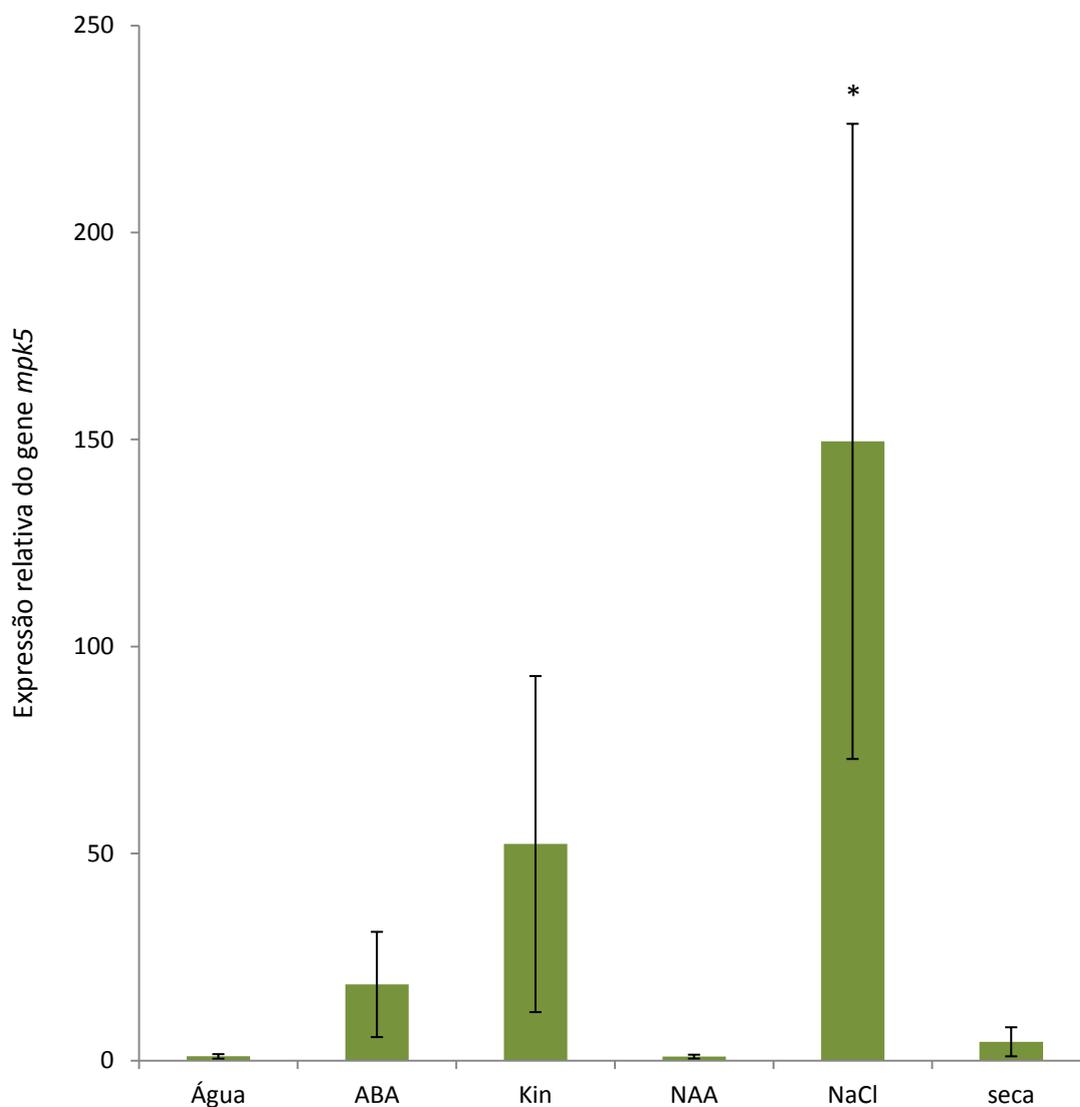


Figura 15: Dados de RT-qPCR da expressão relativa do gene *Egmpk5* em raízes de *E. grandis* sob diferentes tratamentos. Plantas de 90 dias foram irrigadas com 20 mL das soluções e foram coletadas após 3 dias. Para o teste de seca, plantas não foram irrigadas por 7 dias. Os valores estão normalizados com os genes constitutivos *Ec04* e *Ec08* de *E. grandis* e estão expressos em relação ao tratamento com água. O eixo “y” representa valores de expressão relativa calculados pelo método $2^{-\Delta\Delta Ct}$. O eixo “x” representa os tratamentos empregados. Barras de erro representam valores de erro padrão para triplicatas biológicas. Diferenças estatisticamente significativas estão indicadas por (*) $P < 0,05$ segundo teste T de Student.

Destas análises por RT-qPCR, foi possível concluir que *mpk5* é expresso nos órgãos pesquisados, sendo que a expressão é idêntica entre os órgãos quando as plantas são tratadas com água. Já tratamentos com ABA aumentaram a expressão de *mpk5* em folhas e tratamentos com NaCl aumentaram a expressão em todos os órgãos testados. Estes aumentos na expressão indicam possível participação do gene nas respostas celulares causadas pelos tratamentos, e, assim, este trabalho utilizou-se de plantas de tabaco previamente geradas para melhor caracterizar o efeito da expressão heteróloga deste gene frente a tratamentos de estresse.

4.3 Identificação e caracterização de plantas de tabaco expressando *Egmpk5*

4.3.1 Análise por PCR

Com vistas à confirmação do estado transgênico de plantas de tabaco (*N. tabacum* SR1) geneticamente transformadas com o gene *Egmpk5* em trabalho anterior do grupo (KÖRBES, 2006; KIRCH, 2008), análises por PCR foram realizadas utilizando-se *primers* específicos a uma região de 864 pb do gene (**Figura 16**) e DNA genômico extraído de folhas das plantas. As plantas utilizadas nesta análise representam a primeira descendência (T₁) das linhagens originalmente transformadas (T₀) por *A. tumefaciens* LBA4404::pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos}. A análise por PCR resultou na amplificação de fragmentos de tamanho e padrão de bandas que permitiram confirmar o estado transgênico das plantas de tabaco (**Figura 17**). O tamanho dos *amplicons* está de acordo com o esperado conforme o mapa da construção utilizada na transformação (**Figura 16**).

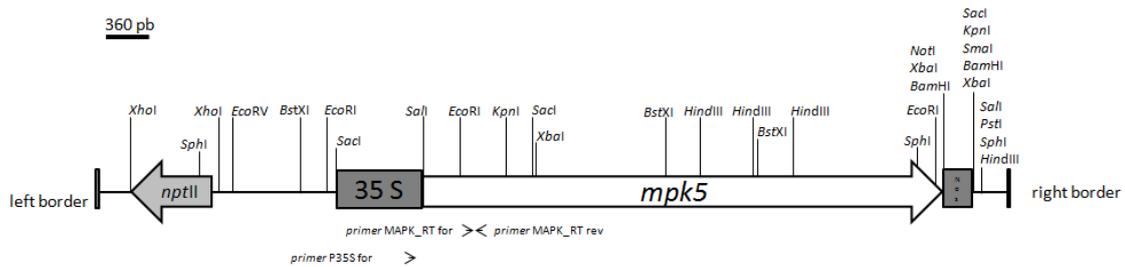


Figura 16: Esquema da região do T-DNA do plasmídeo pCAMBIA2300-P_{35S}-Egmpk5-T_{nos}, construção utilizada na transformação de plantas de tabaco, e indicação do local de anelamento dos primers utilizados para a confirmação do estado transgênico das plantas (ver Materiais e Métodos). As posições dos sítios de reconhecimento por endonucleases estão indicados. As bordas esquerda (*left border*) e direita (*right border*) do T-DNA estão indicadas. Os elementos de DNA indicados nas caixas são *nptII*, gene da neomicina fosfotransferase II; 35S, promotor CaMV 35S, do vírus do mosaico da couve-flor; *mpk5*, cDNA de *Egmpk5* clonado por KÖRBES, 2006; e Nos, terminador do gene da nopalina sintase.

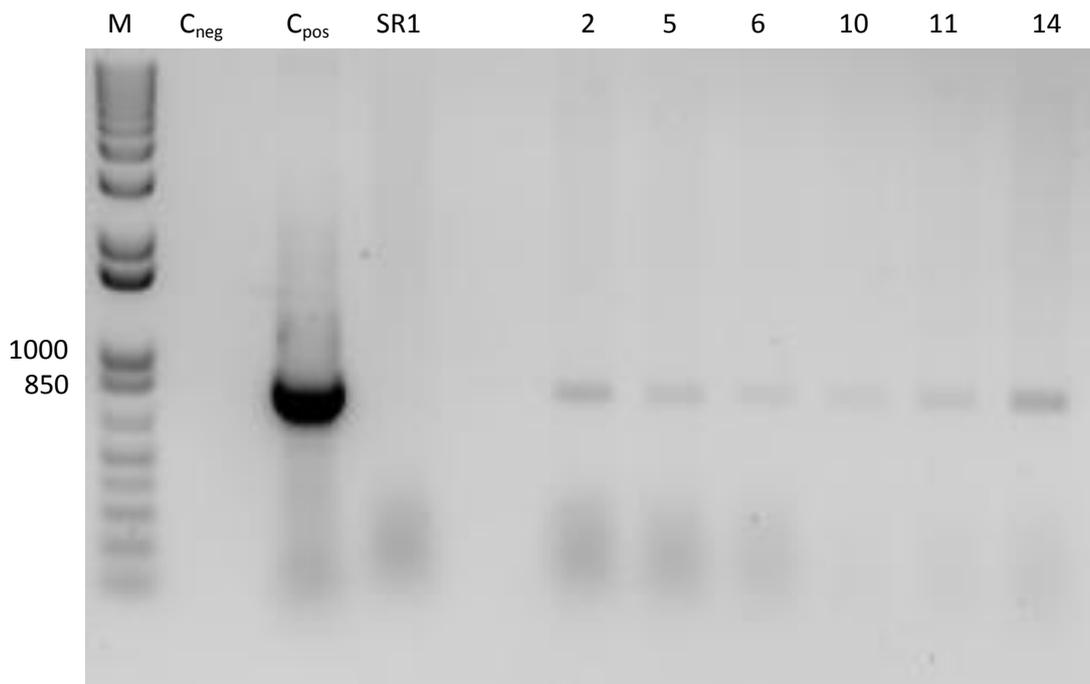


Figura 17: Produtos de PCR com DNA total de plantas transgênicas e controle de tabaco resolvidos em gel de agarose a 1% e corado com brometo de etídeo. As letras sobre as canaletas indicam: M, marcador de tamanho de fragmentos de DNA, em pb *1 kb Plus DNA Ladder* (Invitrogen); C_{neg}, controle negativo, mistura de reação sem a adição de DNA; C_{pos}, controle positivo de reação, utilizando-se DNA do plasmídeo pCAMBIA2300-P_{35S}-Egmpk5-T_{nos} como molde; SR1, reação utilizando DNA de plantas de tabaco SR1, não transformadas; 2, 5, 6, 10, 11 e 14, reações com DNA de plantas transformadas de tabaco.

Após confirmação do estado transgênico das plantas na geração T₁, nas etapas seguintes do trabalho buscou-se classificar o estado de zigose com estudos de germinação de sementes da geração T₂. Após, as plantas transformadas foram submetidas a diferentes tratamentos e tiveram sua resposta avaliada.

4.3.2 Germinação e classificação de genótipos de plantas transgênicas quanto à homo/hemizigose

Com vistas à caracterização biológica e do padrão de herança das plantas transgênicas potencialmente capazes de expressar o gene *Egmpk5*, análises da viabilidade de sementes T₂ e da resistência das plantas ao antibiótico seletor foram realizadas. A viabilidade das sementes T₂ após assepsia foi avaliada pela germinação em até 15 dias em meio de cultura ½MS livre de qualquer agente seletivo. As taxas de germinação das sementes obtidas durante este período estão representadas na **Figura 18**.

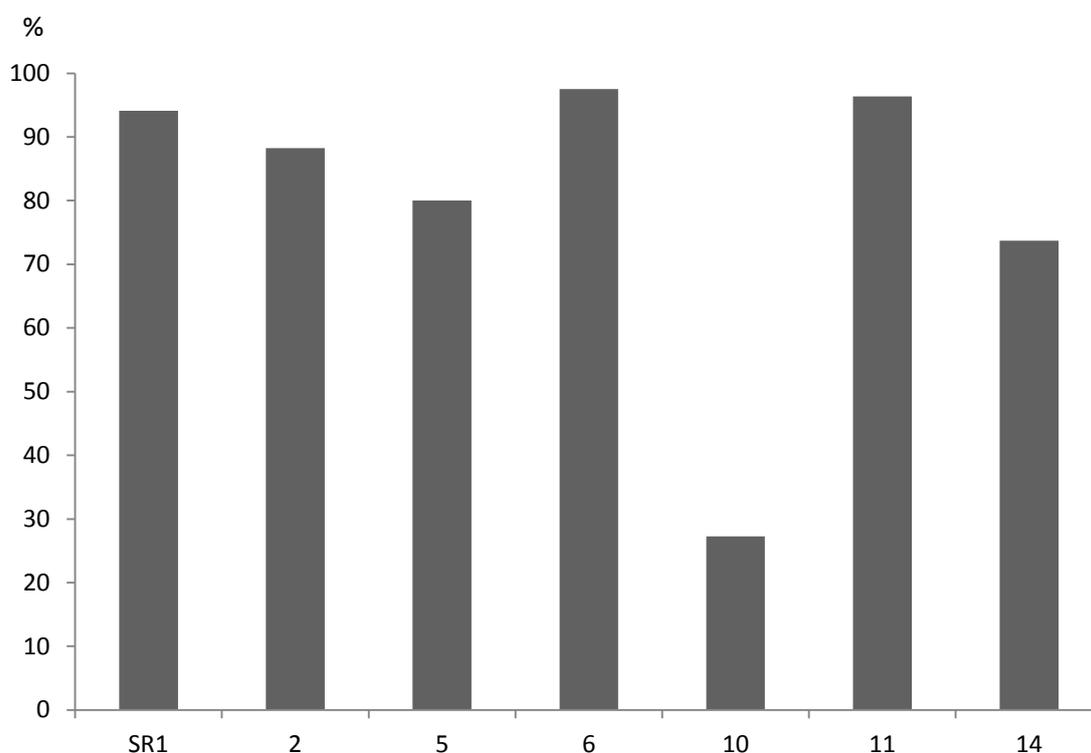


Figura 18: Verificação da viabilidade das sementes T₂ de *N. tabacum* SR1 controle (SR1, não transformada) e transformadas pelo T-DNA do plasmídeo pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos} (2, 5, 6, 10, 11 e 14) após método padrão de assepsia. As sementes foram germinadas *in vitro* em meio ½MS acrescido de ágar a 1,5% e a germinação foi acompanhada por 15 dias.

Na **Figura 18** é mostrado que a maioria das sementes testadas teve taxa de germinação de 80 a 95%, com exceção da planta transformada 10, que demorou mais tempo para germinar e atingiu um máximo de 23% de germinação. A partir deste experimento, foi possível estabelecer o número médio de sementes que germinam nas condições padrão *in vitro*, ou seja, não perdem viabilidade após assepsia. As taxas de germinação assim obtidas serviram como base para comparação de taxas de germinação das plantas transformadas sob diferentes tratamentos como canamicina e salinidade, conforme mostrado a seguir.

Para a classificação de genótipos quanto à homo/hemizigose, plantas da geração T₁ foram autofecundadas e o perfil de resistência à canamicina em plantas da geração T₂ foi avaliado. Sementes T₂ germinadas em meio ½MS sólido acrescido de Kan¹⁰⁰ tiveram o crescimento avaliado por 25 dias (**Figura 19**). Ao longo deste período, foram contadas as sementes que germinaram e, entre as plântulas formadas, as que sobreviveram ao meio com antibiótico e aquelas que morreram. Foram consideradas plantas resistentes, isto é, transgênicas, aquelas que mantiveram cor verde e crescimento de folhas e raízes. As plantas que perderam a cor verde e ficaram amareladas ou brancas ao longo do desenvolvimento foram consideradas sensíveis e, portanto, não transgênicas.

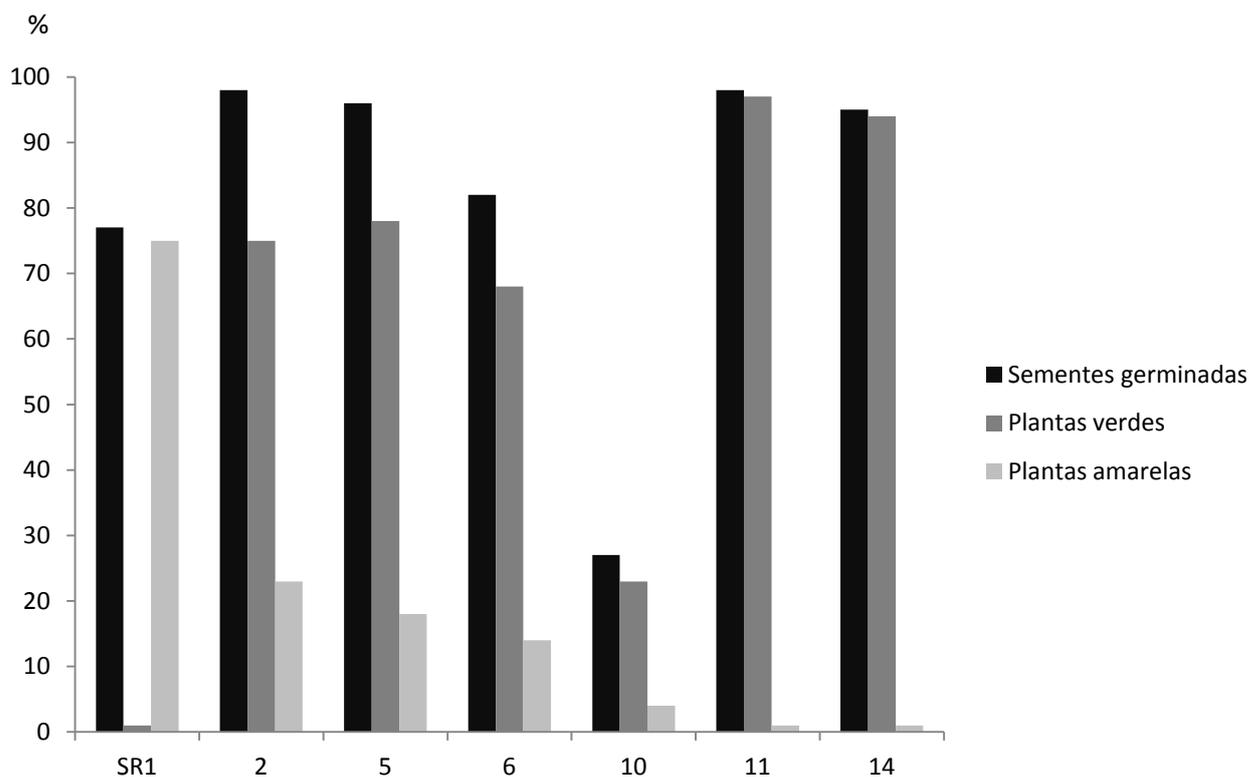


Figura 19: Germinação de sementes e sobrevivência de plantas de *N. tabacum* SR1 controle (SR1, não transformada) e de seis linhagens da geração T₂ de plantas transformadas pelo T-DNA do plasmídeo pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos} (2, 5, 6, 10, 11 e 14). As sementes foram germinadas *in vitro* em meio ½MS acrescido de ágar a 1,5% e Kan¹⁰⁰ durante 25 dias. As barras representam o percentual de sementes germinadas e de plantas resistentes (verdes) ou suscetíveis (amarelas) ao antibiótico.

O método de transformação por *A. tumefaciens*, utilizado na transformação das plantas de tabaco, permite produzir plantas transgênicas com baixo número de cópias do transgene e estes podem ser transmitidos à progênie seguindo padrão Mendeliano (revisado em YIN *et al.*, 2004). Assim, supondo-se partir de uma planta T₁ com uma cópia única do transgene (planta hemizigota), a sua autofecundação geraria plantas T₂ com chances de 25% de serem não-transgênicas, 50% de serem transgênicas hemizigotas e 25% de serem transgênicas homozigotas. Considerando-se que apenas uma cópia do gene de resistência à canamicina é suficiente para a sobrevivência da planta no meio de seleção, tanto indivíduos hemizigotos como homozigotos seriam capazes de se desenvolver, ao passo que plantas não transgênicas não se desenvolveriam neste meio. Conforme a **Figura 19**, as plantas 2, 5, 6 e 10 apresentaram 77 a 85% de indivíduos resistentes à canamicina dentre as sementes que germinaram, o que pode ser interpretado como a soma de indivíduos hemizigotos (50% de chance) e homozigotos

(25% de chance) na geração T₂. Esta situação caracteriza as plantas progenitoras como hemizigotas para o transgene e contendo uma única cópia do T-DNA integrada em seus genomas. Outra situação possível é o caso de a planta T₁ ser homozigota e, neste caso, seriam esperados na geração T₂ 100% de indivíduos resistentes à canamicina e homozigotos. Este padrão pode ser observado nos indivíduos 11 e 14 que, conforme observável na **Figura 19**, tiveram 98% das plântulas germinadas com capacidade de sobreviver em meio com canamicina. Este ensaio de seleção serviu ainda para demonstrar que as plantas SR1, não transformadas, ainda que capazes de germinação no meio de seleção com canamicina, não conseguiram manter seu desenvolvimento e crescimento após 25 dias.

4.3.3 Teste das plantas transgênicas quanto à tolerância ao estresse salino

Ensaio para verificar a capacidade de germinação das sementes transgênicas e controles em NaCl a 100 mM foram realizados e os principais resultados encontra-se na **Figura 20**.

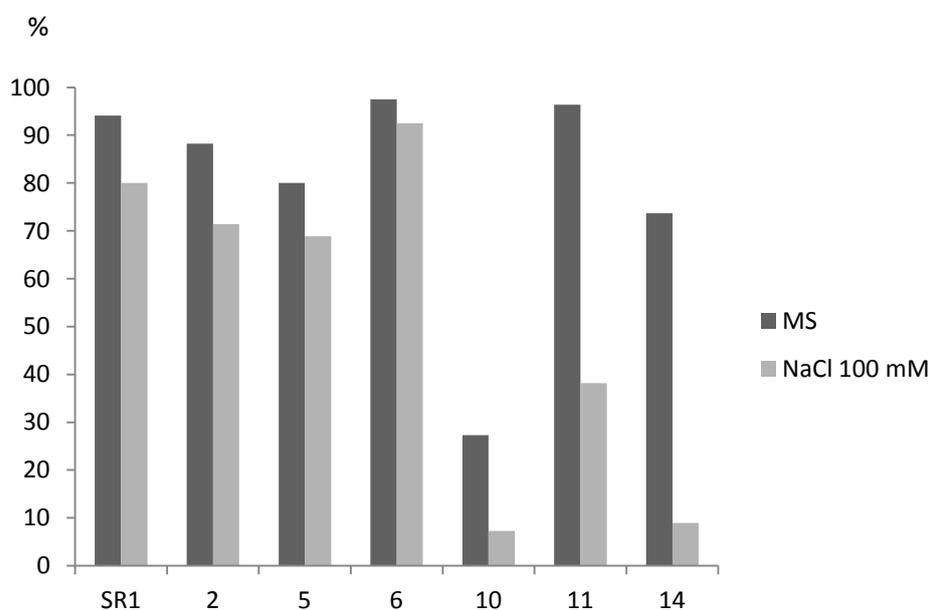


Figura 20: Germinação *in vitro* de sementes de *N. tabacum* SR1 controle (SR1, não transformada) e de seis linhagens da geração T₂ de plantas transformadas pelo T-DNA do plasmídeo pCAMBIA2300-P_{35S}-Egmpk5-T_{nos} (2, 5, 6, 10, 11 e 14) em meio 1/2MS acrescido de ágar a 1,5% e NaCl a 100 mM.

Nos testes de tolerância das plantas transgênicas e controles de tabaco à salinidade, foram também testadas as concentrações de 200 e 300 mM de NaCl, mas estas impediram completamente a germinação das sementes (resultados não apresentados). Conforme mostra a **Figura 20**, a adição de NaCl em concentração final de 100 mM aos meios de cultura reduziu a capacidade de germinação das sementes tanto da linhagem selvagem como das plantas transgênicas. As plantas transformadas 2 e 5 tiveram uma redução da germinação semelhante à planta selvagem, isto é, de cerca de 15%. Já a planta 6 mostrou-se melhor adaptada para germinar em condições de salinidade, tendo uma redução de germinação de 5%. As outras plantas ensaiadas, 10, 11 e 14, tiveram a germinação mais afetada pelo meio com NaCl do que a planta selvagem. Este ensaio demonstrou que sementes de tabaco são capazes de germinar e tolerar quantidades de até 100 mM de NaCl, sendo que a expressão de *mpk5* em cinco das seis plantas de tabaco não permitiu concluir efeitos positivos ou negativos sob estas condições. Possivelmente, concentrações de NaCl na faixa de 150 mM poderiam trazer melhores respostas ao efeito da expressão de *Egmpk5* em tabaco.

4.3.4 Teste das plantas transgênicas quanto à tolerância à seca

Com vistas a avaliar a possível tolerância das plantas transgênicas ao estresse hídrico em virtude da expressão do gene *Egmpk5*, sementes de plantas transformadas e controles foram germinadas e as plantas foram cultivadas em condições de hidroponia. Após pleno estabelecimento das plantas, o meio de irrigação foi removido por quatro dias. Após o quarto dia, a irrigação com meio nutritivo foi reinstalado e a análise das plantas foi realizada até o 36^o dia (**Figura 21**).



Figura 21: Estresse por seca a que foram submetidas plantas de *N. tabacum* SR1 controle (WT, não transformada) e de seis linhagens da geração T₂ de plantas transformadas pelo T-DNA do plasmídeo pCAMBIA2300-P_{35S}-*Egmpk5*-T_{nos} (2, 5, 6, 10, 11 e 14). O estresse hídrico foi imposto por quatro dias e a recuperação das plantas, após restabelecimento da irrigação, foi acompanhado até o 36º dia. Duplicatas de cada linhagem foram submetidas ao tratamento. Uma planta de cada linhagem, representativa do grupo total analisado, está ilustrando a figura.

As plantas selvagens apresentaram folhas extremamente secas já no 2º dia de estresse hídrico, o que perdurou durante todo o período de seca. Mesmo quando as plantas foram recolocadas em meio nutritivo, nenhuma planta selvagem foi capaz de se recuperar, morrendo em 30 dias. Das plantas transformadas, a que teve pior desempenho foi a 14, que perdeu folhas e, mesmo em meio hidropônico, não foi capaz de recuperar-se plenamente. As demais plantas transgênicas testadas (5, 6, 10 e 11) exibiram menores perdas de folhas e plena recuperação, levando-as a uma aparência equivalente e, por vezes, melhor do que as plantas no 36º dia de ensaio em relação a sua aparência no início. Assim, estes resultados mostraram que a superexpressão de *mpk5* nas plantas transformadas pode trazer vantagens para a sobrevivência de plantas adultas sob seca.

Outro teste de estresse hídrico foi realizado com plantas em solo sem, no entanto, ter resultados conclusivos (dados não apresentados). Entre plantas de uma mesma linhagem, sejam plantas transformadas ou controles, houve indivíduos com diferentes graus de suscetibilidade ao estresse. Após 30 dias de seca, a análise da totalidade dos indivíduos mostrou o mesmo perfil de desidratação e morte de folhas, e, após 10 dias de recuperação, as plantas tiveram perfil equivalente de recuperação. Uma possível explicação para a não concordância entre os dois testes de seca é que ambos induzem situações diferentes de estresse hídrico. No sistema hidropônico, o déficit hídrico se dá de forma abrupta, enquanto que em solo esta condição se instala de forma lenta e gradual, permitindo adaptação da planta. Em soja foi demonstrado que estresse hídrico induzido em solo e em hidroponia causa diferentes perfis de transcrição de genes da família RD, que participam na adaptação a condições de seca (NEVES-BORGES *et al.*, 2012).

Assm, das análises com plantas de tabaco expressando de forma heteróloga o gene *mpk5*, foi possível estabelecer que esta condição pode vir a ser vantajosa para plantas sob certas situações de estresse. O ensaio de germinação em sal demonstrou que em pelo menos uma linhagem – possivelmente heterozigota para o transgene –, o gene *mpk5* parece conferir vantagem para germinação de sementes em NaCl a 100 mM. Já no ensaio de seca de plantas sob hidroponia, cinco plantas transformadas com *mpk5* demonstraram melhor capacidade de recuperação do que plantas selvagens a estresse de quatro dias.

5. Discussão

As MAPKs integram um importante mecanismo de sinalização em plantas, o qual contribui para a transmissão de sinais e adaptação dos indivíduos a condições ambientais. Mediante a organização de proteínas em cascatas, MAPKs são ativadas por fosforilação sequencial por MAP3Ks e MAP2Ks, tendo como resultado a fosforilação de substratos nucleares e citoplasmáticos, que incluem fatores de transcrição, outras cinases e proteínas do citoesqueleto (conforme revisões por RODRIGUEZ *et al.*, 2010; ŠAMAJOVÁ *et al.*, 2013a e 2013b). Em plantas, cascatas de MAPKs estão envolvidas na resposta a sinais de desenvolvimento e a estresses bióticos e abióticos, como, por exemplo seca, frio, salinidade, estresse oxidativo e infecção por patógenos (conforme revisão por SINHA *et al.*, 2011).

Dada a participação de MAPKs em diversas respostas da célula vegetal, a análise do gene *mpk5* de *E. grandis* foi conduzida com vistas a melhor entender a participação deste gene e de seu produto na transdução de sinais durante o estresse. Neste trabalho foram feitas análises de estrutura, número de cópias e localização cromossômica deste gene, assim como caracterização de seu perfil de transcrição em situações normais e sob tratamentos. Finalmente, plantas de tabaco expressando *mpk5* de forma heteróloga foram submetidas a estresses de seca e salinidade, e o impacto da presença do transgene em seus genomas foi caracterizado.

5.1 Análises *in silico* do gene *mpk5* de *E. grandis* e proteína deduzida

A análise *in silico* teve como base a sequência do cDNA de *Egmpk5* que fora previamente clonado e sequenciado (KÖRBES, 2006). A busca pelo número de cópias de *mpk5* no genoma de *E. grandis* resultou em dois possíveis genes, sendo a sequência obtida pelo *scaffold* 12 descartada por não conter parte importante da sequência codificadora do domínio catalítico de PKs (HANKS, 2003). Dessa forma, a análise no banco de dados *Phytozome* levou à identificação de um único provável gene codificador de *Egmpk5*, localizado no *scaffold* 7. Na literatura são encontrados relatos da existência tanto de genes *mapks* com cópia única como genes duplicados em genomas vegetais (HAMEL *et al.*, 2006; KONG *et al.*, 2012; ZHANG *et al.*, 2013). O gene *mpk5* apresenta-se como cópia única no genoma de *E. grandis*, assim como os genes de outras

espécies com quais tem similaridade de sequência como, por exemplo, *mpk4* de tabaco (ZHANG *et al.*, 2012), *mpk6* de arroz (LIU *et al.*, 2007) e *mpk6* de *B. distachyon* (CHEN *et al.*, 2012).

A organização do gene *Egmpk5* e de outros quatro genes *mapks* segue um padrão de seis éxons e cinco íntrons nas sequências (**Figura 9**). Dentre todas as sequências analisadas, há maior conservação do tamanho dos éxons internos e variação daqueles localizados nas extremidades dos genes. A importância destas regiões de tamanho conservado pode ser verificada pelo fato de a sequência peptídica de 260 aminoácidos, resultante da tradução destes quatro éxons internos, apresentar o domínio conservado de PKs e os sítios de ligação ao substrato e a ATP e apresentar também os motivos de ativação. Assim, o fato de os éxons de *Egmpk5* terem tamanho conservado reflete uma seleção negativa atuando mais fortemente sobre as sequências codificantes (NICOLE *et al.*, 2006) e, neste caso, a conservação se dá na região catalítica, que é fundamental para a atividade enzimática das proteínas.

A sequência do gene *Egmpk5* identificada *in silico* apresenta sítios canônicos para *splicing* de seu transcrito. Isto foi evidenciado pela presença dos dinucleotídeos 5'-GT e 3'-AG (**Figura 5a**), descritos como altamente conservados nas extremidades dos íntrons e responsáveis pela correta identificação e edição destas regiões. Cerca de 90% dos genes codificadores de proteínas em plantas apresentam íntrons, o que traz consequências para as células como aumento da complexidade de regulação e possibilidade de variação dos produtos gênicos via *splicing* alternativo (conforme revisão em SZCZESNIAK *et al.*, 2013).

A localização dos íntrons em relação aos códons do gene foi também avaliada para *Egmpk5*. Com exceção do primeiro íntron, que se insere após a segunda base de um códon, todos os outros íntrons de *Egmpk5* começam antes ou depois de um códon (formando íntrons de fase 0). Estudos em *A. thaliana* e *P. trichocarpa* permitiram demonstrar que, nestes organismos, os genes *mpk* possuem cerca de 80% dos éxons em fase 0 (NICOLE *et al.*, 2006), assim como acontece no gene estudado de *E. grandis*.

Para melhor entendimento da regulação transcricional e da possível função de *Egmpk5*, foram preditos *cis*-elementos que atuam na região promotora do gene utilizando o banco de dados *PlantCare* (**Tabela 2**). Os elementos presentes na região promotora revelam um potencial papel do gene na regulação de crescimento e

desenvolvimento vegetais. Ademais, a identificação de elementos típicos de promotores eucarióticos é mais um indicativo de que este gene é, de fato, transcrito, e é corroborado por outros dados já abordados por este trabalho: o fato de haver possivelmente uma única sequência no genoma que codifica este gene e também pelo fato de que esta sequência de trabalho deriva de buscas em banco de *ESTs* (KÖRBES, 2006), que foram geradas a partir de uma biblioteca de cDNAs que reflete os transcritos da planta.

No alinhamento das sequências peptídicas deduzidas de EgMPK5 com outras MAPKs de plantas, foram encontrados consensos importantes nas regiões amino- e carbóxi-terminais. Estas regiões foram descritas como imprescindíveis para o arranjo tridimensional que define o sítio catalítico destas enzimas (UBERSAX & FERREL, 2007). Na porção amino-terminal, foram detectados consensos importantes (**Figura 11b-c**) que, segundo a literatura, são resíduos responsáveis por ancorar e orientar o nucleotídeo doador de fosfato às MAPKs. Já na porção carbóxi-terminal, os domínios VI a XI (**Figura 11d-f**), apresentam resíduos responsáveis pela ligação ao substrato peptídico e início da reação de transferência do grupamento fosfato. Outros consensos importantes identificados são o motivo “TEY” (**Figura 11e**), característico de MAPKs do grupo A, e o sítio C-terminal de *docking* (**Figura 11g**), responsável pela ligação a MAP2Ks, substratos e fosfatases e que exerce papel importante na regulação da atividade enzimática (ICHIMURA *et al.*, 2002). Em 1996, HANKS & HUNTER descreveram que membros de uma mesma família de PKs tendem a apresentar funções relacionadas e isso se manifesta por similaridades na topologia das estruturas, no modo de regulação e na especificidade de substratos. Dessa forma, o fato de MAPKs com atividades já descritas e caracterizadas, inclusive em nível proteico, apresentarem estes domínios conservados, é um forte indicativo da provável atividade apresentada por proteínas semelhantes, como a MPK5 de *E. grandis*.

Da literatura, sabe-se que as MAPKs com as mesmas funções tendem a se agrupar nas árvores filogenéticas de acordo com os subgrupos A, B, C, D (ICHIMURA *et al.*, 2002). Na análise filogenética conduzida no presente trabalho com EgMPK5, todas as sequências foram agrupadas como pertencentes à família A. O resultado não poderia ter sido diferente uma vez que a seleção das sequências foi feita por similaridade de sequência peptídica (BLASTp) e os resultados foram filtrados por similaridade (**Figura 10**). Todas as sequências obtidas apresentaram a assinatura do sítio de ativação de PKs ativadas por mitógenos em plantas (*STKc_TEY_MAPK_plant*).

Dentre as proteínas semelhantes à EgMPK5 que foram obtidas nesta análise está MPK6 de *A. thaliana*, que já teve envolvimento comprovado em diversos tipos de estresses como o excesso de sal, frio, seca, fermento e estresse osmótico (ICHIMURA *et al.*, 2002). Outras duas MAPKs, MPK4 de *N. tabacum* (GOMI *et al.*, 2005) e MPK2 de *S. lycopersicum* (XING, *et al.*, 2001), também já tiveram seu envolvimento comprovado na resposta vegetal a fermentos. Outras MAPKs que compuseram a árvore filogenética foram descritas como envolvidas na resistência a patógenos, como a MPK2 de *G. max* (DAXBERGER *et al.*, 2007), a MPK6 de *A. hypogaea* (KUMAR & KIRTI 2010) e a MPK2 de *P. crispum* (KROJ *et al.*, 2003), e na sinalização da divisão celular, como a MSK7 de *M. sativa* (DUERR *et al.*, 1993) e a MAPK D5 de *P. sativum* (STAFSTROM *et al.*, 1993).

5.2 Análises da expressão de *mpk5*

Com vistas a dar início à caracterização funcional do gene *Egmpk5* e da proteína por ele codificada, análises da expressão gênica em nível da quantidade de transcritos presentes foram realizadas por RT-qPCR. Na análise de expressão espacial, foi demonstrado que *Egmpk5* é expresso em níveis equivalentes em folhas, caules e raízes de plantas de 90 dias tratadas com água (**Figura 12**). Outros autores detectaram transcritos de 19 das 20 *mapks* de *A. thaliana* em folhas, caules, raízes e flores, ainda que em baixos níveis. Um padrão semelhante de expressão constitutiva foi observado com componentes da família gênica das *mapks* em *O. sativa* e *P. trichocarpa* (HAMEL *et al.*, 2006). Alguns genes com maior similaridade a *Egmpk5* exibiram expressão diferencial em alguns órgãos, como é o caso do gene *mpk6* de *O. sativa*, o qual é mais expresso em folhas e flores se comparado a raízes e caules, e o gene *mpk6-2* de *P. trichocarpa*, o qual é mais expresso em raízes (HAMEL *et al.*, 2006; NICOLE *et al.*, 2006).

Acredita-se que o acúmulo de transcritos de genes induzidos por estresse possa levar ao acúmulo de seus produtos que, então, participam na amplificação dos sinais de estresse (LIU, 2012). O teste de RT-qPCR permitiu demonstrar que ABA foi capaz de induzir a expressão de *Egmpk5* em folhas (**Figura 13**). Vários autores demonstraram que o tratamento com ABA interfere na regulação transcricional de genes, no acúmulo de proteínas e na atividade de cascatas MAPKs em diversas plantas (revisado em

DANQUAH *et al.*, 2014). A maioria desses genes é ativada por ABA (LIU, 2012) e, assim, MAPKs têm sido implicadas em respostas dependentes de ABA, incluindo defesa antioxidante, sinalização nas células-guarda e na germinação de sementes (revisado em DANQUAH *et al.*, 2014).

O tratamento com NAA levou à uma redução do nível de expressão de *Egmpk5* em caules quando comparado com o controle não tratado (**Figura 14**). Há relatos de respostas negativas e positivas de *mapks* na sinalização por auxinas (MISHRA *et al.*, 2006). A interação entre auxinas e MAPKs influencia a divisão celular (mitose), a transcrição gênica, o rearranjo do citoesqueleto, a atividade de canais iônicos e o desenvolvimento da raiz (MOCKAITIS *et al.*, 2000; MISHRA *et al.*, 2006).

O tratamento de plântulas de *E. grandis* com excesso salino (NaCl a 200 mM) foi o único tratamento capaz de induzir *Egmpk5* nos três órgãos testados (**Figuras 13-15**). Em *A. thaliana*, os genes *mpk3*, *mpk4* e *mpk6* foram ativados por estresse hiperosmótico e por salinidade (DROILLARD *et al.*, 2004; KIM *et al.*, 2012). A proteína MPK6, quando ativada, tem um papel importante na detoxificação de sódio da célula por fazer efluxo deste íon e, também, atua na regulação da transcrição de genes induzíveis (TEIGE *et al.*, 2004; YU *et al.*, 2010). Em raízes de *A. thaliana*, MOCKAITIS *et al.* (2000) demonstrou que 250 mM de NaCl foi capaz de estimular significativamente a atividade de fosforilação de uma MAPK. Todos estes trabalhos corroboram os resultados observados de indução da expressão de *Egmpk5* em resposta a estresse salino. Coerentemente, o conjunto de resultados sugere que o aumento da expressão de *Egmpk5* em resposta ao excesso salino é mediado por ABA e auxina, uma vez que o gene responde positivamente a estes sinalizadores hormonais bem descritos na resposta vegetal a estresses abióticos (LIU, 2012; SMÉKALOVÁ *et al.*, 2013).

Como mencionado anteriormente, elementos de ação *cis* relacionados à resposta a estresse foram encontrados na região promotora de *Egmpk5* (**Tabela 2**). Entre eles, destacam-se ABRE (*ABA-responsive element*) e TGA. ABRE são elementos de ação *cis* importantes na resposta de tecidos vegetativos de *A. thaliana* à desidratação e a ABA (NAKASHIMA *et al.*, 2006). Os elementos TGA foram descritos como sítios promotores de ligação de fatores TGA1, os quais são proteínas do tipo zíper-de-leucinas que provavelmente participam na transcrição tecido-específica de plantas em resposta a hormônios e a sinais de estresse (PASCUZZI *et al.*, 1998).

Em plantas como tomateiro (KONG *et al.*, 2012) e *A. thaliana* (MENGES *et al.*, 2008), transcritos de *mapks* foram detectados em altos níveis em órgãos reprodutivos como estames e pétalas, comparado a tecidos vegetativos como folhas e caules. Embora na presente análise da expressão de *Egmpk5* não tenham sido incluídos órgãos reprodutivos e, portanto, não há dados sobre este tipo de expressão, a análise *in silico* permitiu revelar o *cis* elemento GCN4 que pode participar da expressão tecido-específica em endosperma (WU *et al.*, 1998).

Há relatos na literatura da existência de pares de genes da família *mapk* em uma mesma planta que diferem entre si nos seus perfis de transcrição. Um exemplo é o par *mpk6-mpk3* em *A. thaliana*, em que o primeiro tem expressão constitutiva e o segundo tem a transcrição regulada diferencialmente (MENGES *et al.*, 2008). Dada a característica de plantas de codificarem múltiplos genes da família *mapk*, novos estudos para identificar outros membros desta família gênica em *E. grandis* poderiam complementar os dados de expressão obtidos para *Egmpk5*, principalmente para os tratamentos em que não houve mudança nos níveis do transcrito.

5.3 Análise de plantas de tabaco expressando *Egmpk5*

Para melhor estudo de função de *Egmpk5*, plantas de tabaco previamente transformadas com este gene foram analisadas frente a estresses. O teste de germinação *in vitro* sob NaCl a 100 mM (**Figura 20**) demonstrou que, para a maioria das plantas transformadas, a expressão heteróloga do gene *mpk5* não conferiu vantagem frente a este estresse. No entanto, a linhagem transformada 6 apresentou redução de apenas 5% na sua germinação, o que significa melhor desempenho do que as plantas selvagens (com aproximadamente 15% de redução na germinação). De forma semelhante, a expressão heteróloga de um gene *mapk* de algodoeiro (*Ghmpk2*) em tabaco por Zhang *et al.* (2011) foi capaz de conferir melhores taxas de germinação sob NaCl a 100 mM. Ainda, no trabalho mencionado, Zhang *et al.* (2011) demonstraram que as plantas transformadas foram inclusive mais tolerantes a NaCl a 200 mM, concentrações que inibiram a germinação de todas as plantas testadas no presente trabalho. Na literatura também foi relatada a expressão heteróloga do gene *ZmSIMK1* de milho em *A. thaliana*, que permitiu melhor germinação de plantas transformadas sob NaCl a 200 mM

comparadas a plantas controle, e plantas adultas em solo irrigadas com esta mesma solução também mostraram melhor adaptação (GU *et al.*, 2010).

Aliando-se os resultados do efeito de sal sobre a germinação de sementes com os obtidos no teste de genótipo das plantas transgênicas de tabaco, observou-se que as plantas 11 e 14, possivelmente homozigotas para *mpk5*, tiveram piores taxas de germinação do que as plantas possivelmente hemizigotas ou, mesmo, as selvagens. Isto pode indicar que a expressão de *mpk5* em homozigose não confere vantagem às plantas quando germinadas em meio com adição de sal. A expressão de um gene *mapk* nem sempre traz vantagens para a planta transformada sobreviver a estresses, conforme foi demonstrado para uma MAPK do grupo D de algodoeiro (SHI *et al.*, 2011). Neste trabalho, Shi *et al.* (2011) demonstraram que o gene *Ghmpk16* expresso em *A. thaliana* reduziu as taxas de germinação e de crescimento de raiz sob condições de seca, além de causar um grande acúmulo de peróxido de hidrogênio.

Para testar o efeito de estresse em plantas adultas, desafio de seca foi imposto a plantas crescidas em hidroponia. As cinco linhagens transformadas que foram expostas a seca por quatro dias sofreram menos danos do que plantas selvagens, que foram incapazes de sobreviver mesmo após transferidas para recuperação em meio hidropônico. Em um trabalho recente, foi demonstrado que plantas de tabaco expressando o gene *mpk3* de algodoeiro, quando submetidas a seca em solo por 14 dias, também apresentaram melhor adaptação ao estresse, o que foi verificado por maior retenção de água e menor murcha de folhas (LONG *et al.*, 2014). Os autores deste trabalho mencionado anteriormente, ainda descreveram que a melhor adaptação pôde ser correlacionada com a maior expressão de enzimas antioxidantes pelas plantas transformadas.

6. Conclusão e Perspectivas

Neste trabalho, o gene *Egmpk5* foi melhor caracterizado *in silico* e teve seus perfis de expressão espacial e em resposta a sinais hormonais e de estresse mais bem definidos em plantas de *E. grandis*. O efeito de expressão de *Egmpk5* em plantas de tabaco foi também analisado frente a estresses abióticos como excesso salino e seca.

As análises *in silico* permitiram mostrar semelhança entre o gene *mpk5* de *E. grandis* e outros genes *mapk* de plantas já caracterizados. O peptídeo deduzido EgMPK5 possui importantes regiões estruturais e homologia com MAPKs vegetais com atividade já comprovada frente a estímulos de desenvolvimento e respostas a estresses. Estes resultados permitem afirmar que, teoricamente, a proteína codificada pelo gene *Egmpk5* tem atividade cinásica semelhante às versões existentes em outros vegetais

O perfil de transcrição de *Egmpk5* em nível de acúmulo de transcritos foi caracterizado frente a estresses abióticos e em resposta a hormônios. Os tratamentos com NaCl e ABA causaram aumento da expressão do gene, permitindo-nos sugerir sua possível participação na transdução de sinais em células vegetais causada por estes tratamentos. Para um estudo mais detalhado da participação deste gene na sinalização em *E. grandis*, será importante caracterizar outros genes responsivos a estresses que podem sofrer regulação por *Egmpk5*. Ainda, a detecção do produto do gene *Egmpk5* e da determinação de sua atividade poderão contribuir para o completo entendimento de sua função.

As análises de plantas transgênicas de tabaco capazes de expressar *Egmpk5* permitiram mostrar que uma das linhagens transformadas teve melhor taxa de germinação sob estresse salino e que cinco plantas adultas tiveram melhor tolerância ao estresse por seca. A análise de genótipos quanto à hemi/homozigose sugere que o transgene em homozigose pode não ser vantajoso para plantas sob estresse, e que futuros ensaios de quantificação da expressão gênica e detecção da proteína e de possíveis compostos acumulados por estas plantas podem detalhar o mecanismo de tolerância ao estresse mediado por *Egmpk5*.

7. Referências bibliográficas

- ABRAF – Associação Brasileira de Produtores de Florestas Plantadas. Anuário estatístico ABRAF 2013 ano base 2012 148 p.. Brasília: 2013. Disponível em http://www.abraflor.org.br/estatisticas/ABRAF13/ABRAF13_BR.pdf
- Asai T, Tena G, Plotnikova J, Willmann M, Chiu W, Gomez-Gomez L, Boller T, Ausubel, F, Sheen, J. MAP kinase signaling cascade in *Arabidopsis* innate immunity. *Nature* 415:977–983, 2002
- Avruch J. MAP kinase pathways: The first twenty years. *Biochimica et Biophysica Acta* 1773 1150–1160, 2007
- Bardwell L, Shah K. Analysis of mitogen-activated protein kinase activation and interactions with regulators and substrates. *Methods* 40:213–223, 2006
- Bedon F, Majada J, Feito I, Chaumeil P, Dupuy J, Lomenech A, Barre A, Gion J, Plomion C. Interaction between environmental factors affects the accumulation proteins in hydroponically grown *Eucalyptus globulus* (Labill.). *Plant Physiology and Biochemistry* 49:69-76, 2011
- Bedon F, Villar E, Vincent D, Dupuys J, Lomenech A, Mabialangoma A, Chaumeil P, Barré A, Plomion C, Gion J. Proteomic plasticity of two *Eucalyptus* genotypes under contrasted water regimes in the field. *Plant Cell and Environment* 35, 790-805, 2012
- Berg J, Tymoczko J, Stryer L. Catalytic Strategies. In: *Biochemistry* 5th ed, New York: W H Freeman, 2002
- BRACELPA – Associação Brasileira de Celulose e Papel. Apresentação Dados do Setor, 30 fl., Março/2014. Disponível em <http://www.bracelpa.org.br/bra2/sites/default/files/estatisticas/booklet.pdf>
- Burnett E, Desikan R, Moser R, Neill S. ABA activation of an MBP kinase in *Pisum sativum* epidermal peels correlates with stomatal responses to ABA. *Journal of Experimental Botany* 51:197-205, 2000
- Chen L, Hu W, Tan W, Wang M, Ma Z, Zhou S, Deng X, Zhang Y, Huang C, Yang G, He G. Genome-wide identification and analysis of MAPK and MAPKK gene families in *Brachypodium distachyon*. *PLoS ONE* 7:e46744, 2012
- Costa e Silva F, Shvaleva A, Broetto F, Ortuño MF, Rodrigues ML, Almeida M. Acclimation to short-term low temperatures in two *Eucalyptus globulus* clones with contrasting drought tolerance. *Tree Physiology* 29:77–86, 2008
- Crooks G, Hon G, Chandonia J, Brenner S. WebLogo: A sequence logo generator. *Genome Research* 14:1188-119, 2004
- Danquah A, Zelicourt A, Colcombet J, Hirt H. The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnology Advances* 32:40–52, 2014

- Daxberger A, Nemark A, Mithöfer A, Fliegmann J, Ligterink W, Hirt H, Ebel J. Activation of members of a MAPK module in beta-glucan elicitor-mediated non-host resistance of soybean. *Planta* 225:1559-1571, 2007
- De Block M, Herrera-Estrella L, Van Montagu M, Schell J, Zambryski P. Expression of foreign genes in regenerated plants and their progeny. *The EMBO Journal* 3:1681-1689, 1984
- Decroocq-Ferrant V, Decroocq S, Vanwent J, Schmidt E, Kreis M. A homologue of the MAP/ERK family of protein kinase genes is expressed in vegetative and in female reproductive organs of *Petunia hybrida*. *Plant Molecular Biology* 27:339-350, 1995
- DeWald D, Torabinejad J, Jones C, Shope J, Cangelosi A. Rapid accumulation of phosphatidylinositol 4,5-bisphosphate and inositol 1,4,5-trisphosphate correlates with calcium mobilization in salt-stressed *Arabidopsis*. *Plant Physiology* 126:759-769, 2001
- Dóczy R, Okrész L, Romero A, Paccanaro A, Bögre L. Exploring the evolutionary path of plant MAPK networks. *Trends in Plant Science* 17:518-525, 2012
- Doyle J, Doyle J. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 19:11-15, 1987
- Droillard M, Boudsocq M, Barbier-Brygoo H, Lauriere C. Involvement of MPK4 in osmotic stress response pathways in cell suspensions and plantlets of *Arabidopsis thaliana*: activation by hypoosmolarity and negative role in hyperosmolarity tolerance. *FEBS Letters* 574:42-48, 2004
- Duerr B, Gawienowski M, Ropp T, Jacobs T. MsERK1: a mitogen-activated protein kinase from a flowering plant. *Plant Cell* 5:87-96, 1993
- Edgar, R. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792-1797, 2004
- Fu S, Chou W, Huang D, Huang H. Transcriptional regulation of a rice mitogen-activated protein kinase gene, OsMAPK4, in response to environmental stresses. *Plant and Cell Physiology* 43: 958-963, 2002
- Gomi K, Ogawa D, Katou S, Kamada H, Nakajima N, Saji H. A mitogen-activated protein kinase NtMPK4 activated by SIPKK is required for jasmonic acid signaling and involved in ozone tolerance via stomatal movement in tobacco. *Plant Cell Physiology* 46:1902-1914, 2005
- Gu L, Liu Y, Zong X, Liu L, Li D, Li D. Overexpression of maize mitogen-activated protein kinase gene, ZmSIMK1 in *Arabidopsis* increases tolerance to salt stress. *Molecular Biology Reports* 37:4067-4073, 2010
- Hamel LP, Nicole MC, Sritubtim S, Morency MJ, Ellis M, *et al.* Ancient signals: comparative genomics of plant MAPK and MAPKK gene families. *Trends in Plant Science* 11:192-198, 2006
- Hanks S, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase catalytic domain structure and classification. *The FASEB Journal* 9:576-596, 1995
- Hanks S. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biology* 4:111-118, 2003

- Hoehenwarter W, Thomas M, Nukarinen E, Egelhofer V, Röhrig H, Weckwerth W, Conrath U, Beckers G. Identification of novel in vivo MAP kinase substrates in *Arabidopsis thaliana* through use of tandem metal oxide affinity chromatography. *Mol Cell Proteomics* 12:369-380, 2013
- Horsch RB, Fraley RT, Rogers SG, Sanders PR, Lloyd A, Hoffmann N. Inheritance of functional genes in plants. *Science* 223:496-498, 1984
- Huang G, Ma S, Bai L, Zhang L, Ma H, Jia P, Liu J, Zhong M, Guo Z. Signal transduction during cold, salt, and drought stresses in plants. *Molecular Biology Reports* 39:969–987, 2012
- Huttly A, Pillips A. Gibberellin-regulated expression in oat aleurone cells of two kinases that show homology to MAP kinase and a ribosomal protein kinase. *Plant Molecular Biology* 27:1043-1052, 1995
- Ichimura K *et al* (MAPK Group). Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends in Plant Science* 7:301-308, 2002
- Jammes F, Yang X, Xiao S, Kwak J. Two *Arabidopsis* guard cell-preferential MAPK genes, MPK9 and MPK12, function in biotic stress response. *Plant Signaling & Behavior* 6:1875-1877, 2011
- Ji H, Pardo J, Batelli G, Van Oosten M, Bressan R, Li X. The Salt Overly Sensitive (SOS) pathway: established and emerging roles. *Molecular Plant* 6:275-86, 2013
- Jonak C, Ökrész L, Bögre L, Hirt H. Complexity, cross talk and integration of plant MAP kinase signaling. *Current Opinion in Plant Biology* 5:415–424, 2002
- Jonak C, Nakagami H, Hirt H. Heavy Metal Stress. Activation of distinct mitogen-activated protein kinase pathways by copper and cadmium. *Plant Physiology* 136:3276–3283, 2004
- Kim J, Woo D, Kim S, Lee S, Park H, Seok H. *Arabidopsis* MKKK20 is involved in osmotic stress response via regulation of MPK6 activity. *Plant Cell Reports* 31:217–224, 2012
- Kirch, Rochele Patrícia. Análise da expressão gênica de *mapk* de *Eucalyptus grandis* e em plantas transgênicas de tabaco. Monografia, 49 f (Trabalho de Conclusão da Disciplina de Estágio Curricular em Farmácia). Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, 2008
- Knetsch M, Wang M, Snaar-Jagalska B, Heimovaara-Dijkstra S. Abscisic acid induces mitogen-activated protein kinase activation in barley aleurone protoplasts. *The Plant Cell* 8:1061-1067, 1996
- Kong F, Wang J, Cheng L, Liu S, Wu J, Peng Z, Lu G. Genome-wide analysis of the mitogen-activated protein kinase gene family in *Solanum lycopersicum*. *Gene* 499:108–120, 2012
- Körbes, Ana Paula. Avaliação da capacidade de genes de *Eucalyptus grandis* em conferir tolerância à deficiência hídrica. Dissertação, 143f (Mestrado em Genética e Biologia Molecular). Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 2006

- Kroj T, Rudd J, Nürnberger T, Gäbler Y, Lee J, Scheel D. Mitogen-activated protein kinases play an essential role in oxidative burst-independent expression of pathogenesis-related genes in parsley. *The Journal of Biological Chemistry* 278:2256-2264, 2003
- Kumar K, Kirti P. A mitogen-activated protein kinase, AhMPK6 from peanut localizes to the nucleus and also induces defense responses upon transient expression in tobacco. *Plant Physiology and Biochemistry* 48:481-486, 2010
- Lalle M, Visconti S, Marra M, Camoni L, Velasco R, Aduc P. ZmMPK6, a novel maize MAPK kinase that interacts with 14-3-3 proteins. *Plant Molecular Biology* 59:713-722, 2005
- Lee J, Huh K, Bhargava A, Ellis B. Comprehensive analysis of protein-protein interactions between *Arabidopsis* MAPKs and MAPK kinases helps define potential MAPK signalling modules. *Plant Signaling & Behavior* 3:1037-1041, 2008
- Lescot M, Dhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouz P, Rombauts S. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research* 30:325-327, 2002
- Li Y, Feng D, Zhang D, Su J, Zhang Y, Li Z, Mu P, Liu B, Wang H, Wang J. Rice MAPK phosphatase IBR5 negatively regulates drought stress tolerance in transgenic *Nicotiana tabacum*. *Plant Science* 188-189:10-8, 2012
- Lian W, Tang Y, Gao S, Zhang Z, Zhao X, Zhao C. Phylogenetic analysis and expression patterns of the MAPK gene family in wheat (*Triticum aestivum* L.). *Journal of Integrative Agriculture* 11:1227-1235, 2012
- Lieberherr D, Thao N, Nakashima A, Umemura K, Kawasaki T, Shimamoto K. A sphingolipid elicitor-inducible mitogen-activated protein kinase is regulated by the small GTPase OsRac1 and heterotrimeric G-protein in rice. *Plant Physiology* 138:1644-1652, 2005
- Liu Y, Jin H, Yang K, Kim C, Baker B, Zhang S. Interaction between two mitogen-activated protein kinases during tobacco defense signaling. *The Plant Journal* 34:149-160, 2003
- Liu Q, Xue Q. Computational identification and phylogenetic analysis of the MAPK gene family in *Oryza sativa*. *Plant Physiology and Biochemistry* 45:6-14, 2007
- Liu Y. Roles of mitogen-activated protein kinase cascades in ABA signaling. *Plant Cell Reports* 31:1-12, 2012
- Liu X, Nguyen X, Kim K, Han H, Yoo J, Lee K. Phosphorylation of the zinc finger transcriptional regulator ZAT6 by MPK6 regulates *Arabidopsis* seed germination under salt and osmotic stress. *Biochemical and Biophysical Research Communications* 430:1054-1059, 2013
- Livak K, Schmittgen T. Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402-408, 2001
- Long L, Gao W, Xu L, Liu M, Luo X, He X, Yang X, Zhang X, Zhu L. GbMPK3, a mitogen-activated protein kinase from cotton, enhances drought and oxidative stress tolerance in tobacco. *Plant Cell, Tissue and Organ Culture* 116:153-162, 2014

- Lumbreras V, Vilela B, Irar S, Solé M, Capellades M, Valls M, Coca M, Pagès M. MAPK phosphatase MKP2 mediates disease responses in *Arabidopsis* and functionally interacts with MPK3 and MPK6. *The Plant Journal* 63:1017-1030, 2010
- Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends in Biochemical Sciences* 27:514-520, 2002
- Marshall C. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80:179-185, 1995
- Mayrose M, Bonshtien A, Sessa G. LeMPK3 is a mitogen-activated protein kinase with dual specificity induced during tomato defense and wounding responses. *The Journal of Biological Chemistry* 279:14819-14827, 2004
- Menges M, Dóczy R, Ökrész L, Morandini P, Mizzi L, Soloviev M, Murray J, Bögre J. Comprehensive gene expression atlas for the *Arabidopsis* MAP kinase signalling pathways. *New Phytologist* 179:643–662, 2008
- Mishra N, Tuteja R, Tuteja N. Signaling through MAP kinase networks in plants. *Archives of Biochemistry and Biophysics* 452:55–66, 2006
- Mizoguchi T, Hayashida N, Yamaguchi-Shinozaki K, Kamada H, Shinozaki K. AtMPKs: a gene family of plant MAP kinases in *Arabidopsis thaliana*. *FEBS Letters* 336:44044, 1993
- Mockaitis K, Howell S. Auxin induces mitogenic activated protein kinase MAPK activation in roots of *Arabidopsis* seedlings. *The Plant Journal* 24:785-796, 2000
- Munnik T, Meijer H. Osmotic stress activates distinct lipid and MAPK signalling pathways in plants. *FEBS Letters*, 498:172-178, 2001
- Murashige T, Skoog F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* 15:437-498, 1963
- Murphy L, Smith S, Chen R, Fingar C, Blenis J. Molecular interpretation of ERK signal duration by immediate early gene products. *Nature Cell Biology*, 4: 556-564, 2002
- Nakagami H, Pitzschke A, Hirt H. Emerging MAP kinase pathways in plant stress signaling. *Trends in Plant Science* 10:339-346, 2005
- Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of *Arabidopsis*. *Plant Molecular Biology* 60:51-68, 2006
- Navarrete-Campos D, Bravo L, Rubilar R, Emhart V, Sanhueza R. Drought effects on water use efficiency, freezing tolerance and survival of *Eucalyptus globulus* and *Eucalyptus globulus x nitens* cuttings. *New Forests* 44:119–134, 2013
- Neves-Borges A, Guimarães-Dias F, Cruz F, Mesquita R, Nepomuceno A, Romano E, Loureiro M, Grossi-de-Sá M, Alves-Ferreira M. Expression pattern of drought stress marker genes in soybean roots under two water deficit systems. *Genetics and Molecular Biology*, 35:212-221, 2012

- Nicole M, Hamel L, Morency M, Beaudoin N, Ellis B, Séguin A. MAP-ping genomic organization and organ-specific expression profiles of poplar MAP kinases and MAP kinase kinases. *BMC Genomics* 7:223-245, 2006
- Nishihama R, Banno H, Kawahara E, Irie K, Machida Y. Possible involvement of differential splicing in regulation of the activity of *Arabidopsis* ANP1 that is related to mitogen-activated protein kinase kinase kinases (MAPKKKs). *The Plant Journal* 12:39-48, 1997
- Oliveira L, Breton M, Bastolla F, Camargo S, Margis R, Frazzon J, Pasquali G. Reference genes for the normalization of gene expression in *Eucalyptus* species. *Plant Cell Physiology* 53: 405-422, 2012
- Pascuzzi P, Hamilton D, Bodily K, Arias J. Auxin-induced stress potentiates trans-activation by a conserved plant basic/leucine-zipper factor. *The Journal of Biological Chemistry* 273:26631-26637, 1998
- Pita P, Pardos J. Growth, leaf morphology, water use and tissue water relations of *Eucalyptus* globulus clones in response to water deficit. *Tree Physiology* 21:599-607, 2001
- Plotnikov A, Zehorai E, Procaccia S, Seger R. The MAPK cascades: Signaling components, nuclear roles and mechanisms of nuclear translocation. *Biochimica et Biophysica Acta* 1813:1619-1633, 2011
- Popescu S, Popescu G, Bachan S, Zhang Z, Gerstein M, Snyder M, Dinesh-Kumar S. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Development* 23:80-92, 2009
- Rodriguez M, Petersen M, Mundy J. Mitogen-activated protein kinase signaling in plants. *Annual Review of Plant Biology* 61:621–49, 2010
- Romeis, T. Protein kinases in the plant defence response. *Current Opinion in Plant Biology* 4:407-414, 2001
- Šamajová O, Komis G, Šamaj J. Emerging topics in the cell biology of mitogen-activated protein kinases. *Trends in Plant Science* 18:140-148, 2013a
- Šamajová O, Plíhal O, Al-Yousif M, Hirt H, Šamaja J. Improvement of stress tolerance in plants by genetic manipulation of mitogen-activated protein kinases. *Biotechnology Advances* 31:118-128, 2013b
- Samuel M, Miles G, Ellis B. Ozone treatment rapidly activates MAP kinase signalling in plants. *The Plant Journal* 22:367-376, 2000
- Shen H, Liu C, Zhang Y, Meng X, Zhou X, Chu C. OsWRKY30 is activated by MAP kinases to confer drought tolerance in rice. *Plant Molecular Biology* 80:241–253, 2012
- Shi J, Zhang L, An H, Wu C, Guo X. *GhMPK16*, a novel stress-responsive group D MAPK gene from cotton, is involved in disease resistance and drought sensitivity. *BMC Molecular Biology* 12:22-37, 2011
- Shou H, Bordallo P, Wang K . Expression of the *Nicotiana* protein kinase (NPK1) enhanced drought tolerance in transgenic maize. *Journal of Experimental Botany* 55:1013-1019, 2004

- Sinha A, Jaggi M, Raghuram B, Tuteja N. Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signaling & Behavior* 6:196-203, 2011
- Smékalová V, Doskočilová A, Komis G, Samaj J. Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. *Biotechnol Advances* 32:2-11, 2013
- Souza C, Rossi L, Azevedo C, Lima R. Behaviour of *Acacia mangium* and clones of *Eucalyptus grandis* x *E. urophylla* in experimental plantations in Central Amazonia. *Scientia Forestalis* 65:95-101, 2004
- Soyano T, Nishihama R, Morikiyo K, Ishikawa M, Machida Y. NQK1/NtMEK1 is a MAPKK that acts in the NPK1 MAPKKK-mediated MAPK cascade and is required for plant cytokinesis. *Genes Development* 17:1055-1067, 2003
- Stafstrom J, Altschuler M, Anderson D. Molecular cloning and expression of a MAP kinase homologue from pea. *Plant Molecular Biology* 22:83-90, 1993
- Szczesniak M, Kabza M, Pokrzywa R, Gudys A, Makalowska I. ERISdb: A database of plant splice sites and splicing signals. *Plant Cell Physiology* 54:1-8, 2013
- Taiz L, Zeiger E. Stress physiology. In: *Plant Physiology*, 4th edition. Sinauer Associates Inc. Sunderland, 2006
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731-2739, 2011
- Tang N, Zhang H, Li X, Xiao J, Xiong L. Constitutive activation of transcription factor OsbZIP46 improves drought tolerance in rice. *Plant Physiology* 158:1755-1768, 2012
- Teige M, Scheikl E, Eulgem T, Doczi R, Ichimura K, Shinozaki K, Dangl J, Hirt H. The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Molecular Cell* 15:141-152, 2004
- Tena G, Asai T, Chiu W, Sheen J. Plant mitogen-activated protein kinase signaling cascades. *Current Opinion in Plant Biology* 4:392-400, 2001
- Thompson J, Higgins D, Gibson T. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22:4673-4680, 1994
- Thumma B, Sharma N, Southerton S. Transcriptome sequencing of *Eucalyptus calmadulensis* seedlings subjected to water stress reveals functional single nucleotide polymorphisms and genes under selection. *BMC Genomics* 13:364-385, 2012
- Tsugama D, Liu S, Takano T. Drought-induced activation and rehydration-induced inactivation of MPK6 in *Arabidopsis*. *Biochemical and Biophysical Research Communications* 426: 626-629, 2012
- Tuteja N, Mahajan S. Cold, salinity and drought stresses: An overview. *Archives of Biochemistry and Biophysics* 444:139-158, 2005

- Ubersax A, Ferrel J. Mechanisms of specificity in protein phosphorylation. *Nature Reviews Molecular Cell Biology* 8:530–541, 2007
- Valdés A, Irar S, Majadad J, Rodríguez A, Fernández B, Pagès M. Drought tolerance acquisition in *Eucalyptus globulus* (Labill.): A research on plant morphology, physiology and proteomics. *Journal of Proteomics* 79: 263-276, 2013
- Villar E, Klopp C, Noirot C, Novaes E, Kirst M, Plomion C, Gion J. RNA-Seq reveals genotype-specific molecular responses to water deficit in *Eucalyptus*. *BMC Genomics*, 12:538-556, 2011
- Voorrips, R. MapChart: Software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity* 93:77-78, 2002
- Voronin V, Touraev A, Kieft H, van Lammeren A, Heberle-Bors E, Wilson C. Temporal and tissue-specific expression of the tobacco ntf4 MAP kinase. *Plant Molecular Biology* 45:679-689, 2001
- Voronin V, Aionesei T, Limmongkon A, Barinova I, Touraev A, Lauriere C, Coronado M, Testillano P, Risueno M, Heberle-Bors E, Wilson C. The MAP kinase kinase NtMEK2 is involved in tobacco pollen germination. *FEBS Letters* 560:86-90, 2004
- Wang J, Ding H, Zhang A, Ma F, Cao J, Jiang M. A novel mitogen-activated protein kinase gene in maize (*Zea mays*), ZmMPK3, is involved in response to diverse environmental cues. *Journal of Integrative Plant Biology* 52:442–452, 2010
- Wilson C, Eller N, Gartner A, Vicente O, Heberle-Bors E. Isolation and characterization of a tobacco cDNA clone encoding a putative MAP kinase. *Plant Molecular Biology* 23:543–51, 1993
- Wu C, Suzuki A, Washida H, Takaiwa F. The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *The Plant Journal* 14:673-83, 1998
- Xing T, Malik K, Martin T, Miki B. Activation of tomato PR and wound-related genes by a mutagenized tomato MAP kinase kinase through divergent pathways. *Plant Molecular Biology* 46:109-120, 2001
- Xu H, Li K, Yang F, Shi Q, Wang X. Overexpression of CsNMAPK in tobacco enhanced seed germination under salt and osmotic stresses. *Molecular Biology Reports* 37:3157-3163, 2010
- Yin, Z; Plyder W, Malepszy S. Transgene inheritance in plants. *Journal of Applied Genetics* 45:127-144, 2004
- Yu L, Nie J, Cao C, Jin Y, Yan M, Wang F. Phosphatidic acid mediates salt stress response by regulation of MPK6 in *Arabidopsis thaliana*. *New Phytologist* 188:762–773, 2010
- Yung Y, Yao Z, Hanoch T, Seger R. ERK1b, a 46-kDa ERK isoform that is differentially regulated by MEK. *The Journal of Biological Chemistry* 275:15799-15808, 2000
- Zhang S, Klessig D. MAPK cascades in plant defense signaling. *Trends in Plant Science* 6:520-527, 2001

Zhang T, Liu Y, Xue L, Xu S, Chen T, Yang T, Zhang L, An L. Molecular cloning and characterization of a novel MAP kinase gene in *Chorispora bungeana*. *Plant Physiology and Biochemistry* 44:78-84, 2006

Zhang S (2008) Mitogen-activated protein kinase cascades in plant intracellular signaling. In: Yang Z (ed) *Intracellular signaling in plants*. *Annual Plant Reviews* 33:100-136. Blackwell Publishing.

Zhang L, Xi D, Li S, Gao Z, Zhao S, Shi J, Wu C, Guo X. A cotton group C MAP kinase gene, GhMPK2, positively regulates salt and drought tolerance in tobacco. *Plant Molecular Biology* 77:17–31, 2011

Zhang S, Xu R, Luo X, Jiang Z, Shu H. Genome-wide identification and expression analysis of MAPK and MAPKK gene family in *Malus domestica*. *Gene* 531:377–387, 2013

Zhou Y, Zhang D, Pan J, Kong X, Liu Y, Sun L, Wang L, Li D. Overexpression of a multiple stress-responsive gene, ZmMPK4, enhances tolerance to low temperature in transgenic tobacco. *Plant Physiology and Biochemistry* 58:174-181, 2012

8. ANEXO I – Alinhamento da sequência codificante do clone de cDNA de *mpk5* com as duas CDS obtidas dos *scaffolds* do genoma

```

          10      20      30      40      50      60
          |      |      |      |      |      |
CDS_scf07 ATGGACGGCGGGG-CTCCTC-AGCCGG-CGGATACCGTCATGTTCGGAGGCGGCGCC-GGC
CDS_cDNA  ATGGACGGCGGGG-CTCCTC-AGCCGG-CGGATACCGTCATGTTCGGAGGCGGCGCC-GGC
CDS_scf12 ATGGACGGCGGGGGCTCCTCCAGCCGGGCGGATACTGTTCATGTTCGGAGGTGGCAGCCCGGC
*****
Prim.cons. ATGGACGGCGGGGGCTCCTCCAGCCGGGCGGATAACCGTCATGTTCGGAGGCGGCGCCCGGC

          70      80      90      100     110     120
          |      |      |      |      |      |
CDS_scf07 GCAGCAGCAGCAGCAGCC-GCAGCAGGCGCAGCC-GCAGGGGATCGAGAACATCCCGGCG
CDS_cDNA  GCAGCAGCAGCAGCAGCC-GCAGCAGGCGCAGCC-GCAGGGGATCGAGAACATCCCGGCG
CDS_scf12 GCAGCAGCAGCAGCAGCCCGCAGCAGGCGCAGCCAGCAGGGGATCGAGA-TATCC-GGCG
*****
Prim.cons. GCAGCAGCAGCAGCAGCCCGCAGCAGGCGCAGCCAGCAGGGGATCGAGAACATCCCGGCG

          130     140     150     160     170     180
          |      |      |      |      |      |
CDS_scf07 ACGCTCAGCCACGGGGGCGCCTTCATCC-AGTACAACATCTTCGGCAACATCTTCGAGGT
CDS_cDNA  ACGCTCAGCCACGGGGGCGCCTTCATCC-AGTACAACATCTTCGGCAACATCTTCGAGGT
CDS_scf12 ACGCTCAGCCACGGGGGCGCCTTCATCCAGTACAACATCTTCGGCAACATCTTCGAGGT
*****
Prim.cons. ACGCTCAGCCACGGGGGCGCCTTCATCCAGTACAACATCTTCGGCAACATCTTCGAGGT

          190     200     210     220     230     240
          |      |      |      |      |      |
CDS_scf07 CACCGCCAAGTACAAGCCCCCATCATGCCATCGGCAAGGGCGCCTACGGCATCGTCTG
CDS_cDNA  CACCGCCAAGTACAAGCCCCCATCATGCCATCGGCAAGGGCGCCTACGGCATCGTCTG
CDS_scf12 CACCGCCAAGTACAAGCCCCCATCATGCCATCGGCAAGGGCGC-TACGGCATCGTCTG
*****
Prim.cons. CACCGCCAAGTACAAGCCCCCATCATGCCATCGGCAAGGGCGCCTACGGCATCGTCTG

          250     260     270     280     290     300
          |      |      |      |      |      |
CDS_scf07 CTCAGCTTTGAATTTCGGAGACGAACGAGCACGTGGCCATAAAGAAGATTGCTAATGCTTT
CDS_cDNA  CTCAGCTTTGAATTTCGGAGACGAACGAGCACGTGGCCATAAAGAAGATTGCTAATGCTTT
CDS_scf12 CTCAGCTTTGAATTTCGGAGACGAACGAGCACGTGGCCATAAAGAAGATTGCTAATGCTTT
*****
Prim.cons. CTCAGCTTTGAATTTCGGAGACGAACGAGCACGTGGCCATAAAGAAGATTGCTAATGCTTT

          310     320     330     340     350     360
          |      |      |      |      |      |
CDS_scf07 CGATAACAAGATCGATGCGAAGAGGACTCTCCGTGAGATCAAGCTTCTCCGGCACATGGA
CDS_cDNA  CGATAACAAGATCGATGCGAAGAGGACTCTCCGTGAGATCAAGCTTCTCCGGCACATGGA
CDS_scf12 CGATAACAAGATCGATGCGAAGAGGACTCTCCGTGAGATCAAGCTTCTCCGGCACATGGA
*****
Prim.cons. CGATAACAAGATCGATGCGAAGAGGACTCTCCGTGAGATCAAGCTTCTCCGGCACATGGA

          370     380     390     400     410     420
          |      |      |      |      |      |
CDS_scf07 CCATGAAAACGTCGTGGCAATTAGGGATATTATTCCACCGCCACAGAGAGAGGTGTTCAA
CDS_cDNA  CCATGAAAACGTCGTGGCAATTAGGGATATTATTCCACCGCCACAGAGAGAGGTGTTCAA
CDS_scf12 CCATGAAAACGTTGTGGCAATTAGGGATATTATTCCACCGCCACAGAGAGAGGTGTTCAA
*****
Prim.cons. CCATGAAAACGTCGTGGCAATTAGGGATATTATTCCACCGCCACAGAGAGAGGTGTTCAA

```

	430	440	450	460	470	480
CDS_scf07	TGATGTTTATATTGCATATGAGCTTATGGACACTGATCTGCATCAAATTATTCGTTCCAA					
CDS_cDNA	TGATGTTTATATTGCATATGAGCTTATGGACACTGATCTGCATCAAATTATTCGTTCCAA					
CDS_scf12	TGATGTTTATATTGCATATGAGCTTATGGACACTGATCTGCATCAAATTATTCGTTCCAA					
Prim.cons.	*****					
Prim.cons.	TGATGTTTATATTGCATATGAGCTTATGGACACTGATCTGCATCAAATTATTCGTTCCAA					
	490	500	510	520	530	540
CDS_scf07	CCAAGCATTGTCTGAGGAGCATTGTCAGTATTTTCTATATCAGATCTTGCAGGATTAAA					
CDS_cDNA	CCAAGCATTGTCTGAGGAGCATTGTCAGTATTTTCTATATCAGATCTTGCAGGATTAAA					
CDS_scf12	CCAAGCATTGTCTGAGGAGCATTGTCAGTATTTTCTATATCAGATCTTGCAGGATTAAA					
Prim.cons.	*****					
Prim.cons.	CCAAGCATTGTCTGAGGAGCATTGTCAGTATTTTCTATATCAGATCTTGCAGGATTAAA					
	550	560	570	580	590	600
CDS_scf07	ATACATACATTCTGCAAATGTTCTGCATAGAGACTTGAAGCCAGCAATCTTCTCCTAAA					
CDS_cDNA	ATACATACATTCTGCAAATGTTCTGCATAGAGACTTGAAGCCAGCAATCTTCTCCTAAA					
CDS_scf12	ATACATACATTCTGCAAATGTTCTGCATAGAGACTTGAAGCCAGCAATCTTCTCCTAAA					
Prim.cons.	*****					
Prim.cons.	ATACATACATTCTGCAAATGTTCTGCATAGAGACTTGAAGCCAGCAATCTTCTCCTAAA					
	610	620	630	640	650	660
CDS_scf07	TGCAAATTGCGATTTGAAAATATGTGATTTTGGACTAGCTCGTGTCACTTCTGAAACTGA					
CDS_cDNA	TGCAAATTGCGATTTGAAAATATGTGATTTTGGACTAGCTCGTGTCACTTCTGAAACTGA					
CDS_scf12	TGCAAATTGCGATTTGAAAATATGTGATTTTGGTCTAGCTCGTGTCACTTCTGAAACTGA					
Prim.cons.	*****					
Prim.cons.	TGCAAATTGCGATTTGAAAATATGTGATTTTGGACTAGCTCGTGTCACTTCTGAAACTGA					
	670	680	690	700	710	720
CDS_scf07	TTTTATGACAGAATATGTTGTCACAAGATGGTACCGTGCACCAGAGCTATTGTTAAATTC					
CDS_cDNA	TTTTATGACAGAATATGTTGTCACAAGATGGTACCGTGCACCAGAGCTATTGTTAAATTC					
CDS_scf12	TTTTATGACAGAATATGTTGTCACAAGATGGTACCGTGCACCAGAGCTACTGTAAATTC					
Prim.cons.	*****					
Prim.cons.	TTTTATGACAGAATATGTTGTCACAAGATGGTACCGTGCACCAGAGCTATTGTTAAATTC					
	730	740	750	760	770	780
CDS_scf07	TTCAGACTATACGGCGGCAATAGATGTATGGTCTGTAGGCTGTATCTTTATGGAACTAAT					
CDS_cDNA	TTCAGACTATACGGCGGCAATAGATGTATGGTCTGTAGGCTGTATCTTTATGGAACTAAT					
CDS_scf12	TTCAGACTATACGGCGGCAATAGATGTATGGTCTGTAGGCTGTATCTTTATGGAACTAAT					
Prim.cons.	*****					
Prim.cons.	TTCAGACTATACGGCGGCAATAGATGTATGGTCTGTAGGCTGTATCTTTATGGAACTAAT					
	790	800	810	820	830	840
CDS_scf07	GGATCGGAAACCCTTGTTTCCTGGCAGAGACCATGTGCAACAGCTGCGTTTGTGATGGA					
CDS_cDNA	GGATCGGAAACCCTTGTTTCCTGGCAGAGACCATGTGCAACAGCTGCGTTTGTGATGGA					
CDS_scf12	GGATCGGAAACCCTTGTTTCCTGGCAGAGACCATGTGCAACAGCTGCGTTTGTGATGGA					
Prim.cons.	*****					
Prim.cons.	GGATCGGAAACCCTTGTTTCCTGGCAGAGACCATGTGCAACAGCTGCGTTTGTGATGGA					
	850	860	870	880	890	900
CDS_scf07	GCTGATTGGCACCCCATCAGAGGCAGAGTTGGGGTTCCTTAAATGAAAATGCTAAGAAGTA					
CDS_cDNA	GCTGATTGGCACCCCATCAGAGGCAGAGTTGGGGTTCCTTAAATGAAAATGCTAAGAAGTA					
CDS_scf12	G-----					
Prim.cons.	*					
Prim.cons.	GCTGATTGGCACCCCATCAGAGGCAGAGTTGGGGTTCCTTAAATGAAAATGCTAAGAAGTA					

```

          910      920      930      940      950      960
CDS_scf07   TATCAGACAGCTTCCCTCTGTACCGTCGGCAATCTTTCACTGAAAAGTTTCCCCATGTCCA
CDS_cDNA    TATCAGACAGCTTCCCTCTGTACCGTCGGCAATCTTTCACTGAAAAGTTTCCCCATGTCCA
CDS_scf12   -----

Prim.cons.  TATCAGACAGCTTCCCTCTGTACCGTCGGCAATCTTTCACTGAAAAGTTTCCCCATGTCCA

          970      980      990      1000     1010     1020
CDS_scf07   CCCGCTTGCAATCGATCTCGTTGAGAAGATGTTAACGTTTCGATCCCAGGCTGAGGCTCAC
CDS_cDNA    CCCGCTTGCAATCGATCTCGTTGAGAAGATGTTAACGTTTCGATCCCAGGCTGAGGCTCAC
CDS_scf12   -----

Prim.cons.  CCCGCTTGCAATCGATCTCGTTGAGAAGATGTTAACGTTTCGATCCCAGGCTGAGGCTCAC

          1030     1040     1050     1060     1070     1080
CDS_scf07   AGTTGAAGAGGCATTGGCTCATCCCTACCTAAACTCACTGCACGACATCAGCGATGAGCC
CDS_cDNA    AGTTGAAGAGGCATTGGCTCATCCCTACCTAAACTCACTGCACGACATCAGCGATGAGCC
CDS_scf12   --TTGAAGAGGCATTGGCTCATCCCTACCTAAACTCACTGCATGACATCAGCGATGAGCC
          *****
Prim.cons.  AGTTGAAGAGGCATTGGCTCATCCCTACCTAAACTCACTGCACGACATCAGCGATGAGCC

          1090     1100     1110     1120     1130     1140
CDS_scf07   AACTTGATGAATCCATTCAACTTCGACTTTGAGCAGCATGCACTCACGGAGGAACAGAT
CDS_cDNA    AACTTGATGAATCCATTCAACTTCGACTTTGAGCAGCATGCACTCACGGAGGAACAGAT
CDS_scf12   GACTTGATGAATCCATTCAACTTCGACTTTGAGCAGCATGCACTCACGGAGGAACAGAT
          *****
Prim.cons.  AACTTGATGAATCCATTCAACTTCGACTTTGAGCAGCATGCACTCACGGAGGAACAGAT

          1150     1160     1170     1180     1190
CDS_scf07   GAGGGAGTTAATTTATAGGGAAGCGCTTGCATTTAATCCCAGATATCTACAGTAA
CDS_cDNA    GAGGGAGTTAATTTATAGGGAAGCGCTTGCATTTAATCCCAGATATCTACAGTAA
CDS_scf12   GAGGGAGTTAATTTATAGGGAAGCGCTTGCATTTAATCCCAGATATCTACAGTAA
          *****
Prim.cons.  GAGGGAGTTAATTTATAGGGAAGCGCTTGCATTTAATCCCAGATATCTACAGTAA

```

Alinhamento das sequências de DNA codificante (CDS) após edição de regiões não codificadoras de sequências nucleotídicas obtidas por clonagem de cDNA de *Egmpk5* (representado por CDS_cDNA) e por busca por BLASTn no genoma de *E. grandis* no *scaffold* 7 (representado por CDS_scf07) e no *scaffold* 12 (representado por CDS_scf12). Consenso no alinhamento está indicado pela linha “prim. cons”. Regiões que melhor alinharam as três sequências estão representadas sem destaque, e regiões que indicam consenso de sequência apenas entre CDS_cDNA e CDS_scf07 estão marcadas com um retângulo. Este alinhamento foi construído com as três sequências para melhor representação gráfica, no entanto os dados apresentados na seção 4.1.1 (número e percentual de nucleotídeos similares) foram obtidos por alinhamento par a par de CDS_cDNA com cada uma das sequências obtidas pelos *scaffolds*.

9. ANEXO II - The Mitogen-Activated Protein Kinase 5 (MPK5) from *Eucalyptus grandis* (manuscrito a ser submetido para a revista *Molecular Biology Reports*)

The Mitogen-Activated Protein Kinase 5 (MPK5) from *Eucalyptus grandis*

Krás Borges, JD¹; Kirch, RP¹; Körbes, AP²; Maraschin, FS³; Margis, R¹; Bodanese-Zanettini, MH²; Pasquali, G^{1*}

¹Graduate Program in Cell and Molecular Biology, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91.501-970, Brazil ²Graduate Program in Genetics and Molecular Biology, Inst. Biociências, UFRGS.

³Dept. Botânica, Inst. Biociências, UFRGS.

*Corresponding author: E-mail, pasquali@cbiot.ufrgs.br;

Keywords: *Eucalyptus grandis*; mitogen-activated protein kinase; MAPK; signal transduction; stress tolerance; transgenic tobacco.

Abstract

Mitogen-activated protein kinases (MAPKs) are modules in phosphorylation cascades found in all eukaryotes and are responsible for transducing extracellular stimuli into intracellular responses. In plants, MAPK cascades are involved in growth and development processes and have important roles in the response to hormones and biotic and abiotic stresses. In this study, we cloned and characterized *EgMPK5*, a group A MAPK gene found in *Eucalyptus grandis*. Southern blot and *in silico* analyses suggest the presence of the gene as a single copy in *E. grandis* genome. RT-qPCR analysis showed equal expression of *EgMPK5* in roots, stems and leaves of plants treated with water and we detected increase of expression in leaves treated with ABA and also increase in expression in all three organs upon treatment with NaCl. Analysis of the promoter region of *EgMPK5* revealed a group of putative cis-acting elements related to stress responses, and the deduced product of this gene presents homology to known functional MAPKs and shares the protein kinase domain signature, ATP-binding site and dual phosphorylation motifs. The gene was overexpressed in *Nicotiana tabacum*

SR1 plants under the control of the CaMV 35S promoter and individuals showed increased tolerance to salt and drought stresses.

Introduction

Plants in their natural environment are exposed to different biotic and abiotic stresses and have thus evolved mechanisms to perceive external signals and translate them into intracellular responses that lead to the organism adaptation [1]. The mitogen-activated protein kinases (MAPKs) are members of a signaling pathway that is highly conserved in eukaryotes and mediates adaptive processes in plants [2]. MAPKs have been reported to have multiple roles in the cells, such as biotic- and abiotic-stress responses [3,4], hormone signaling [5], and regulation of cell division, growth and differentiation [6, 7].

MAPKs are organized in modules within the cell and function downstream of sensors, acting to convert perceived external signals into cellular responses that alter gene expression [8]. The main components of a MAPK cascade are three interconnected kinases, MAP kinase kinase kinases (MAP3Ks/MAPKKKs), MAP kinase kinases (MAP2Ks/MAPKKs) and MAP kinases (MAPKs/MPKs) [9]. Such members are activated through sequential phosphorylation of specific amino acid residues, and activated MAPKs target different substrates in the cytoplasm or nucleus [2], such as transcription factors, other protein kinases, and cytoskeleton-associated proteins [4]. Members of the MAPK pathway are activated in response to many stimuli, including pathogens, drought, salinity, cold, wounding, O₃, reactive oxygen species, and hormone stimuli [3,8]. Plant MAPKs can be grouped in four distinct families that have characteristic sequence signatures, and MAPKs in family A are mostly involved in environmental and hormonal responses [10].

A variety of genes encoding members of the MAPK pathway have been reported [2,5] and the completion of genome sequencing projects aided the identification of possibly all members of the MAPK family in several plant species. Among eudicot species, *Arabidopsis thaliana* presents 20 MAPK genes in its genome [11], *Solanum lycopersicum* presents 16 [12], and, among woody species, there are 21 MAPKs in *Populus trichocarpa* [13] and 26 in *Malus domestica* [14]. Monocot species have

similar high numbers, as is the case with 15 MAPKs in *Oryza sativa* [15] and 16 in *Brachypodium distachyon* [16]. Recently, *Eucalyptus* genomic and transcriptomic data have been decoded, thus providing a tool for genome analysis and study of genes in this plant as well [17]. Plants of the genus *Eucalyptus* are mainly planted for industrial applications such as cellulose pulp and paper [18], but in spite of its economic importance, the mechanisms by which plants of this genus overcome stresses is less known than those of model species [19].

Research on MAPKs has been mainly focused in model and crop plants, and to our knowledge there is no report of characterization of MAPKs in *Eucalyptus*. In the present study, we isolated and characterized *EgMPK5*, a novel MAPK gene in *Eucalyptus grandis*, using database analyses and PCR methods. The gene was then characterized for its exon/intron structure, chromosomal location, sequence homology to other plant species and phylogenetic relationships. We also investigated *EgMPK5* expression profile under physiological conditions and stresses by quantitative real-time PCR (RT-qPCR). To further analyze the function of *EgMPK5* in response to stresses, transgenic tobacco (*Nicotiana tabacum*) over expressing the gene under the control of CaMV 35S promoter was produced and assayed against stress conditions.

Materials and methods

Plant materials

Seeds of *E. grandis* and *N. tabacum* cv. SR1 were grown in soil-vermiculite mixture (1:1) and kept in a light chamber under a photoperiod cycle of 16h light/8h dark at a temperature of (25 °C). For hydroponic culture, tobacco plants were transferred to plastic tanks containing 2 L ½ Murashige-Skoog nutrient solution (MS) [20] with pH 5,6-5,7 and kept under the same conditions described above.

Surface sterilization of seeds and *in vitro* growth conditions

Seeds of *N. tabacum* cv. SR1 were surface sterilized in 70% ethanol for 2 minutes, followed by 1,5% sodium hypochlorite solution for 15 min and four rinses with autoclaved distilled water. For *in vitro* culture, plants were placed either in Petri

dishes or culture bottles containing 30 mL of ½ MS salts and 1,5% agar. Seeds assayed for kanamycin resistance were sown on *in vitro* culture media containing such antibiotic at a final concentration of 100 mg/L.

Stress treatments

For assessment of salt stress on tobacco seed germination, NaCl was added to *in vitro* germination media to final concentrations of 100, 200 and 300 mM. For drought stress, water was withdrawn from hydroponic-grown tobacco plants for 4 days.

For characterization of *EgMPK5* expression profile under treatments, ninety-day old *E. grandis* seedlings grown in soil were treated in triplicates with 20 mL of the following solutions: 50 µM naphthaleneacetic acid (NAA), 100 µM abscisic acid (ABA), 10 µM kinetin (KIN) and 200 mM sodium chloride. Plants were sampled at 3 days after each treatment, except for drought assay that had water withdrawn from plants for a week. After treatments, roots, stems and leaves were collected individually from each plant, frozen in liquid nitrogen and stored at -80 °C for RNA extraction.

Total DNA extraction and PCR

Total DNA was isolated from wild-type (SR1) and transformed *N. tabacum* plants according to previously described protocol [21]. Extracted DNA was amplified by PCR using the following primers: 35Sfor, forward primer 5' CGACACTCTCGTCTACTCCAAG 3' and MAPK-RTrev, reverse primer 5' TCTGTGGCGGTGGAATAATA 3' (**Fig 5a**). Each reaction had a total volume of 50 µL, containing either 5 µL of plant DNA or 0,5 µL of plasmid DNA as templates, 10 µL of 5x reaction buffer (Promega), 2 µL of dNTP's solution (dATP, dGTP, dCTP e dTTP at 5 mM each), 2 µL of MgCl₂ (25 mM) (Promega), 2.5 µL of each primer (10 µM), 0.25 µL of GoTaq DNA polymerase (5 U/µL, Promega) and ultrapure water to final volume. The PCR conditions included a 3 minute initial denaturation at 96 °C, followed by 35 cycles of denaturation at 96 °C, annealing at 50 °C and extension at 72 °C for 30 s each, and a final 5 minute extension at 72 °C

Southern blot analysis

Genomic DNA was extracted from young *E. grandis* leaves as described above. Twenty µg of genomic DNA were digested with *EcoRI*, *EcoRV*, *HindIII*, *BamHI*, *Sall* and *PstI* and fractionated by electrophoresis on a 0.8% agarose gel. DNA fragments were transferred to Hybond N⁺ membranes (Amersham Biosciences). Transfer, hybridization and detection were performed using the Gene Images Random Prime Labelling and Gene Images CDP-Star Detection Modules (Amersham Biosciences) following the manufacturer's instructions. The 1,555 bp *XbaI/HindIII* fragment of *EgMPK5* cDNA was used as probe (**Fig 5a**).

Total RNA extraction and Quantitative Real Time PCR

Roots, stems and leaves of *E. grandis* were ground in liquid nitrogen and total RNA was extracted according to the CTAB protocol as described in the literature [21]. Extraction quality was analyzed in ethidium bromide-stained 1% agar gels and by UV absorbance. To remove the remaining genomic DNA, the RNA was treated with RQ1 RNase-Free DNaseI (Promega) according to the manufacturer's instructions. cDNA was synthesized with M-MuLV reverse transcriptase (New England Biolabs) according to the manufacturer's instructions.

For RT-qPCR analysis, the primers MAPK_RT forward 5' GATAACAAGATCGACGCGAA 3' and reverse 5' TCTGTGGCGGTGGAATAATA 3' were used (**Fig 5a**). The amplification of *EgMPK5* was performed using 10 µL of a 1:25 dilution of cDNA as a template, which was added to 10 µL of reaction mix. The final reaction media was composed of: 1x reaction buffer (Promega), dNTP's solution (dATP, dGTP, dCTP e dTTP at 25 µM each), MgCl₂ at 3 mM (Promega), SYBR green 0.1x, GoTaq DNA polymerase 0.25 U (Promega) and each primer at 0.2 µM.

The analyses were performed on ABI 7500 Real-Time PCR System (Applied Biotechnology) programmed to heat for 300 s at 94 °C (hot start), followed by 40 cycles of denaturation (15 s at 94 °C), annealing (10s at 60 °C) and extension (35s at 72 °C). Melting curve was obtained by heating from 50 to 94 °C (Δ 0.4 °C/s). Three biological and four technical replicates were performed for each sample. To normalize the total amount of cDNA present in each reaction, two constitutively expressed *E. grandis*

genes were amplified using the following primers: Ec8 forward 5' TCCAATCCGAGTCGCTGTCATTG 3', reverse 5' TGATGAGCC TCTCTGGTTTGACCT 3' and Ec04 5' TACAAGCGCTGTTGATATGTGGGC 3', reverse TTGCCAATGAGGCGGATTCACAAG 3' [22]. The expression level of different samples relative to constitutive genes was calculated using the $2^{-\Delta\Delta CT}$ method [23] and statistical analyses were performed with Student's t-test.

Definition of the sequence of work

The primary data source for this work were clustered DNA sequences of the GENOLYPTUS project database, derived from expressed sequence tags (ESTs) obtained from cDNAs generated from different tissues, organs or growth conditions. A putative full-length *MAPK* cDNA was identified among the ESTs and its identity was further confirmed by BLASTp and BLASTn in GenBank database [24]. Multiple fragments of *EgMPK5* cDNA cloned into pSport1 (Invitrogen) were generated with internal restriction enzymes and subcloned into SK⁺ pBluescript (Stratagene) to allow full sequencing. Internal primers were also designed (G1for, 5'-CACGAGGGATTTGCTCAATT-3'; G1rev, 5'-AGTGACCTGCATGAACCTCC-3'; G2for, 5'-GGGATTGGGACATGAAATTG-3'; G2rev, 5'-AAAAGAGCAGCCAAGGATCA-3'; G3rev, 5'-CTGACAAAACGCTGCAAGAA-3') and, along with universal SK⁺ primers, employed to complete sequencing of the insert.

Conserved domain detection and phylogenetic analysis

The putative amino acid sequence of *EgMPK5* was obtained by *in silico* translation of the coding sequence of *EgMPK5* cDNA. This sequence was queried on GenBank database using the tool BLASTp to identify proteins of the MAPK family in other plants. Sequences sharing at least 80% similarity with *EgMPK5* were selected to perform a multiple sequence alignment with MUSCLE algorithm. A phylogenetic tree including thirteen plants MAPKs was built with Maximum Likelihood method in the program MEGA5 [25] with default parameters and the reliability of the tree topology was tested by a bootstrap analysis with 500 replicates. Conserved motifs within the sequences were first identified by GenBank annotation and the consensus sequences of

the MAPK motif were analyzed on the MUSCLE alignment and the result was illustrated using the program WebLogo [26].

Gene structure and position on *E. grandis* chromosomes

For exon/intron determination, the full-length *EgMPK5* cDNA sequence was compared with the corresponding genomic DNA sequence obtained by BLASTn on Phytozome [17] database and to other *mapk* genes. Gene position on *E. grandis* genome was obtained by mapping gene location to scaffolds data using the MapChart program [27].

Analysis of *cis*-elements in promoter sequences

Genomic sequences 2 kb upstream of the start codon of the *EgMPK5* gene were browsed on PlantCare database [28] for putative *cis*-elements.

Tobacco genetic transformation

The cDNA cloned into pSport1 was transferred to vector pMOG463 (MOGEN) for acquisition of promoter (CaMV 35S) and terminator (*nos*) regions. The resulting construct was inserted into the binary vector pCAMBIA2300 (CAMBIA), which was introduced into the *Agrobacterium tumefaciens* strain LBA4404. Tobacco was transformed using the leaf disc co cultivation method [29, 30].

Results

Isolation and sequence analysis of *EgMPK5* cDNA

As part of the GENOLYPTUS Project: The Brazilian Research Network on the *Eucalyptus* Genome, our group screened a database of 120,000 expressed sequence tags (ESTs) for sequences related to MAPKs. A total of 3 clusters and 12 singlets with

deduced protein homology to MAPKs were identified in the GENOLYPTUS database, all derived from *E. grandis*. Among these ESTs, one putative full-length clone was identified based on the presence of ATG translational start codon aligned to other known plant MAPK sequences (result not shown) and due to the homology with stress related MAPKs. This 5,068 bp long cDNA sequence was named *EgMPK5* (**Fig. 1b**) and submitted to GenBank under the accession number _____ (to be defined).

Genomic organization of *EgMPK5* gene

The completion of *E. grandis* genome sequencing and assembly makes it possible to analyze the genomic organization of the plant. BLASTn search against *E. grandis* genomic sequences resulted in two sequences belonging to scaffolds 7 (6,936 bp) and 12 (6,550 bp), which showed high scores of alignment and low E-values. The cDNA sequence of *EgMPK5* was then aligned to the retrieved sequences and compared to known structures of MAPK genes described in the literature for the assignment of exons e introns. The sequence retrieved from scaffold 7 presents six exons, and its five introns belong to the GT-AG type (**Fig. 1a**). The 1,188 bp long coding DNA sequence (CDS) obtained by joining the exons is predicted to encode a protein of 396 aminoacids that possesses characteristic motifs known to the MAPK protein domain. The nucleotidic sequence retrieved from scaffold 12 has only five exons and the lack of one exon and one intron explains the length difference of studied sequences. The genomic sequence obtained from scaffold 12 fails to provide a potential functional MAPK, as its 1,011 bp CDS codes several stop codons and leads to a very small peptide with no homology to any known protein domains. Therefore it was concluded that *EgMPK5* is coded by a gene present in scaffold 7, which is designated chromosome G according to the main genome assembly *E. grandis* (**Fig. 1c**). Additionally, sequence comparison revealed that the cloned *EgMPK5* cDNA (5,068 bp) corresponds to an immature transcript of the gene. The sequence length difference between the cloned cDNA and the predicted CDS is due to the presence of two exons and one intron in the cloned sequence (**Fig. 1b**). *In silico* data is confirmed by our Southern blot result that detected a single copy of *EgMPK5* in the *E. grandis* genome (**Fig. 1d**).

Phylogenetic analysis of MAPKs

To investigate the phylogenetic relationships among MAPKs in plants, the full-length amino acid sequences of MAPKs were aligned by MUSCLE and analyzed with MEGA 5 employing the Maximum Likelihood method. The resulting phylogenetic tree grouped plant MAPKs in three main ramifications corresponding to monocots and eudicot species (**Fig. 2**). Among the eudicots, the phylogenetic tree was able to group MAPKs of plants belonging to the same family, as is the case with Fabaceae and Solanaceae species. Because the sequences used to build the tree were selected based on highest homology, they all belong to the family A of plant MAPKs and present characteristic motifs (**Fig. 3a**).

Domain analysis of EgMPK5 and similar plant MAPKs

Analysis of conserved motifs in all of the above sequences was performed by multiple sequence alignment of their conserved protein kinase domain. A BLASTn search in Genbank detected two domains in the sequences, STKc_Tey_MAPK_plant (cd07858) and PKc_like superfamily (cl09925). A more detailed search for sub domains present in the conserved domain showed that all of the sequences have a highly conserved TEY signature in their activation loop (**Fig. 3e**), which is characteristic of MAPKs belonging to group A. Besides that, important consensus sequences were found within the N- and C- terminal extensions of the kinase domain (**Fig. 3b-f**). The studied sequences also present a conserved CD domain (common docking site) at the C-terminal ends of the peptides (**Fig. 3g**).

EgMPK5 promoter analysis

To further understanding the transcriptional regulation and the potential function of *EgMPK5*, cis-elements in the gene promoter sequence were predicted using the PlantCare database. A total of 25 *cis*-elements were found in the region 2-kb upstream of the start codon and they can be categorized in three main groups: cis-elements involved in plant growth and development, cis-elements related to plants hormones and

stress responses, and common *cis*-elements related to promoter and enhancer regions (**Table 1**).

Expression profiles of *EgMPK5* under normal conditions and in response to signaling molecules/stress conditions

The expression pattern of *EgMPK5* in roots, leaves and stems of *E. grandis* under normal conditions and treatments was evaluated by RT-qPCR. *EgMPK5* mRNA was detected in leaves, roots, and stems of ninety-day old plants grown in soil and irrigated with water. Calculation of relative expression revealed that no statistical difference occurs in the accumulation of *EgMPK5* mRNA in the studied organs (**Fig 4a**). To determine if signaling molecules are capable of triggering differential gene expression, ninety-day old plants were treated with NAA, ABA, kinetin, NaCl and also exposed to drought conditions. In leaves, treatments with ABA and NaCl resulted in increased transcript accumulation of *EgMPK5* (**Fig. 4b**). Relative expression of *EgMPK5* decreased in stems of plants treated with NAA, but there was increase in expression under NaCl treatment (**Fig. 4c**). Additionally, NaCl was the only treatment that led to an increase in relative expression of *EgMPK5* in roots (**Fig. 4d**).

Identification of transgenic tobacco plants

To test the function of *EgMPK5* in response to stresses, transgenic tobacco plants over expressing the gene under the control of CaMV 35S promoter were produced. Transgenic lines were assayed for kanamycin resistance and confirmed by PCR with genomic DNA (**Fig. 5b**). Six lines (2; 5; 6; 10; 11; 14) among the transgenic plants were selected for further experiments. To assess transgene inheritance in selected plants, seeds of the T₁ generation were germinated in kanamycin-containing media and germination/survival rates were measured. This assay resulted in the detection of four possible hemizygous (individuals 2; 5; 6; 10) and two possible homozygous (individuals 11 and 14) T₁ plants (**Fig. 6a**).

Over expression of *EgMPK5* in transgenic tobacco and responses to stress conditions

Because the expression of *EgMPK5* was induced by salt stress in *E. grandis* (**Fig. 4**), we tested whether the gene could increase tolerance to NaCl in transgenic tobacco. Stress tolerance was assessed on the germination rate of seeds and the growth of tobacco seedlings under 100, 200 and 300 mM NaCl. The criterion used to define a plant as tolerant was the emergence of green cotyledon and continued growth of the seedling under stress conditions. Under normal growth condition (medium without addition of salt), no great difference in germination rate was detected among WT and three transgenic lines, which showed rates of at least 80% of germination. Two transgenic lines, though, had lower germination rates on regular medium. At a concentration of 100 mM NaCl, the transgenic line 6 showed higher germination rates than the WT, and other two transformed plants showed germination rates at least 10% lower than the WT plants (**Fig. 6b**). Higher concentrations of NaCl inhibited germination of all tested lines. Plants were also assayed for the capability of survival under drought stress for 4 days followed by 30 days of recovery in hydroponic culture system. In this assay it was observed the number of dead and wilted leaves and the ability of the plants for recovering a green appearance and sustaining growth. The transformed lines had more success at recovering from the stress and were able to maintain vigorous growth afterwards (**Fig. 7**).

Discussion

In this study, we cloned and characterized *EgMPK5*, a gene whose predicted protein belongs to the group A of plant MAPKs. The gene has a single copy in the genome and is located on the G chromosome of *E. grandis*. Works in the literature have described the existence of group A MAPKs genes as both single copy and duplicated in many plant species [11,12,14]. Even though we found two similar genomic sequences for *EgMPK5*, *in silico* and Southern blot analyses discarded the possibility of a duplicated gene. *EgMPK5* thus is a single copy gene as are its closely sequence-related MAPKs found in tobacco (*mpk4*) [31], rice (*mpk6*) [15] and *B. distachyon* (*mpk6*) [16].

Comparison of *EgMPK5* with four group A MAPK genes that have already had their intron/exon structure characterized allowed us to detect conserved size and

distribution of exons and introns. The introns of *EgMPK5* possess GT and AG terminal dinucleotides at their 5' and 3' ends, which characterize the major class of introns subject to splicing in plants [32]. The four internal exons of analyzed sequences are *in silico* translated to a 260 amino acid peptide that has motifs of the protein kinase domain of plants. The highly conserved length of internal exons of group A MAPK genes thus can be explained by a negative selection for alterations in sequences that code for the protein catalytic and regulatory sites. Intron phase was also analyzed for *EgMPK5*. Five of the six introns present in the gene are within phase 0 (with the intron inserted before the first base of a codon), but the first intron is within phase 1. This result shows that *EgMPK5* gene organization is similar to other MAPK genes such as those from *A. thaliana* and *P. trichocarpa*, each showing almost 80% of their introns within phase 0 [13].

Because the putative protein *EgMPK5* has a conserved TEY motif in its activation T-loop, it is defined as a member of the group A of plant MAPKs according to the classification proposed in literature [10]. The amino acid sequence also presents residues characteristic to the eleven subdomains (I–XI) of the eukaryotic protein kinase domain and a CD-domain for docking of regulatory proteins at N-terminal portion [33]. Studies on protein kinases have elucidated common structural features present in the domain. The conserved motifs found in *EgMPK5* includes regions important to correct function of MAPKs such as those used for anchoring and orientating the nucleotide triphosphate during the phosphate transfer, binding the peptide substrate, and forming a cleft between C- and N-terminal lobes that is recognized as the site of catalysis [34].

The phylogenetic tree was used to analyze MAPK proteins similar to *EgMPK5* aiming at the identification of possible paralogous. Because genes belonging to the same group might have similar functions, the homologous sequences detected may help the functional annotation of an uncharacterized MAPK such as *EgMPK5*. Among the sequences used for tree construction is MPK6 from *A. thaliana*, which has been already characterized in the response to stresses such as salt, cold, dehydration, touch and wounding. Tobacco MPK4 [35] and tomato MPK2 [36] have been implicated in responses to wounding. Several members of the phylogenetic tree also have been shown to be involved in the response to pathogens, such as MPK2 from *Glycine max* [37], MPK6 from *Arachis hypogaea* [38] MPK2 from *Petroselinum crispum* [39] and also involved in signaling of cellular division, such as MSK7 from *Medicago sativa* [40] and

MAPK D5 from *Pisum sativum* [41]. Thus we found MAPKs from other plants that share homology to EgMPK5 and have been reported to be associated in a variety of cellular processes and biotic and abiotic stress responses that might aid us to identify possible functions of EgMPK5.

It is thought that accumulation of gene transcripts induced by stress conditions might lead to accumulation of their products in order to participate in amplification of the stress signals [42]. Our RT-qPCR data showed that ABA was capable of increasing the expression of *EgMPK5* in leaves. It has been already demonstrated in many plant species that ABA treatment interferes in transcriptional regulation of genes, in protein accumulation, and in the activity of MAPK signaling cascades. Most of these genes are up-regulated by ABA [42], and thus MAPK cascades have been implicated in ABA-mediated responses, which includes antioxidant defense, guard cell signaling and seed germination as reviewed in [43].

NAA treatment decreased the level of *EgMPK5* expression in stems compared to the untreated control. In the literature there are reports of both positive and negative MAPK pathways in auxin signaling [44]. Interaction of auxins and MAPKs has been shown to influence root development, mitosis, gene transcription, cytoskeletal rearrangement and ion channel activity [44, 45].

NaCl 200 mM was the only treatment capable of inducing *EgMPK5* expression in all tested organs. In *A. thaliana*, *MPK3*, *MPK4* and *MPK6* are activated by hyperosmotic stress and salinity [46, 47]. *MPK6*, when activated, plays an important role in sodium detoxification by inducing the ion efflux off the cell and also in the transcriptional regulation of inducible genes [48, 49]. In roots of *A. thaliana*, it has been shown that 250 mM NaCl stimulated phosphorylation activity of a MAPK significantly [45].

Taken together, these data suggest that expression of *EgMPK5* is influenced by treatments with salt, ABA and auxin, and thus the gene is possibly involved in responses to abiotic stresses and hormone treatments.

Several cis-acting elements related to stress responses were found in the promoter region of *EgMPK5*. Among these, ABRE (ABA-responsive element) and TGA (auxin-responsive element) relate to our obtained RT-qPCR expression data.

ABREs are *cis*-acting elements that function in vegetative tissues of *A. thaliana* in response to dehydration and ABA treatment [50]. TGA1 are basic/leucine-zipper proteins thought to aid tissue-specific transcription in plants in response to hormone and chemical stress cues [51].

In MAPKs such as tomato [12] and *A. thaliana* [52] higher levels of MAPK transcripts were detected in reproductive organs such as stamen and petal compared to vegetative organs such as leaf and stem. As we did not test for these organs we cannot be sure if the activity of *EgMPK5* is relevant to the flower or pollen development. However, we found a GCN4_motif in the promoter region of *EgMPK5* that might be important for specific gene expression in endosperm.

The literature describes that in plants it has been detected the existence of paralogous gene pairs that differ in the level of transcriptional control. Example of such gene pairs are *MPK6–MPK3* in *A. thaliana*, where the first member of the gene pair is constitutive and the second is differentially expressed [52]. Because of the characteristic of plants to present several MAPK genes, we could expect *E. grandis* to encode such paralogous gene pairs as well. Future studies might aid us detect other *mapk* genes whose transcription profiles relate to those observed for *EgMPK5* expression.

For assessment of transgenic tobacco stress tolerance, we performed tests with seeds and adult plants. The transformed lines were first tested for their ability to germinate on regular media and three lines showed germination rates comparable to the WT plants. Under 100 mM NaCl, the super expression of *EgMPK5* enhanced the ability of one transgenic line to survive the salt stress compared to the WT. Similarly, tobacco plants over expressing cotton *GhMPK2* achieved better germination rates than non transgenic plants on media containing 100 mM NaCl [53]. Also, over expression of maize *ZmSIMK1* in *A. thaliana* has led to higher germination rates under NaCl 200 mM and adult plants in soil showed better adaptation to this treatment as well [54]

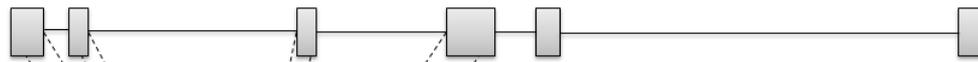
We observed that the two lines with lower germination rates in this assay are possible homozygous for the transgene, which suggests that such a state is not an advantage for plants exposed to salt stress during germination. The expression of a *MAPK* transgene might not always be beneficial for overcoming some stress treatments, as was recently demonstrated for a cotton group D MAPK [55]. In the present study, drought assay with adult plants showed that four transgenic lines were able to overcome

the stress treatment in contrast to the non-transformed lines, which did not survive. Thus we found lines of *EgMPK5*-overexpressing tobacco plants that seemed to have higher tolerance to salt stress during germination and better ability to survive drought treatment. Recently, tobacco plants over expressing a cotton *MPK3* gene showed enhanced drought tolerance compared to control plants when subjected to 14 days of water deprivation in soil [56]. The mentioned work also found higher expression of antioxidant enzymes in transformed plants, which may positively regulate drought tolerance through reactive oxygen species scavenging. For further understanding the stress tolerance mechanisms of our tobacco plants over expressing *EgMPK5*, it would be desirable to analyze gene transcription levels and the expression profile of stress-responsive marker genes that might be affected by it.

In conclusion, a novel group A MAPK gene named *EgMPK5* was cloned and characterized. Southern blot and *in silico* analyses suggested that a single copy of *EgMPK5* is present in the *E. grandis* genome and that the gene structure presents exons of conserved length and position. Analysis of promoter sequences and RT-qPCR data showed the involvement of *EgMPK5* in response to ABA, NaCl and auxin. The deduced aminoacid sequence coded by *EgMPK5* shows high homology to plant group A MAPKs involved in biotic, abiotic and hormonal responses. Tobacco plants over expressing *EgMPK5* under the control of the CaMV 35S promoter presented increased tolerance to salt and drought treatments, suggesting a possible role of *EgMPK5* in the regulation of abiotic stresses in plants.

Figure 1

a) *EgMPK5*, scaffold 7

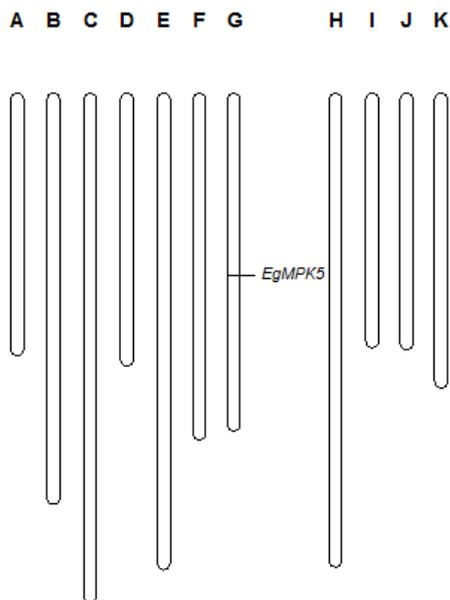


b) *EgMPK5*, cloned cDNA



500 bp

c)



d)

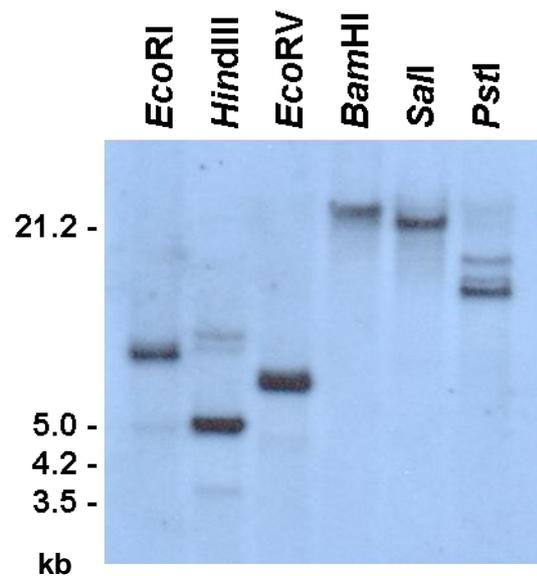


Figure 1: Analysis of gene structure and genomic location of *EgMPK5*. **A)** Organization of *EgMPK5* in 6 exons (boxes) and 5 introns (lines). Intron regions belong to the GT-AG type, a major class of introns in plants. The four internal exons code a 260 amino acid polypeptide that presents motifs of the protein kinase domain of plants. **B)** Structure of our cloned cDNA based on sequencing and comparison with genomic data. The sequence still has two introns, yet to be removed to form a mature transcript. **C)** *EgMPK5* mapped to the G chromosome of *E. grandis* according to main genome assembly. It reflects the best match among our performed sequence alignments and suggests a single copy of the gene **D)** Southern blot experiment showing strong hybridization signals in all lanes using a 1.5 kb *EgMPK5* cDNA fragment as a probe (see figure 5a for details).

Figure 2

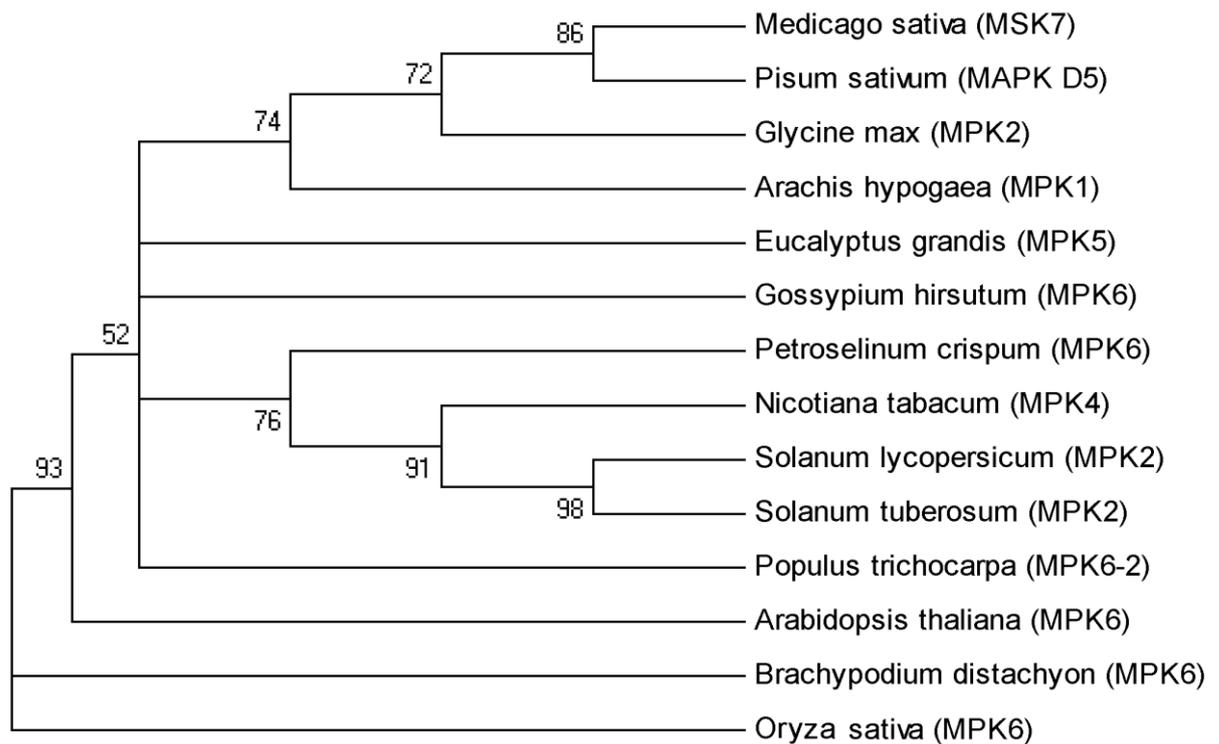
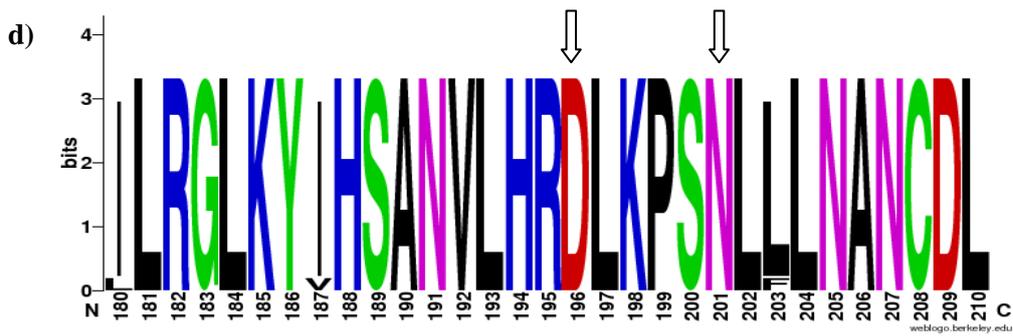
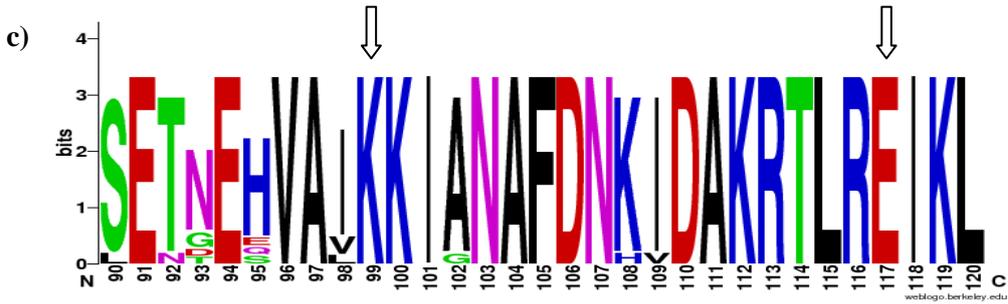
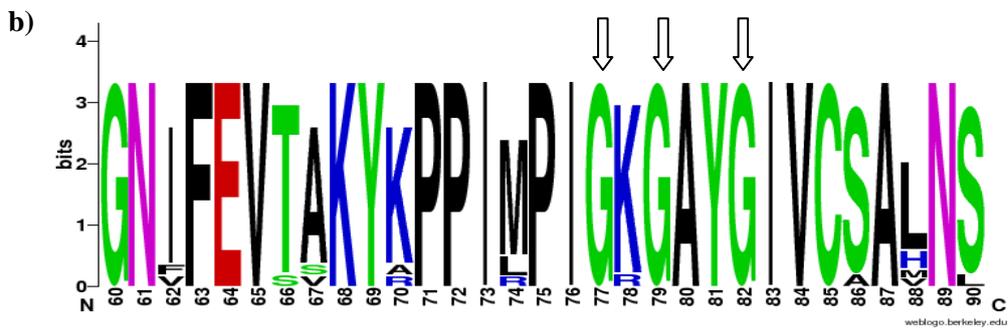
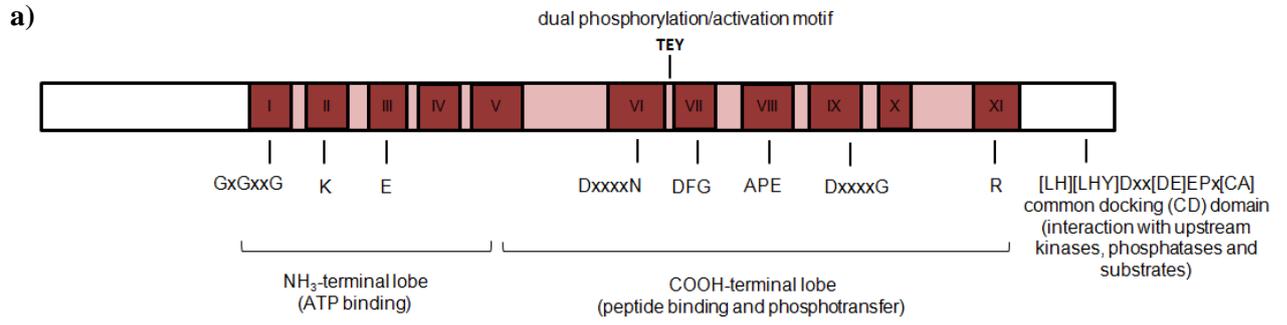


Figure 2: Phylogenetic tree built with the deduced product of the gene EgMPK5 and thirteen plant MAPKs sharing at least 80% similarity. The selected sequences are: *Glycine max* MPK2 (NP_001235426), *Medicago sativa* MSK7 (Q07176), *Pisum sativum* MAPK D5 (Q06060), *Arachis hypogaea* MAPK1 (AAZ23128.1); *Gossypium hirsutum* MPK6 (ADT65137.1), *Solanum lycopersicum* MPK2 (NP_001234355.1), *Solanum tuberosum* MPK2 (BAB93530.1), *Petroselinum crispum* MPK6 (AAN65179.1), *Oryza sativa* MPK6 (ACD76439.1), *Populus trichocarpa* MAPK6-2 (XP_002310398.2), *Brachypodium distachyon* MPK6 (NP_001266884.1), *Nicotiana tabacum* MPK4 (Q40532.1), *Arabidopsis thaliana* MPK6 (NP_181907.1). A multiple sequence alignment was performed with the MUSCLE algorithm employing the Maximum-Likelihood method on MEGA5 with a bootstrap of 500 replicates. Several members of the tree have been implicated in responses to stimuli such as salt, cold, dehydration, touch, wounding and pathogens. Such sequences that share sequence homology may help the functional annotation EgMPK5.

Figure 3



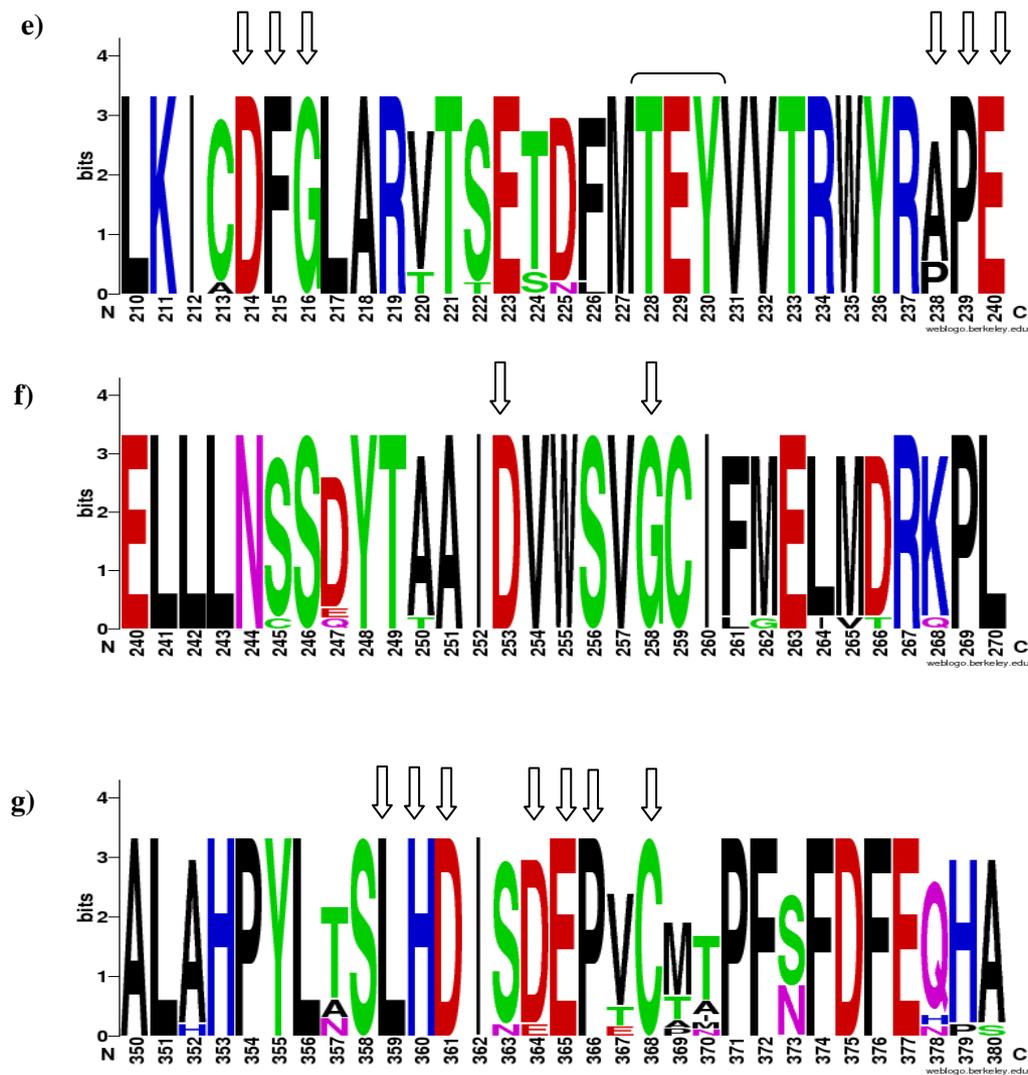


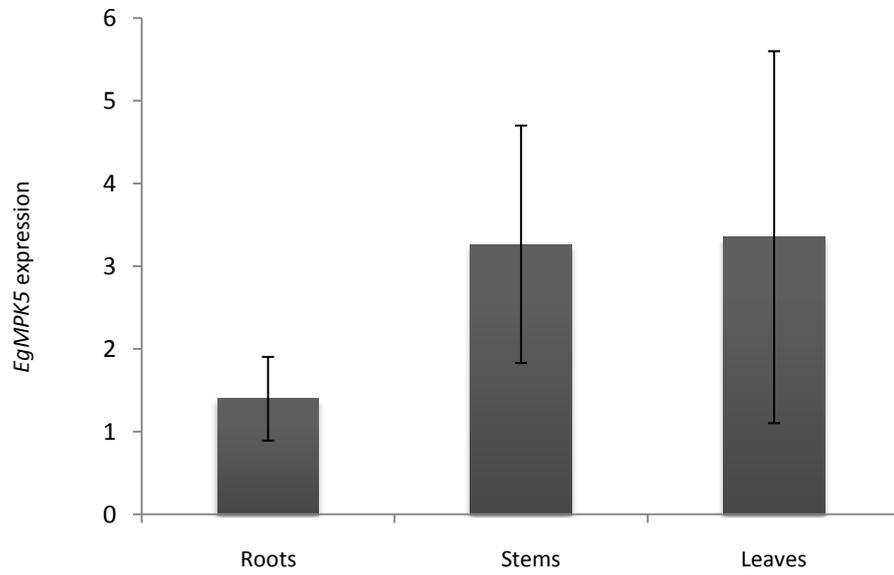
Figure 3: Conserved residues found in the protein kinase domain after alignment of full-length sequences of fourteen plant MAPKs. **A)** Schematic representation of domains and sub domains present in group A of plant MAPKs, adapted from eukaryotic protein kinase catalytic domain structure proposed in literature [32]. **B)** consensus GxGxxG of sub domain I (white arrows). **C)** consensus K (sub domain II) and consensus E (sub domain III). **D)** consensus DxxxxN (sub domain VI). **E)** consensus DFG (sub domain VII) and consensus APE (sub domain VIII) and dual phosphorylation/activation motif TEY (between sub domains VI and VII). **F)** consensus DxxxxG (sub domain IX). **G)** COOH-terminal *docking* site (CD) domain [LH][LHY]Dxx[DE]EPx[CA]

Table 1: *cis*-acting elements and their motifs found in the region 2-kb upstream of *EgMPK5* start codon using the PlantCare database.

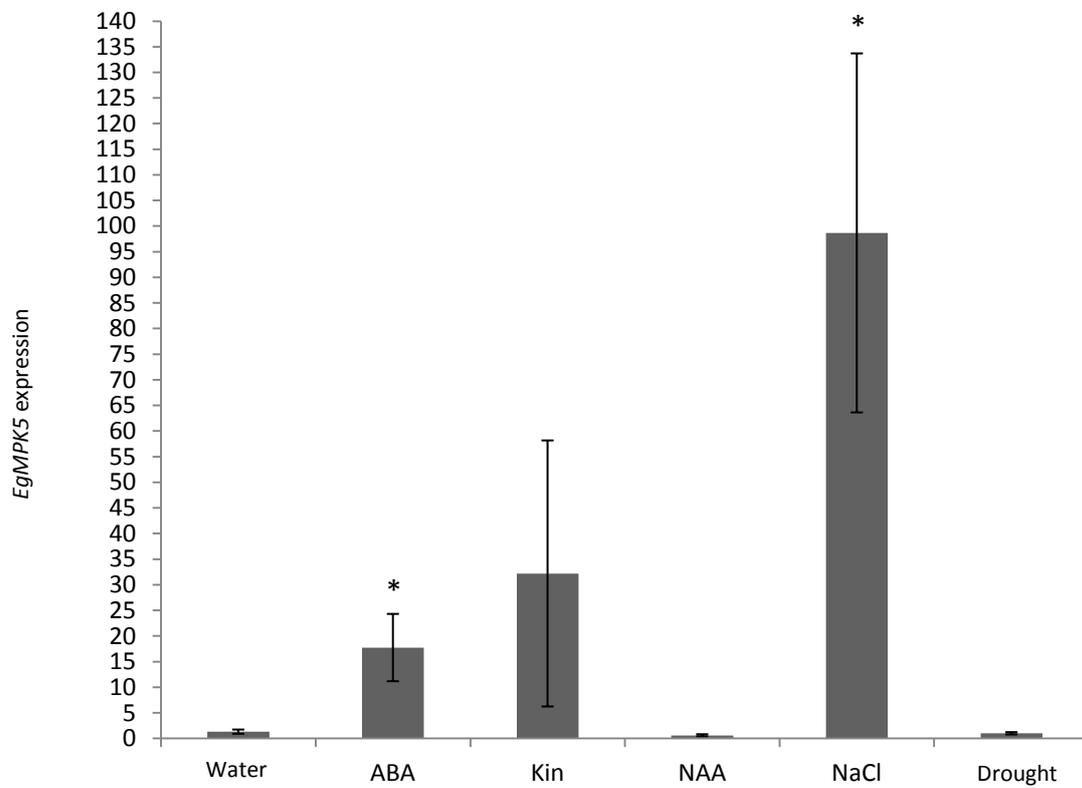
Process		Element	Motif
cis-elements involved in plant growth and development	cis-acting elements involved in light responsiveness	Box 4	ATTAAT
		ACE	AAAACGTTTA
		ATCT-motif	AATCTAATCC
		CATT-motif	GCATTC
		G-Box	CACGTA
		G-box	TACGTG
		GA-motif	AAGGAAGA
		GT1-motif	GGTTAA
		MRE	AACCTAA
		TCT-motif	TCTTAC
cis-elements involved in plant growth and development	cis-regulatory element involved in endosperm expression	GCN4_motif	CAAGCCA
	MYB binding site involved in flavonoid biosynthetic genes regulation	MBSII	AAAAGTTAGTTA
cis-elements related to plants hormones and stress responses	cis-acting element involved in the abscisic acid responsiveness	ABRE	TACGTG
	cis-acting regulatory element involved in the MeJA-responsiveness	CGTCA-motif	CGTCA
		TGACG-motif	TGACG
	cis-acting element involved in salicylic acid responsiveness	TCA-element	GAGAAGAATA
	auxin-responsive element	TGA-element	AACGAC
	ethylene-responsive element	ERE	ATTTCAAA
	cis-acting element involved in heat stress responsiveness	HSE	AAAAAATTC
	cis-acting element involved in low-temperature responsiveness	LTR	CCGAAA
	gibberellin-responsive element	GARE-motif	TCTGTTG
	fungal elicitor responsive element	Box-W1	TTGACC
common cis-elements related to eukaryotic promoters	cis-acting element conferring high transcription levels	5UTR Py-rich stretch	TTTCTTCTCT
	common cis-acting element in promoter and enhancer regions	CAAT-box	CCAAT
	core promoter element around -30 of transcription start	TATA-box	ATATAT

Figure 4

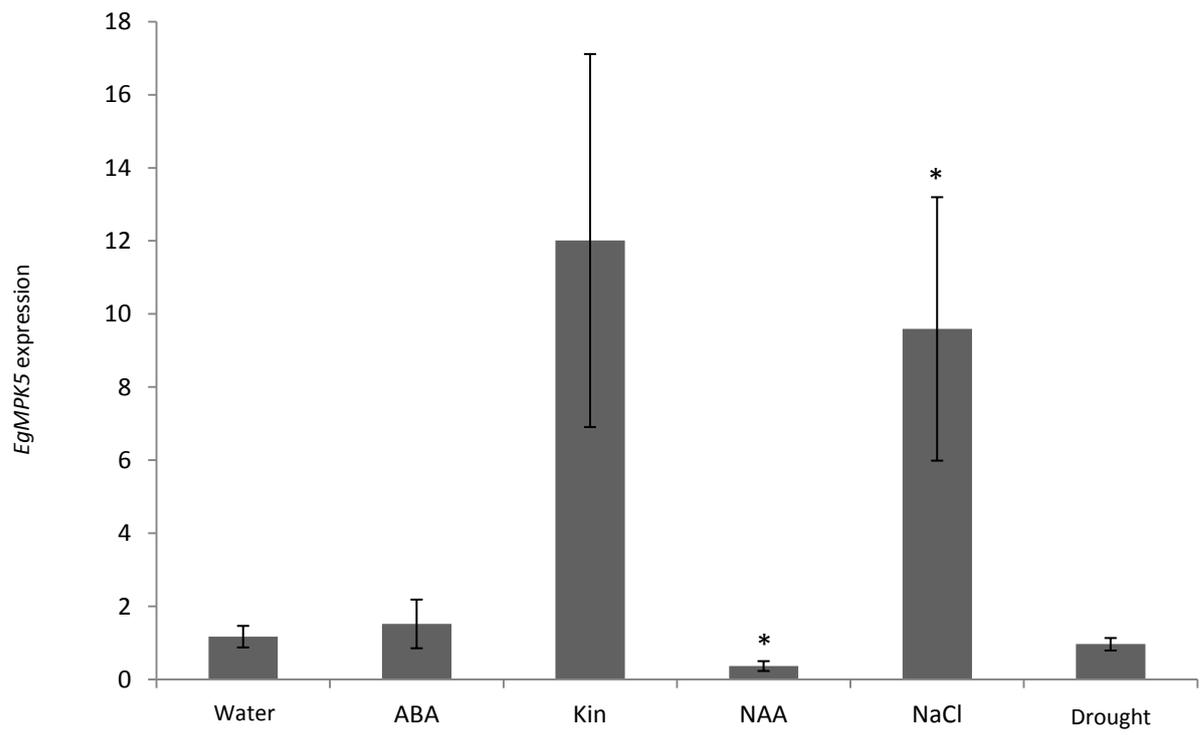
a)



b)



c)



d)

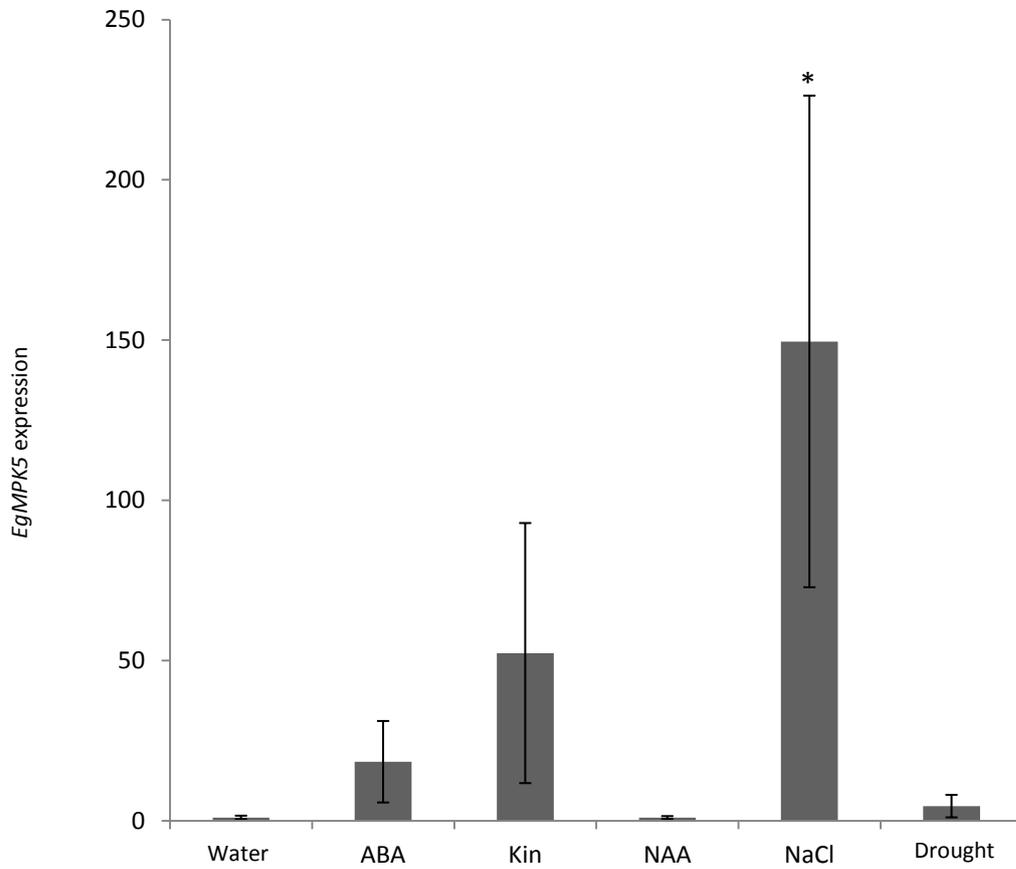
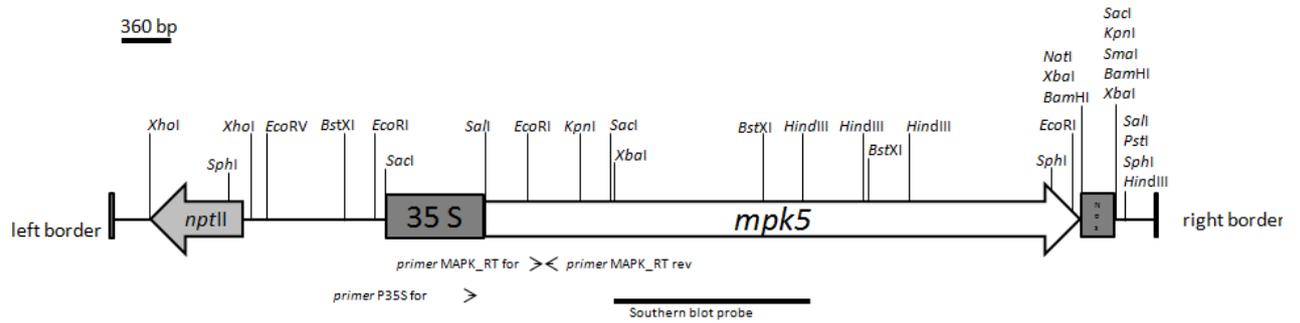


Figure 4: Expression profiles of *EgMPK5* in three organs of *E. grandis* under different treatments. Error bars represent standard errors of biological triplicates **A)** Spatial expression of *EgMPK5* in roots, stems and leaves of plants treated with water. The graphic shows results relative to expression data of roots. **B)** Differential expression of *EgMPK5* in leaves of plants exposed to several stimuli. The graphic shows results relative to expression data of water treatment. **C)** Differential expression in stems exposed to several stimuli. The graphic shows results relative to expression data of water treatment. **D)** Differential expression in roots exposed to several stimuli. The graphic shows results relative to expression data of water treatment. Asterisks mark statistically significant differences ($P < 0.05$) between a treatment condition and its untreated control (water).

Figure 5

a)



b)

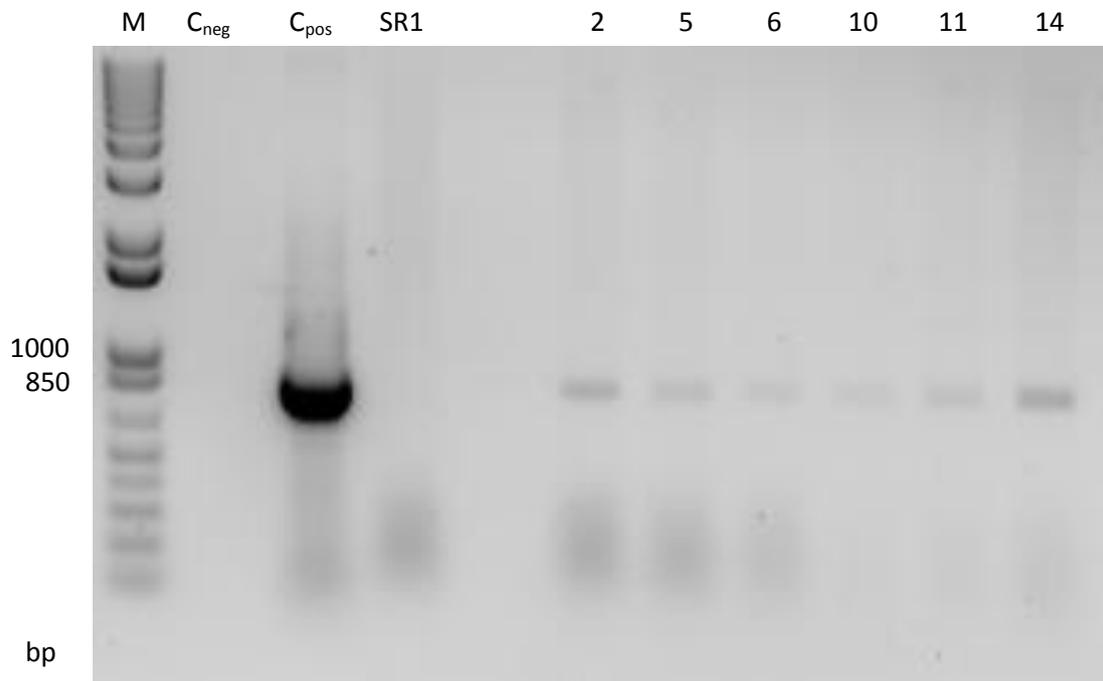


Figure 5: **A)** A map of the construction pCAMBIA2300-*mpk5* (which contains the cloned *EgMPK5* cDNA) is presented along with annealing regions of primers, restriction sites and region used for Southern blot probe. **B)** PCR for detection of transgenic tobacco plants. A band of approximately 900 bp was detected in transformed lines and in the positive control lanes. M-molecular size marker 1 kb plus DNA ladder (Invitrogen), C_{neg} – negative control, reaction mix without DNA; C_{pos} – positive control using pCAMBIA2300-*mpk5* as DNA template, SR1 – reaction with wild-type, non transgenic tobacco line DNA, and 2,5,6,10,11,14 – reaction with transformed plants DNA.

Figure 6

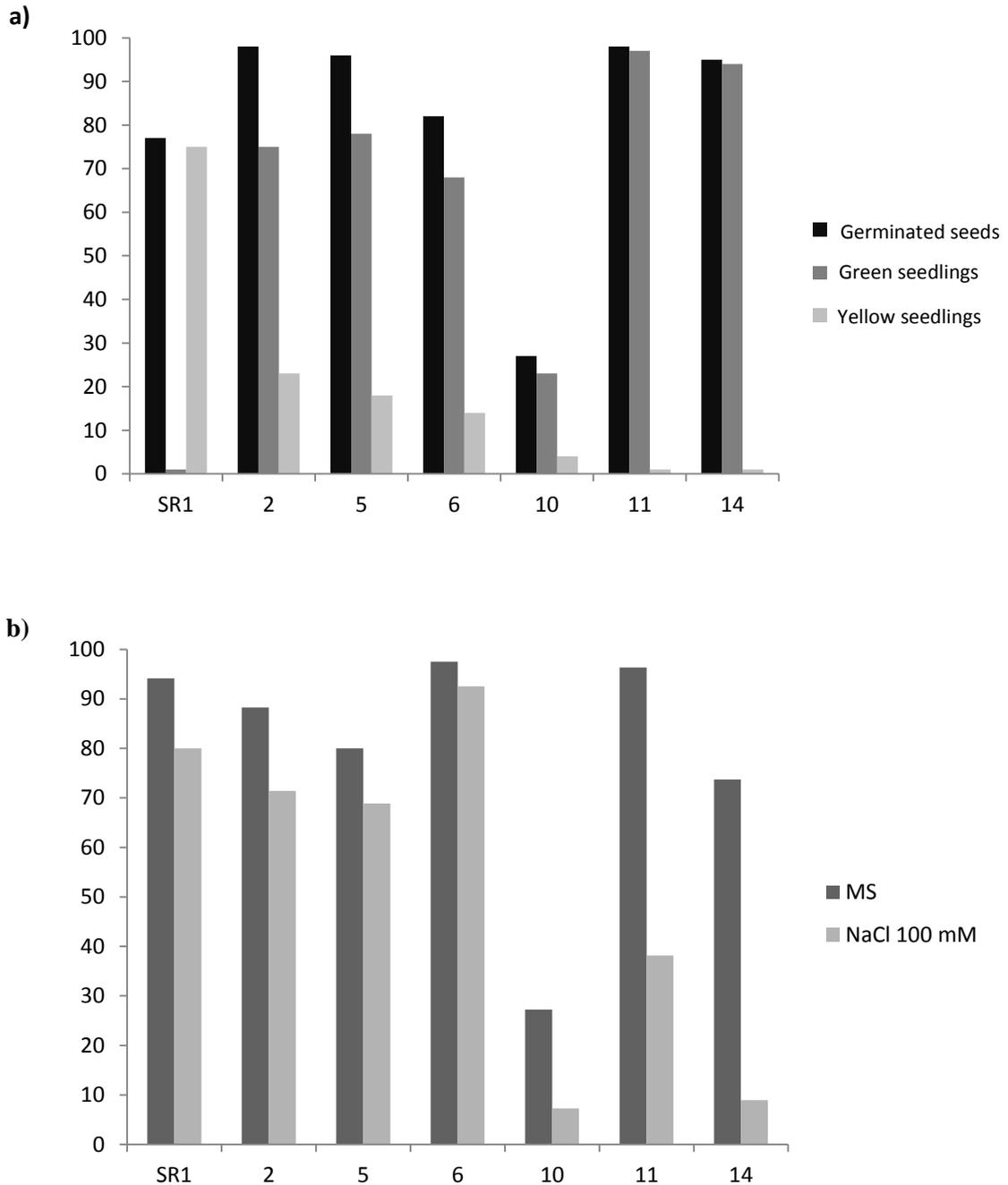


Figure 6: Seeds of tobacco T₁ generation were germinated on different media. **A)** Kanamycin selection of transgenic plants. Among the seeds that germinated, seedlings were evaluated according to their ability to survive the antibiotic (green seedlings) or incapacity of maintaining growth (yellow seedlings). The survival rate obtained on this experiment also served to assign transgene hemi- and homozygous genotypes of plants. **B)** Germination of seeds was compared between regular growth medium and with addition of NaCl 100 mM. Concentrations of 200 and 300 mM inhibited germination of all lines.

Figure 7

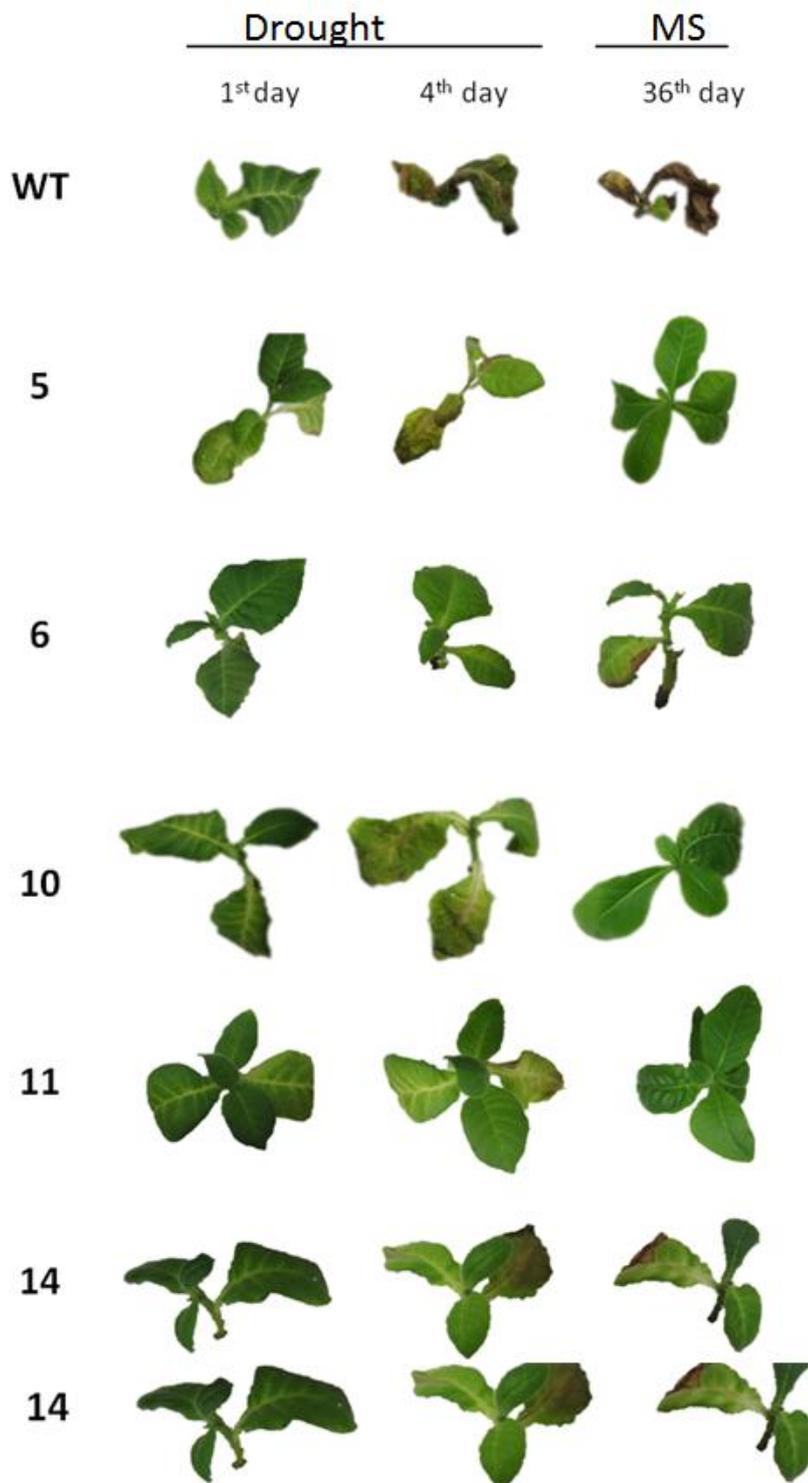


Figure 7: Appearance of non-transformed (WT) and transformed plants during drought assay. All the plants wilted and lost leaves due to desiccation up to four days of the experiment. Transformed lines showed higher capability of recovery after 32 days in water.

References

1. Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K (2006) Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Curr Opin Plant Biol* 9:436–442
2. Rodriguez M, Petersen M, Mundy J (2010) Mitogen-Activated Protein Kinase Signaling in Plants. *Annu Rev Plant Biol* 61:621–49
3. Jonak C, Ökrész L, Bögre L, Hirt H (2002) Complexity, cross talk and integration of plant MAP kinase signaling. *Curr Opin Plant Biol* 5:415–424
4. Nakagami H, Pitzschke A, Hirt H (2005) Emerging MAP kinase pathways in plant stress signaling. *Trends Plant Sci* 10:339-346
5. Sinha A, Jaggi M, Raghuram B, Tuteja N (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signal Behav* 6:196-203
6. Müller J, Beck M, Mettbach U, Komis G, Hause G, Menzel D (2010) *Arabidopsis* MPK6 is involved in cell division plane control during early root development, and localizes to the pre-prophase band, phragmoplast, trans-Golgi network and plasma membrane. *Plant J* 61:234–248
7. Sasabe M, Machida Y (2012) Regulation of organization and function of microtubules by the mitogen-activated protein kinase cascade during plant cytokinesis. *Cytoskeleton* 69:913–918
8. Smékalová V, Doskočilová A, Komis G, Samaj J (2013) Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. *Biotechnol Adv* 32:2-11
9. Zhang S (2008) Mitogen-Activated Protein Kinase Cascades in Plant Intracellular Signaling. In: Yang Z (ed) *Intracellular Signaling in Plants*. Annual Plant Reviews 33:100-136. Blackwell Publishing.
10. MAPK Group (Ichimura K *et al.*) (2002) Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends Plant Sci* 7:301-308
11. Hamel LP, Nicole MC, Sritubtim S, Morency MJ, Ellis M, *et al.* (2006) Ancient signals: comparative genomics of plant MAPK and MAPKK gene families. *Trends Plant Sci* 11:192–198
12. Kong F, Wang J, Cheng L, Liu S, Wu J, Peng Z, Lu G (2012) Genome-wide analysis of the mitogen-activated protein kinase gene family in *Solanum lycopersicum*. *Gene* 499:108–120

13. Nicole M, Hamel L, Morency M, Beaudoin N, Ellis B, Séguin A (2006) MAP-ping genomic organization and organ-specific expression profiles of poplar MAP kinases and MAP kinase kinases. *BMC Genomics* 7:223-245
14. Zhang S, Xu R, Luo X, Jiang Z, Shu H (2013) Genome-wide identification and expression analysis of MAPK and MAPKK gene family in *Malus domestica*. *Gene* 531:377–387
15. Liu Q, Xue Q (2007) Computational identification and phylogenetic analysis of the MAPK gene family in *Oryza sativa*. *Plant Physiol Biochem* 45:6-14
16. Chen L, Hu W, Tan W, Wang M, Ma Z, Zhou S, Deng X, Zhang Y, Huang C, Yang G, He G (2012) Genome-wide identification and analysis of MAPK and MAPKK gene families in *Brachypodium distachyon*. *PLoS ONE* 7: e46744
17. *Eucalyptus grandis* Genome Project 2010, <http://www.phytozome.net/eucalyptus>
18. Ugalde L, Pérez O (2001) Mean annual volume increment of selected industrial forest plantation species. In: Mead D (ed) *Forest Plantation Thematic Papers, Working Paper 1*. Forest Resources Development Service, Forest Resources Division. FAO, Rome
19. Valdés A, Irar S, Majadad J, Rodríguez A, Fernández B, Pagès M. (2013) Drought tolerance acquisition in *Eucalyptus globulus* (Labill.): A research on plant morphology, physiology and proteomics. *J Prot* 79: 263-276
20. Murashige T, Skoog F (1963) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* 15:437-498
21. Doyle J, Doyle J (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochem Bull* 19:11-15
22. Oliveira L, Breton M, Bastolla F, Camargo S, Margis R, Frazzon J, Pasquali G (2012) Reference genes for the normalization of gene expression in *Eucalyptus* species. *Plant Cell Physiol* 53:405–422
23. Livak K, Schmittgen T (2001) Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402–408
24. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. *Nucleic Acids Res* 41:36-42
25. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Mol Biol Evol* 28:2731-2739
26. Crooks G, Hon G, Chandonia J, Brenner S (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188-119

27. Voorrips, R (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77-78
28. Lescot M, Dhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouz P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res* 30:325-327
29. De Block M, Herrera-Estrella L, Van Montagu M, Schell J, Zambryski P (1984) Expression of foreign genes in regenerated plants and their progeny. *EMBO J* 3:1681-1689
30. Horsch RB, Fraley RT, Rogers SG, Sanders PR, Lloyd A, Hoffmann N (1984) Inheritance of functional genes in plants. *Science* 223:496-498
31. Zhang X, Cheng T, Wang G, Yan Y, Xia Q (2012) Cloning and evolutionary analysis of tobacco MAPK gene family. *Mol Biol Rep* 40:1407-1415
32. Szczesniak M, Kabza M, Pokrzywa R, Gudys A, Makalowska I (2013) ERISdb: A database of plant splice sites and splicing signals. *Plant Cell Physiol* 54:1-8
33. Hanks S (2003) Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biol* 4:111-118
34. Hanks S, Hunter T (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 9:576-596
35. Gomi K, Ogawa D, Katou S, Kamada H, Nakajima N, Saji H (2005) A mitogen-activated protein kinase NtMPK4 activated by SIPKK is required for jasmonic acid signaling and involved in ozone tolerance via stomatal movement in tobacco. *Plant Cell Physiol* 46:1902-1914
36. Xing T, Malik K, Martin T, Miki B (2001) Activation of tomato PR and wound-related genes by a mutagenized tomato MAP kinase kinase through divergent pathways. *Plant Mol Biol* 46:109-120
37. Daxberger A, Nemark A, Mithöfer A, Fliegmann J, Ligterink W, Hirt H, Ebel J (2007) Activation of members of a MAPK module in beta-glucan elicitor-mediated non-host resistance of soybean. *Planta* 225:1559-1571
38. Kumar K, Kirti P (2010) A mitogen-activated protein kinase, AhMPK6 from peanut localizes to the nucleus and also induces defense responses upon transient expression in tobacco. *Plant Physiol Biochem* 48:481-486
39. Kroj T, Rudd J, Nürnberger T, Gäbler Y, Lee J, Scheel D (2003) Mitogen-activated protein kinases play an essential role in oxidative burst-independent expression of pathogenesis-related genes in parsley. *J Biol Chem* 278:2256-2264

40. Duerr B, Gawienowski M, Ropp T, Jacobs T (1993) MsERK1: a mitogen-activated protein kinase from a flowering plant. *Plant Cell* 5:87-96
41. Stafstrom J, Altschuler M, Anderson D (1993) Molecular cloning and expression of a MAP kinase homologue from pea. *Plant Mol Biol* 22:83-90
42. Liu Y (2012) Roles of mitogen-activated protein kinase cascades in ABA signaling. *Plant Cell Rep* 31:1–12
43. Danquah A, Zelicourt A, Colcombet J, Hirt H (2014) The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnol Adv* 32:40–52
44. Mishra N, Tuteja R, Tuteja N (2006) Signaling through MAP kinase networks in plants. *Arch Biochem Biophys* 452:55–66
45. Mockaitis K, Howell S (2000) Auxin induces mitogenic activated protein kinase (MAPK) activation in roots of *Arabidopsis* seedlings. *Plant J* 24:785-796
46. Droillard M, Boudsocq M, Barbier-Brygoo H, Lauriere C (2004) Involvement of MPK4 in osmotic stress response pathways in cell suspensions and plantlets of *Arabidopsis thaliana*: activation by hypoosmolarity and negative role in hyperosmolarity tolerance. *FEBS Lett* 574:42–48
47. Kim J, Woo D, Kim S, Lee S, Park H, Seok H (2012) *Arabidopsis* MKKK20 is involved in osmotic stress response via regulation of MPK6 activity. *Plant Cell Rep* 31:217–224
48. Yu L, Nie J, Cao C, Jin Y, Yan M, Wang F (2010) Phosphatidic acid mediates salt stress response by regulation of MPK6 in *Arabidopsis thaliana*. *New Phytol* 188:762–773
49. Teige M, Scheikl E, Eulgem T (2004) The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Mol Cell* 15:141–152
50. Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of *Arabidopsis*. *Plant Mol Biol* 60:51–68
51. Pascuzzi P, Hamilton D, Bodily K, Arias J (1998) Auxin-induced Stress Potentiates trans-activation by a Conserved Plant Basic/Leucine-zipper Factor. *J Biol Chem* 273:26631-26637
52. Menges M, Dóczi R, Ökrész L, Morandini P, Mizzi L, Soloviev M, Murray J, Bögre J (2008) Comprehensive gene expression atlas for the *Arabidopsis* MAP kinase signalling pathways. *New Phytol* 179: 643–662

53. Zhang L, Xi D, Li S, Gao Z, Zhao S, Shi J, Wu C, Guo X (2011) A cotton group C MAP kinase gene, GhMPK2, positively regulates salt and drought tolerance in tobacco. *Plant Mol Biol* 77:17–31
54. Gu L, Liu Y, Zong X, Liu L, Li D, Li D (2010) Overexpression of maize mitogen-activated protein kinase gene, *ZmSIMK1* in *Arabidopsis* increases tolerance to salt stress. *Mol Biol Rep* 37:4067–4073
55. Wang J, Ding H, Zhang A, Ma F, Cao J, Jiang M (2010) A novel mitogen-activated protein kinase gene in maize (*Zea mays*), *ZmMPK3*, is involved in response to diverse environmental cues. *J Integr Plant Biol* 52: 442–452
56. Long L, Gao W, Xu L, Liu M, Luo X, He X, Yang X, Zhang X, Zhu L (2014) *GbMPK3*, a mitogen-activated protein kinase from cotton, enhances drought and oxidative stress tolerance in tobacco. *Plant Cell Tiss Organ Cult* 116:153–162

CURRICULUM VITÆ

Krás Borges, J.

DADOS PESSOAIS

Nome: Juliana Dametto Krás Borges

E-mail: julianadkb@gmail.com

Local e data de nascimento: Porto Alegre, RS, Brasil, 14/03/1987

FORMAÇÃO ACADÊMICA

Graduação em Farmácia Industrial: UFRGS. Início em março/2010 e conclusão em dezembro/2011.

Intercâmbio de graduação: The Ohio State University, Columbus, Estados Unidos. Início em setembro/2010 e conclusão em dezembro/2010.

Graduação em Farmácia: UFRGS. Início em agosto/2005 e conclusão em dezembro/2009.

ESTÁGIOS DE INICIAÇÃO CIENTÍFICA E APREFEIÇOAMENTO CIENTÍFICO

Centro de Biotecnologia – UFRGS: Maio a Julho de 2010 e Fevereiro a Dezembro de 2011; bolsista CNPq, Orientador Dr. Giancarlo Pasquali; estudo de genes envolvidos na biossíntese do alcaloide ricinina

Department of Horticulture and Crop Science – The Ohio State University (Columbus, EUA): Setembro de 2010 a Dezembro de 2010; bolsista CAPES, Orientador Dr. David Mackey; estudo de proteínas patogênicas de *Pseudomonas syringae*

Centro de Biotecnologia – UFRGS: Julho de 2008 a Abril de 2010; bolsista PIBIC UFRGS, Orientador Dr. Giancarlo Pasquali; cultura *in vitro* e transformação de *Ricinus communis*

Inst de Ciências Básicas da Saúde – UFRGS: Agosto de 2007 a Junho de 2008; bolsista FAPERGS, Orientador Dr. Luis Valmor Cruz Portela; estudo dos efeitos neuroquímicos da dexametasona em camundongos

RESUMOS E TRABALHOS APRESENTADOS EM CONGRESSOS

Krás Borges, J; Kirch, RP; Körbes, AP; Maraschin, FS; Margis, R; Bodanese-Zanettini, MH; Pasquali, G. The mitogen-activated protein kinase 5 (*mapk5*) from *Eucalyptus grandis*. IV Simpósio Brasileiro de Genética Molecular de Plantas, 8 a 12 de abril de 2013.