



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
DEPARTAMENTO DE GENÉTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E  
BIOLOGIA MOLECULAR

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DOS GENES DA  
FAMÍLIA DE FATORES TRANSCRICIONAIS E2F EM ARROZ  
(*ORYZA SATIVA* L.)**

**RAFAEL RAUBER**

**PORTO ALEGRE – RS**

**Março, 2012**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
DEPARTAMENTO DE GENÉTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DOS GENES DA FAMÍLIA DE FATORES  
TRANSCRICIONAIS E2F EM ARROZ (*ORYZA SATIVA* L.)**

**Rafael Rauber**

**Dissertação submetida ao Programa de Pós-graduação em Genética e Biologia Molecular como um dos requisitos para a obtenção do título de mestre em genética e biologia molecular.**

**Orientadora: Dra. Márcia M. A. N. Pinheiro Margis**

**PORTO ALEGRE – RS**

**Março, 2012**

## **INSTITUIÇÃO E FONTES FINANCIADORAS**

Este trabalho foi desenvolvido no Núcleo de Genômica Funcional de Plantas situado no Laboratório de Genética e Biologia Molecular Vegetal do Departamento de Genética do Instituto de Biociências da UFRGS.

As agências de financiamento do projeto de pesquisa em que se enquadra o presente trabalho de mestrado foram o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## **AGRADECIMENTOS**

Agradeço em primeiro lugar ao CNPq e CAPES por fornecerem o suporte financeiro ao projeto e a minha bolsa de mestrado.

A Professora Doutora Márcia Márgis-Pinheiro por me orientar neste mestrado, por ser firme quando precisou e por estar sempre disponível quando precisei (assim como para qualquer aluno) além de ajudar com o meu crescimento profissional e muitas vezes pessoal.

A minha grande amiga Andréia Caverzan por ajudar a tornar este trabalho realidade desde o começo com seu conhecimento e também, é claro, as nossas horas de conversa pipetando.

A doutora Andréia Turchetto por toda a ajuda e conhecimento a respeito de filogenias além da calma e paciência para me ensinar tantas coisas.

Aos colegas do grupo de genômica funcional de plantas, sempre dispostos a ajudar e a todos os meus colegas de laboratório por tornarem o ambiente de trabalho muito divertido e produtivo.

Aos doutores Antonio Costa de Oliveira, Júlio Cesar de Lima e Loreta Brandão de Freitas por aceitarem ser a banca avaliadora deste trabalho.

Ao meu pai Francisco, por estar sempre presente, apesar da distância, dando-me conselhos profissionais e emocionais, além de ser um dos meus exemplos de vida. Junto com ele, a minha mãe Teresa, por ser outro exemplo de vida e luta e por também estar presente quase todos os dias de distância por telefone e nunca me deixar desistir e conseguir me fazer seguir em frente, lutando.

Ao meu irmão Renan, por estar sempre por perto quando precisei, é muito bom ter alguém com quem sei que posso contar sempre, independente do que o mundo diga, me apoiando, me fazendo rir e me enchendo o saco, obviamente.

A minha namorada Cristina, por estar ao meu lado me apoiando e lutando por um objetivo maior, e por ter uma participação especial na reta final deste trabalho com horas de estudo juntos ou apenas me apoiando e dando força com um torpedão, que pode conter poucas palavras mas sempre me animaram.

Sem o apoio de todos vocês alguma parte deste trabalho não existiria ou não teria sido completada. As pessoas dizem: “eu venci, eu fiz este trabalho” eu digo: “nós vencemos, nós fizemos este trabalho se tornar realidade”.

A todos o meu MUITO OBRIGADO

## RESUMO

Devido ao constante ataque à camada de ozônio, as plantas, nos próximos anos, terão que lidar cada vez mais com uma maior quantidade de luz UV atingindo a atmosfera terrestre. O arroz é considerado uma planta modelo para as monocotiledôneas por possuir um genoma relativamente pequeno (aproximadamente 390 Mb) e devido a sua grande sintonia com outros cereais. Recentemente, as proteínas E2F foram descritas, especialmente em vertebrados, como fatores transcricionais que podem estar envolvidos com a resposta ao dano de DNA. Tendo como objetivo principal estudar como a família de fatores transcricionais E2F responde ao dano genotóxico em arroz, experimentos com luz UV-B foram conduzidos para avaliar o envolvimento destes com as respostas ao dano de DNA. Além disso, foram realizadas análises filogenéticas dessa família gênica com representantes de plantas. Em arroz, estudos prévios de nosso grupo identificaram seis genes pertencentes à família E2F, sendo quatro E2F típicos e dois E2F atípicos. Em resposta à luz UV, dois dos quatro genes que codificam E2F típicos e os dois genes de E2F atípicos, tiveram a expressão relativa aumentada após o estresse em relação a plantas mantidas em condição controle. Nesse experimento, além dos genes E2F e dos genes de reparo, também foram avaliados os níveis de transcritos de genes sabidamente envolvidos com ciclo celular e apoptose, permitindo-se verificar se o aumento de expressão dos genes E2F em resposta ao estresse está efetivamente relacionado com a resposta a danos no DNA e não a estes dois outros processos celulares. As análises filogenéticas revelaram que um desses dois E2F típicos, que apresentou sua expressão diferencial nas condições testadas, provavelmente atue como um repressor de transcrição. Esta afirmação está baseada em estudos comparativos com sequências da planta modelo para as dicotiledôneas, *Arabidopsis thaliana*. O conjunto de resultados obtidos sugere que alguns membros da família E2F de arroz estão envolvidos especificamente com as repostas ao dano ao DNA, além de permitir a proposição de um possível mecanismo de ação.

## ABSTRACT

Due to the constant attack to the ozone layer, the plants, in the coming years, will have to deal with an increasing amount of UV light reaching the earth's atmosphere. Rice is considered a model plant for monocots due to its relatively small genome (approximately 390 Mb) and because its high synteny with other cereals. E2F proteins have recently been reported, especially in vertebrates, as transcriptional factors involved in the response to DNA damage. To study how the family of transcription factors E2F responds to genotoxic damage in rice, experiments with UV-B were conducted to evaluate their involvement with the responses to DNA damage. In addition, phylogenetic analyzes of this gene family in plants were performed. In rice, previous studies of our group identified six genes belonging to the E2F family, four typical and two atypical E2F. In response to UV light, two of four genes encoding typical E2F and the two atypical E2F genes increased their relative expression after stress compared to the plants grown in the control condition. In this experiment, in addition to the E2F gene and repair genes, we have also analyzed the transcriptional level of genes known to be involved with cell cycle and apoptosis, in order to verify if the increases in E2F gene expression, in response to stress, are actually related to the response to DNA damage, rather than to other two cellular processes. Phylogenetic analysis revealed that one of these two typical E2F, which showed differential expression under the conditions tested, probably acts as a transcriptional repressor. This analysis is based on comparisons with sequences from the model plant for dicots, *Arabidopsis thaliana*. These set of results suggest that some members of the E2F family of rice are involved specifically with responses to DNA damage and allowed the proposition of a possible mechanism of action.

## SUMÁRIO

<b>LISTA DE ABREVIATURAS.....</b>	<b>9</b>
<b>LISTA DE FIGURAS E TABELAS.....</b>	<b>11</b>
<b>1 INTRODUÇÃO .....</b>	<b>14</b>
1.1 O ARROZ.....	14
1.2 O ESTRESSE POR LUZ ULTRA VIOLETA .....	15
1.3 OS FATORES DE TRANSCRIÇÃO E2F.....	17
<b>2 OBJETIVOS.....</b>	<b>21</b>
2.1 OBJETIVOS GERAIS.....	21
2.2 OBJETIVOS ESPECÍFICOS .....	21
<b>3 MATERIAL E MÉTODOS .....</b>	<b>22</b>
3.1 MATERIAL VEGETAL.....	22
3.2 ESTRESSE POR LUZ UV-B .....	22
3.3 EXTRAÇÃO DE RNA TOTAL.....	23
3.4 SÍNTESE DE cDNA A PARTIR DE RNA TOTAL .....	23
3.5 DETERMINAÇÃO DE OLIGONUCLEOTÍDEOS INICIADORES ESPECÍFICOS (PRIMERS) PARA PCR .....	24
3.6. ANÁLISE POR PCR QUANTITATIVA EM TEMPO REAL (RT-qPCR).....	24
3.7 ANÁLISE DA EXPRESSÃO DOS GENES E2F DE ARROZ E DE GENES ENVOLVIDOS NOS MECANISMOS DE REPARO DO DNA. ....	25
3.8 BUSCA DE SEQUÊNCIAS DE E2F .....	25
3.9 ANÁLISES FILOGENÉTICAS .....	26
3.9.1 Análise A .....	27
3.9.2 Análise B .....	27
3.9.3 Análise C .....	27
3.9.4 Análise D .....	28
<b>4 RESULTADOS.....</b>	<b>29</b>
4.1 ESTRESSE POR LUZ UV-B .....	29
4.2 ANÁLISES FILOGENÉTICAS .....	45
<b>5 DISCUSSÃO .....</b>	<b>51</b>
6 REFERÊNCIAS BIBLIOGRÁFICAS .....	59
<b>ANEXO I.....</b>	<b>62</b>

**ANEXO II.....63**



## LISTA DE ABREVIATURAS

ACT – actina

BLAST – ferramenta de busca por alinhamento local básico, do inglês: *Basic Local Alignment Search Tool*

CDK – Ciclina Dependente de Kinase

cDNA – DNA complementar

CDS – sequência codificadora, do inglês, *Coding Sequence*

cv – cultivar

CYC – ciclina

DNA – Ácido Desoxirribonucléico

dNTP - desoxirribonucleotídeos trifosfatados

DP – Parceiro de dimerização, do inglês, *Dimerization Partner*

DSB – Quebra de fita dupla, do inglês, *Double Strand Brake*

EF1 – Fator de Elongação 1

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária

EST – Etiqueta de sequência expressa, do inglês, *Expressed Sequence Tag*

FAO – Organização de Alimentos e Agricultura, do inglês, *Food and Agriculture Organization*

FDH – Formaldeído Desidrogenase

G1 – fase de crescimento 1 do ciclo celular

G2 – fase de crescimento 2 do ciclo celular

genBank – banco de dados de sequencias genéticas do NCBI

h – horas

H4 – histona H4

MCMC – Monte Carlo via Cadeia de Markov

MEGA – *Molecular Evolutionary Genetics Analysis Program*

min - minutos

M-MLV - Vírus Moloney da leucemia de murinos, do inglês “*Moloney Murine Leukemia Virus*”

MMR – reparo por malpareamento, do inglês, *mismatch repair*

NCBI – Centro Nacional para informação biotecnológica, do inglês, *National Center for Biotechnology Information*

pb – pares de bases

PCD – morte celular programada, do inglês, *Programmed Cell Death*

PCR – reação em cadeia da polimerase, do inglês, *Polymerase Chain Reaction*

PFAM – banco de dados PFAM, do inglês, Protein Family

Primer – oligonucleotídeo iniciador específico

RB – proteína do retinoblastoma

RNA – Ácido Ribonucléico

RNAm – RNA mensageiro

RT-qPCR – Reação em cadeia da polimerase quantitativa associada à transcrição reversa, do inglês, *Reverse Transcript and quantitative Polymerase Chain Reaction*

ssp - subespécie

U - unidades

UBI - ubiquitina

UGLY – uracyl DNA glycosilase

UV – Ultra Violeta

## LISTA DE FIGURAS E TABELAS

<b>Figura 1 Representação dos motivos conservados presentes em E2F típicos, proteínas DP e E2F atípicos.....</b>	<b>19</b>
<b>Figura 2. Análises por RT-qPCR do nível de expressão de um gene de fotoliase (UVR3)(A) e de um gene envolvido com a via de reparo de DNA por excisão de nucleotídeos (RAD1)(B) em resposta ao estresse com luz UV-B. ....</b>	<b>30</b>
<b>Figura 3. Análises por RT-qPCR do nível de expressão de dois genes envolvidos com a morte celular programada: BAG (A) e MXK3 (B), em resposta ao estresse com luz UV-B.....</b>	<b>31</b>
<b>Figura 4. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 04611 (A) e E2F 05530 (B), em resposta ao estresse com luz UV-B.....</b>	<b>32</b>
<b>Figura 5. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 08436 (A) e E2F 08489(B), em resposta ao estresse com luz UV-B.....</b>	<b>33</b>
<b>Figura 6. Análises por RT-qPCR do nível de expressão de um gene de fotoliase (UVR3)(A) e um gene envolvido com a via de reparo de DNA por excisão de nucleotídeos (RAD1)(A) em resposta ao estresse com luz UV-B.....</b>	<b>35</b>
<b>Figura 7. Análises por RT-qPCR do nível de expressão de um gene envolvido na rota de reparo ao DNA por excisão de bases (UGLY)(A) e um gene envolvido na progressão do ciclo celular (CYC B1_1)(A) em resposta ao estresse com luz UV-B.....</b>	<b>36</b>
<b>Figura 8. Análises por RT-qPCR do nível de expressão de dois genes envolvidos na progressão do ciclo celular: CDKB2 (A) e H4 (B), em resposta ao estresse com luz UV-B.....</b>	<b>37</b>

**Figura 9. Análises por RT-qPCR do nível de expressão de dois genes envolvidos na progressão do ciclo celular: CYC D4 (A) e CYC D6 (B), em resposta ao estresse com luz UV-B.....38**

**Figura 10. Análises por RT-qPCR do nível de expressão de dois genes envolvidos com a morte celular programada (PCD): BAG (A) e MXK3 (B), em resposta ao estresse com luz UV-B.....40**

**Figura 11. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 08436 (A) e E2F 08489 (A), em resposta ao estresse com luz UV-B.....41**

**Figura 12. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 04611 (A) e E2F 05530 (B), em resposta ao estresse com luz UV-B.....42**

**Figura 13. Análises por RT-qPCR do nível de expressão dos membros atípicos da família de fatores transcricionais de E2F: DEL 06082 (A) e DEL 05745 (B), em resposta ao estresse com luz UV-B.....43**

**Figura 14. Análises por RT-qPCR do nível de expressão de dois membros da família DP: DP 06098 (A) e DP 02619 (B), em resposta ao estresse com luz UV-B.....44**

**Figura 15. Análise filogenética A: construída por inferência bayesiana das sequência protéicas dos domínios de ligação ao DNA E2F/DP de algas e plantas.....46**

**Figura 16. Árvore filogenética B: construída por inferência bayesiana utilizando-se sequências E2F de algas, plantas e animais. ....48**

**Figura 17. Árvore filogenética C: construída por inferência bayesiana utilizando-se sequências DEL de plantas e animais. ....49**

**Figura 18. Árvore filogenética D: construída por inferência bayesiana utilizando-se sequências DP de algas, plantas e algumas espécies de animais.....50**

**Figura 19. Expressão relativa dos genes em função do tempo após estresse com luz UV-B sob plantas de arroz.....55**

**Figura 20. Representação esquemática do modelo de atuação dos genes da família E2F em resposta a danos causados pela luz UV-B.....58**

# 1 INTRODUÇÃO

## 1.1 O ARROZ

Devido a sua ampla produção, área de cultivo e consumo, o arroz (*Oryza sativa* L.) destaca-se como um dos cereais mais importantes do mundo, sendo significativa fonte alimentícia em todos os continentes. Segundo dados da FAO (*Food and Agriculture Organization*), o arroz, em 2010, apresentou produção mundial de aproximadamente 700 milhões de toneladas. As estimativas realizadas em julho de 2011, para a safra de 2011, apontam um aumento de aproximadamente 18 milhões de toneladas (2,5%), resultando em um novo recorde de produção mundial. Em 2010, a área ocupada mundialmente pelo plantio dessa cultura foi de, aproximadamente, 165 milhões de hectares e foi previsto para 2011 um aumento para quase 170 milhões (FAO, 2012).

Este cereal é a cultura alimentar de maior importância na maioria dos países em desenvolvimento, sendo o alimento básico para cerca de 2,4 bilhões de pessoas. O arroz possui um excelente balanço nutricional, fornecendo 20% da energia e 15% da proteína *per capita* necessária à dieta humana. A média estimada de consumo anual de arroz em 2007 foi de 52,96 kg por pessoa (FAO, 2007). Estima-se que até 2050 a demanda pelo cereal deverá dobrar (EMBRAPA, 2007).

Uma característica marcante da cultura de arroz é sua versatilidade. Por possuir uma facilidade de adaptação a diferentes condições de solo e clima, o arroz é considerado uma das espécies que apresenta um dos maiores potenciais de combate à fome no mundo. No entanto, a produtividade das lavouras é afetada pelas condições ambientais adversas que podem provocar grandes perdas. Portanto, o aumento da produtividade na agricultura e a segurança alimentar global estão diretamente relacionadas com os estresses ambientais enfrentados pelas culturas (Dhaliwhal G.S, 1999).

O arroz (*Oryza sativa*) é considerado a planta modelo de estudos entre as monocotiledôneas, assim como *Arabidopsis thaliana* é para as dicotiledôneas. Aquele é um organismo diplóide com 24 cromossomos e com genoma haplóide de 390 megabases (Mb), o menor entre os cereais, que está seqüenciado (Goff *et*

*al.*, 2002) e disponível para consulta em bancos de dados públicos (Ouyang *et al.*, 2007). A disponibilidade de etiquetas de genes expressos ou ESTs (“*Expressed Sequence Tags*”) e de mapas moleculares saturados, o progresso nos protocolos de transformação mediada por *Agrobacterium tumefaciens*, além da sintonia do genoma de arroz com o de outros cereais como aveia, milho e trigo, fazem do arroz um organismo modelo para estudos fisiológicos, genéticos e de evolução dentre as plantas, principalmente em relação às monocotiledôneas (Gale & Devos, 1998; Shimamoto & Kyozuka, 2002). Tendo em vista que o arroz é uma espécie modelo para estudos de genômica nas gramíneas, torna-se evidente que o estudo do seu genoma terá um grande impacto na agricultura do século 21.

## 1.2 O ESTRESSE POR LUZ ULTRAVIOLETA

A radiação ultravioleta (UV) que atinge a superfície da terra tem implicações importantes na saúde humana, nos ecossistemas terrestres e aquáticos, nos ciclos biogeoquímicos, na qualidade do ar (McKenzie *et al.*, 2011). A revisão realizada por McKenzie *et al.*, em 2011, mostra que inúmeros trabalhos estão sendo realizados nos últimos anos visando uma melhor compreensão das respostas biológicas a esta radiação.

A luz UV, que possui ondas de comprimento entre 220 e 400 nm, é emitida pela luz solar. A luz UV-B de comprimento de onda entre 280 e 320 nm, corresponde ao componente mais energético da luz solar que chega à superfície terrestre (Mpoloka, 2008).

Pesquisas sobre esse tema foram estimuladas pela percepção, há mais de 30 anos, que a camada de ozônio estratosférico estaria em risco e que haveria consequentes aumentos de radiação UV-B atingindo a superfície terrestre. Esses aumentos devem-se à diminuição da quantidade de ozônio durante os anos 1980 e 1990, particularmente em altas latitudes (>~60°), onde a destruição deste gás foi mais pronunciada (McKenzie *et al.*, 2011).

Quase toda a radiação UV-C danosa e maior parte da radiação UV-B são filtradas pela camada de ozônio. O aumento dos níveis de UV-B atingindo a superfície terrestre se deve à destruição acelerada da camada de ozônio estratosférico devido à poluição do ar (Mpoloka, 2008).

Uma vez que esta radiação é prontamente absorvida, ela pode provocar uma série de efeitos, tal como a fotoexcitação de um grande número de biomoléculas (i. e. ácidos nucleicos, proteínas e lipídios), resultando em mudanças de funções genéticas, bioquímicas e fisiológicas dentro das células, com impactos significativos sobre muitos processos biológicos que podem ter importância regulatória (Tian & Yu, 2009).

Em uma meta-análise realizada em 2011, Ballaré et al observaram que existe uma tendência de resposta da planta no sentido de diminuição da biomassa vegetal quando submetida à irradiação UV-B. Os autores indicam que pode existir uma redução aproximada de 1% no crescimento para um aumento de 3% na incidência de irradiação UV-B. Se essa tendência for verificada no contexto das mudanças na quantidade de radiação UV-B que ocorreram nos últimos 30 anos, é possível produzir uma estimativa do tamanho das mudanças geradas no crescimento dos produtores primários (seres autotróficos), que, portanto, pode ter ocorrido como consequência da destruição do ozônio (Ballaré *et al.*, 2011).

Wargent et al, 2011 verificaram que folhas de alface (*Lactuca sativa*) estressadas por luz UV-B possuíam a área foliar diminuída em 18% quando comparadas com folhas de plantas não-tratadas. Essa redução ocorreu para outros parâmetros medidos pelos autores como: largura, espessura e comprimento, além de peso fresco (Wargent *et al.*, 2011). Outro estudo mostrou que as trocas gasosas nas folhas de arroz foram severamente afetadas pela exposição à UV-B. A soma dos pigmentos do ciclo da xantofila (violaxanthin, antheraxanthin, zeaxantina) foi moderadamente afetada um dia após o término do tratamento UV-B, mas sofreram uma queda muito mais forte (de 61%) pelo 7º dia após o final da irradiação. Em folhas totalmente desenvolvidas foi observado um impacto bem menor, mas alguns efeitos posteriores significativos ainda estavam presentes em alguns dos parâmetros testados como, por exemplo, uma queda na condução de vapor de água pelo estômato. Os autores verificaram ainda que após o estresse por UV-B, as folhas de arroz irradiado tornaram-se gradualmente cloróticas, seguida por necrose progressiva que pode ser detectado 14 dias após o fim da irradiação UV-B (Lidon & Ramalho, 2011).

A luz do tipo UV-B interage com o DNA, modificando suas bases ou provocando a interação entre estas, bloqueando, assim, os processos no qual o



DNA encontra-se envolvido (i. e. replicação do DNA). Esta interação da luz UV-B com o DNA pode afetar a integridade do genoma (Mpoloka, 2008). Em resposta a esse estresse, são ativadas complexas redes de sinalização visando bloquear o efeito provocado por esses estressores.

As plantas terrestres possuem caráter sésbil e dependência da luz solar para a fotossíntese. É provável que elas tenham que lidar com os níveis aumentados de UV-B nas próximas décadas, o que leva à necessidade de pesquisas contínuas para melhor compreensão a respeito das estratégias de aclimatação a esse tipo de estresse (Lidon & Ramalho, 2011).

### 1.3 OS FATORES DE TRANSCRIÇÃO E2F

A família gênica que codifica os fatores de transcrição E2F tem sido extensivamente estudada em mamíferos. Esta família possui membros que desempenham funções diversas tais como: progressão do ciclo celular, transição G1/S, segregação e condensação do cromossomo, replicação e reparo do DNA, interesse principal deste trabalho, além de diferenciação e apoptose. Os genes dessa família apresentam similaridades entre plantas e mamíferos, tanto em nível estrutural, quanto funcional (Ramirez-Parra *et al.*, 1999). No contexto das respostas ao dano de DNA, alguns dados estabeleceram uma ligação entre a resposta à quebra na fita dupla (*double strand break* – DSB) e a ativação de genes alvos de E2F. No entanto, o envolvimento de E2F em morte celular programada (*programmed cell death* – PCD) em plantas não é conhecido (Blais & Dynlacht, 2004; Inze, 2005).

Em mamíferos, foi descrito um novo papel para E2F7 e E2F8 durante a resposta a danos no DNA. Observou-se que ambos os genes são induzidos em células tratadas com agentes danosos ao DNA. Além disso, foi demonstrado que esses genes regulam a expressão de um gene da própria família, o E2F1, sendo essencial para as alterações que ocorrem no ciclo celular em decorrência do dano. Juntos, estes resultados indicam que E2F7 e E2F8 estão intimamente envolvidos com a resposta celular ao dano no DNA (Zalmas *et al.*, 2008).

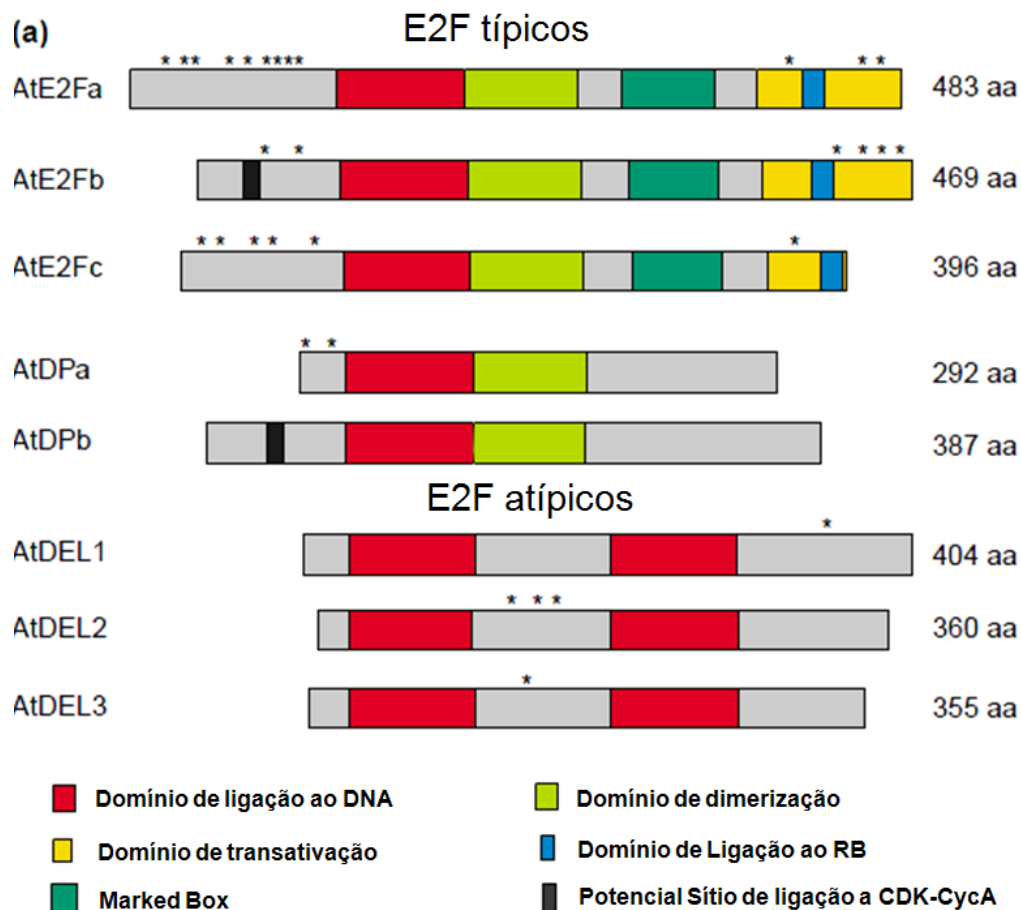
MSH6, é um gene componente da rota de reparo de bases mal pareadas em *Arabidopsis thaliana*, foi identificado como um alvo dos fatores de transcrição E2F. Lario *et al.*, 2011 demonstraram que o promotor de MSH6 contém três sítios de

ligação de E2F, dois deles compreendendo motivos duplos de ligação a E2F. Uma investigação posterior da regulação da expressão de MSH6 mostrou um nível aumentado dos transcritos de MSH6 em plantas de *Arabidopsis* que superexpressam os ativadores E2Fa e E2Fb. Estes resultados sugerem que fatores de transcrição E2F regulam a expressão de MSH6 e, conseqüentemente, a atividade MMR (mismatch repair) durante a progressão do ciclo celular (Lario *et al.*, 2011).

A família E2F é composta por dois tipos de proteínas: os E2F típicos e os E2F atípicos. Os genes E2F da planta modelo *Arabidopsis thaliana* foram previamente caracterizados (Mariconti *et al.*, 2002). Para os E2F típicos, doravante denominados somente de E2F, os domínios característicos são: um domínio de ligação ao DNA responsável pela interação da proteína com os genes alvos, um domínio de heterodimerização responsável pela dimerização com a proteína DP (Dimerization Partner), que será descrita posteriormente, uma região descrita como “*marked Box*” que é necessária para interação proteína-proteína e que confere especificidade à resposta transcricional, um domínio de transativação e um domínio de ligação à proteína do Retinoblastoma (RB). As proteínas E2F atípicos, descritos neste trabalho como proteínas DEL, possuem uma organização diferente, com apenas dois domínios de ligação ao DNA (Kosugi & Ohashi, 2002) (figura 1).

As proteínas DP (Dimerization Partner) também possuem um domínio de ligação ao DNA e um domínio de heterodimerização, o que condiz com o possível mecanismo de atuação deste complexo protéico (figura 1). Segundo este modelo, as proteínas E2F interagem e formam heterodímeros com a proteína DP, formando, assim, um complexo protéico com dois domínios de ligação ao DNA. Esta ligação ao DNA promove a ativação da transcrição. Por outro lado, as proteínas DEL são competidores diretos dos E2F/DP e teriam um efeito antagônico, uma vez que possuem dois domínios de ligação ao DNA e não possuem os domínios responsáveis pela ativação da transcrição. Como estas proteínas ocupam a mesma sequência dos genes alvo, ao qual se liga o heterodímero E2F-DP, elas bloqueiam o efeito desse complexo, produzindo um efeito repressor por competição pela sequência consenso de ligação (Lammens *et al.*, 2009) que foi caracterizada como sendo: TTTCCCGC (de Jager *et al.*, 2001) . Este motivo de ligação é bem descrito e conservado entre plantas e animais, inclusive em humanos. .

Existe ainda outra rota de regulação da transcrição pelas proteínas da família E2F. Esta via é controlada pela proteína retinoblastoma (RB). Experimentos *in vitro* demonstraram alta afinidade entre a proteína RB e o complexo E2F/DP. Em animais, a proteína RB se liga ao heterodímero E2F/DP em um local correspondente ao domínio de transativação C-terminal. Esta ligação promove a inibição do complexo com consequente inibição da transcrição. Para o desligamento da proteína RB do complexo E2F/DP, a proteína RB deve ser fosforilada por diversas quinases dependentes de ciclinas (CDK), fazendo com que esta perca sua afinidade pelo complexo e o torne livre para ativar a transcrição dos genes alvo. Este sistema é bastante conservado e está presente na maioria dos eucariotos, inclusive em plantas superiores (Shen, 2002).



\* Potencial sítio de fosforilação CDK-ciclina

**Figura 1** Representação dos motivos conservados presentes em E2F típicos, proteínas DP e E2F atípicos. Domínios caracterizados para *Arabidopsis thaliana*. Figura modificada (Shen, 2002).

Esta família gênica possui nove membros em *Arabidopsis thaliana*, sendo três E2F, três DEL e dois DP (Mariconti *et al.*, 2002). Até o momento, para *Nicotiana tabacum* foi descrito apenas um membro da família E2F, sendo descrito como NtE2F (Sekine *et al.*, 1999). Em trigo foi também descrito apenas um E2F, tendo sido denominado TmE2F, além de um gene DP (Ramirez-Parra *et al.*, 1999; Ramirez-Parra & Gutierrez, 2000). Para mamíferos, já foram caracterizados oito membros da família gênica E2F, sendo seis E2F e dois DEL, além de duas proteínas DP (Dyson, 1998; Zalmas *et al.*, 2008). Em trabalhos prévios, nosso grupo identificou no genoma de arroz, 9 sequências codificando proteínas E2F, sendo 4 E2F, 2 DEL e 3 DP (Rauber, 2009).

Análises genéticas e bioquímicas e revisões bibliográficas recentes têm permitido um avanço no entendimento dos mecanismos de percepção e da sinalização de estresses genotóxicos (Sakamoto *et al.*, 2009; Tuteja *et al.*, 2009; Lario *et al.*, 2011; Lidon & Ramalho, 2011). Entretanto, os componentes específicos de sinalização que são ativados nas plantas em resposta a esses estresses são ainda pouco conhecidos. A comparação entre os diferentes modelos eucarióticos e a integração entre análises moleculares, bioquímicas e fisiológicas fornecerá uma visão mais abrangente a respeito das respostas geradas pelas plantas ao dano no DNA.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

O objetivo deste trabalho foi caracterizar o padrão de expressão dos genes da família E2F em arroz em resposta a um estresse genotóxico visando entender as relações entre estes fatores de transcrição e as respostas ao dano de DNA em plantas além de aumentar o conhecimento a respeito da história evolutiva desses fatores transcricionais.

### 2.2 OBJETIVOS ESPECÍFICOS

- 1 ) Caracterizar o padrão de expressão dos genes da família E2F em folhas de arroz em resposta ao tratamento com radiação UV-B.
- 2 ) Relacionar o padrão de expressão dos genes E2F e os genes de reparo ao dano do DNA, ciclo celular e morte celular programada em arroz em função do tratamento com UV.
- 3 ) Obter as sequencias gênicas das famílias E2F/DEL e DP para 23 espécies de plantas e 2 espécies de algas.
- 4 ) Fazer uso de métodos de inferência bayesiana para estudos evolutivos das sequencias identificadas.

### 3 MATERIAL E MÉTODOS

#### 3.1 MATERIAL VEGETAL

Para os experimentos com luz UV-B, foram utilizadas plantas de arroz (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare) crescidas em casa de vegetação à temperatura de 28°C e foto-período de 12/12h em potes com terra e permanentemente submersos em bandejas com água.

#### 3.2 ESTRESSE POR LUZ UV-B

Foram realizados dois experimentos utilizando uma lâmpada que emite luz no comprimento de 280nm (UV-B) como única fonte luminosa durante o período de estresse.

No primeiro experimento, denominado experimento sem aclimação, as plantas de arroz foram submetidas à radiação com luz Ultravioleta B (UV-B) conforme descrito a seguir:

- Para o tratamento, um lote de 24 plantas com idade aproximada de quatro semanas foi submetido a uma aplicação de 4 horas de luz UV-B (plantas tratadas);
- Como controle, 24 plantas (não-tratadas) foram mantidas nas mesmas condições das plantas estressadas, porém submetidas à iluminação normal, sem a presença de luz UV;
- Prosseguiu-se a coleta da parte aérea de duas plantas e estas formaram uma amostra correspondente a uma das 4 replicatas biológicas que foram utilizadas para cada tempo de coleta distribuídos em: 0 hora (logo após o estresse), 6 horas e 24 horas após o tratamento, denominados de 0h, 6h e 24h, respectivamente;
- As amostras de tecido vegetal (parte aérea) foram coletadas, imediatamente congeladas em nitrogênio líquido e estocadas em ultra freezer a -80°C até o momento da extração de RNA.

No segundo experimento, denominado experimento com aclimação, procederam-se algumas alterações devido aos resultados obtidos no primeiro experimento:

- As plantas possuíam 7 semanas de vida;
- Estas foram submetidas a um pré-estresse de 4 horas de luz UV-B, realizado uma semana antes do experimento;
- Para o tratamento, as plantas foram submetidas a um período de 4 horas de luz UV-B, seguido de um intervalo de 20 horas de iluminação normal. Em seguida, um segundo período de 4 horas com luz UV-B foi aplicado às plantas tratadas;
- Logo após o segundo período de estresse, as amostras biológicas foram coletadas nos seguintes tempos: 0 (logo após o estresse), 30 minutos, 2, 6 e 24 horas após o estresse e estes foram denominados 0h, 30min, 2h, 6h e 24h, respectivamente.

### 3.3 EXTRAÇÃO DE RNA TOTAL

A extração do RNA total foi realizada com o reagente TRIZOL (Invitrogen) seguindo as instruções do fabricante.

### 3.4 SÍNTESE DE cDNA A PARTIR DE RNA TOTAL

A síntese de cDNA a partir de RNA total foi realizada utilizando-se a enzima M-MLV Reverse Transcriptase (Promega) segundo as recomendações do fabricante, com as seguintes adaptações: em um volume final de 20  $\mu\text{L}$ , foram utilizados 3  $\mu\text{L}$  de RNA total de arroz (1 $\mu\text{g}$  de RNA total aproximadamente), 1  $\mu\text{L}$  do oligonucleotídeo poli-T (1  $\mu\text{g}/\mu\text{L}$ ), 1 $\mu\text{L}$  de desoxirribonucleotídeos (dNTP) a 10 mM, 7  $\mu\text{L}$  de água Milli-Q ultra-pura. A solução foi incubada à 65°C por 5 minutos e imediatamente incubada em gelo por mais 5 minutos. A seguir, 10  $\mu\text{L}$  de tampão 2,5X (First Strand Buffer 2,5X + DTT 0,1 M), 27  $\mu\text{L}$  de H<sub>2</sub>O, 1  $\mu\text{L}$  de M-MLV Reverse Transcriptase (200 U/ $\mu\text{L}$ ) foram adicionados à solução. A reação foi mantida a 37°C por 2 horas e em seguida incubada a 70°C por 15 minutos. A qualidade dos cDNAs

foi verificada por PCR usando primers para o gene constitutivo Formaldeído Desidrogenase (FDH - código de acesso no GenBank: U77637.1).

### 3.5 DETERMINAÇÃO DE OLIGONUCLEOTÍDEOS INICIADORES ESPECÍFICOS (*PRIMERS*) PARA PCR

As seqüências CDS dos genes avaliados nesses experimentos foram utilizadas para a construção de oligonucleotídeos iniciadores específicos (*primers*) usando a ferramenta *primer3* (Rozen & Skaletsky, 2000). Os *primers* foram projetados para produzir fragmentos de 180 a 250 pb para as análises de PCR quantitativo em tempo real. Os primers foram delineados com uma sequencia que fornecesse uma temperatura de fusão de todos os *primers* equivalente a 60 °C.

### 3.6. ANÁLISE POR PCR QUANTITATIVA EM TEMPO REAL (RT-qPCR)

A determinação da expressão relativa de cada um dos genes foi realizada pela técnica de PCR quantitativa em tempo real (RT-qPCR). As reações foram realizadas em aparelho Step One Plus (Applied Biosystem). As condições de reação seguiram os seguintes parâmetros: uma etapa de desnaturação inicial de 5 min a 94 °C, seguida por 40 ciclos de 10 s a 94 °C, 15 s a 60 °C, e 15 s a 72 °C. Após a amplificação, as amostras foram mantidas por 2 min a 40 °C para reanelamento e então aquecidas de 55 até 99 °C com uma taxa de incremento de 0.4 °C/s para aquisição de dados e obtenção da curva de desnaturação. As reações foram feitas em um volume final de 20 µl contendo 10 µl do produto da transcrição reversa (diluída 100 vezes em água estéril), 2 µl de tampão PCR 10x (Invitrogen), 1.2 µl de MgCl<sub>2</sub> 50 mM, 0.4 µl de dNTPs 5 mM, 0.2 µl de cada par de *primers* específicos 10 µM, 3.95 µl de água ultra pura, 2.0 µl de SYBR green (1:100,000, Molecular Probes), e 0.05 µl de Platinum Taq DNA polimerase (5 U/µl, Invitrogen). Pares de *primers* que amplificam genes constitutivos como o fator de alongação (Ef1; código de acesso no GenBank: AJ252135); Ubiquitina (Ubi; código de acesso no GenBank: AK103230) e actina 1 (ACT; código de acesso no GenBank: NC\_008396) foram usados como controle interno para normalizar a quantidade de mRNA presente em cada amostra. Todos os resultados foram expressos através de



quantificação relativa a um conjunto desses genes constitutivos, utilizando-se a metodologia  $2^{-\Delta\Delta Ct}$  descrita por Livak & Schmittgen (2001). Os experimentos foram realizados em quadruplicatas técnicas (Livak & Schmittgen, 2001).

### 3.7 ANÁLISE DA EXPRESSÃO DOS GENES E2F DE ARROZ E DE GENES ENVOLVIDOS EM TRÊS MECANISMOS CELULARES.

As reações de RT-qPCR foram realizadas utilizando-se primers específicos para todos os genes da família gênica E2F identificados em arroz e mais dois dos três genes identificados para proteínas DP. As sequências dos *primers* utilizados em todos os experimentos estão descritas no anexo 1.

Além disso, foram projetados *primers* para amplificar os genes envolvidos em três diferentes processos celulares: reparo de DNA, ciclo celular e apoptose. O gene UVR3 é uma fotoliase, RAD1 está envolvido no sistema de reparo por excisão de nucleotídeos, enquanto o gene Uracil DNA glicosilase (UGLY) responde ao reparo por excisão de base (anexo 1). Para os genes envolvidos no ciclo celular, foram selecionados CYCB1\_1 e CDKB2, que estão relacionados com a passagem da fase G2 para M (Lee *et al.*, 2003; La *et al.*, 2006), os genes para CYC D4 e CYCD6, que são ciclinas específicas da fase G1 do ciclo celular (La *et al.*, 2006), e o gene da Histona H4 (Reichheld *et al.*, 1998). Os genes marcadores de morte celular programada são BAG (BAG-domain protein 1 / regulator of cell death - NCBI) e MXK3 (cell death 2 C-terminal domain-containing protein – Phytozome).

### 3.8 BUSCA DE SEQUÊNCIAS DE E2F

A busca dos genes pertencentes à família gênica E2F foi realizada utilizando o bancos de dados Phytozome (Goodstein *et al.*, 2012) que disponibiliza publicamente o genoma completo de diversas espécies de plantas e algas (disponível em: <http://www.phytozome.net/>).

As sequências homólogas a E2F presentes nos genomas disponíveis foram obtidas através da ferramenta BLAST (*Basic Local Alignment Search Tool*), utilizando como isca (*query*) sequências E2F descritas para *Arabidopsis thaliana* (Mariconti *et al.*, 2002) e *Oryza sativa* (Rauber, 2009). Nas sequências que

retornaram como respostas da busca avaliou-se o “valor de *e*” (valor determinado por um algoritmo que informa a probabilidade desde alinhamento ser aleatório), e o ponto de corte estabelecido foi  $1.0e^{-5}$ . A presença dos domínios protéicos característicos dos E2F/DEL/DP, verificado através da ferramenta PFAM (Sonnhammer *et al.*, 1997; Punta *et al.*, 2012), foi utilizado como critério de seleção para inclusão da sequência na análise. Os domínios utilizadas nessa seleção foram: i) a presença de um “*E2F/DP family winged-helix DNA-binding domain*” ao longo da proteína (para proteínas E2F típicas); ii) a presença de dois “*E2F/DP family winged-helix DNA-binding domain*” (para proteínas DEL) e, iii) a presença de um “*E2F/DP family winged-helix DNA-binding domain*” e um domínio “*Transcription factor DP*” (para proteínas DP). Após a confirmação dessas características em cada grupo de proteínas, realizou-se a coleta dos seguintes dados: informação da sequência (número de acesso); obtenção das sequências codificadoras (CDS) e das sequências protéicas correspondentes. Todas essas informações foram inseridas em um banco de dados, que encontra-se disponível no anexo 2.

As sequências foram sendo nomeadas após a confirmação da existência dos domínios citados acima. Cada proteína encontrada no banco do JGI (Phytozome) era nomeada pelo nome da proteína e uma letra. Esta letra apenas representa a ordem em que essa proteína foi encontrada, sendo a proteína E2FA, portanto, a primeira a ser encontrada, a segunda a ser encontrada foi denominada E2FB, e assim por diante. A única espécie que possuía nomes próprios no banco do JGI é da planta modelo *Arabidopsis thaliana*, sendo o nome fornecido pelo banco e confirmado na literatura. No banco do NCBI todas as proteínas já possuíam o devido número e estes também foram caracterizados por Peng *et al* em artigo publicado em 2010.

### 3.9 ANÁLISES FILOGENÉTICAS

As sequências obtidas através das buscas foram utilizadas para a reconstrução filogenética de dessas proteínas. Para isso, as sequências protéicas foram alinhadas utilizando o programa MUSCLE (Edgar, 2004) implementado no MEGA 5.0 (Molecular Evolutionary Genetics Analysis) (Tamura *et al.*, 2011). O alinhamento foi codificado manualmente para eliminar regiões ambíguas.

As análises filogenéticas foram realizadas após o alinhamento das sequências protéicas usando o método de análise bayesiana. Essa análise foi realizada utilizando-se o programa MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003). O algoritmo de Monte Carlo via Cadeia de Markov (MCMC) foi iniciado a partir de uma árvore aleatória e processado por 5.000.000 de gerações em duas corridas independentes. Amostras de cada 100 gerações foram coletadas, sendo que 25% das amostras iniciais (*burn in*) coletadas foram descartadas e as remanescentes foram utilizadas para determinar as distribuições dos valores de probabilidade posterior. O modelo utilizado nessa análise foi o modelo misto de substituição de aminoácidos. A convergência das duas corridas foi acessada checando a média do desvio padrão das frequências (*split frequencies*) menores que 0,01. Nos casos onde a média estava acima deste valor, foi adicionado mais 5.000.000 de gerações à corrida e assim sucessivamente até alcançar o valor abaixo de 0,01.

As quatro análises rodadas foram nomeadas como A, B, C e D.

### 3.9.1 Análise A

A análise denominada letra A está representada na figura 17. Para esta análise foram utilizados os domínios de ligação ao DNA completos para todas as proteínas encontradas neste estudo. Para as proteínas DEL, que possuem dois domínios de ligação ao DNA, o domínio que estava presente no começo da proteína foi denominado DBDA enquanto o segundo domínio de ligação ao DNA foi denominado DBDB. Para esta análise foram utilizadas as sequências pertencentes a plantas e algas presentes no banco Phytozome. Um total de 226 sequências foram utilizadas, perfazendo 106 caracteres.

### 3.9.2 Análise B

Esta análise contou com a sequência protéica de 109 proteínas E2F de plantas, algas, metazoários, amebas e coanoflagelados. A árvore consenso gerada por esta análise está representada na figura 18. As sequências alinhadas tiveram regiões não informativas eliminadas manualmente, restando um conjunto de 252 aminoácidos para a análise.

### 3.9.3 Análise C

Para a construção da árvore consenso contendo as proteínas DEL (figura 19) foram utilizadas todas as proteínas DEL encontradas na busca em plantas, algas, metazoários, amebas e coanoflagelados, somando 54 sequencias. O alinhamento foi corrigido manualmente e regiões não informativas foram excluídas, restando ainda 298 aminoácidos em cada proteína para a análise.

### 3.9.4 Análise D

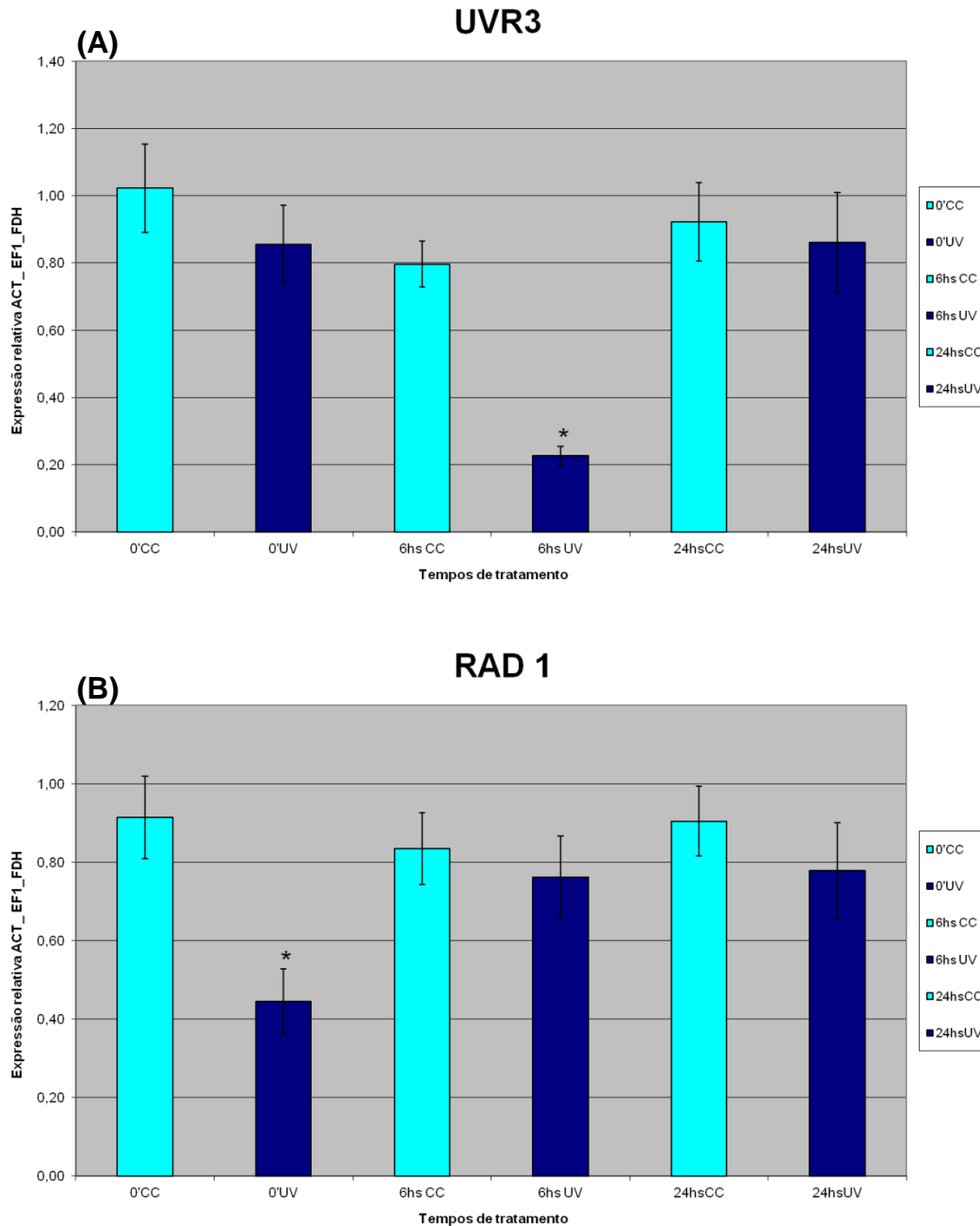
Esta análise contou com 67 sequencias encontradas para as proteínas DP em plantas, algas, metazoários, amebas e coanoflagelados. Ao longo do alinhamento foram eliminadas regiões não informativas e sobraram 265 caracteres a serem analisados. A árvore consenso gerada por esta análise está representada na figura 20.

## 4 RESULTADOS

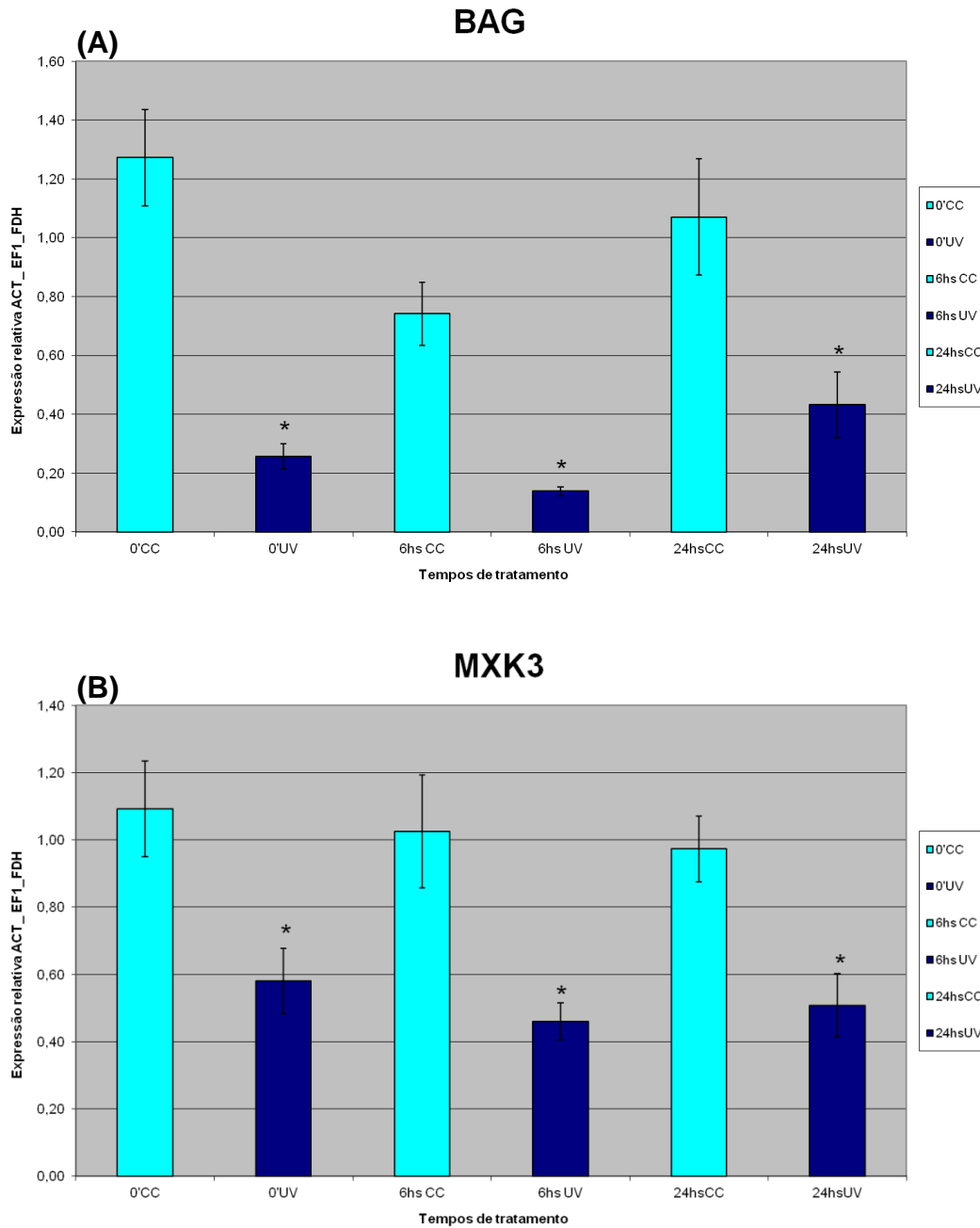
### 4.1 ESTRESSE POR LUZ UV-B

Para verificar como a família gênica E2F responde a possíveis danos no DNA, dois experimentos com luz UV-B foram conduzidos. Através da análise do padrão de expressão de genes envolvidos em mecanismos celulares como o ciclo celular, a morte celular programada e o dano ao DNA, buscou-se distinguir a resposta desses fatores transcricionais a um dano genotóxico.

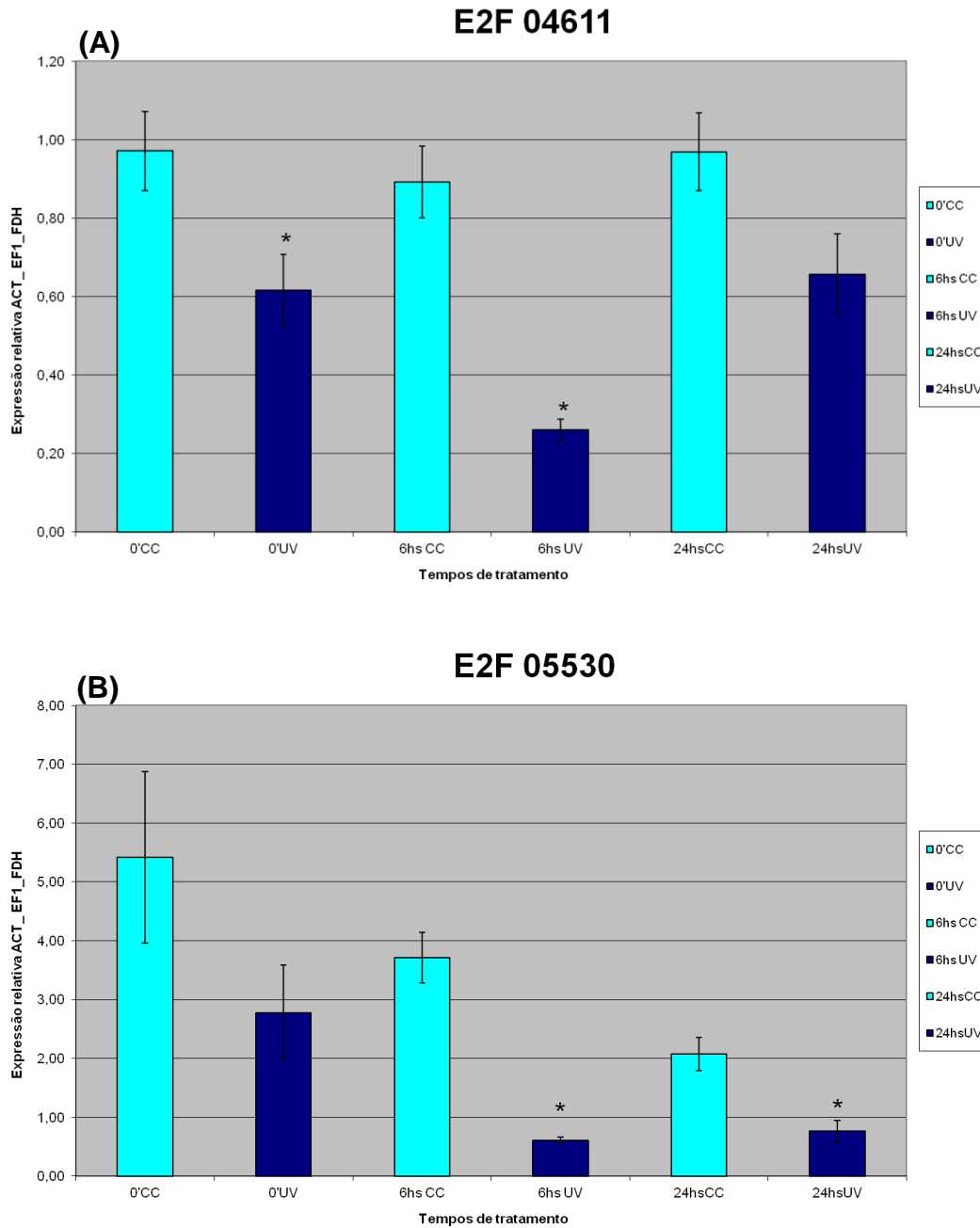
Na figura 2 estão quantificadas as expressões relativas de dois genes de reparo. Este experimento sem aclimatação foi constituído de três tempos de coleta das amostras biológicas: logo após o estresse de 4 horas de luz UV-B (0h), 6 horas após o estresse (6h) e 24 horas após o estresse (24h). É possível observar que os genes UVR-3 e RAD-1, envolvidos na rota de reparo, possuem expressão relativa diminuída nos tempos 6h e 0h respectivamente, indicando uma repressão na rota. O mesmo acontece com os genes BAG e MXK3, relacionados com a rota de apoptose, (figura 3). Apesar destes resultados de repressão gênica divergirem do esperado devido ao dano genotóxico, os quatro genes E2F típicos foram analisados. As figuras 4 e 5 mostram que todos os genes E2F típicos também possuem expressão reduzida ou equivalente à amostra controle.



**Figura 2. Análises por RT-qPCR do nível de expressão de um gene de fotoliase (UVR3)(A) e de um gene envolvido com a via de reparo de DNA por excisão de nucleotídeos (RAD1)(B) em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por um período de 4 horas e coletadas 0, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.**

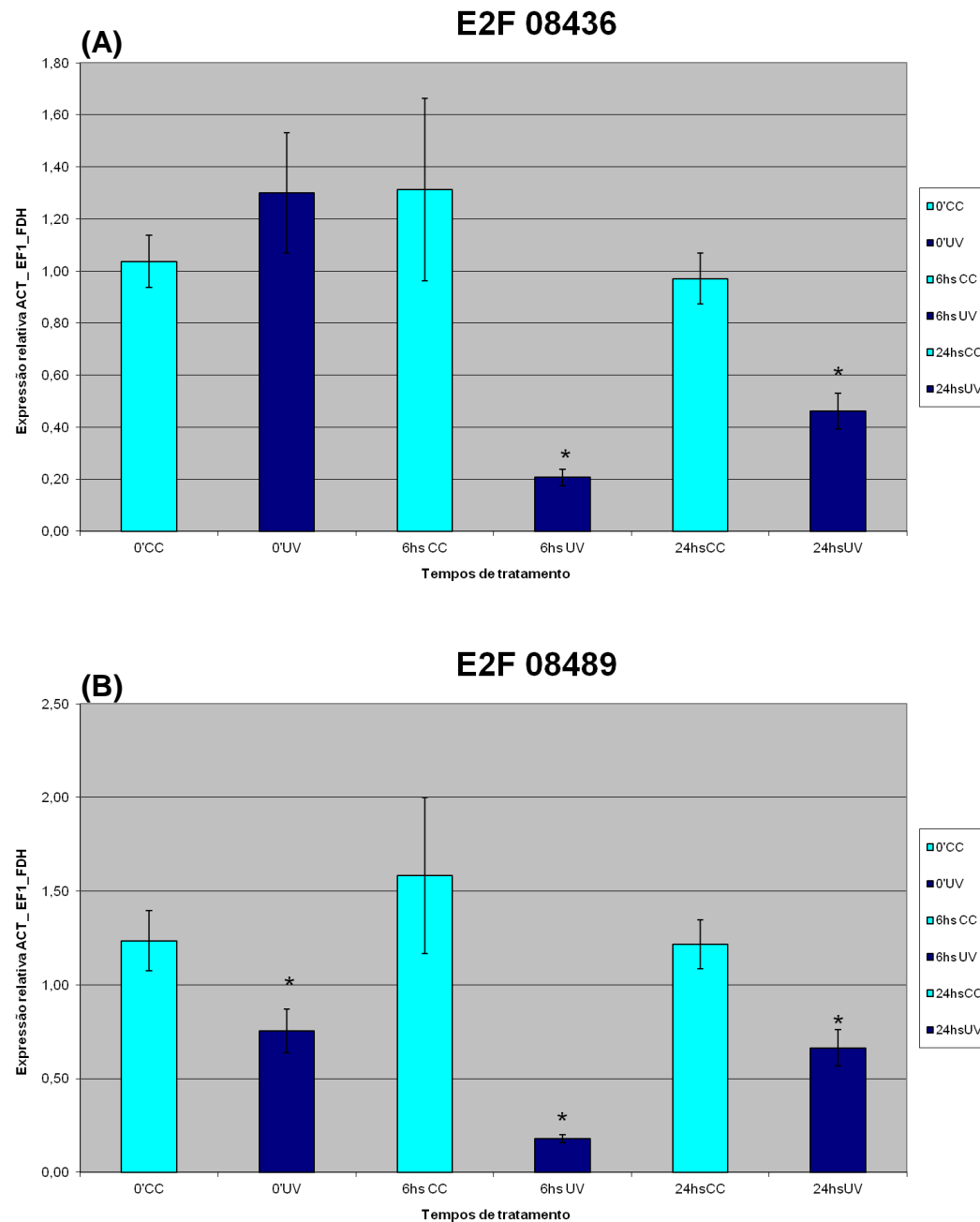


**Figura 3. Análises por RT-qPCR do nível de expressão de dois genes envolvidos com a morte celular programada: BAG (A) e MXK3 (B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por um período de 4 horas e coletadas 0, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.



**Figura 4. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 04611 (A) e E2F 05530 (B), em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por um período de 4 horas e coletadas em 0, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.**





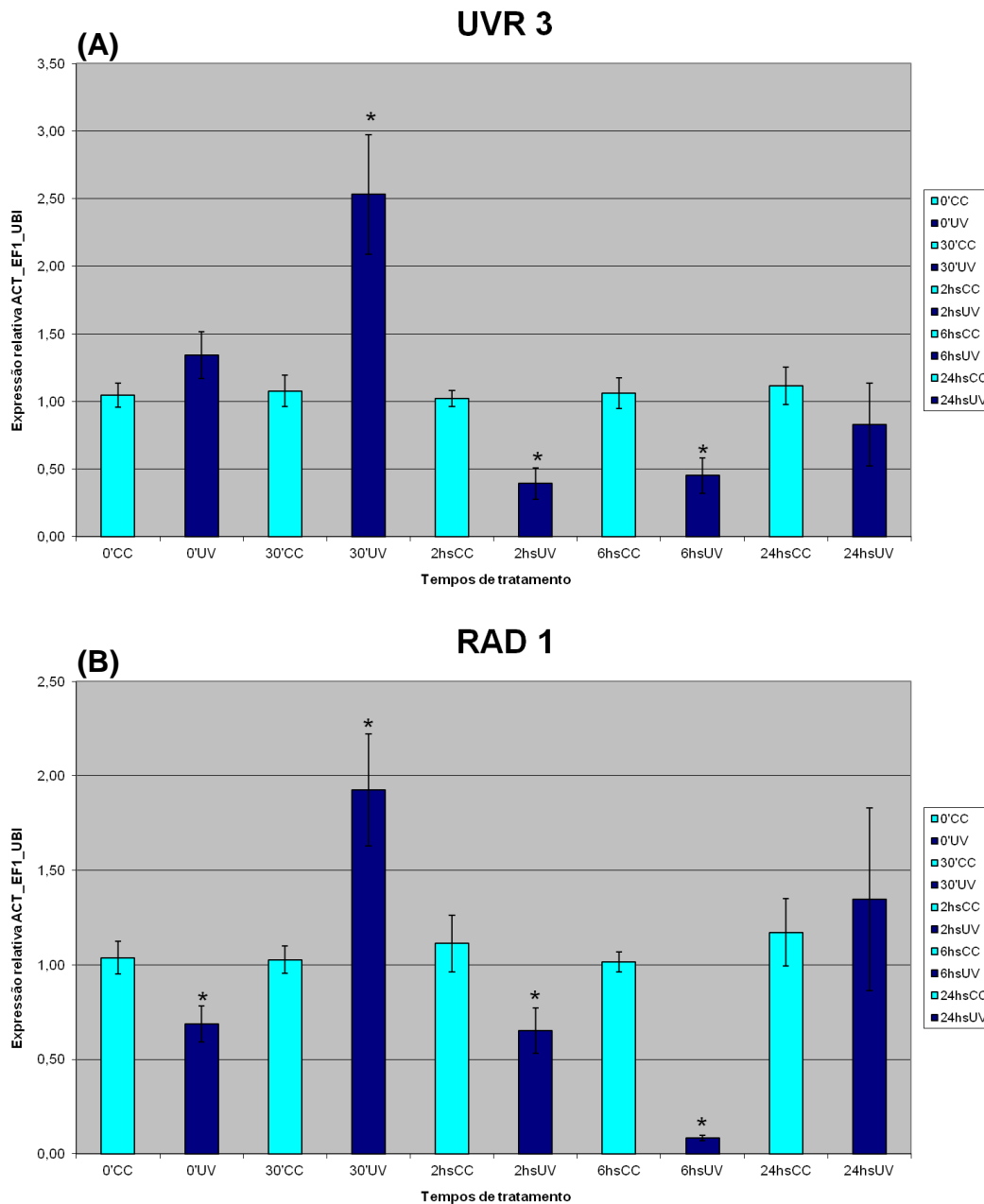
**Figura 5. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 08436 (A) e E2F 08489(B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por um período de 4 horas e coletadas em 0, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.

Tendo em vista os resultados do experimento sem aclimação, um segundo lote de plantas foi preparado para um experimento, o qual foi projetado para permitir uma aclimação das plantas ao estresse (ver Material e Métodos). As figuras 6 e 7 mostram os padrões de expressão dos genes de reparo em resposta ao tratamento com radiação UV-B neste novo experimento constituído de cinco tempos de coleta: logo após o estresse com luz UV-B (0h), 30 minutos após o estresse (30min), 2 horas após o estresse (2h), 6 horas após o estresse (6h) e 24 horas após o estresse (24h). Os resultados mostram que dois dos três genes testados, que previamente foram identificados pela literatura como envolvidos na rota de reparo ao DNA, apresentaram expressão diferenciada. O gene UV3 possui a expressão relativa aumentada 30 minutos após o estresse, aumento esse que foi quantificado como sendo aproximadamente 2,5 vezes maior do que na situação controle. O mesmo acontece com o gene RAD1 que possui seus níveis de mRNA aumentados significativamente em aproximadamente duas vezes no tempo de 30min. O gene da Uracyl DNA glicosilase respondeu de maneira diferente, com a expressão diminuída nos tempos 0h e 30min, retornando aos níveis normais<sup>1</sup> nos tempos de coleta seguintes.

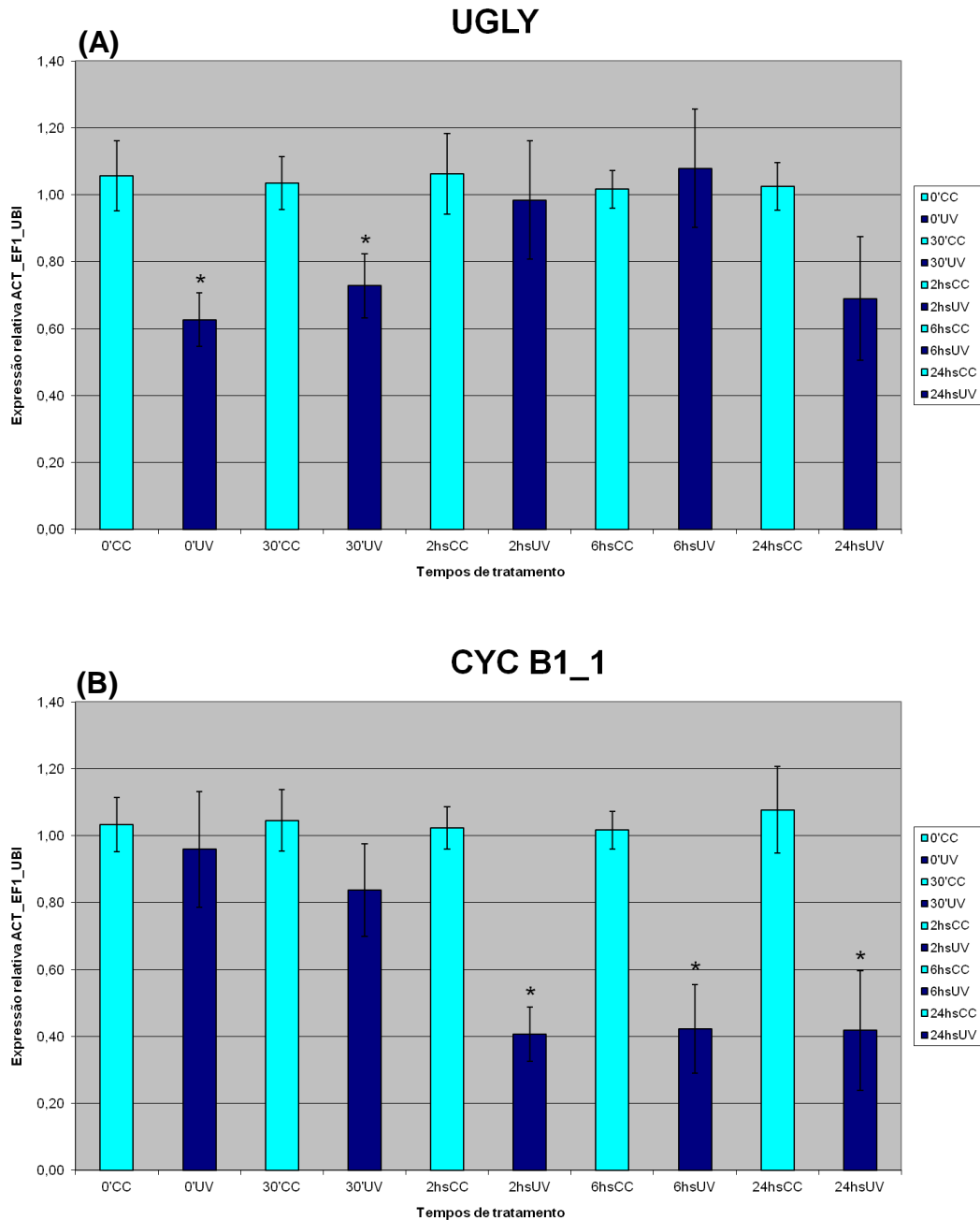
O padrão de expressão do gene da CYC B1\_1: para os tempos 0h e 30m a expressão não foi alterado pelo tratamento e se manteve em níveis iguais aos da situação controle. No entanto, nos tempos de coleta de 2h, 6h e 24h após o estresse, este gene apresentou expressão diminuída (Figura 7). Os genes CDKB2 e H4 (figura 8), tiveram diminuição da expressão relativa no tempo 0h, retorno a patamares normais no tempo 30min. No tempo 2h e 6h ocorre uma nova diminuição na quantidade de mRNA retornando a níveis controle no tempo 24h. Os genes CYC D4 e CYC D6 (figura 9) apresentaram padrões de expressão semelhantes, estando diminuídos nos tempos 2h e 6h e expressão não diferencial nos outros tempos.

---

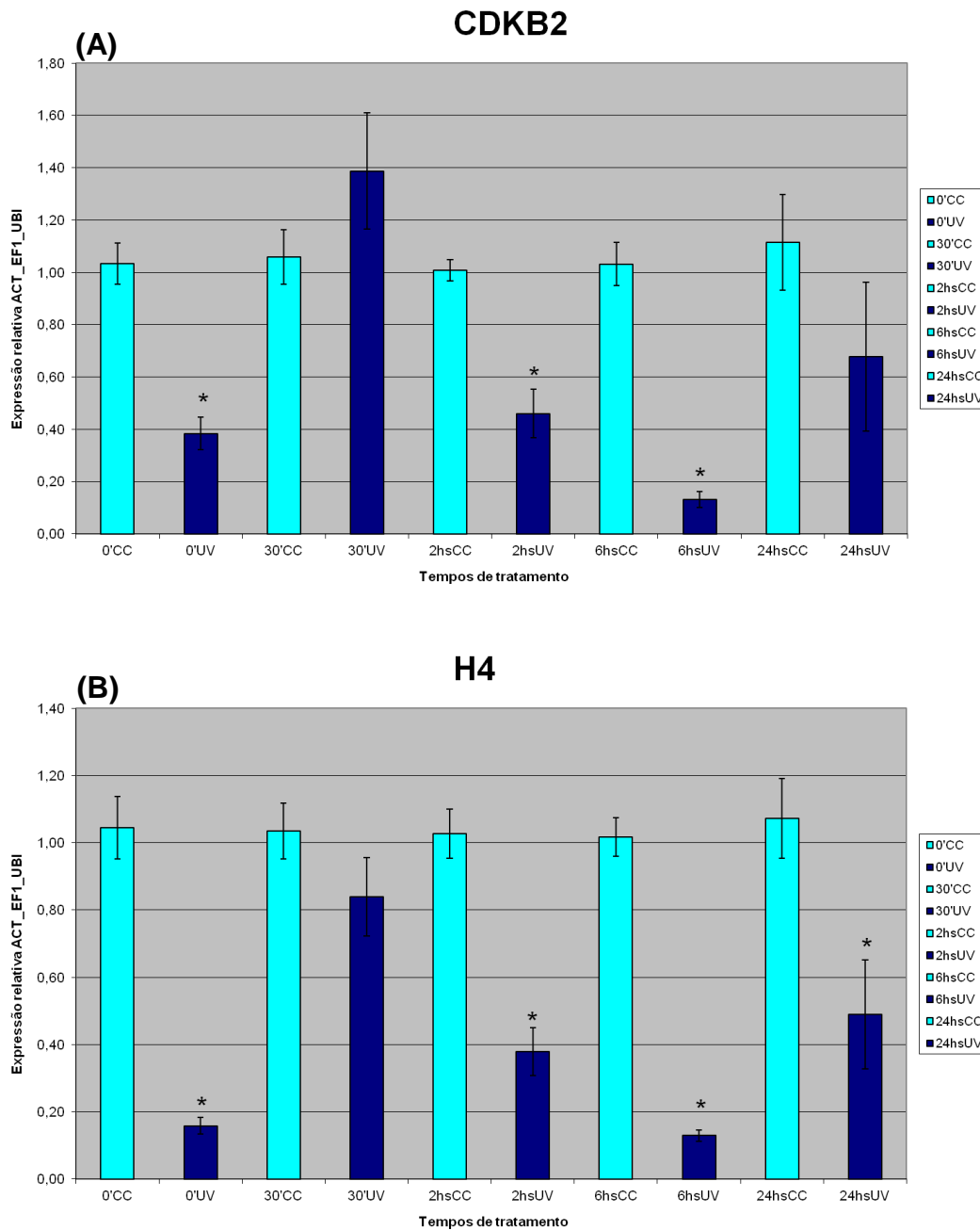
<sup>1</sup> Para níveis normais entenda-se níveis de expressão relativa que não desviam significativamente da situação controle.



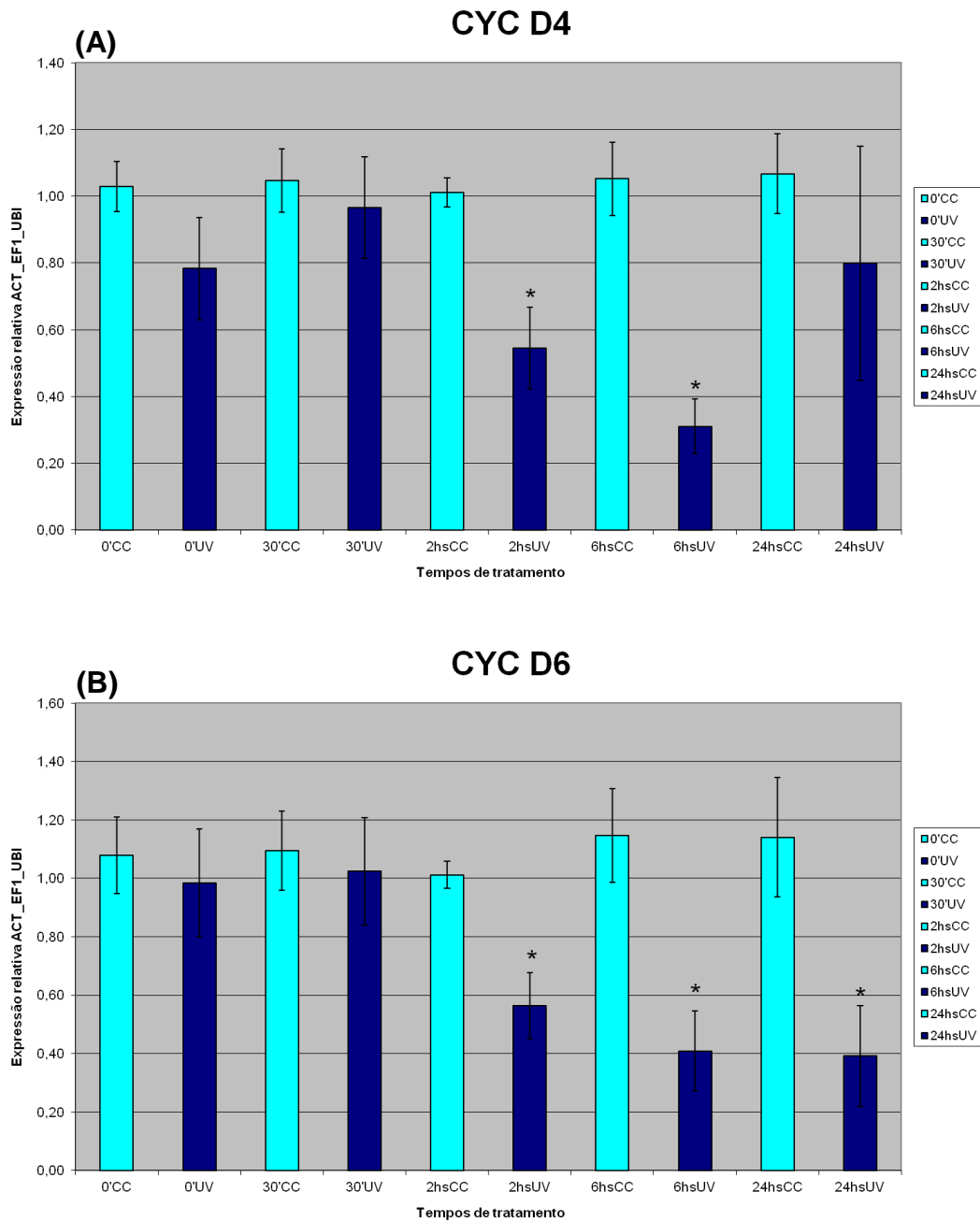
**Figura 6. Análises por RT-qPCR do nível de expressão de um gene de fotoliase (UVR3)(A) e um gene envolvido com a via de reparo de DNA por excisão de nucleotídeos (RAD1)(A) em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.**



**Figura 7. Análises por RT-qPCR do nível de expressão de um gene envolvido na rota de reparo ao DNA por excisão de bases (UGLY)(A) e um gene envolvido na progressão do ciclo celular (CYC B1\_1)(A) em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.**



**Figura 8. Análises por RT-qPCR do nível de expressão de dois genes envolvidos na progressão do ciclo celular: CDKB2 (A) e H4 (B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.



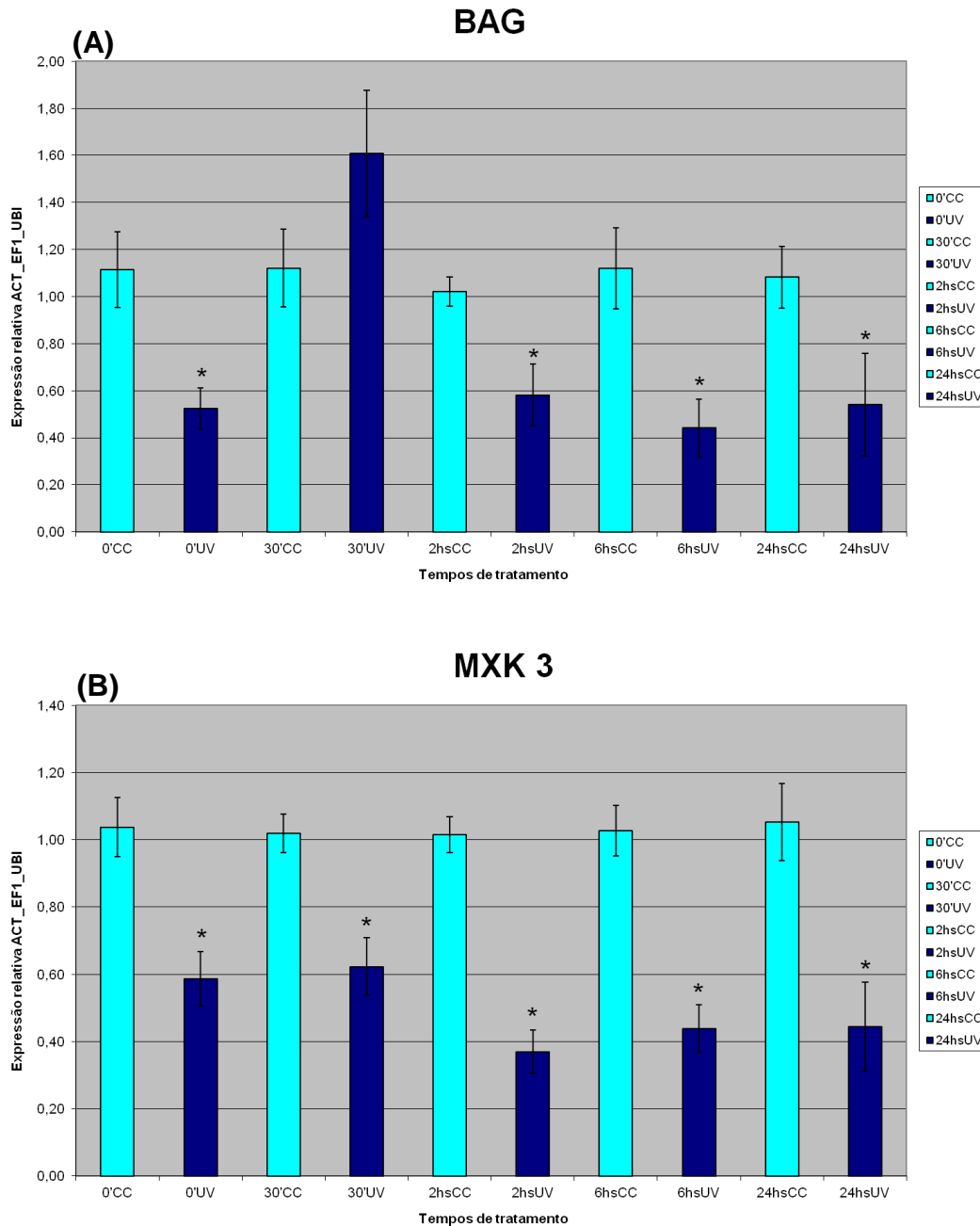
**Figura 9.** Análises por RT-qPCR do nível de expressão de dois genes envolvidos na progressão do ciclo celular: CYC D4 (A) e CYC D6 (B), em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.

Para os genes BAG e MXK3, envolvidos na rota metabólica da apoptose (figura 10), houve uma repressão ao longo de todos os tempos de tratamento, exceto no tempo 30min para o gene BAG, no qual os níveis ficaram iguais aos das plantas não tratadas.

E2F 08436 e E2F 08489 apresentaram cinética de resposta muito semelhante, com níveis de expressão similares ao controle nos tempos 0h e 30min e repressão nos tempos seguintes (Figura 11). O gene E2F 04611 possui uma resposta diferenciada, tendo níveis de expressão normais nos tempos 0h, 30min, 2h e 24h e um aumento significativo de aproximadamente 4 vezes no tempo de coleta de 6h (Figura 12). O Gene E2F 05530 teve sua expressão aumentada no tempo 30min em aproximadamente 3 vezes, enquanto nos tempos 2h e 24h seus níveis de expressão retornaram aos da situação controle (figura 12).

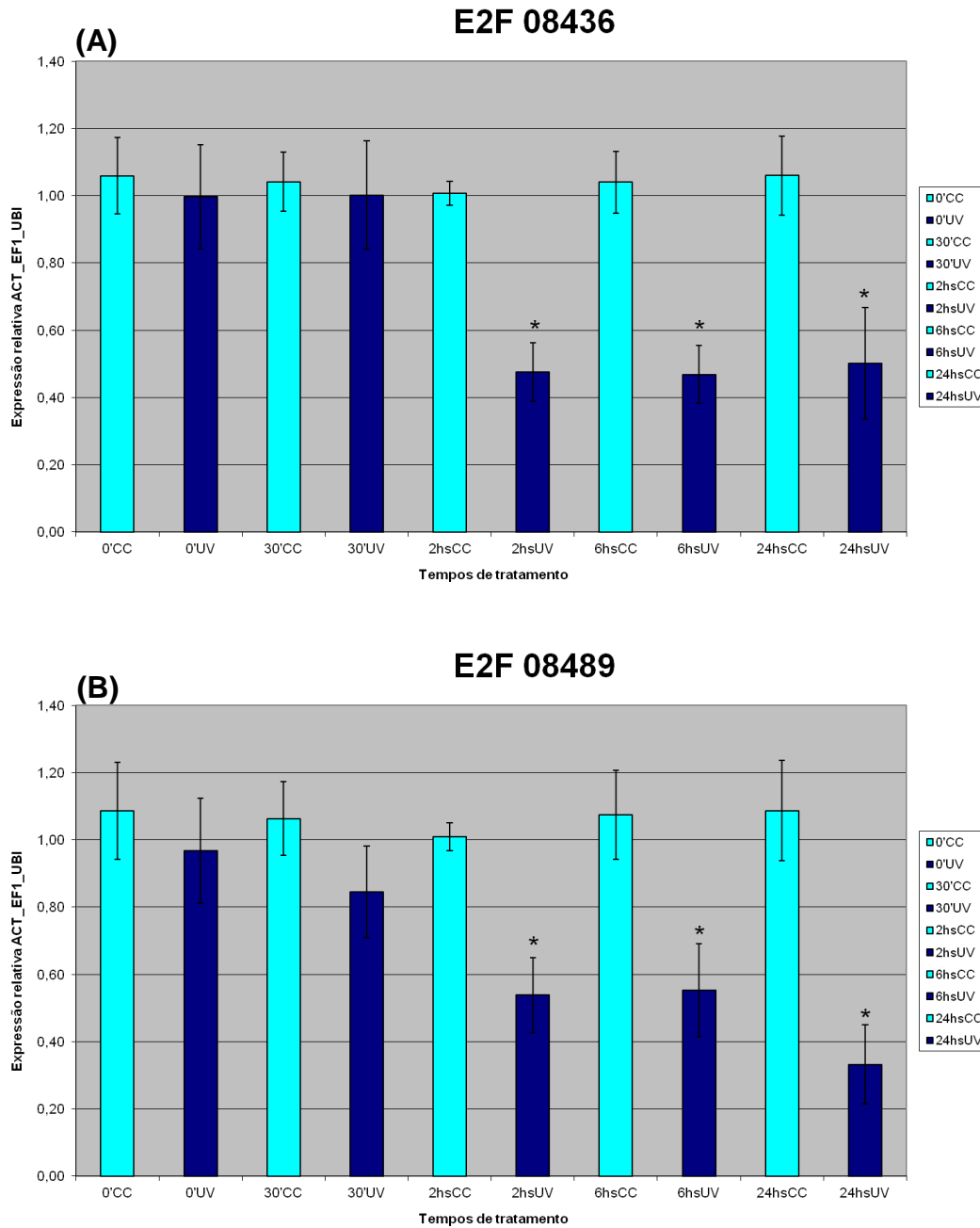
Para os genes de DEL, foram obtidos dois padrões de resposta. O gene DEL 05745 teve sua expressão aumentada nos 4 primeiros tempos de coleta, voltando a patamares normais no tempo 24h. Seu aumento de expressão variou de 2 (tempo 2h) a 4 vezes (tempo 6h). Para o gene E2F 06082, o aumento de expressão ocorreu apenas no tempo 30min. Nos tempos 0h e 6h a expressão foi reduzida, permanecendo nos mesmos níveis dos controles nos outros tempos de tratamento (figura 13).

Os genes correspondentes a proteínas DP testados apresentaram resultados semelhantes (figura 14). Ambos (DP 06098 e DP 02619) tiveram expressão reduzida nos tempos 2h e 6h, enquanto nos outros tempos os níveis de transcritos retornaram aos níveis do controle.

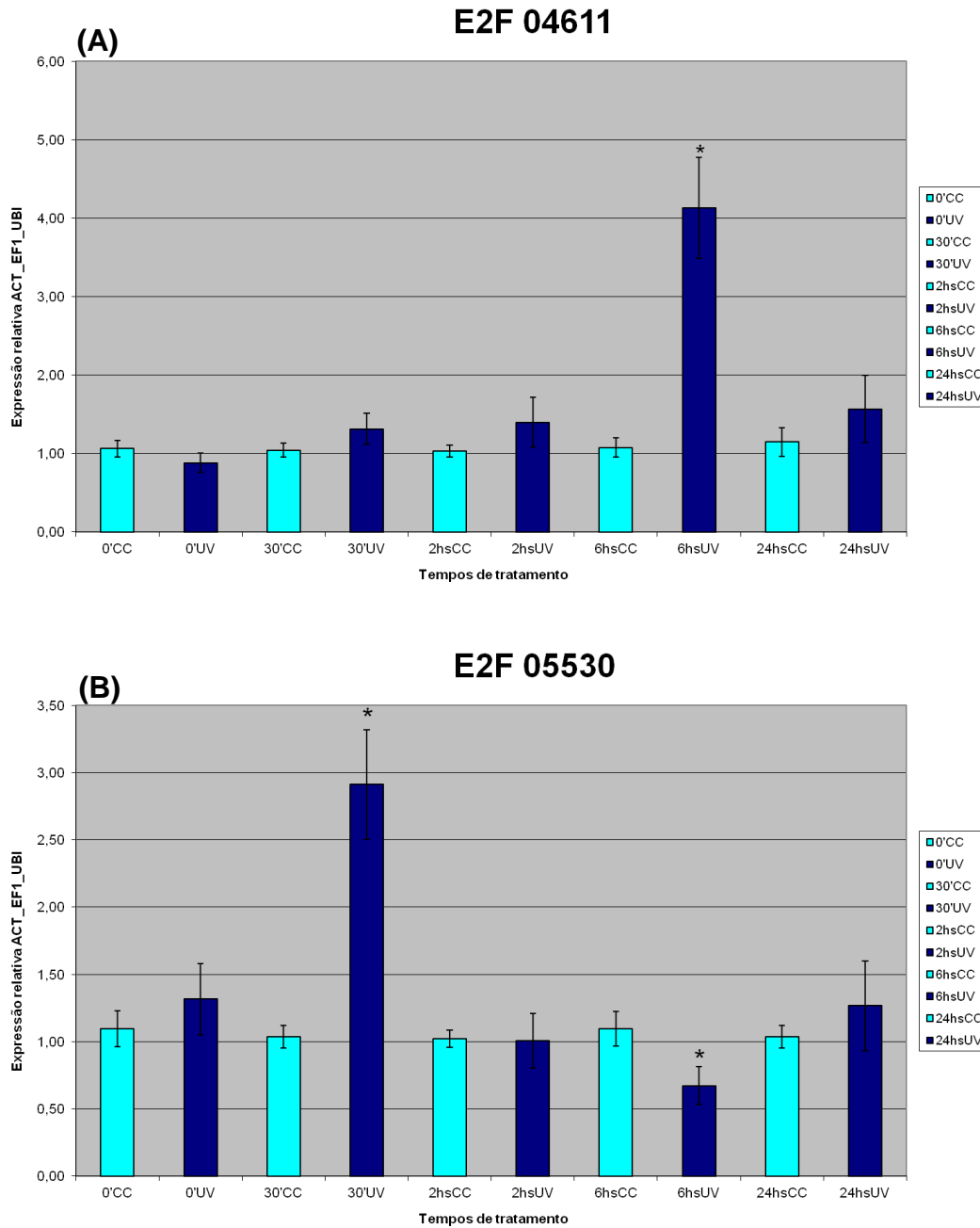


**Figura 10. Análises por RT-qPCR do nível de expressão de dois genes envolvidos com a morte celular programada (PCD): BAG (A) e MXK3 (B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.

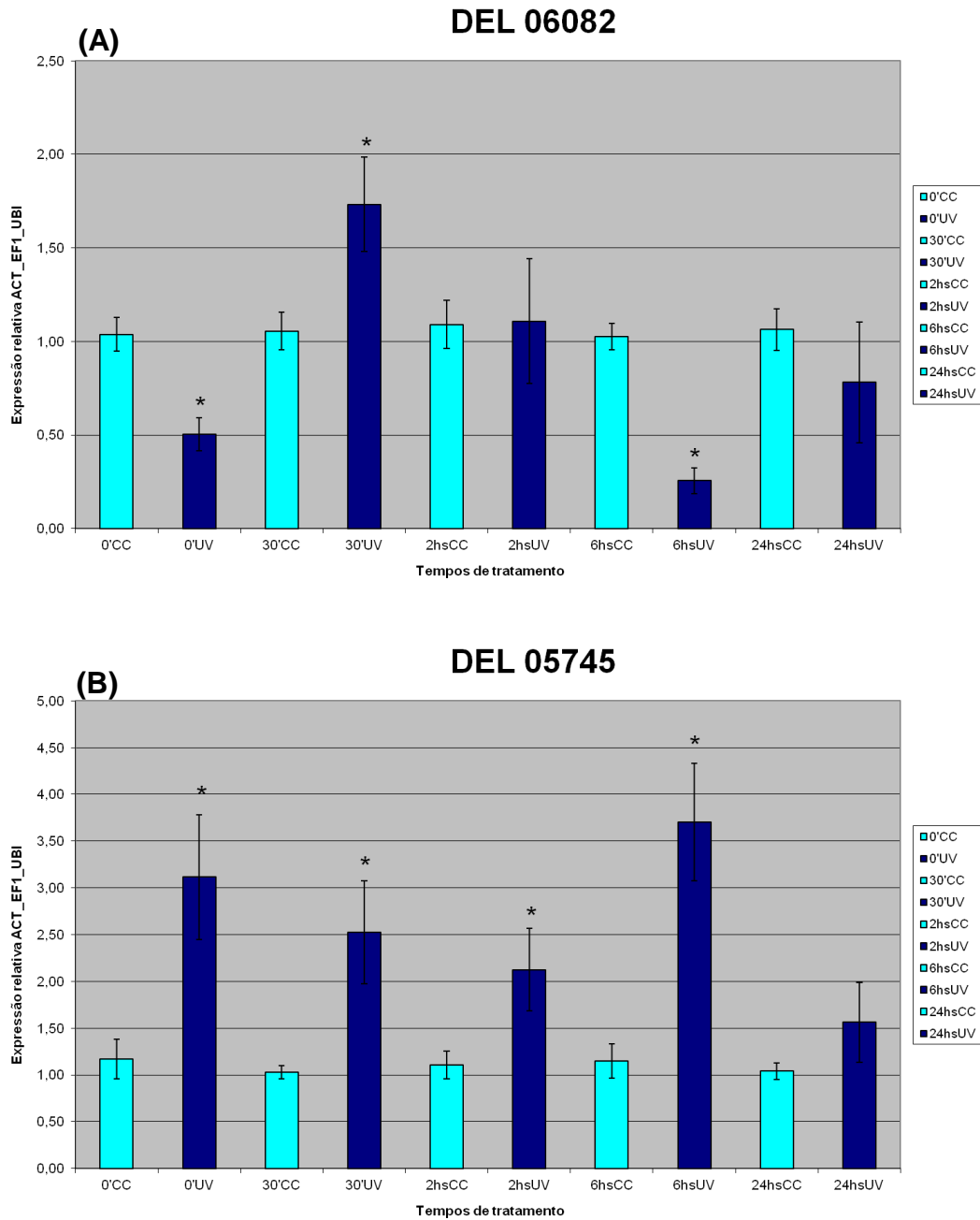




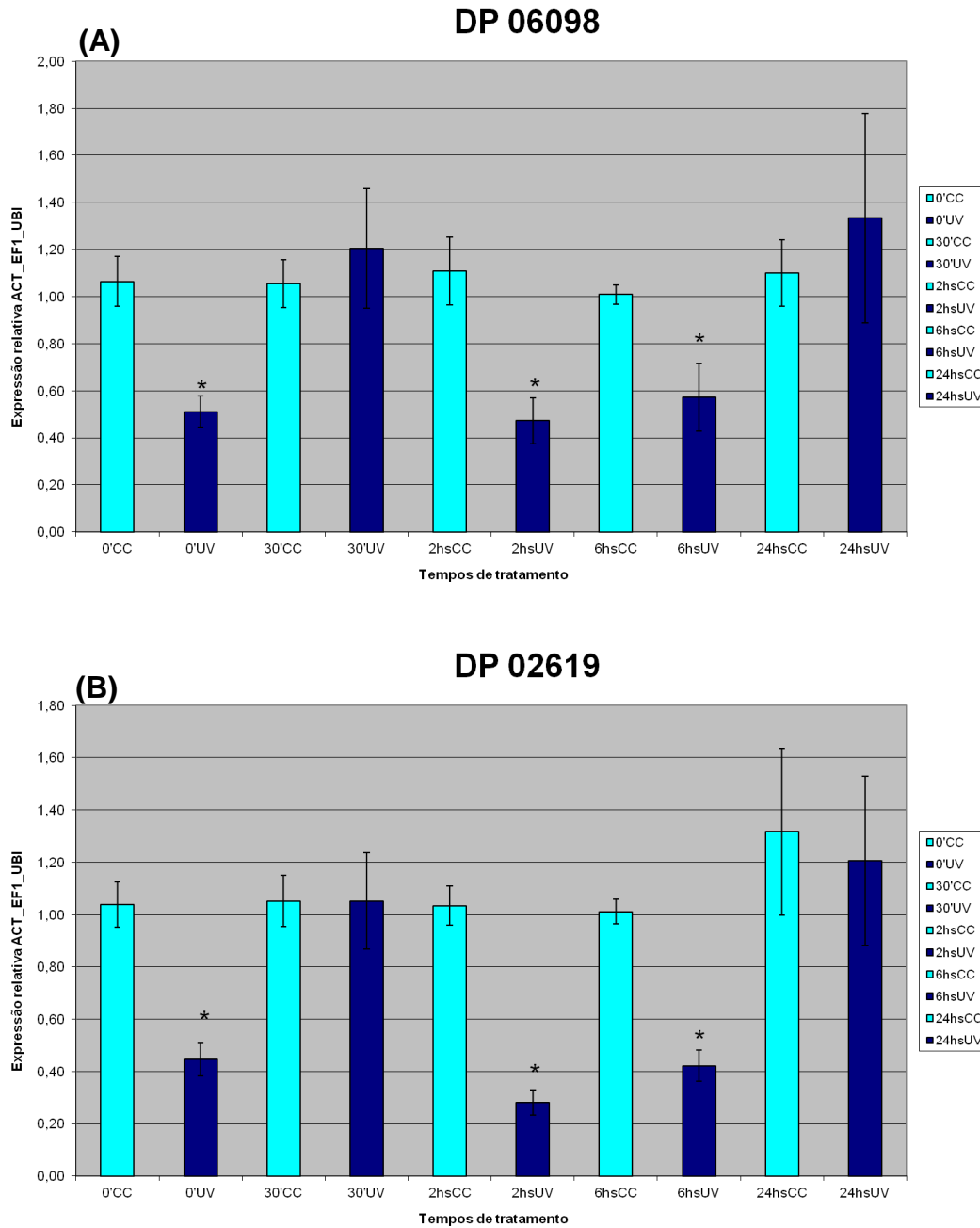
**Figura 11. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 08436 (A) e E2F 08489 (A), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.



**Figura 12. Análises por RT-qPCR do nível de expressão dos membros típicos da família de fatores transcricionais de E2F: E2F 04611 (A) e E2F 05530 (B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.



**Figura 13. Análises por RT-qPCR do nível de expressão dos membros atípicos da família de fatores transcricionais de E2F: DEL 06082 (A) e DEL 05745 (B), em resposta ao estresse com luz UV-B.** As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student



**Figura 14. Análises por RT-qPCR do nível de expressão de dois membros da família DP: DP 06098 (A) e DP 02619 (B), em resposta ao estresse com luz UV-B. As análises foram realizadas em plantas submetidas a luz ultravioleta por dois períodos de 4 horas com intervalo de 20 horas e coletadas em 0h, 30 minutos, 2, 6 e 24 horas após o estresse. O nível de expressão foi normalizado em relação ao nível de expressão dos genes *Act*, *Ef1* e *Ubi*. \* :  $p < 0,05$  no teste-t de Student.**

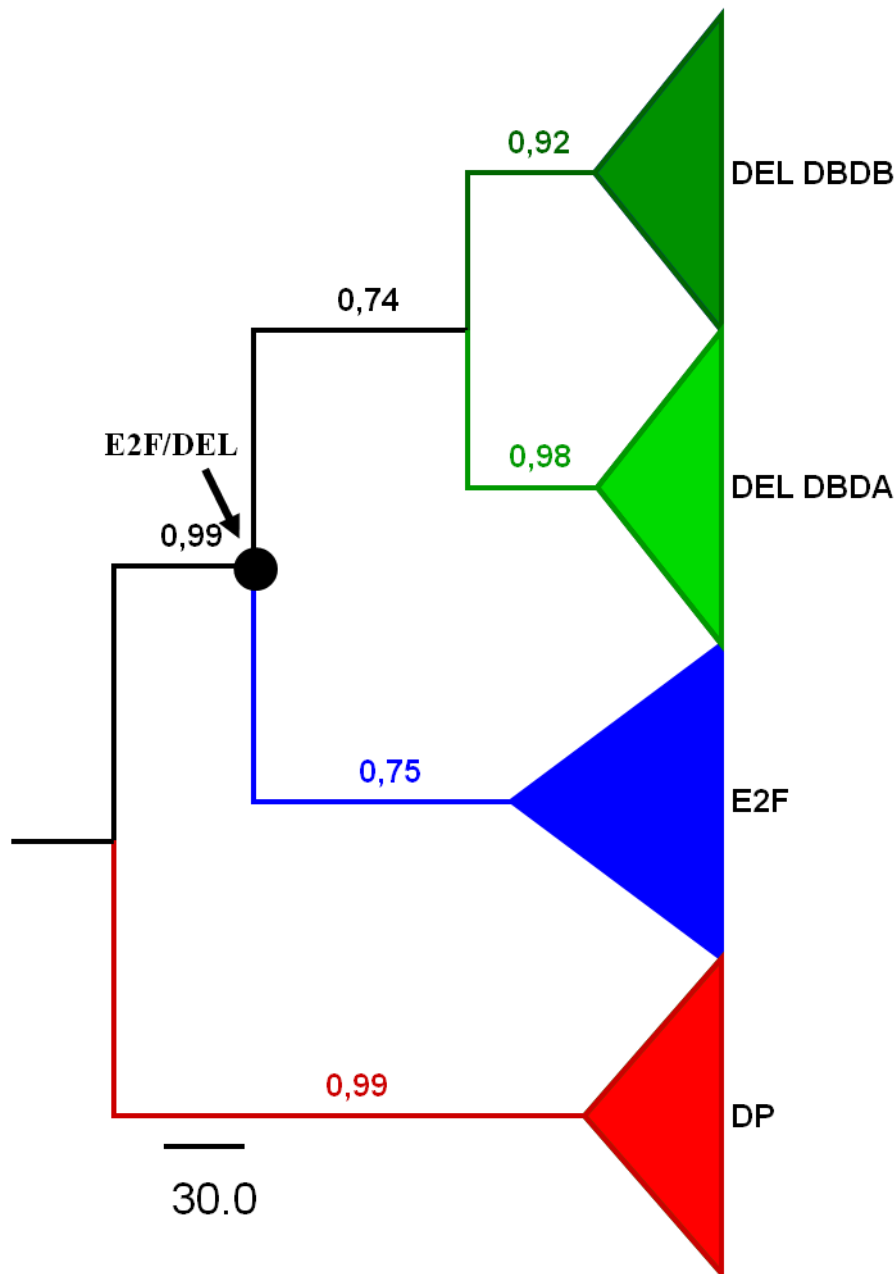
## 4.2 ANÁLISES FILOGENÉTICAS

As buscas por sequências E2F/DEL/DP permitiram a identificação de 246 sequências presentes em diferentes espécies que englobam desde amebas, coanoflagelados, algas, metazoários e plantas. São 23 plantas e 2 algas que estavam disponíveis nos bancos do phytozome, além de 1 alga, 1 ameba 1 coanoflagelado e 6 metazoários (representativos deste grupo) presentes no banco do NCBI. Todos estes dados estão presentes no anexo 1.

A árvore filogenética mostrada na figura 15 apresenta dois principais grupos, ambos com probabilidade posterior de 0.99. O primeiro grupo é formado por sequências de proteínas E2F/DEL enquanto o segundo reúne sequências de proteínas DP, mostrando uma clara separação da família de proteínas DP em relação aos membros da família de proteínas E2F/DEL. Dentro do grupo E2F/DEL, verificou-se a separação dessas proteínas em dois subgrupos, um contendo sequências E2F e o outro contendo sequências DEL, com probabilidade posterior de 0.75 e 0.74, respectivamente. As proteínas DEL foram divididas em dois grupos distintos, um identificado como DEL DBDA, contendo apenas as sequências presentes no primeiro domínio de ligação ao DNA das proteínas DEL e outro como DEL DBDB, por conter apenas sequências do segundo domínio de ligação ao DNA das proteínas DEL. A fim de entender melhor as relações evolutivas dentro dessas classes de proteínas de ligação ao DNA (E2F, DEL e DP), sequências proteicas de E2F, DEL e DP de diferentes espécies foram utilizadas para a reconstrução filogenética dentro de cada uma delas separadamente (Figuras 16, 17 e 18).

A figura 16 mostra a árvore consenso gerada para as proteínas E2F de animais, algas e plantas. Observou-se a formação de três clados distintos e bem suportados pela probabilidade posterior (100%). O primeiro agrupa as proteínas E2F de animais, o segundo contém as sequências de algas e o terceiro reúne as sequências de plantas. Este último foi dividido em dois grupos distintos (Figura 16), um com proteínas mais semelhantes à sequência E2FC de *Arabidopsis thaliana* e outro com proteínas semelhantes as outras duas sequências de E2Fs de *Arabidopsis*. Ambos os grupos contém tanto sequências de eudicotiledôneas como de monocotiledôneas, mostrando que eventos de duplicação dentro dessa classe de proteínas ocorreram cedo na divisão entre esses dois grupos taxonômicos. No grupo

que reúne as proteínas não semelhante a E2FC observou-se que as sequências de monocotiledôneas agruparam com as sequências de *Physcomitrella patens* e *Selaginella moellendorffii*, enquanto as sequências de eudicotiledôneas formaram um grupo separado.



**Figura 15. Análise filogenética A: construída por inferência bayesiana das sequência protéicas dos domínios de ligação ao DNA E2F/DP de algas e plantas.** A análise foi realizada no programa mrBayes utilizando-se 10 milhões de gerações. As probabilidades posteriores estão indicadas acima dos ramos. O círculo preto representa o clado formado por proteínas E2F e DEL de plantas. Os quatro clados principais foram colapsados para permitir uma melhor visão da topologia.

A árvore filogenética construída com sequências DEL de animais e plantas é mostrada na Figura 17. Essa análise foi realizada somente com sequências DEL de animais e plantas, pois não foram encontradas proteínas DEL em algas com genomas disponíveis. Em plantas a espécie *Physcomitrella patens* foi o organismo mais basal apresentando estas proteínas. Nessa análise observou-se a formação de dois clados principais, um contendo proteínas DEL de plantas e outro de animais. Dentro do clado das plantas, verificou-se a separação das sequências DEL pertencentes às espécies de monocotiledôneas e eudicotiledôneas daquelas de *Physcomitrella patens*. As Sequências DEL de espécies de monocotiledôneas agruparam separadamente das eudicotiledôneas, mostrando que essas proteínas divergiram antes da separação desses dois grupos taxonômicos (Figura 17).

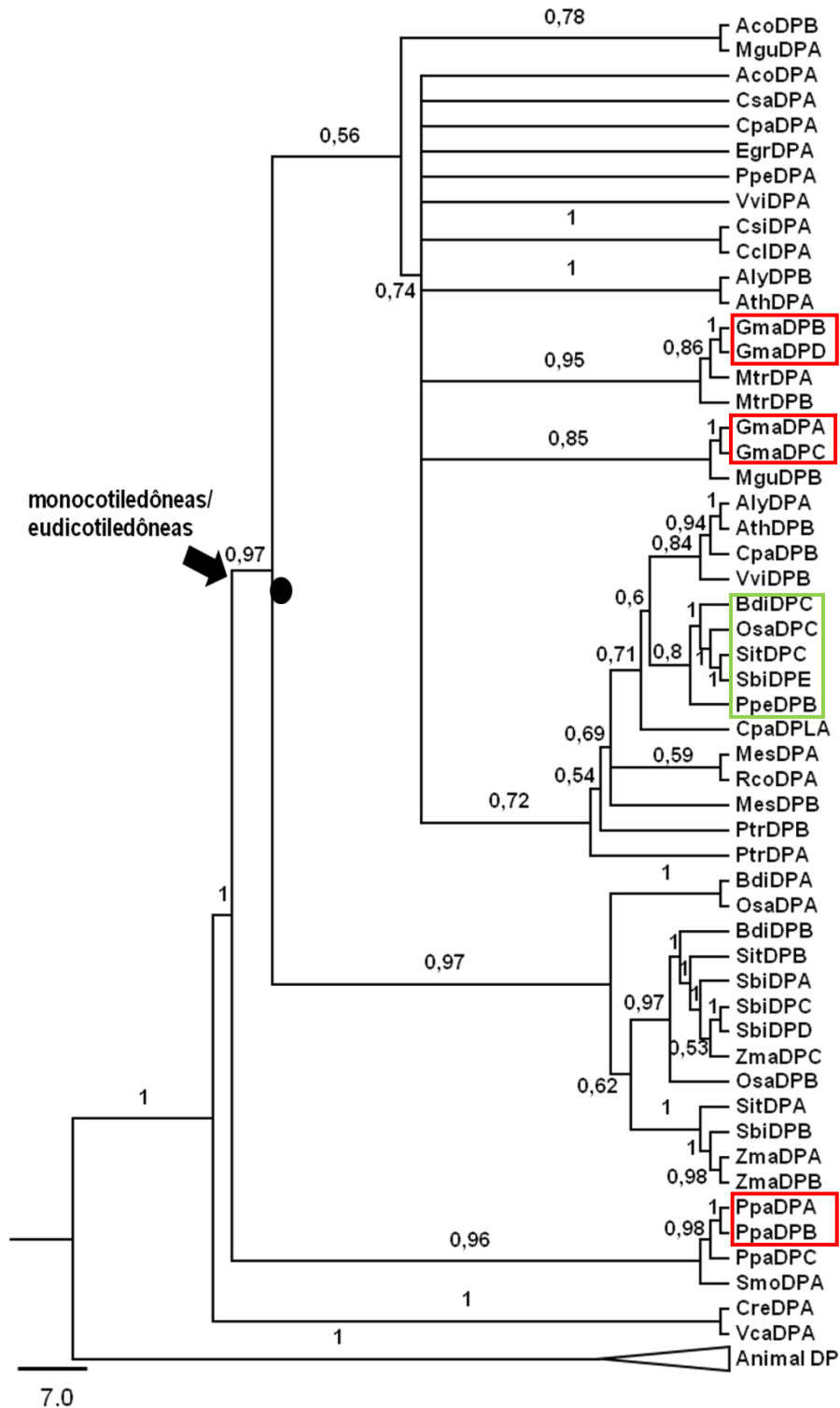
Na família gênica das proteínas DP, obteve-se um padrão semelhante ao encontrado para as proteínas E2F. Observou-se a formação de um grupo contendo sequências DP de animais, outro grupo reunindo sequências de algas, e um com sequências de plantas. Dentro do clado que reúne mono e eudicotiledoneas, observa-se que sequências de monocotiledoneas agruparam separadamente das dicotiledôneas (figura 18), exceto algumas sequencias (BdiDPC, BdiDPE, PpeDPB, OsaDPC e SitDPC), as quais agruparam junto com as sequências das espécies de eudicotiledôneas (figura 18).



Figura 16. Árvore filogenética B: construída por inferência bayesiana utilizando-se seqüências E2F de algas, plantas e animais. A análise foi realizada no programa MrBayes utilizando-se 30 milhões de gerações. Os valores de probabilidade posterior estão acima dos ramos. Caixas vermelhas: prováveis duplicações recentes.







**Figura 18. Árvore filogenética D: construída por inferência bayesiana utilizando-se sequências DP de algas, plantas e algumas espécies de animais.** A análise foi realizada no programa MrBayes utilizando-se 20 milhões de gerações. Os valores de probabilidade posterior estão acima dos ramos. Dentro das caixas vermelhas estão indicadas as prováveis duplicações recentes e dentro das caixas verdes está o grupo de monocotiledôneas que se agrupou junto às sequências de dicotiledôneas.

## 5 DISCUSSÃO

Os fatores transcricionais E2F são importantes reguladores de diversos processos biológicos e vêm sendo extensivamente estudados. Esta família gênica está presente desde os eucariotos mais simples até os organismos mais complexos, como plantas e vertebrados. Porém, para algumas vias regulatórias exercidas por estas proteínas ainda existe uma falta de conhecimento. O envolvimento dos E2F com a resposta ao dano de DNA é uma descoberta relativamente recente e ainda pouco se sabe a respeito dela para as plantas.

Na tentativa de aumentar o conhecimento esta família gênica Vandepoele, *et al* identificaram os genes-alvo de E2F de *Arabidopsis* em uma escala genômica comparando o transcriptoma de plantas selvagem e plantas expressando ectopicamente os genes E2Fa-DPa usando microarranjos contendo 22.750 genes anotados do genoma *Arabidopsis thaliana*. Entre os genes diferencialmente expressos, os autores buscaram na sequência um sítio de ligação a E2F e possíveis ortólogos em arroz. Nesse estudo foram encontrados 70 genes com estas características. Dentre os genes identificados para a função de reparo ao dano de DNA se destacam 4 sequências contendo expressão aumentada e presentes em ambos os organismos. São eles: Uracil DNA glycosilase, com aumento de expressão relativa aproximado de 17 vezes, MSH6-1, uma proteína de reparo de pareamentos imperfeitos, com expressão aumentada de aproximadamente 16 vezes, uma proteína semelhante a RAD-1 com um aumento de 10 vezes e uma 6-4 fotoliase (UVR3) com aumento de quase 7 vezes (Vandepoele *et al.*, 2005). Estes quatro genes foram os escolhidos como marcadores de ativação da rota de reparo ao DNA para os experimentos com UV-B que foram realizados no presente trabalho. A hipótese de trabalho foi que se algum destes genes apresentasse expressão aumentada seria um indicativo de que algum mecanismo de reparo estaria sendo ativado, o que tornaria possível verificar o envolvimento dos genes E2F nessa resposta.

A fim de descartar que a ativação dos genes E2F fosse consequência de outra função que esta família protéica possua, foram testados genes que sabidamente estão envolvidos com a progressão de ciclo celular e morte celular programada (PCD). Os genes testados foram: CycB1\_1 e CDKB2, relacionados

com a passagem da fase G2 para M do ciclo celular (Lee *et al.*, 2003; La *et al.*, 2006), os genes *CYC D4* e *CYCD6*, que são ciclinas específicas da fase G1 do ciclo celular (La *et al.*, 2006) e os genes *MXK3* (Phytozome) e *BAG* (NCBI) envolvidos na morte celular programada.

Os resultados obtidos no experimento sem aclimação para os marcadores de dano de DNA *UVR3* e *RAD1* demonstraram que, a princípio, neste desenho experimental, o estresse não produziu o efeito esperado. Após o experimento as plantas estressadas foram mantidas em casa de vegetação para observação de seus fenótipos. Logo nos primeiros dias observou-se uma leve clorose e necrose nas folhas. Após 7 dias quase todas as plantas estavam mortas. Esse efeito sugeriu que o nível de estresse aplicado foi extremo e que a planta não estaria respondendo ao estresse de maneira eficiente, tendo os genes sido reprimidos ao invés de ativados. Para tentar elucidar essa divergência nos dados foram testados os genes para morte celular programada e todos os genes da família *E2F*. O que se verificou foi uma repressão de todos os genes em algum momento do experimento. Isso reforçou a hipótese de que o estresse havia sido muito forte e que teria levado essas plantas a uma repressão geral da maquinaria celular com posterior morte.

Em revisão publicada em 2007 Bruce *et al.* descreveram que a exposição a um agente primário pode ativar um gene ou grupo de genes. Ao fim deste estresse ao invés de reverter para um estado de silenciamento transcricional, uma vez que o estímulo é removido, uma marca epigenética (como a acetilação da histona) pode ficar, mantendo uma região em um estado permissivo provavelmente facilitando uma resposta mais rápida e potente a ataques subsequentes. Eles verificaram que experimentos com diversas plantas, como por exemplo *Arabidopsis thaliana*, *Nicotiana Tabacum*, *Vitis vinifera*, entre outras, mostrou que os efeitos de um estresse primário podem durar pelo menos alguns dias variando de acordo com o estresse e com a planta utilizada (Bruce *et al.*, 2007).

Para o experimento com aclimação um estresse prévio com 4hrs de luz UV-B foi realizado a fim de potencializar a resposta e evitar a repressão de todos os genes e morte celular. As plantas para este segundo experimento possuíam 7

semanas pós-germinação. Após o estresse primário elas foram deixadas por 7 dias em condição controle e então o estresse definitivo foi realizado.

Conforme verificado nos resultados, foi observado um aumento na expressão de genes envolvidos em mecanismos de reparo, UVR3 e RAD1 indicando assim que este mecanismo estava sendo ativado. Observou-se que essa resposta foi rápida, estando presente 30 minutos após a segunda irradiação. Os resultados para os genes de progressão do ciclo celular ou de morte celular programada indicaram um estado de repressão ou uma quantidade de mRNA igual as plantas controle. Isto mostra que tanto o ciclo celular quanto a morte celular programada estão inativas, em níveis equivalentes às plantas controle, ou com a rota sendo reprimida. Portanto pode-se afirmar que os dados de indução referentes aos genes E2F são relacionados com as respostas ao dano no DNA e com o processo de reparo do dano, e não a respostas que os genes E2F tradicionalmente possam estar ligados (como progressão do ciclo celular).

Na avaliação da expressão dos genes E2F no experimento com aclimatação obteve-se uma resposta diferencial específica para cada gene. Para os genes E2F 08436 e E2F08489 houve uma resposta inicial igual a situação controle e 2 horas após o estresse uma repressão que se manteve até 24 horas após o estresse. Para o gene E2F05530 houve uma ativação deste gene 30 minutos após o estresse e posterior queda a níveis da situação controle. E2F04611 foi o gene que respondeu com aumento de sua expressão 6 horas após o estresse e nos outros tempos estava semelhante aos níveis do controle.

Uma resposta diferente para cada gene da família transcricional E2F já era esperada, visto que na literatura um padrão semelhante já fora relatado em vários trabalhos. Em um deles os autores avaliaram a expressão da família de fatores de transcrição E2F em linhagens celulares humanas tratadas com diferentes agentes nocivos ao DNA (doxorubicina, etoposida, e neocarzinostatina). Neste experimento E2F1 foi induzido por todas as três drogas. Observou-se que a expressão de E2F3a também foi induzida pela doxorubicina e etoposida, enquanto que os níveis de proteína E2F3b diminuiu em resposta a estes medicamentos. A droga radiomimética neocarzinostatina não reproduziu os níveis alterados de E2F3a enquanto reduziu a expressão de E2F3b. Em

contraste, nenhum dos tratamentos quimioterápicos alterou a expressão de E2F4 e E2F5 (Martinez *et al.*, 2009).

Os DEL de arroz, conforme indicado nos resultados, tiveram aumento significativo da expressão nos tempos 0h, 30min, 2h e 6h para o gene DEL05745 e no tempo 30min para DEL06082. Estes resultados para proteínas repressoras não difere do relatado na literatura. Os resultados de um estudo publicado em 2008 indicam que E2F7 e E2F8 (os representantes DEL em humanos) têm um papel importante na regulação da atividade de outros E2F durante a resposta a danos no DNA. Neste mesmo estudo os autores demonstram que E2F7 e E2F8 existem como um complexo sobre o promotor E2F1, e este tende a ser mediado através da capacidade de E2F7 e E2F8 para reprimir a transcrição de E2F1 (Zalmas *et al.*, 2008). Portanto além de competirem diretamente pelas sequencias promotoras dos genes alvo os E2F podem regular uns aos outros.

Além da expressão dos genes da família E2F, foi verificado o padrão de expressão dos genes da família DP. Dos três identificados em arroz só conseguimos avaliar o perfil de expressão de dois deles, estes possuem um padrão de expressão muito semelhante aos E2F 08436 e E2F 08489 o que pode indicar que estes são seus parceiros de dimerização.

Um panorama dos resultados da expressão obtidas nesse trabalho está representado esquematicamente na figura 19. Nesta figura os aumentos de expressão são mostrados na cor vermelha, e a redução da expressão na cor azul. As células em branco indicam que não houve alteração significativa neste ponto de coleta.

Para suprir a falta de conhecimento a respeito da história evolutiva da família gênica dos E2F em plantas foi realizado um estudo filogenético com as sequências E2F, DEL e DP obtidas em bancos de dados públicos. Para animais foi recentemente publicado um trabalho que descreve a história evolutiva dessa família. Estes autores mostraram onde possivelmente ocorreram eventos de duplicação partindo-se de um provável eucarioto ancestral. Para plantas apenas duas espécies foram utilizadas em seu estudo, apenas para indicar que esta família gênica está também presente em plantas (Cao *et al.*, 2010).

A análise filogenética realizada com as famílias gênicas de proteínas DP e E2F/DEL de plantas mostrou a separação desses dois grupos em clados bem

suportados mostrando que são duas famílias gênicas separadas e corroborando com o que foi descrito para animais (Cao *et al.*, 2010).

	1º EXP			2º EXP				
Loci E2F/DEL/DP	0h	6h	24h	0h	30 min	2h	6h	24h
E2F 04611								
E2F 05530								
E2F 08436								
E2F 08489								
DEL 06082	nt	nt	nt					
DEL 05745	nt	nt	nt					
DP 02619	nt	nt	nt					
DP 06098	nt	nt	nt					
Loci Genes reparo								
UGLY	nt	nt	nt					
RAD1								
UVR3								
Loci genes ciclo celular								
CYCB1_1	nt	nt	nt					
CDKB2	nt	nt	nt					
CYC D4	nt	nt	nt					
CYC D6	nt	nt	nt					
H4	nt	nt	nt					
Loci genes apoptose								
BAG								
MXK3								

**Figura 19. Expressão relativa dos genes em função do tempo após estresse com luz UV-B sob plantas de arroz.** No experimento sem aclimação houve apenas uma irradiação de 4 horas nas plantas. No experimento com aclimação foram realizadas duas irradiações de 4 horas com um intervalo de 20 horas. Neste último houve também um estresse primário uma semana antes do tratamento final. Os quadros em azul indicam uma redução na expressão relativa, as células em vermelho um aumento, as células em branco nenhuma mudança significativa foi observada e as letras “nt” indicam que a expressão destes genes não foram testados nesta condição. Todos os valores em azul ou vermelho foram avaliados estatisticamente com o teste-t de Student e possuem valor de  $p < 0,05$ .

O objetivo inicial da análise filogenética A (figura 15) foi identificar e caracterizar as famílias gênicas E2F/DEL e DP. As proteínas DEL são formadas por dois domínios de ligação: DBDA (DNA Binding Domain A) e DBDB (DNA Binding Domain B) (Kosugi & Ohashi, 2002), os quais formaram grupos separados na análise filogenética (Figura 15). O grupo identificado como DEL DBDA é composto apenas por sequências que estavam presentes no primeiro domínio de cada proteína DEL, e o grupo DEL DBDB possui apenas as sequências do segundo domínio das proteínas. Outro dado interessante é que os dois domínios de ligação ao DNA presentes nas proteínas DEL são mais semelhantes entre si, formando um grupo distinto bem suportado por um valor posterior de 0,74, do que o domínio presente nas proteínas E2F, mas ficando em um grupo separado (figura 15) dos domínios de ligação das proteínas DP. Esses resultados e os dados de Cao, Peng, et al sugerem que a proteína ancestral possa ter sofrido uma duplicação de seu domínio de ligação ao DNA e que a partir daí surgiu a família de proteínas DEL. Uma possibilidade é que uma proteína E2F seja o ancestral das proteínas DEL devido a características funcionais semelhantes dessas duas proteínas, como por exemplo, a competição pelo mesmo sítio de ligação ao DNA (Kosugi & Ohashi, 2002), funções nas mesmas rotas bioquímicas (revisado por Lammens, *et al*, 2009), mas para confirmar esta hipótese mais estudos filogenéticos e funcionais são necessários.

A análise do grupo E2F apresentou dois grupos distintos: Os E2F que são mais semelhantes com a proteína AtE2FC e o grupo das proteínas semelhantes a AtE2FA/B (figura 16). Conforme Del Pozo, et al 2002, as proteínas AtE2FC não possuem o domínio de transativação, o que faz com que elas atuem como repressores de ativação ao invés de ativadores como os outros dois E2Fs de Arabidopsis. No mesmo grupo que AtE2FC está a proteína OsE2FD. Esta proteína nos estudos de expressão, realizados neste trabalho, ela é identificada como E2F05530 podendo este também ser um repressor da expressão gênica em conjunto com os DEL de arroz. Os dados de expressão para o experimento com aclimação dão indícios de que essa hipótese é verdadeira pois no tempo 30 min após o estresse ela possui um aumento de expressão e nos tempos posteriores a esse, 2h, 6h e 24h, dois dos três E2F ativadores (E2F 08436 e E2F 08489) possuem sua expressão diminuída.



Na busca realizada por proteínas DEL no banco de dados nota-se a ausência destas proteínas nas espécies de algas *Chlamydomonas reinhardtii* e *Volvox carteri* e na licófito *Selaginella moellendorffii*. Peng et al 2010 em seu estudo filogenético, também notou a ausência destas proteínas em alguns organismos, porém encontrou esta proteína no protozoário *Monosiga brevicollis*. A sugestão destes autores é de que esta proteína tenha surgido antes da separação de protistas, animais e plantas e especulam que sua presença não seja tão importante quanto os E2Fs e os DPs presentes em todos os grupos. A análise filogenética das proteínas DEL mostrou com clareza a separação do grupo de sequências de espécies animais daquelas de plantas. As sequências de plantas foram divididas em dois grupos: o primeiro é formado por sequências de espécies de plantas avasculares (musgos) e o segundo por sequências de espécies de plantas vasculares (monocotiledôneas e eudicotiledôneas), sendo observado a separação entre monocotiledôneas e eudicotiledôneas em subgrupos separados (figura 17).

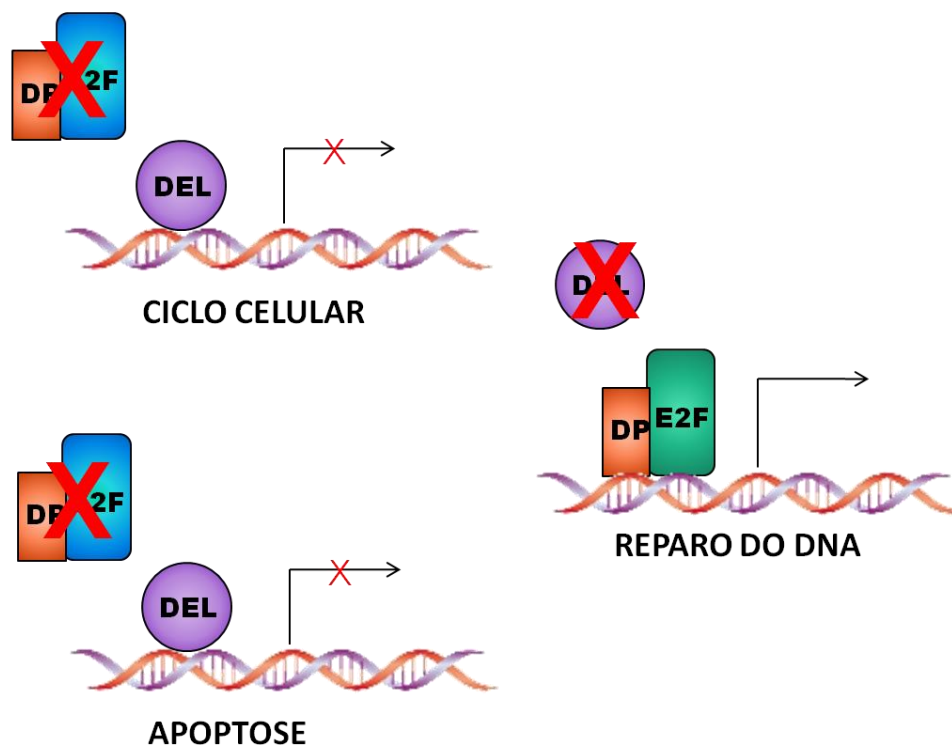
Na análise das proteínas DP, representada pela análise D (figura 18) estão destacadas em verde as sequências BdiDPC, BdiDPE, PpeDPB OsaDPC e SitDPC que pertencem as monocotiledôneas. Este grupo está separado do outro grupo maior de proteínas de monocotiledôneas e está inserido junto com as proteínas pertencentes às eudicotiledôneas. Uma possível explicação é a ocorrência de eventos de duplicação posterior a divergência entre monocotiledôneas e eudicotiledôneas, enquanto a origem das outras sequências provavelmente tenha ocorrido antes, com a emergência desses dois grupos taxonômicos. Para os animais Cao, Peng et al 2010 também caracterizaram duplicações antigas, porém nenhuma duplicação recente foi encontrada.

Em alguns grupos taxonômicos observou-se que as sequências de um mesmo organismo não se agrupam no mesmo clado, indicando a presença de prováveis eventos de duplicações mais antigos durante a evolução.

Diversos eventos de duplicação foram observados dentro de cada uma dessas classes de proteínas, algumas tendo ocorrido mais cedo e outras mais recentes no decorrer da história evolutiva dos eucariotos. Alguns exemplos de duplicação recentes podem ser observados nos genomas da soja (*Glycine max*), *Populus trichocarpa*, *Manihot esculenta* e *Zea mays* (Figuras 16, 17 e 18). A soja

possui duplicações recentes em seu genoma que já foram relatadas (Shoemaker *et al.*, 2006).

Levando em conta os resultados obtidos nos experimentos de estresse genotóxico, associado aos dados da literatura, um esquema representando o modelo de atuação dos fatores transcricionais E2F em plantas sob o estresse de luz UV-B é proposto na figura 20. Este modelo propõe que aos genes alvo dos E2F, envolvidos com as rotas de progressão do ciclo celular e apoptose, estejam ligadas proteínas DEL, para que estas reprimam estes processos. Com os resultados obtidos na filogenia pode-se adicionar mais uma etapa regulatória nesta hipótese, onde o E2F 05530, possível repressor, estaria auxiliando os genes DEL na tarefa de reprimir os genes envolvidos no ciclo celular e apoptose. Nos genes alvo envolvidos com o reparo do dano de DNA estaria ligado um gene E2F específico, provavelmente E2F 04611, para que esta rota estivesse ativa.



**Figura 20. Representação esquemática do modelo de atuação dos genes da família E2F em resposta a danos causados pela luz UV-B. Os E2Fs representados em na cor azul são específicos a ciclo e apoptose (E2F 08489 e E2F 08436) O E2F representado em verde é específico a reparo ao dano de DNA (E2F 04611).**

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

- Ballare C L, Caldwell M M, Flint S D, Robinson S A and Bornman J F. 2011.** Effects of solar ultraviolet radiation on terrestrial ecosystems. Patterns, mechanisms, and interactions with climate change. *Photochem Photobiol Sci*, **10**:226-241.
- Blais A and Dynlacht B D. 2004.** Hitting their targets: an emerging picture of E2F and cell cycle control. *Curr Opin Genet Dev*, **14**:527-532.
- Bruce T J A, Matthes M C, Napier J A and Pickett J A. 2007.** Stressful "memories" of plants: Evidence and possible mechanisms. *Plant Science*, **173**:603-608.
- Cao L, Peng B, Yao L, Zhang X, Sun K, Yang X and Yu L. 2010.** The ancient function of RB-E2F pathway: insights from its evolutionary history. *Biol Direct*, **5**:55.
- de Jager S M, Menges M, Bauer U M and Murra J A. 2001.** Arabidopsis E2F1 binds a sequence present in the promoter of S-phase-regulated gene AtCDC6 and is a member of a multigene family with differential activities. *Plant Mol Biol*, **47**:555-568.
- Dhaliwhal G.S A R. 1999.** *Stresses in agroecosystems: concepts and approaches*: Ajay Verma Commonwealth Publishers.
- Dyson N. 1998.** The regulation of E2F by pRB-family proteins. *Genes Dev*, **12**:2245-2262.
- Edgar R C. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, **32**:1792-1797.
- EMBRAPA. 2007.** Empresa Brasileira de Pesquisa Agropecuária.
- FAO. 2012.** Food and Agriculture Organization.
- Gale M D and Devos K M. 1998.** Comparative genetics in the grasses. *Proc Natl Acad Sci U S A*, **95**:1971-1974.
- Goff S A, Ricke D, Lan T H, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange B M, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W L, Chen L, Cooper B, Park S, Wood T C, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller R M, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A and Briggs S. 2002.** A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science*, **296**:92-100.
- Goodstein D M, Shu S, Howson R, Neupane R, Hayes R D, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N and Rokhsar D S. 2012.** Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*, **40**:D1178-1186.

- Inze D. 2005.** Green light for the cell cycle. *EMBO J*, **24**:657-662.
- Kosugi S and Ohashi Y. 2002.** E2Ls, E2F-like repressors of Arabidopsis that bind to E2F sites in a monomeric form. *J Biol Chem*, **277**:16553-16558.
- La H, Li J, Ji Z, Cheng Y, Li X, Jiang S, Venkatesh P N and Ramachandran S. 2006.** Genome-wide analysis of cyclin family in rice (*Oryza Sativa* L.). *Mol Genet Genomics*, **275**:374-386.
- Lammens T, Li J, Leone G and De Veylder L. 2009.** Atypical E2Fs: new players in the E2F transcription factor family. *Trends Cell Biol*, **19**:111-118.
- Lario L D, Ramirez-Parra E, Gutierrez C, Casati P and Spampinato C P. 2011.** Regulation of plant MSH2 and MSH6 genes in the UV-B-induced DNA damage response. *J Exp Bot*, **62**:2925-2937.
- Lee J, Das A, Yamaguchi M, Hashimoto J, Tsutsumi N, Uchimiya H and Umeda M. 2003.** Cell cycle function of a rice B2-type cyclin interacting with a B-type cyclin-dependent kinase. *Plant J*, **34**:417-425.
- Lidon F C and Ramalho J C. 2011.** Impact of UV-B irradiation on photosynthetic performance and chloroplast membrane components in *Oryza sativa* L. *J Photochem Photobiol B*, **104**:457-466.
- Livak K J and Schmittgen T D. 2001.** Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**:402-408.
- Mariconti L, Pellegrini B, Cantoni R, Stevens R, Bergounioux C, Cella R and Albani D. 2002.** The E2F family of transcription factors from Arabidopsis thaliana. Novel and conserved components of the retinoblastoma/E2F pathway in plants. *J Biol Chem*, **277**:9911-9919.
- Martinez L A, Goluszko E, Chen H Z, Leone G, Post S, Lozano G, Chen Z and Chauchereau A. 2009.** E2F3 is a mediator of DNA damage-induced apoptosis. *Mol Cell Biol*, **30**:524-536.
- McKenzie R L, Aucamp P J, Bais A F, Bjorn L O, Ilyas M and Madronich S. 2011.** Ozone depletion and climate change: impacts on UV radiation. *Photochem Photobiol Sci*, **10**:182-198.
- Mpoloka S W. 2008.** Effects of prolonged UV-B exposure in plants. *African Journal of Biotechnology*, **7**:4874-4883.
- Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek R L, Lee Y, Zheng L, Orvis J, Haas B, Wortman J and Buell C R. 2007.** The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res*, **35**:D883-887.

**Punta M, Coggill P C, Eberhardt R Y, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer E L, Eddy S R, Bateman A and Finn R D. 2012.** The Pfam protein families database. *Nucleic Acids Res*, **40**:D290-D301.

**Ramirez-Parra E and Gutierrez C. 2000.** Characterization of wheat DP, a heterodimerization partner of the plant E2F transcription factor which stimulates E2F-DNA binding. *FEBS Lett*, **486**:73-78.

**Ramirez-Parra E, Xie Q, Boniotti M B and Gutierrez C. 1999.** The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G(1)/S regulators. *Nucleic Acids Res*, **27**:3527-3533.

**Rauber R. 2009.** A família gênica do fator transcricional E2F e as respotas ao dano de DNA em plantas. In *Departamento de genética*, pp. 36.

**Reichheld J P, Gigot C and Chaubet-Gigot N. 1998.** Multilevel regulation of histone gene expression during the cell cycle in tobacco cells. *Nucleic Acids Res*, **26**:3255-3262.

**Ronquist F and Huelsenbeck J P. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**:1572-1574.

**Rozen S and Skaletsky H. 2000.** Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*, **132**:365-386.

**Sakamoto A N, Lan V T, Puripunyanich V, Hase Y, Yokota Y, Shikazono N, Nakagawa M, Narumi I and Tanaka A. 2009.** A UVB-hypersensitive mutant in *Arabidopsis thaliana* is defective in the DNA damage response. *Plant J*, **60**:509-517.

**Sekine M, Ito M, Uemukai K, Maeda Y, Nakagami H and Shinmyo A. 1999.** Isolation and characterization of the E2F-like gene in plants. *FEBS Lett*, **460**:117-122.

**Shen W H. 2002.** The plant E2F-Rb pathway and epigenetic control. *Trends Plant Sci*, **7**:505-511.

**Shimamoto K and Kyojuka J. 2002.** Rice as a model for comparative genomics of plants. *Annu Rev Plant Biol*, **53**:399-419.

**Shoemaker R C, Schlueter J and Doyle J J. 2006.** Paleopolyploidy and gene duplication in soybean and other legumes. *Curr Opin Plant Biol*, **9**:104-109.

**Sonnhammer E L, Eddy S R and Durbin R. 1997.** Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins*, **28**:405-420.

**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.

**Tian J and Yu J. 2009.** Changes in ultrastructure and responses of antioxidant systems of algae (*Dunaliella salina*) during acclimation to enhanced ultraviolet-B radiation. *J Photochem Photobiol B*, **97**:152-160.

**Tuteja N, Ahmad P, Panda B B and Tuteja R. 2009.** Genotoxic stress in plants: shedding light on DNA damage, repair and DNA repair helicases. *Mutat Res*, **681**:134-149.

**Vandepoele K, Vlieghe K, Florquin K, Hennig L, Beemster G T, Gruissem W, Van de Peer Y, Inze D and De Veylder L. 2005.** Genome-wide identification of potential plant E2F target genes. *Plant Physiol*, **139**:316-328.

**Wargent J J, Elfadly E M, Moore J P and Paul N D. 2011.** Increased exposure to UV-B radiation during early development leads to enhanced photoprotection and improved long-term performance in *Lactuca sativa*. *Plant Cell Environ*, **34**:1401-1413.

**Zalmas L P, Zhao X, Graham A L, Fisher R, Reilly C, Coutts A S and La Thangue N B. 2008.** DNA-damage response control of E2F7 and E2F8. *EMBO Rep*, **9**:252-259.

**ANEXO I. Códigos de acesso dos genes identificados e *primers* utilizados para estudos de expressão.** Dados provenientes do banco de dados TIGR (<http://rice.plantbiology.msu.edu/>). Para alguns genes os *primers* não foram desenhados e são indicados como “nd”.

<b>Loci E2F /DEL</b>	<b>CDS number / name</b>	<b>Genomic</b>	<b>Protein</b>	<b>PrimerF</b>	<b>PrimerR</b>
<u>LOC_Os04g33950</u>	CDS 12004.m08489 / E2F	12004.t03070	12004.m08489	GGTTCGGATGACTCAGGAAC	CCACCTATGGTTGCTTTCGT
<u>LOC_Os12g06200</u>	CDS 12012.m04611 / E2F	12012.t00506	12012.m04611	GGGCGAGTACCACCAGTTC	CACTAATCCAGCAGGCAGC
<u>LOC_Os02g33430</u>	CDS 12002.m8436 / E2F	12002.t02986	12002.m08436	GGTTGGGCAATGTAGCAACT	AATGCTGACATCCCCTTCTG
<u>LOC_Os04g02140</u>	CDS 12004.m05530 / E2F	12004.t00111	12004.m05530	CCGGCTACGACTATGGCTAC	TTTTCTAGGCCGTTTGATGC
<u>LOC_Os06g02459</u>	CDS 12006.m06082 / DEL	12006:7583781.7589606	12006.m06082	GTACAGCCGCAAGCAGAAGT	ATCGTCAACGTGCTCGAGAG
<u>LOC_Os02g50630</u>	CDS 13102.m05745 / DEL	13102.t04690	13102.m05745	CTGCTGGTGCAAGAAAACAA	TACCACGATTTCCAGGCTTC
<u>LOC_Os03g05760</u>	CDS 12003.m06098 / DP	12003.t00447	12003.m06098	CCTCATGCAACTGTGGAAGT	CAAATCGACGGTATTCATGATC
<u>LOC_Os10g30420</u>	CDS 13110.m02619 / DP	13110.t02375	13110.m02619	GCAGACACGTCCTCATGCTA	GGTTGAGCACTCACCTCCAT
<u>LOC_Os01g48700</u>	CDS 13101.m05054 / DP	13101.t04235	13101.m05054	nd	nd
<b>Loci Genes Reparo</b>					
<u>LOC_Os04g57730</u>	CDS 12004.m10658 / UGLY	12004.t05239	12004.m10658	CACGAGAGGTTGCATAGCC	GGAATTTTGATCCCCTCTGG
<u>LOC_Os09g24220</u>	CDS12009.m05578 / MSH6-1	12009.t02105	12009.m05578	nd	nd
<u>LOC_Os06g04190</u>	CDS 12006.m31828 / RAD1	12006.t00308	12006.m31828	CTCCTGTGCGAGATCCGCTAC	TGTATTCCCAGCATGCTCAA
<u>LOC_Os02g10990</u>	CDS 12002.m06397 / UVR3	12002.t00947	12002.m06397	ATCATCTTGCACGCCACTCT	CAGGAGAGCCACATCCAGTT
<b>Loci Genes Ciclo celular</b>					
<u>LOC_Os08g40170</u>	CDS 13108.m04293 / CDKB 2_1	13108.t03759	13108.m04293	GCTCGTTCACTGTCCCTCTC	CTGCTGAACCTCGGAATCTC
<u>LOC_Os01g59120</u>	CDS 13101.m06279 / CYC B1_1	13101.t05228	13101.m06279	TGCTGCCTCTGCTGTCTATG	AACCCTCAGCTTGCTCTCAG
<u>LOC_Os10g39410</u>	CDS 13110.m03607 / H4	13110.t03178	13110.m03607	GACAACATCCAGGGGATCA	GTGCTCGGTGTAGGTGACG
<u>LOC_Os09g29100</u>	CDS 13109.m02880 / CYC D6	13109.t02595	13109.m02880	AGATGGAGGAAACCGATGTG	CGACGTAGGAGAAGGGAGTG
<u>LOC_Os08g37390</u>	CDS 13108.m03935 / CYC D4	13108.t03487	13108.m03935	GCCGTGACTCCATTCTCCTA	GCCTCTGCGATGTCTTCTTC
<b>Loci Genes Apoptose</b>					
<u>LOC_Os04g54820</u>	CDS 13104.m05675 / MXK3	13104.t04960	13104.m05675	ATGCCCTGTAGGAAAGCAGA	GAACGTTCTGTCAGCACCAA
<u>LOC_Os09g35630</u>	CDS 13109.m03487 / BAG	13109.t03048	13109.m03487	TGATGAACGAGCTGGTCAAG	GTCCAGCGACTCCACGTA

**ANEXO II. Tabela de genes encontrados nas buscas em bancos de dados de acesso público.** Os genes foram nomeados pelo tipo de gene (E2F, DEL e DP), logo após foi incorporada uma letra (A, B, C) que indica a ordem que foram encontrados na busca, exceto para *Arabidopsis thaliana* por ser organismo modelo e possuir estas famílias caracterizadas funcionalmente. As ordens e famílias estão em ordem alfabética.

Order	Family	Specie	Taxa terminologies	Gene	Database	Access	Length(aa)
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F1	NCBI	ENSXETT00000046332	381
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F3	NCBI	gi 58331835 ref NP_001011109.1	427
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F4	NCBI	gi 167560905 ref NP_001107966.1	374
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F5	NCBI	gi 188528909 ref NP_001120880.1	371
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F7	NCBI	ENSXETT00000031323	862
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	E2F8	NCBI	ENSXETG00000004436	694
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	DP1	NCBI	gi 58332126 ref NP_001011211.1	411
Anura	Pipidae	<i>Xenopus tropicalis</i>	Xtr	DP2	NCBI	gi 166158068 ref NP_001107442.1	374
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	E2FA	JGI	Aly489072	457
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	E2FB	JGI	Aly482574	486
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	E2FC	JGI	Aly314229	415
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	E2FLA	JGI	Aly477381	1,378
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	DELA	JGI	Aly893401	356
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	DELB	JGI	Aly485177	397
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	DELLA	JGI	Aly899871	146
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	DPA	JGI	Aly487025	290
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	Aly	DPB	JGI	Aly487116	389
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	E2FA	JGI	AT5G22220.2	469
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	E2FB	JGI	AT2G36010.2	514
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	E2FC	JGI	AT1G47870.1	396
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	DELA	JGI	AT3G01330.1	354
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	DELB	JGI	AT5G14960.1	359
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	DELC	JGI	AT3G48160.2	403
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	DPA	JGI	AT5G03415.1	385
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Ath	DPB	JGI	AT5G02470.1	292



Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	E2FA	JGI	evm.model.supercontig_195.40	437
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	E2FB	JGI	evm.model.supercontig_104.59	488
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	E2FLA	JGI	evm.TU.contig_39412.1	188
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	DELA	JGI	evm.model.supercontig_35.34	357
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	DELB	JGI	evm.model.supercontig_14.76	357
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	DPA	JGI	evm.model.supercontig_738.1	157
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	DPLA	JGI	evm.model.supercontig_738.2	213
Brassicales	Caricaceae	<i>Carica papaya</i>	Cpa	DPB	JGI	evm.TU.contig_32800.1	314
Choanoflagellate	Choanoflagellate	<i>Monosiga brevicollis</i>	Mbr	E2F1	NCBI	gi 167523471 ref NP_001746072.1	413
Choanoflagellate	Choanoflagellate	<i>Monosiga brevicollis</i>	Mbr	E2F7	NCBI	gi 167517423 ref NP_001743052.1	550
Choanoflagellate	Choanoflagellate	<i>Monosiga brevicollis</i>	Mbr	DP1	NCBI	gi 167516980 ref NP_001742831.1	310
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	E2FA	JGI	Cucsa.263480.1	283
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	E2FB	JGI	Cucsa.185850.1	462
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	E2FC	JGI	Cucsa.083720.1	478
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	E2FD	JGI	Cucsa.159820.1	306
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	DELA	JGI	Cucsa.098190.1	384
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	DELB	JGI	Cucsa.197250.1	318
Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	Csa	DPA	JGI	Cucsa.397930.1	348
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F3	NCBI	gi 220673319 emb cax13650.1	429
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F4	NCBI	gi 47087407 ref NP_998597.1	393
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F5	NCBI	gi 68533607 gb aah98598.1	363
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F6	NCBI	gi 71892405 ref NP_001025315.1	405
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F7	NCBI	gi 169234759 ref NP_001038612.2	723
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	E2F8	NCBI	gi 189521060 ref NP_694311.3	860
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	DP1	NCBI	gi 41152118 ref NP_957070.1	386
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	DP2	NCBI	gi 38016161 ref NP_937851.1	409
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Dre	DP3	NCBI	gi 189523624 ref NP_001918928.1	409
Dictyosteliida	Dictyosteliidae	<i>Dictyostelium discoideum</i>	Ddi	E2F1	NCBI	gi 66810101 ref NP_638774.1	863
Dictyosteliida	Dictyosteliidae	<i>Dictyostelium discoideum</i>	Ddi	DP1	NCBI	gi 66818451 ref NP_642885.1	657
Diptera	Drosophilidae	<i>Drosophila melanogaster</i>	Dme	E2F1	NCBI	gi 24648770 ref NP_524437.2	805
Diptera	Drosophilidae	<i>Drosophila melanogaster</i>	Dme	E2F2	NCBI	gi 17137542 ref NP_477355.1	370
Diptera	Drosophilidae	<i>Drosophila melanogaster</i>	Dme	DP1	NCBI	gi 17136994 ref NP_477039.1	445
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FA	JGI	Glyma11g18400.1	296
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FB	JGI	Glyma12g09860.1	450
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FC	JGI	Glyma10g09410.1	466

Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FD	JGI	Glyma02g35850.1	468
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FE	JGI	Glyma06g02740.1	325
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FF	JGI	Glyma04g02720.1	343
Fabales	Fabaceae	<i>Glycine max</i>	Gma	E2FLA	JGI	Glyma16g19460.1	224
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DELA	JGI	Glyma17g10130.1	380
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DELB	JGI	Glyma05g01760.1	403
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DELC	JGI	Glyma06g19840.1	213
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DELLA	JGI	Glyma04g34870.1	340
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DPA	JGI	Glyma11g01270.1	337
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DPB	JGI	Glyma12g02050.1	333
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DPC	JGI	Glyma01g43800.1	338
Fabales	Fabaceae	<i>Glycine max</i>	Gma	DPD	JGI	Glyma11g09750.1	335
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	E2FA	JGI	AC146557_13.1	460
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	E2FB	JGI	Medtr3g156310.1	390
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	E2FLA	JGI	Medtr3g005120.1	410
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	DELA	JGI	Medtr8g124980.1	395
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	DPA	JGI	Medtr4g107190.1	342
Fabales	Fabaceae	<i>Medicago truncatula</i>	Mtr	DPB	JGI	Medtr7g132600	462
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	E2FA	JGI	Pp1s38_356V6.1	433
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	E2FB	JGI	Pp1s22_60V6.1	400
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	E2FC	JGI	Pp1s97_96V6.1	428
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	E2FD	JGI	Pp1s364_42V6.1	463
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DELA	JGI	Pp1s172_68V6.1	665
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DELB	JGI	Pp1s121_120V6.1	677
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DELC	JGI	Pp1s132_65V6.1	672
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DPA	JGI	Pp1s74_252V6.1	307
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DPB	JGI	Pp1s37_43V6.1	304
Funariales	Funariaceae	<i>Physcomitrella patens</i>	Ppa	DPC	JGI	Pp1s219_101V6.1	241
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	E2FA	JGI	mgv1a006165m	454
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	E2FB	JGI	mgv1a026657m	414
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	E2FC	JGI	mgv1a026529m	406
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	DELA	JGI	mgv1a008382m	375
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	DELB	JGI	mgv1a009071m	354
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	DPA	JGI	mgv1a008778m	362
Lamiales	Scrophulariaceae	<i>Mimulus guttatus</i>	Mgu	DPB	JGI	mgv1a009628m	336

Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	E2FA	JGI	cassava4.1_007247m	460
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	E2FB	JGI	cassava4.1_007936m	435
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	E2FC	JGI	cassava4.1_030363m	438
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	E2FD	JGI	cassava4.1_026018m	356
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DELA	JGI	cassava4.1_008673m	412
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DELB	JGI	cassava4.1_009139m	396
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DELC	JGI	cassava4.1_009539m	385
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DPA	JGI	cassava4.1_033513m	269
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DPLA	JGI	cassava4.1_020268m	88
Malpighiales	Euphorbiaceae	<i>Manihot esculenta</i>	Mes	DPB	JGI	cassava4.1_010769m	352
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	E2FA	JGI	30055.m001619	414
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	E2FB	JGI	28324.m000396	451
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	E2FC	JGI	30170.m014171	386
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	DELA	JGI	29822.m003421	328
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	DELB	JGI	30147.m014015	389
Malpighiales	Euphorbiaceae	<i>Ricinus communis</i>	Rco	DPA	JGI	30078.m002288	360
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FA	JGI	POPTR_0009s16350.1	473
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FB	JGI	POPTR_0004s21100.1	474
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FC	JGI	POPTR_0006s22150.1	424
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FD	JGI	POPTR_0016s07310.1	445
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FE	JGI	POPTR_0005s25930.1	111
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	E2FF	JGI	POPTR_0002s02500.1	476
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	DELA	JGI	POPTR_0015s08160.1	384
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	DELB	JGI	POPTR_0012s07700.1	385
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	DPA	JGI	POPTR_0006s12620.1	299
Malpighiales	Salicaceae	<i>Populus trichocarpa</i>	Ptr	DPB	JGI	POPTR_0016s09640.1	355
Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	Egr	E2FA	JGI	Egrandis_v1_0.011384m	471
Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	Egr	E2FB	JGI	Egrandis_v1_0.011166m	476
Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	Egr	DELA	JGI	Egrandis_v1_0.014805m	401
Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	Egr	DELB	JGI	Egrandis_v1_0.015435m	389
Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	Egr	DPA	JGI	Egrandis_v1_0.016718m	368
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	E2FA	JGI	Bradi3g45130.1	466
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	E2FLA	JGI	Bradi5g09640.1	341
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	E2FB	JGI	Bradi5g01920.1	392
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	E2FC	JGI	Bradi4g42050.1	449

Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	DELA	JGI	Bradi1g44520.1	431
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	DELB	JGI	Bradi3g59740.1	433
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	DPA	JGI	Bradi1g74890.1	383
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	DPB	JGI	Bradi3g27520.1	349
Poales	Poaceae	<i>Brachypodium distachyon</i>	Bdi	DPC	JGI	Bradi2g46530.1	194
Poales	Poaceae	<i>Oryza sativa</i>	Osa	E2FA	JGI	LOC_Os02g33430.1	475
Poales	Poaceae	<i>Oryza sativa</i>	Osa	E2FB	JGI	LOC_Os04g33950.1	319
Poales	Poaceae	<i>Oryza sativa</i>	Osa	E2FC	JGI	LOC_Os12g06200.1	446
Poales	Poaceae	<i>Oryza sativa</i>	Osa	E2FD	JGI	LOC_Os04g02140.1	494
Poales	Poaceae	<i>Oryza sativa</i>	Osa	DELA	JGI	LOC_Os02g50630.1	441
Poales	Poaceae	<i>Oryza sativa</i>	Osa	DELB	JGI	LOC_Os06g13670.1	425
Poales	Poaceae	<i>Oryza sativa</i>	Osa	DPA	JGI	LOC_Os03g05760.1	379
Poales	Poaceae	<i>Oryza sativa</i>	Osa	DPB	JGI	LOC_Os10g30420.1	346
Poales	Poaceae	<i>Oryza sativa</i>	Osa	DPC	JGI	LOC_Os01g48700.1	294
Poales	Poaceae	<i>Setaria italica</i>	Sit	E2FA	JGI	Si017108m	469
Poales	Poaceae	<i>Setaria italica</i>	Sit	E2FB	JGI	Si010044m	462
Poales	Poaceae	<i>Setaria italica</i>	Sit	E2FC	JGI	Si009650m	599
Poales	Poaceae	<i>Setaria italica</i>	Sit	E2FLA	JGI	Si011009m	223
Poales	Poaceae	<i>Setaria italica</i>	Sit	DELA	JGI	Si006412m	451
Poales	Poaceae	<i>Setaria italica</i>	Sit	DELB	JGI	Si019635m	396
Poales	Poaceae	<i>Setaria italica</i>	Sit	DPA	JGI	Si036124m	384
Poales	Poaceae	<i>Setaria italica</i>	Sit	DPB	JGI	Si036476m	340
Poales	Poaceae	<i>Setaria italica</i>	Sit	DPC	JGI	Si003799m	236
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	E2FA	JGI	Sb08g003670.1	377
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	E2FB	JGI	Sb04g021970.1	423
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	E2FC	JGI	Sb06g000890.1	350
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	E2FD	JGI	Sb06g016190.1	441
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DELA	JGI	Sb04g028440.1	387
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DELB	JGI	Sb10g008800.1	436
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DPA	JGI	Sb01g021000.1	340
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DPB	JGI	Sb01g046810.1	392
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DPC	JGI	Sb05g024630.1	334
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DPD	JGI	Sb09g000900.1	293
Poales	Poaceae	<i>Sorghum bicolor</i>	Sbi	DPE	JGI	Sb03g031130.1	246
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FA	JGI	GRMZM2G052515_T01	446

Poales	Poaceae	<i>Zea mays</i>	Zma	E2FB	JGI	AC233850.1_FGT005	479
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FC	JGI	GRMZM2G041701_T01	462
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FD	JGI	GRMZM2G050590_T01	468
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FE	JGI	GRMZM2G333591_T01	265
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FF	JGI	GRMZM2G351232_T01	434
Poales	Poaceae	<i>Zea mays</i>	Zma	E2FG	JGI	AC214168.3_FGT001	346
Poales	Poaceae	<i>Zea mays</i>	Zma	DELA	JGI	GRMZM2G361659_T01	444
Poales	Poaceae	<i>Zea mays</i>	Zma	DELB	JGI	GRMZM2G378665_T01	448
Poales	Poaceae	<i>Zea mays</i>	Zma	DELC	JGI	GRMZM2G060000_T02	426
Poales	Poaceae	<i>Zea mays</i>	Zma	DELLA	JGI	GRMZM2G077336_T01	335
Poales	Poaceae	<i>Zea mays</i>	Zma	DELLB	JGI	GRMZM2G169709_T01	354
Poales	Poaceae	<i>Zea mays</i>	Zma	DPA	JGI	GRMZM2G063387_T01	387
Poales	Poaceae	<i>Zea mays</i>	Zma	DPB	JGI	GRMZM2G086072_T01	386
Poales	Poaceae	<i>Zea mays</i>	Zma	DPC	JGI	GRMZM2G139024_T01	260
Prasinophyceae	Mamiellales	<i>Ostreococcus tauri</i>	Ota	E2F1	NCBI	gi 55978010 gb AAV68605.1	312
Prasinophyceae	Mamiellales	<i>Ostreococcus tauri</i>	Ota	DP1	NCBI	gi 116060017 emb CAL56076.1	361
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F1	NCBI	gi 12669911 ref NP_005216.1	437
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F2	NCBI	gi 4758226 ref NP_004082.1	437
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F3	NCBI	gi 4503433 ref NP_001940.1	465
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F4	NCBI	gi 12669915 ref NP_001941.2	413
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F5	NCBI	gi 134142811 ref NP_001942.2	346
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F6	NCBI	gi 109637795 ref NP_937987.2	281
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F7	NCBI	gi 145580626 ref NP_976328.2	911
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	E2F8	NCBI	gi 38505226 ref NP_078956.2	867
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	DP1	NCBI	gi 6005900 ref NP_009042.1	410
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	DP2	NCBI	gi 5454112 ref NP_006277.1	386
Primates	Hominidae	<i>Homo sapiens</i>	Hsa	DP3	NCBI	gi 189409125 ref NP_057605.3	405
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	E2FA	JGI	AcoGoldSmith_v1.021644m	405
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	E2FB	JGI	AcoGoldSmith_v1.027212m	472
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	DELA	JGI	AcoGoldSmith_v1.016387m	384
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	DELB	JGI	AcoGoldSmith_v1.017550m	367
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	DPA	JGI	AcoGoldSmith_v1.006960m	371
Ranunculales	Ranunculaceae	<i>Aquilegia Coerulea</i>	Aco	DPB	JGI	AcoGoldSmith_v1.006776m	378
Rhabditida	Rhabditidae	<i>Caenorhabditis elegans</i>	Cel	E2F1	NCBI	gi 17559226 ref NP_507289.1	342
Rhabditida	Rhabditidae	<i>Caenorhabditis elegans</i>	Cel	E2F2	NCBI	gi 71998283 ref NP_496825.3	412

Rhabditida	Rhabditidae	<i>Caenorhabditis elegans</i>	Cel	DP1	NCBI	gi 17532739 ref NP_495957.1	598
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F1	NCBI	gi 6681243 ref NP_031917.1	430
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F2	NCBI	gi 29244208 ref NP_808401.1	443
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F3	NCBI	gi 83523736 ref NP_034223.1	457
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F4	NCBI	gi 22507329 ref NP_683754.1	410
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F5	NCBI	gi 31982405 ref NP_031918.2	335
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F6	NCBI	gi 237681138 ref NP_150373.2	272
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F7	NCBI	gi 40254337 ref NP_848724.2	904
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	E2F8	NCBI	gi 67972650 ref NP_001013386.2	860
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	DP1	NCBI	gi 6678305 ref NP_033387.1	410
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	DP2	NCBI	gi 182765448 ref NP_848782.2	385
Rodentia	Muridae	<i>Mus musculus</i>	Mmu	DP3	NCBI	gi 149260219 ref NP_001481321.1	447
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	E2FA	JGI	ppa005262m	470
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	E2FB	JGI	ppa014783m	398
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	E2FC	JGI	ppa024272m	402
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	DELA	JGI	ppa015063m	358
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	DPA	JGI	ppa005481m	459
Rosales	Rosaceae	<i>Prunus persica</i>	Ppe	DPB	JGI	ppa021369m	299
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	E2FA	JGI	clementine0.9_009852m	456
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	E2FB	JGI	clementine0.9_010950m	427
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	E2FC	JGI	clementine0.9_008982m	479
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	E2FD	JGI	clementine0.9_028348m	286
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	DELA	JGI	clementine0.9_035370m	376
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ccl	DPA	JGI	clementine0.9_014461m	350
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	E2FA	JGI	orange1.1g012778m	456
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	E2FB	JGI	orange1.1g015878m	399
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	E2FC	JGI	orange1.1g015627m	403
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	E2FD	JGI	orange1.1g042344m	302
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	DELA	JGI	orange1.1g018790m	350
Sapindales	Rutaceae	<i>Citrus sinensis</i>	Csi	DPA	JGI	orange1.1g018827m	350
Selaginellales	Selaginellaceae	<i>Selaginella moellendorffii</i>	Smo	E2FA	JGI	Smo118926	175
Selaginellales	Selaginellaceae	<i>Selaginella moellendorffii</i>	Smo	E2FB	JGI	Smo136738	352
Selaginellales	Selaginellaceae	<i>Selaginella moellendorffii</i>	Smo	DPA	JGI	Smo80923	276
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	E2FA	JGI	GSVIVT01033931001	466
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	E2FB	JGI	GSVIVT01036059001	446

Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	E2FC	JGI	GSVIVT01009890001	436
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	DELA	JGI	GSVIVT01007832001	815
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	DELB	JGI	GSVIVT01032583001	372
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	DPA	JGI	GSVIVT01025739001	349
Vitales	Vitaceae	<i>Vitis vinifera</i>	Vvi	DPB	JGI	GSVIVT01034190001	378
Volvocales	Chlamydiminadaceae	<i>Chlamydomonas reinhardtii</i>	Cre	E2FA	JGI	Cre01.g052300.t1.1	513
Volvocales	Chlamydiminadaceae	<i>Chlamydomonas reinhardtii</i>	Cre	DPA	JGI	Cre07.g323000.t1.1	646
Volvocales	Volvocaceae	<i>Volvox carteri</i>	Vca	E2FA	JGI	Vca127253	468
Volvocales	Volvocaceae	<i>Volvox carteri</i>	Vca	DPA	JGI	Vca121369	721