

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Evolução convergente da protease FtsH5 de *Arabidopsis thaliana* e seu regulador negativo putativo FIP (*FtsH5 interacting protein*)

Marcos Araújo Castro e Silva

Dissertação apresentada para obtenção de título de Mestre em Ciências. Área de concentração: Genética e Melhoramento de Plantas

**Piracicaba
2015**

Marcos Araújo Castro e Silva
Bacharel em Ciências Biológicas

Evolução convergente da protease FtsH5 de *Arabidopsis thaliana* e seu regulador negativo putativo FIP (*FtsH5 interacting protein*)
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **MARCIO DE CASTRO SILVA FILHO**

Dissertação apresentada para obtenção de título de Mestre em Ciências. Área de concentração: Genética e Melhoramento de Plantas

**Piracicaba
2015**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Silva, Marcos Araújo Castro e
Evolução convergente da protease FtsH5 de *Arabidopsis thaliana* e seu regulador
negativo putativo FIP (*FtsH5 interacting protein*) / Marcos Araújo Castro e Silva. - - versão
revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2015.
91 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz".

1. Metaloproteases FtsH 2. Interação proteína-proteína 3. Evolução convergente
I. Título

CDD 581.1925
S586e

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Dedico:

*À minha família, Pedro, Maria Vanda e Lucas,
pelo amor e incentivo.*

AGRADECIMENTOS

Agradeço primeiramente ao meu orientador Prof. Dr. Marcio de Castro Silva Filho, pela oportunidade e confiança, mas principalmente pelo exemplo de entusiasmo e amor pela ciência;

Ao Dr. Marcelo Mendes Brandão e a Dr^a. Renata Dias de Oliveira pelos valiosos ensinamentos e colaboração;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de Mestrado concedida para a execução deste trabalho;

Agradeço aos amigos do Laboratório de Biologia Molecular: Renata Oliveira, Lucas Lopes, Flávia Franco, Wiliane Garcia, Karina Lopes, Larissa Spoladore, Thais Souza, Marcelo Brandão, Aline Borges, Poliene Costa, Elaine Castelhamo, Wellington Campos, Horácio Montenegro, Fabiana Mingossi e Lúgia Macedo, pela amizade, apoio e convivência tão agradável ao longo destes dois anos;

Agradeço também aos amigos do Laboratório de Bioquímica de Proteínas, em especial à Tabatá Bergonci, ao Juan Guerrero e à Bianca Ribeiro, pela disponibilidade em ajudar e pela amizade;

Ao Rafael Columbi por sua dedicação na manutenção e organização do Laboratório de Biologia Molecular de Plantas;

À Ana Paula Mendes pelo apoio nos momentos difíceis, amizade e motivação tão importantes;

Aos amigos de república André Herman, Luciano Silveira e Elenilson Moreira pelos bons momentos compartilhados e muitos pensamentos e experiências trocados.

Aos demais amigos de Viçosa, Piumhi e espalhados pelo Brasil, por me lembrar dos meus verdadeiros propósitos e pela essencial amizade, que mesmo distante se faz presente.

À minha família pelo carinho, suporte e amor ao longo de todos estes anos, sem os quais nada teria sido possível;

Agradeço especialmente aos meus pais, Pedro e Maria Vanda pelo amor incondicional e exemplo de vida. E ao meu irmão pelo incentivo e amizade mais que necessários;

A todos o meu mais profundo e sincero obrigado!

“Tenho a impressão de ter sido uma criança brincando à beira-mar, divertindo-me em descobrir uma pedrinha mais lisa ou uma concha mais bonita que as outras, enquanto o imenso oceano da verdade continua misterioso diante de meus olhos”. - Isaac Newton

SUMÁRIO

RESUMO.....	11
ABSTRACT	13
LISTA DE FIGURAS	15
LISTA DE TABELAS	17
1 INTRODUÇÃO	19
2 REVISÃO BIBLIOGRÁFICA.....	23
2.1 – Compartimentalização celular e controle de qualidade de proteínas	23
2.2 – AAA/FtsH proteases	24
2.3 – Proteases FtsH plastidiais	26
2.4 – Regulação da atividade das proteases FtsH	30
2.5 – FIP (FtsH5 Interacting Protein).....	32
3 MATERIAL E MÉTODOS	35
3.1 – Análise de expressão	35
3.1.1 – Expressão tecido-específica e temporal.....	35
3.1.2 – Expressão em resposta a estresses abióticos	35
3.2 – Análises filogenéticas	36
3.2.1 – Filogenia dos homólogos de FIP e das proteases do complexo FtsH tilacoidal e inferência sobre o processo de coevolução	36
3.2.1.1 – Identificação e obtenção das sequências para a análise	36
3.2.1.2 – Alinhamento múltiplo das sequências e modelo de substituição de DNA ..	37
3.2.1.3 – Proposição de modelo evolutivo e inferência sobre a coevolução	38
3.2.2 – Filogenia representativa dos homólogos das FtsH do tipo A.....	39
3.3 – Detecção de sítios de pressão de seleção	40
3.4 – Análise de conservação de sequência e estrutural do complexo FtsH	42
4 RESULTADOS E DISCUSSÃO.....	45

4.1 – FIP como um regulador negativo putativo da atividade do complexo FtsH tilacoidal.....	45
4.1.1 – Perfil de expressão tecido específico e ao longo do desenvolvimento	45
4.1.2 – Perfil de expressão de FIP e das proteases do complexo FtsH em resposta a diversos estresses abióticos	47
4.2 – História evolutiva da proteína FIP e das proteases do complexo FtsH tilacoidal	52
4.2.1 – Inferência filogenética e classificação dos homólogos das proteases do complexo FtsH.....	52
4.2.2 – Evolução convergente entre as FtsHs do tipo A e FIP.....	57
4.2.3 – A proteína FIP no contexto da história evolutiva das FtsHs.....	61
4.3 – Pressões seletivas e conservação de sequência	63
5 CONSIDERAÇÕES FINAIS.....	69
REFERÊNCIAS	71
ANEXOS.....	81

RESUMO

Evolução convergente da protease FtsH5 de *Arabidopsis thaliana* e seu regulador negativo putativo FIP (*FtsH5 interacting protein*)

As metaloproteases AAA/FtsH são componentes chave do controle da qualidade das proteínas inseridas nas membranas de mitocôndrias e cloroplastos. Em *Arabidopsis thaliana*, as proteases FtsH presentes nas membranas dos tilacóides formam um complexo heterohexamérico composto pelas subunidades FtsH1/FtsH5 (tipo A) e FtsH2/FtsH8 (tipo B). Este complexo está envolvido na reciclagem de proteínas foto-danificadas, especialmente da proteína D1, centro de reação do PSII. Algumas linhas de evidências indicam ainda que existe um limiar de concentração das proteases FtsH, necessário para a correta formação e desenvolvimento dos cloroplastos. Apesar da extensiva caracterização genética e molecular das proteases FtsH, o mecanismo regulatório do complexo FtsH dos cloroplastos não foi totalmente elucidado até o momento, contudo existem evidências de que a sua ativação pode estar relacionada a alta incidência luminosa e a outras condições de estresse. A presença de fatores proteicos auxiliares, foi testada como hipótese alternativa por nosso grupo, através do uso da protease FtsH5 como isca em um ensaio de duplo híbrido de leveduras. Este ensaio identificou uma proteína interagente putativa, nomeada FIP (*FtsH5 Interacting Protein*), a qual comprovadamente interage com FtsH5 e está localizada nas membranas dos tilacóides. De modo a investigar o papel regulatório putativo de FIP sobre a atividade do complexo FtsH, nós analisamos os padrões de expressão em uma ampla gama de condições de estresse a partir de dados públicos de microarranjos de DNA. Os perfis de expressão indicam que FIP pode ser um regulador negativo da atividade do complexo. Os resultados também sugerem que o complexo pode estar envolvido na resposta do cloroplasto a diferentes tipos de condições de estresse. O estudo da história evolutiva das proteínas interagentes FtsH5 e FIP evidenciou que as sequências homólogas a FIP são encontradas exclusivamente em musgos e plantas superiores, sugerindo assim que a origem de FIP pode estar relacionada a colonização terrestre. Todos os genes codificantes das proteases FtsH do complexo foram usados como “*query*” na busca por sequências homólogas, permitindo a classificação das proteases FtsH nos tipos A e B por inferência filogenética Bayesiana. Análises filogenéticas Bayesianas também foram feitas para FIP e as proteases FtsH tipos A e B, independentemente. A análise *Mirrortree* suportou a existência de coevolução entre FIP e as proteases FtsH tipo A. Por outro lado, nenhuma correlação foi encontrada entre FIP e as proteases FtsH tipo B, o que corrobora nossas observações experimentais anteriores. Além disso, o agrupamento portador de homólogos FIP pôde ser recuperado em uma filogenia mais completa das proteases FtsH do tipo A. Análises subsequentes mostraram que ambas as proteínas interagentes estão extensivamente sobre seleção negativa e que proteases FtsH tipo A são bastante conservadas, principalmente nos seus domínios internos.

Palavras-chave: Metaloproteases FtsH; Interação proteína-proteína; Evolução convergente

ABSTRACT

Convergent evolution of *Arabidopsis thaliana* FtsH5 protease and its putative negative regulator FIP (FtsH5 interacting protein)

Eukaryotic AAA/FtsH metalloproteases display a key role in the protein quality control of membrane-inserted proteins in mitochondria and chloroplasts. In *Arabidopsis thaliana*, chloroplast thylakoidal membranes FtsH proteases form a heterohexameric complex made by FtsH1/FtsH5 (type A) and FtsH2/FtsH8 (type B) subunits. This complex is involved in protein turnover of photo-damaged proteins, in particular the D1 protein at the PSII reaction center. Several lines of evidence also indicate that a FtsH threshold level is necessary for the proper formation and development of chloroplasts. Despite extensive genetic and molecular characterization of the FtsH proteases, the regulatory mechanism of the FtsH complex in chloroplasts has not yet been fully elucidated, however, there are evidences that its activation might be related to high light incidence and other stress conditions. The presence of auxiliary protein factors, as an alternative hypothesis, was tested by our group, through the use of the protease FtsH5 as bait in a yeast two-hybrid assay. This essay identified a putative interacting protein named FIP (*FtsH5 Interacting Protein*), which has been proved to interact with FtsH5 and be located at the thylakoid membranes. In order to investigate a putative regulatory role of FIP on FtsH complex activity, we analyzed gene expression patterns in a wide range of stress conditions from public DNA microarray data. The expression profiles indicate that FIP could be a negative regulator of the FtsH complex activity. The results also suggest that the complex may be involved in the chloroplast response to different types of stress conditions. In order to shed some light on the evolutionary history of FtsH5 and FIP interacting proteins, we have shown that FIP's homologous sequences were exclusively found in mosses and higher plants, suggesting that FIP origin might be related to the plant terrestrial colonization. All *Arabidopsis* FtsH complex-encoding genes were used as "query" sequences in search for homologous sequences, allowing us to classify the FtsH proteases in type A and B by Bayesian phylogenetic inference. Bayesian phylogenetic analyses were also run for FIP and FtsH types A and B proteases, independently. Mirrortree analysis supported coevolution between FIP and type A FtsH proteases. On the other hand, no correlation was found between FIP and type B FtsH homologues, which support our previous experimental observations. In addition, the FIP bearing cluster could be recovered in a more complete type A FtsH phylogeny. Subsequent analyzes have shown that both interacting proteins are extensively under negative selection and that type A FtsH are very conserved, mainly in its inner domains.

Keywords: FtsH metalloproteases; Protein-protein interaction; Convergent evolution

LISTA DE FIGURAS

Figura 1 – Proposição filogenética para os homólogos de FtsH.....	28
Figura 2 – Expressão de FIP e das proteases do complexo FtsH ao longo do desenvolvimento de <i>Arabidopsis thaliana</i>	46
Figura 3 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. A) Estresse a frio. B) Estresse a seca. C) Estresse a calor.....	49
Figura 4 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. A) Estresse osmótico. B) Estresse salino. C) Estresse oxidativo.....	50
Figura 5 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. A) Estresse a UV-B. B) Estresse por ferimento. C) Estresse genotóxico.....	51
Figura 6 – Inferência filogenética bayesiana das sequências homólogas de FIP.....	54
Figura 7 – Inferência filogenética bayesiana das sequências homólogas das proteases do complexo FtsH.....	56
Figura 8 – Inferência filogenética bayesiana das sequências homólogas as proteases FtsHs do tipo A, do grupo de espécies portador de FIP.....	57
Figura 9 – Inferência filogenética bayesiana das sequências homólogas as proteases FtsHs do tipo B, do grupo de espécies portador de FIP.....	58
Figura 10 – Inferência sobre a existência de coevolução pelo método Mirrortree.....	60
Figura 11 – Inferência filogenética bayesiana das sequências homólogas as FtsHs do tipo A.....	62
Figura 12 – Representação gráfica tridimensional das posições estruturais conservadas das FtsHs do tipo A, do grupo de espécies portador de FIP.....	67

LISTA DE TABELAS

Tabela 1 – Parâmetros utilizados nos modelos de substituição por sítio.....	42
Tabela 2 – Sequências homólogas a FtsHs do tipo A agrupadas no táxon monofilético das sequências provenientes de espécies portadoras de homólogos de FIP utilizadas na análise da Figura 8.....	63
Tabela 3 – Descrição dos parâmetros estimados e os escores de verossimilhança para os homólogos de FIP sob diferentes modelos de seleção com razões de ω variáveis entre os sítios.....	65
Tabela 4 – Descrição dos parâmetros estimados e os escores de verossimilhança para os homólogos de FtsH do tipo A (apenas das espécies portadoras de FIP) sob diferentes modelos de seleção com razões de ω variáveis entre os sítios.....	65

1 INTRODUÇÃO

A compartimentalização das células eucarióticas através do seu sistema de endomembranas possibilita a separação de reações químicas e vias metabólicas específicas, pela criação de ambientes com características únicas e diferenciadas entre si. Muitas proteínas participam do estabelecimento e manutenção destas características e governam os processos celulares desenvolvidos nestes compartimentos. Por isso esses compartimentos apresentam um rigoroso sistema de controle de qualidade de suas proteínas, que nos cloroplastos auxiliam na manutenção das condições organelares ótimas para o desempenho de suas funções. Nos cloroplastos, mecanismos do sistema de controle de qualidade eliminam proteínas mal-dobradas ou danificadas por danos foto-oxidativos, causados tanto pelo processo natural de fotossíntese quanto pelo estresse luminoso, e danos causados por outros tipos de estresse, principalmente abióticos. Neste sentido, compreender o funcionamento de tais mecanismos possibilita a aquisição de um maior conhecimento sobre a função organelar e sobre a sua manutenção.

As AAA/FtsH proteases se localizam nas membranas de organelas e tem um papel fundamental na regulação das proteínas presentes nessas estruturas. As AAA/FtsH proteases são metaloproteases com atividades proteolíticas e de chaperonas (*chaperone-like*), pertencem a família das proteínas AAA (ATPases Associadas à Diversas Atividades Celulares) e portanto utilizam a energia da hidrólise do ATP para realizar suas atividades biológicas. As FtsHs formam homo ou heterohexâmeros funcionais e possuem três domínios funcionais em cada cadeia polipeptídica, são eles: um domínio transmembrana, um domínio ATPásico conservado (AAA) e um domínio metalopeptidase (SUNO *et al.* 2006). As FtsHs foram primeiramente descritas em *Escherichia coli* (OGURA *et al.* 1991) formando um complexo hexamérico na membrana plasmática e desempenhando um papel importante no controle de qualidade de proteínas intracelulares (ITO; AKIYAMA, 2005).

Recentes estudos realizados com os complexos FtsH, permitem assumir que eles estão presentes em todos os organismos, exceto naqueles pertencentes ao domínio Archaea. Nos cloroplastos de plantas superiores, FtsHs apresentam um domínio transmembrana e formam um complexo heterohexamérico presente na

membrana dos tilacóides. O complexo FtsH localizado na membrana do tilacóide realiza a degradação da proteína D1 foto-danificada, centro de reação do fotossistema II (PSII) e permite a reciclagem do PSII (LINDAHL *et al.* 2000, BAILEY *et al.* 2002). Algumas linhas de evidências, indicam ainda que um nível limiar de FtsHs é necessário para a correta formação e desenvolvimento dos cloroplastos (YU *et al.* 2005; ZALTSMAN *et al.* 2005a; KATO; SAKAMOTO, 2007).

As FtsHs foram extensivamente estudadas nos anos recentes e têm sido muito bem caracterizadas em termos genéticos e moleculares, entretanto o mecanismo de regulação do complexo FtsH em cloroplastos ainda não foi elucidado. Não obstante, algumas linhas de evidências sugerem que a ativação do complexo possa estar relacionada a alta incidência luminosa ou a outras condições de estresse. Uma hipótese alternativa para o mecanismo de ativação do complexo é a presença de fatores proteicos auxiliares, que estariam modulando a atividade do complexo. Neste sentido, nosso grupo de pesquisa realizou anteriormente um ensaio de duplo híbrido de leveduras com a proteína FtsH5 (At5G42270) como isca, componente do complexo heterohexamérico presente nos tilacóides, de modo a identificar possíveis fatores proteicos auxiliares. Um candidato em potencial denominado FIP (*FtsH5 Interacting Protein/At5g02160*) foi identificado. FIP é uma proteína localizada na membrana dos tilacóides e mostrou interação com FtsH5 *in vivo* (duplo híbrido de levedura) e *in vitro* (*Pull-Down*). Além disso experimentos de importação *in-vitro* utilizando FIP marcada radiotivamente, seguidos de eletroforese em géis nativos, mostraram que FIP está associada a um complexo de 440 kDa (tamanho esperado para o complexo FtsH presente nos tilacóides) e a localização plastidial foi determinada pela fusão da proteína FIP a proteína verde fluorescente (GFP) (RODRIGUES, 2011).

Recentemente nosso grupo de pesquisa mapeou a região de interação de FIP com FtsH5 utilizando o sistema de duplo híbrido de leveduras, mostrando que a interação é mantida por 46 aminoácidos que compreendem dois motivos ricos em cisteína e a região localizada entre eles e que substituições em qualquer uma das cisteínas presentes nestes motivos por alanina, desfazem a interação (BRAGA, 2013). Experimentos de análise de expressão de FIP e FtsH5 sobre as condições de estresse salino e estresse a frio, levantaram a hipótese de que FIP seja um regulador negativo do complexo FtsH (BRAGA, 2013).

Este trabalho visou analisar os padrões de expressão das proteínas do complexo FtsH dos tilacóides e da proteína FIP frente a diversas condições de estresse. Assim, foi possível investigar o papel de FIP na regulação da atividade do complexo e avaliar a resposta destas proteínas de maneira a evidenciar as suas relações com os diferentes tipos de estresse, a fim de lançar luz sobre o papel da interação de FIP com o complexo FtsH. Além disso buscou propor modelos evolutivos para a proteína FIP e as proteínas do complexo FtsH, por meio da análise e comparação das histórias evolutivas destas, além de elucidar as relações entre as mesmas, buscando inferir e avaliar a existência e magnitude dos processos evolutivos envolvidos, tais como coevolução e pressões de seleção. E, finalmente, de modo a compor um quadro mais completo da interação e buscando modelos para esta, objetivou-se a construção de modelos estruturais capazes de acessar informações como conservação de sequências.

2 REVISÃO BIBLIOGRÁFICA

2.1 – Compartimentalização celular e controle de qualidade de proteínas

As membranas biológicas tiveram sua origem em primitivas vesículas lipídicas, as quais evoluíram para o sofisticado sistema de endomembranas atual, formando os limites intra e intercelulares de todas as células conhecidas. Tais membranas permitem a separação e organização das células em compartimentos bem estruturados e capazes de estabelecer e manter características próprias, como valor de pH, força e composição iônica, estado redox e constituição de proteínas, carboidratos e muitas outras moléculas. Além disso, cada membrana celular possui a sua composição específica de lipídeos, carboidratos e proteínas, estas últimas associadas assimetricamente, externamente ou de maneira integral a bicamada lipídica. Todos estes atributos possibilitam a especialização dos compartimentos para a execução de conjunto bem definido e específico de reações, o que leva a uma maior eficiência dos processos bioquímicos e celulares como um todo.

Acredita-se que as mitocôndrias assim como os cloroplastos, provavelmente evoluíram de procariotos de vida livre e que em algum momento, possivelmente a cerca de 1,5 bilhões de anos, bactérias de vida livre foram endocitadas por uma célula eucariótica (ou pré-eucariótica) e se tornaram organismos endossimbiontes (MARGULIS, 1970). Primeiramente, um procarioto aeróbico foi endocitado e estabeleceu o primeiro evento endossimbiótico, evoluindo para o que hoje conhecemos como as mitocôndrias. Posteriormente células eucarióticas já contendo mitocôndrias, teriam endocitado procariotos fotossintéticos, os quais constituem o segundo evento endossimbiótico e teriam dado origem aos cloroplastos. Com o processo evolutivo a maior parte dos genes procarióticos teria migrado para o genoma nuclear, levando a uma grande redução dos genomas das mitocôndrias e dos cloroplastos e perda da autonomia destes. Atualmente o genoma de cloroplastos possui apenas de 50 a 150 genes, codificando somente para as funções centrais da atividade da organela, como proteínas de tradução e transcrição e subunidades dos complexos fotossintéticos (LYSKA; MEIERHOFF; WESTHOFF, 2013). Contudo, o cloroplasto apresenta um vasto proteoma, contando com 2700 proteínas, evidenciando um robusto sistema de importação de proteínas, bem como um sofisticado sistema de controle de qualidade das mesmas.

O sistema de controle de qualidade de proteínas determina o destino das proteínas celulares, prevenindo a acumulação de proteínas indesejadas devido a mutações genéticas, erros biossintéticos, síntese desbalanceada de subunidades e diferentes tipos de danos causados por condições de estresse. É evidente que o acúmulo de proteínas não-nativas pode alterar o estado correto de homeostase das proteínas e conduzir a formação de organelas não-funcionais e a uma reduzida viabilidade celular. Na maioria dos compartimentos subcelulares, sistemas de controle de proteínas foram identificados, incluindo chaperonas moleculares e proteases.

As proteínas de membrana, assim como as outras proteínas, também estão sujeitas à constante reciclagem, removendo-se proteínas mal dobradas ou danificadas e substituindo-as por novas proteínas nativas. Entretanto, a bicamada lipídica impede a acessibilidade destas proteínas. Portanto, proteases de membrana especializadas tem um papel fundamental na regulação de proteínas de membrana. Um grupo bem estudado de proteases de membrana são as AAA/FtsH proteases, as quais são enzimas com atividades proteolíticas e de chaperonas (*chaperone-like*) e atuam como elementos principais dos sistemas de controle de qualidade de proteínas de mitocôndrias e cloroplastos.

2.2 – AAA/FtsH proteases

As AAA proteases são assim nomeadas por serem membros da superfamília das AAA (ATPases associadas a adiversas atividades celulares), sendo que são metaloproteases dependentes de ATP, estando presentes em eubactérias assim como nas organelas de origem procariótica, as mitocôndrias e os cloroplastos. As AAA proteases são também chamadas de FtsHs e tal nome se refere a uma enzima de *Escherichia coli* (*Filamentous temperature sensitive H*). O nome deriva do hábito de crescimento do mutante de *E. coli* Y16 (SANTOS; DE ALMEIDA, 1975) deficiente para o gene *ftsH*. Baseado na literatura é visível a preferência do uso do termo AAA proteases para enzimas mitocondriais, enquanto que o termo FtsH é mais usado para as enzimas cloroplastidiais e bacterianas (JANSKA *et al.* 2013).

A maioria das bactérias possui apenas um único gene para FtsH (AKIYAMA *et al.* 1995), leveduras e humanos possuem três genes (ARNOLD; LANGER 2002), enquanto que são encontrados 17 sequências para FtsH no genoma de *Arabidopsis*

thaliana, doze desses genes codificando para isoformas ativas (GARCIA-LORENZO *et al.* 2006) e cinco para isoformas inativas (FtsHi; SOKOLENKO *et al.* 2002). As isoformas inativas não possuem o motivo dedo de zinco responsável pela atividade proteolítica. Oito das FtsH's, assim como as cinco FtsHi's são exclusivamente direcionadas para os cloroplastos (FtsH 1, 2, 5-9 e 12; FERRO *et al.* 2010), três FtsH proteases são localizadas nas mitocôndrias (FtsH 3, 4 e 10; JANSKA *et al.* 2005) e a FtsH 11, excepcionalmente apresenta duplo direcionamento para ambas as organelas (URANTOWKA *et al.* 2005). A evolução da fotossíntese correlaciona notavelmente com a multiplicação dos genes FtsH, muito provavelmente relacionado a maior quantidade de danos a proteínas causados pelo aumento do estresse oxidativo (ADAM *et al.*, 2005).

Os membros da família FtsH formam homo ou heterohexâmeros funcionais, consistindo de seis FtsH proteases individuais com três domínios funcionais em cada cadeia polipeptídica. O domínio N-terminal transmembrana que ancora a enzima na membrana é seguido por um domínio ATPásico da superfamília AAA e um domínio metalopeptidase M41 C-terminal (<http://merops.sanger.ac.uk/>, SUNO *et al.* 2006). Acredita-se que o domínio AAA atue como uma chaperona molecular e ele é caracterizado por um cassete de 200-250 resíduos altamente conservados. Este domínio contém os motivos Walker A e B relacionados à ligação ao ATP, bem como um motivo SRH (*Second Region of Homology*) o qual carrega alguns resíduos de arginina conservados (“Dedo de arginina”) importante na oligomerização e hidrólise de nucleotídeos (TOMOYASU *et al.* 1993; OGURA & WILKINSON, 2001; BIENIOSSEK *et al.* 2009). Baseado em análises cristalográficas de raio-X, o domínio ATPásico das FtsH de *E. coli* constituem um anel hexamérico que contém resíduos aromáticos conservados, voltados para a estrutura em anel, sendo que a energia de quebra do ATP é usada para translocação dos substratos e carregamento até a câmara proteolítica (KRZYWDA *et al.*, 2002; BIENIOSSEK *et al.*, 2009). As FtsH plastidiais são quase idênticas às de *E. coli*, o que implica uma estrutura hexamérica similar.

Ainda a partir de estudos de cristalografia sabe-se que as subunidades do complexo FtsH são capazes de alternar entre um estado aberto e fechado. A hidrólise do ATP promove uma mudança estrutural no domínio AAA que move uma área hidrofóbica conservada (motivo de poro FGV; Bieniossek *et al.* 2009) em direção ao interior da câmara proteolítica. O motivo FGV se liga ao substrato e o

puxa para o interior da câmara proteolítica. O anel proteásico é uma “região estacionária” e os movimentos dos domínios AAA realizam o desdobramento do substrato e a translocação de sua cadeia peptídica para o poro central. A hipótese corrente mais aceita é a de que a energia proveniente da hidrólise do ATP é convertida em força mecânica pelas mudanças conformacionais envolvendo os resíduos do poro (BIENIOSSEK *et al.* 2009).

O domínio metalopeptidase M41 localizado na extremidade C-terminal abriga o centro ativo da atividade proteolítica, portando o motivo de ligação a zinco HEXXH. Além dos dois resíduos de histidina do motivo HEXXH, as FtsHs contêm um terceiro ligante ao zinco (ácido aspártico) e portanto pertence a classe das Asp-zincinas (BIENIOSSEK *et al.* 2006). As FtsHs apresentam duas topologias possíveis em suas respectivas membranas, dependendo da presença de uma ou duas regiões transmembrana. A FtsH de *E. coli* atravessa a membrana plasmática duas vezes e expõe ambos os domínios ATPásico e proteolítico para o citoplasma (TOMOYASU *et al.* 1993). Diferentemente, existem dois tipos de topologias diferentes para as FtsHs de mitocôndrias de leveduras (LEONHARD *et al.* 1996), humanos (CASARI *et al.* 1998; BANFI *et al.* 1999; COPPOLA *et al.* 2000) e plantas (URANTOWKA *et al.* 2005; PIECHOTA *et al.* 2010). As proteases i-AAA atravessam a membrana interna das mitocôndrias uma vez e expõe o seu domínio catalítico para o espaço intermembranas, enquanto que as proteases m-AAA com dois domínios transmembrana perfuram a membrana duas vezes posicionando seus domínios catalíticos na matrix mitocondrial. Nos cloplastos, quatro FtsHs (AtFtsH 1, AtFtsH 2, AtFtsH 5 e AtFtsH 8) apresentam um domínio transmembrana e foram detectadas no mesmo complexo presente na membrana dos tilacóides, com os seus domínios catalíticos voltados para o estroma (LINDHAL *et al.* 1996). As AtFtsHs 7, 9 e 12 com dois domínios transmembrana e a AtFtsH 11 com um domínio, foram recentemente encontrados no envelope de cloroplastos (FERRO *et al.* 2010), sendo que as proteínas AtFtsH 7 e AtFtsH 9 também foram encontradas nos tilacóides (RODRIGUES, 2011), mas a topologia destas é desconhecida. A localização subplastidial de AtFtsH 6 ainda não foi demonstrada.

2.3 – Proteases FtsH plastidiais

Estudos genômicos e proteômicos realizados com os complexos FtsH na última década, permitiram assumir que eles estão presentes em todos os organismos que realizam fotossíntese oxigênica. Tal importância fica clara pela grande conservação de sequência, pelo número e função dos genes ortólogos e pelo fato de que nocautes para o complexo FtsH tilacoidal são letais.

O genoma de *Arabidopsis* codifica para 12 isoformas funcionais de FtsH. As FtsH tilacoidais tem os maiores níveis de expressão quando comparadas com as outras FtsH proteases. Alguns pares de FtsHs direcionadas para o cloroplastos constituem pares intimamente relacionados, são eles FtsH2 e FtsH8, FtsH1 e FtsH5 e por fim FtsH7 e FtsH9. Os padrões de expressão das FtsHs 1, 2, 5 e 8 são bastante similares, como mostrado por estudos de fusão de promotores ao gene β -glucuronidase (YU *et al.* 2004; YU *et al.* 2005). Quanto ao nível de proteínas, FtsH2 é a proteína FtsH mais abundante em cloroplastos, seguida por FtsH5, e então por níveis bem menores de FtsH8 e FtsH1 (SINVANY-VILLALOBO *et al.* 2004).

Interessantemente, estas diferenças no acúmulo das proteínas FtsH se refletem nos fenótipos variegados dos mutantes de FtsH. O mutante de FtsH2 tem o fenótipo de variação mais severo (CHEN *et al.* 2000; TAKECHI *et al.* 2000) e apresenta sensibilidade a foto-inibição (SAKAMOTO *et al.* 2002), enquanto que o mutante para FtsH5 apresenta características similares, porém menos pronunciadas (SAKAMOTO *et al.* 2002). Não obstante os mutantes para FtsH8 e FtsH1 se apresentaram como plantas selvagens (SAKAMOTO *et al.* 2003). O padrão variegado e o fato de que a variação diminui ao longo do desenvolvimento da planta (ZALTSMAN *et al.* 2005a), sugere que os genes FtsH possam ser redundantes. De fato a superexpressão de FtsH8 recuperou o fenótipo selvagem do mutante FtsH2 (YU *et al.* 2004), e a superexpressão de FtsH1 complementou o mutante FtsH5 (YU *et al.* 2005).

Mais conhecimento foi adquirido sobre o complexo através de uma série de análises de duplo mutantes (SAKAMOTO *et al.* 2002; ZALTSMAN *et al.* 2005b). A inativação dos pares intimamente relacionados FtsH1 e FtsH5 (Tipo A) e FtsH2 e FtsH8 (Tipo B) resulta em fenótipo albino heterotrófico (ZALTSMAN *et al.* 2005b). Além disso a ausência de um dos pares de genes, os produtos do outro par também não se acumularam como revelado por análises imunológicas (ZALTSMAN *et al.* 2005b).

Alguns autores realizaram estudos na tentativa de esclarecer as relações filogenéticas entre os genes da família gênica FtsH, e têm mostrado que os isômeros de FtsH aparentemente se diversificaram ao longo da evolução dos organismos fotossintéticos, o que resultou no surgimento dos subgrupos tipo A e B de FtsH (MARBACH; COELHO; SILVA-FILHO, 2001; YU; PARK; RODERMEL, 2004; KATO; KOUSO; SAKAMOTO, 2012; ADAM *et al.*, 2005; CHEN *et al.*, 2000). Kato, Kouso e Sakamoto (2012) demonstraram por meio de análise filogenética que os isômeros não estão apenas presentes em plantas superiores, mas também em musgos, algas vermelhas, algas verdes e cianobactérias (Figura 1) e, portanto, é muito provável que eles se localizem nas membranas dos tilacoídes e formem heterocomplexos. De acordo com a informação genômica disponível, o número de isômeros em cada subgrupo aparentemente varia (KATO; KOUSO; SAKAMOTO, 2012).

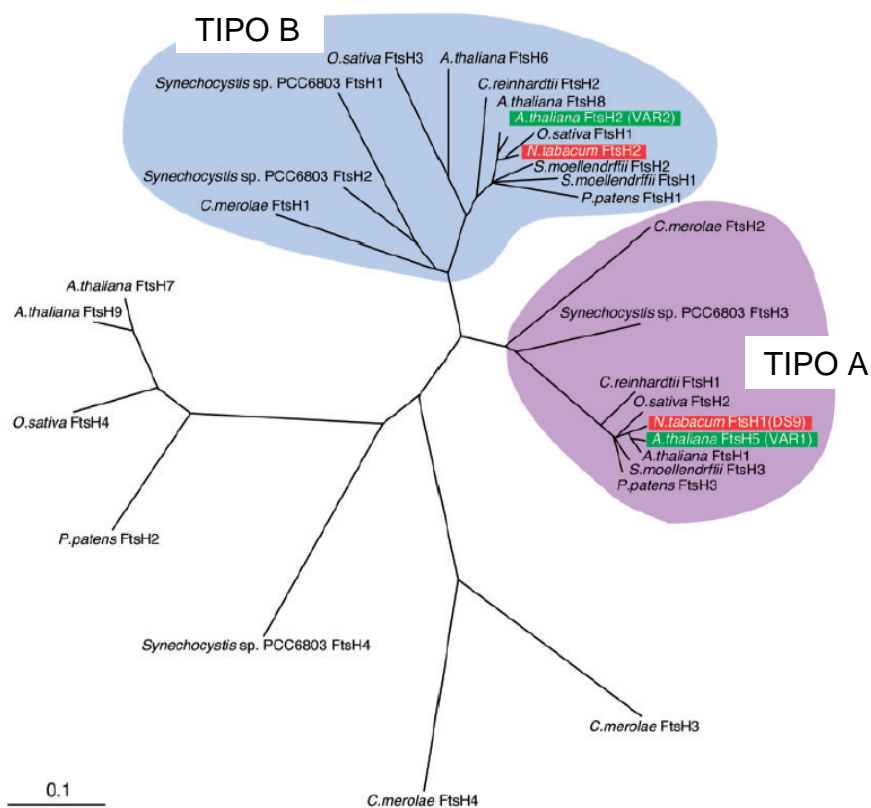


Figura 1 – Proposição filogenética para os homólogos de FtsH. A árvore filogenética foi feita através do método de Neighbor-Joining, baseada em um alinhamento dos homólogos feita por ClustalW. Em roxo as proteases FtsH tipo A e em azul as proteases FtsH tipo B. Extraída e adaptada de: Kato, Kouso & Sakamoto (2012)

Recentemente, nosso grupo demonstrou através de ensaios de importação que as subunidades dos tipos A e B se integram na membrana do tilacoíde por

diferentes vias. As subunidades do tipo A são integradas pela via Sec, enquanto que as subunidades do tipo B são integradas pela via Tat (RODRIGUES *et al.* 2011). Os níveis de transcritos das subunidades A são reduzidos em 50% em comparação aos das subunidades B nos complexos FtsH, de modo que este complexo aparentemente consiste de duas subunidades A e 4 subunidades B (ZALTSMAN *et al.* 2005b).

O complexo FtsH localizado na membrana do tilacóide é responsável pela degradação da proteína D1 fotodanificada, centro de reação do fotossistema II (PSII) (LINDAHL *et al.* 2000, BAILEY *et al.* 2002), em conjunto com a protease Deg, pertencente a uma família de endoproteases do tipo serina-protease, associadas periféricamente às membranas de cloroplastos de *A. thaliana* (KATO; SAKAMOTO, 2009). Recentemente, o mesmo grupo de pesquisa mostrou que fragmentos de D1 gerados por Deg podem ser substratos do complexo FtsH. Estes novos resultados suportam o papel cooperativo das proteases cloroplastidiais no processo de degradação de D1 danificada (KATO *et al.* 2012).

Interessantemente, os mutantes variegados *var1* e *var2* apresentam cotilédones verdes, mas as folhas subsequentes são variegadas com uma razão de decrescente de setores brancos para setores verdes (ZALTSMAN *et al.* 2005a). Enquanto que células com setores verdes nas folhas contêm cloroplastos aparentemente normais, os setores brancos apresentam cloroplastos não diferenciados (CHEN *et al.* 2000; TAKECHI *et al.* 2000). Isto sugere que o fenótipo variegado dos mutantes é resultado de problemas na formação dos tilacóides, o que está separado dos defeitos no *turnover* da proteína D1 (ZALTSMAN *et al.* 2005a; KATO; SAKAMOTO, 2009; MIURA; SAKAMOTO, 2007). Entretanto, eventos recentemente encontrados de mutantes de supressão da variação, *svr4*, sugerem uma ligação entre a variação e o dano fotooxidativo no PSII (YU *et al.* 2011). A identificação de linhas supressoras de *var2* demonstraram múltiplos mecanismos de compensação para a falta de AtFtsH2 durante os primeiros estágios de biogênese dos cloroplastos, revisado por Liu, Yu e Rodermeil (2010).

Foi demonstrado também que duplos mutantes para os isômeros do tipo B (FtsH2 e FtsH8), podem ser resgatados pela expressão de uma FtsH2 com a função proteolítica inativada, indicando que a atividade proteásica destes isômeros é dispensável para o desenvolvimento correto dos cloroplastos (ZHANG *et al.* 2010). Wagner e colaboradores (2011) revisaram o mecanismo de regulação das FtsHs e a

supressão da variegação. O autor especula que aparentemente o fenótipo variegado é independente da função de proteólise e que apenas a função de chaperona do complexo é necessária para a correta formação dos cloroplastos. Contudo o mecanismo de variegação ainda não foi resolvido e existem algumas linhas de evidências indicando que um nível limiar de FtsHs é necessário para a correta formação dos cloroplastos (YU *et al.* 2005; ZALTSMAN *et al.* 2005a; KATO; SAKAMOTO, 2007). Na verdade, este limiar depende da fase de desenvolvimento e de condições ambientais e o destino das células é determinado nos primeiros estágios de desenvolvimento dos plastídeos (YU *et al.* 2005; ZALTSMAN *et al.* 2005a; KATO; SAKAMOTO, 2007).

Adicionalmente foi observado em espinafre que a proteína D1 é degradada pelo complexo FtsH após estresse térmico (YOSHIOKA *et al.* 2006). Além disso, mostrou-se que a protease FtsH degrada a proteína Rieske FeS em ervilha (OSTERSETZER; ADAM, 1997) e regula os níveis de citocromo b_6 em *Chlamydomonas* (MALNOE *et al.* 2011). O complexo FtsH ainda apresenta outras funções, como por exemplo em *Capsicum* onde está envolvido na fusão de vesículas (HUGUENEY *et al.* 1995; SUMMER *et al.* 2000) e em *Nicotiana* onde pode estar relacionado ao desenvolvimento da reação de hipersensibilidade causado pelo TMV (*Tobacco mosaic virus*; SEO *et al.* 2000).

2.4 – Regulação da atividade das proteases FtsH

O controle da qualidade de proteínas possui uma relação com o tempo gasto pela proteína para chegar a sua conformação adequada. Um polipeptídeo recém sintetizado deve se dobrar corretamente, além de sofrer modificações e se associar a cofatores e outras subunidades a fim de formar uma proteína funcional. Este dobramento pode ser independente ou necessitar do auxílio de chaperonas moleculares, contudo proteínas mal dobradas, assim como as proteínas danificadas e não nativas devem ser degradadas por proteases. Tais proteases apresentam mecanismos de regulação de sua atividade de modo a impedir a sua atuação sobre proteínas nativas e direcionar sua atividade sobre as proteínas adequadas. As FtsHs como componentes do sistema de controle organelar de proteínas, da mesma forma possuem um mecanismo regulatório para a sua atividade proteolítica.

Estes mecanismos regulatórios muitas vezes envolvem a participação de outras proteínas. Em *E. coli* por exemplo o complexo FtsH está associado ao complexo proteico HflKC, o qual não é necessário para a atividade proteásica, como demonstrado em ensaios *in vitro* onde o complexo FtsH isolado apresenta atividade. De fato, nos mutantes hflK e hflC tem-se um aumento na atividade proteásica do complexo sobre os substratos SecY (AKIYAMA *et al.*, 1996) e a proteína cII do bacteriófago λ (KIHARA *et al.* 1997). Ao passo que o complexo HflKC inibe a degradação de SecY por FtsH *in vitro*. Tomados juntos estas evidências sugerem que a atividade proteolítica de FtsH é modulada negativamente pela associação com HflKC em *E. coli* (AKIYAMA 2009; ITO; AKIYAMA, 2005).

O mecanismo de regulação do complexo FtsH em cloroplastos ainda não foi elucidado, mas algumas linhas de evidências sugerem que a ativação do complexo está relacionada a alta incidência luminosa ou a outras condições de estresse. Yoshioka e Yamamoto (2011) propuseram alguns mecanismos, um dos quais que permitiria uma ativação reversível seria a oligomerização das subunidades FtsHs para a formação do hexâmero ativo e a inativação com a sua desorganização. Foi demonstrado pelo mesmo grupo de pesquisa que as proteínas FtsHs não formam os hexâmeros na região estromal das lamelas dos tilacóides, sendo que a oligomerização somente ocorre na região dos grana (YOSHIOKA *et al.* 2010). Entretanto, as membranas dos grana, em condições normais, seriam muito densamente comprimidas para permitir a formação dos hexâmeros, portanto, os autores argumentam que a ação das FtsHs deve ocorrer nas margens dos grana ou em condições em que as membranas dos tilacóides estão menos comprimidas, como por exemplo com alta luminosidade ou estresse térmico.

Uma segunda possibilidade é a da sensibilidade da cauda luminal ao pH do tilacóide. Demonstrou-se anteriormente que em *E. coli* a região periplasmática de FtsH, equivalente a região exposta ao lúmen em tilacóides, apresenta papel crucial na interação entre os monômeros e na modulação da atividade de proteólise (AKIYAMA *et al.* 1998). A acidificação do lúmen dos tilacóides em condições de iluminação deve protonar alguns resíduos da calda luminal, promovendo alterações conformacionais que favorecem a formação do complexo (YOSHIOKA; YAMAMOTO, 2011).

Por fim, também é possível que espécies reativas de oxigênio (ROS), produzidas em condições de alta luminosidade, atuem na oligomerização do

complexo. Sabe-se que as ROS causam danos oxidativos a proteínas e que elas atuam como sinais de ativação para resposta ao estresse. É possível que a síntese das FtsHs seja ativada pela presença de ROS ou mesmo que essas moléculas promovam modificações na estrutura das proteases, induzindo a oligomerização (YOSHIOKA; YAMAMOTO, 2011).

Uma hipótese alternativa a se considerar é a presença de fatores proteicos adicionais, que atuam associados ao complexo na modulação de sua atividade. Nesse sentido nosso próprio grupo de pesquisa, buscando a identificação de novas proteínas interagentes envolvidas, utilizou a protease FtsH5 como isca no sistema de duplo híbrido de leveduras, como um passo inicial no entendimento do mecanismo regulatório da atividade proteásica do complexo FtsH em cloroplastos (RODRIGUES, 2011). Dentre as diversas proteínas plastidiais encontradas, chamou atenção uma pequena proteína hipotética, de localização cloroplastidial e função ainda não determinada, a qual foi denominada FIP (*FtsH5 Interacting Protein*).

2.5 – FIP (*FtsH5 Interacting Protein*)

A análise *in-silico* de FIP revelou que tal proteína pode ser dividida em três regiões distintas: um peptídeo de trânsito clivável N-terminal, uma região hidrofóbica para ancoragem na membrana e um domínio rico em cisteína (CXXCXGXG/ onde X é qualquer aminoácido) C-terminal. Tais domínios são característicos de proteínas DnaJ (SHI *et al.*, 2005), onde estão repetidos quatro vezes formando uma estrutura em “dedo de zinco”, formada por dois centros de ligação ao zinco. Sendo que o primeiro centro é responsável pela atividade chaperona e o segundo pela interação com proteínas DnaK (TANG & WANG, 2001). As proteínas DnaK são chaperonas de 70 KDa que se ligam a regiões hidrofílicas de peptídeos não montados, auxiliando no dobramento por meio de ciclos de ligação e desligamento com a energia proveniente da hidrólise do ATP. As proteínas DnaJ com 40 kDa também são chaperonas e atuam na regulação da atividade das DnaKs, principalmente pela ligação do domínio J com a região ATPásica da DnaK (WALSH *et al.*, 2004).

Também foi mostrado pelo nosso grupo de pesquisa que FIP não interage com FtsH2 pelo sistema de duplo híbrido em leveduras (RODRIGUES, 2011). Como já discutido, a atividade proteásica das proteínas do tipo B é totalmente dispensável, embora sejam necessárias para a montagem do complexo (ZHANG *et al.*, 2010). A

especificidade da interação de FIP com proteases do tipo A pode estar relacionada diretamente com a função proteásica destas.

A interação de FIP e FtsH5 também foi mostrada por um experimento *GST-Pull Down* (RODRIGUES, 2011), além disso experimentos de importação *in-vitro* utilizando FIP marcada radiotivamente, seguidos de eletroforese em géis nativos, mostraram que FIP está associada a um complexo de 440 kDa (RODRIGUES, 2011), que é o tamanho esperado para o complexo FtsH nos tilacóides (SAKAMOTO *et al.*, 2003). A localização plastidial foi determinada pela fusão da proteína FIP a proteína verde fluorescente (GFP; RODRIGUES, 2011). O papel de FIP ainda é incerto, mas os dados indicam que FIP possa regular de alguma forma a atividade ATPásica ou proteásica do complexo FtsH. A metaloprotease de membrana ADAM, presente em *Xenopus*, está envolvida na modificação do comportamento celular através da alteração de superfície celular e matriz extracelular. ADAM, diferentemente de FtsH5, apresenta um domínio rico em cisteína, o qual atua de maneira intramolecular na regulação *in vivo* do domínio proteásico (SMITH *et al.*, 2002).

Recentemente nosso grupo de pesquisa mapeou a região de interação de FIP com FtsH5 utilizando o sistema de duplo híbrido de leveduras (BRAGA, 2013). Para isso utilizou-se deleções da proteína FIP, onde demonstrou-se que a interação foi mantida com os 46 aminoácidos que compreendem os dois motivos ricos em cisteína e a região localizada entre eles. Com o objetivo de verificar se a interação de FtsH5 era mantida mesmo quando as cisteínas eram substituídas, realizou-se 7 combinações de substituições diferentes de cisteína por alanina, resultando que nenhuma substituição foi capaz de manter a interação.

Braga (2013) realizou ainda experimentos de análise de expressão sobre duas condições de estresse abiótico: estresse salino e estresse a frio. Os valores de expressão de FIP e FtsH5 responderam de maneira antagônica ao estresse, com diminuição da expressão de FIP e aumento da expressão de FtsH5 em ambas condições de estresse. Tais resultados podem sugerir que FIP seja um regulador negativo do complexo FtsH.

3 MATERIAL E MÉTODOS

3.1 – Análise de expressão

3.1.1 – Expressão tecido-específica e temporal

Os níveis e padrões de expressão tecido-específica das proteínas do complexo FtsH, bem como de FIP e a magnitude de expressão das mesmas ao longo do desenvolvimento foram obtidos através de dados públicos provenientes dos bancos de dados do *Bio-Array Resource for Arabidopsis Functional Genomics* (BAR) (<http://bar.utoronto.ca>) e do *Genevestigator* (<http://www.genevestigator.ethz.ch>), respectivamente. As imagens pictográficas representando os padrões e níveis da expressão tecido-específica foram criados utilizando a ferramenta eFP browser (WINTER *et al.*, 2007; <http://bar.utoronto.ca>), enquanto que os níveis de expressão ao longo do desenvolvimento foram acessados através da utilização de todos os dados disponíveis para plantas selvagens no *Genevestigator* e com as ferramentas do próprio *webserver* (ZIMMERMANN *et al.*, 2004; <http://www.genevestigator.ethz.ch>).

3.1.2 – Expressão em resposta a estresses abióticos

Para a análise dos padrões de expressão de FIP e das proteases do complexo FtsH, foram utilizados dados públicos de experimentos de microarranjo de DNA, referentes ao projeto AtGenExpress (KILIAN *et al.* 2007; <https://www.arabidopsis.org/portals/expression/microarray/ATGenExpress.jsp>) depositados no banco de dados do TAIR (*The Arabidopsis Information Resource* – www.arabidopsis.org). Este projeto desenvolveu um estudo da expressão do genoma de *A. thaliana* bastante compreensível e completo, através da plataforma Affymetrix ATH1, com a utilização de plantas de genótipo idêntico, crescidas sobre condições idênticas e em diferentes estresses ambientais (KILIAN *et al.* 2007).

Na análise dos dados, os arquivos de leitura de intensidade de célula (CEL) foram analisados com o pacote R/Bioconductor (LI; WONG, 2001), normalizados pelo método de análise robusta de multiarranjo (RMA) (IRIZARRY *et al.*, 2003), que está disponível no pacote R/Bioconductor para arranjos da Affymetrix (KERR *et al.*,

2000). Foi utilizada a biblioteca 'limma' do Bioconductor para análise dos dados de expressão, o pacote 'hclust' para agrupamento dos genes com perfis de expressão mais correlacionados (em termos de *fold change*, contraste entre controle e tratamento) e os dados foram plotados em formato de histograma, através do software Gnuplot (<http://www.gnuplot.info/>). Foi utilizado o valor do \log_2fc (logaritmo na base 2 do contraste entre tratamento e controle) para comparação dos dados.

3.2 – Análises filogenéticas

3.2.1 – Filogenia dos homólogos de FIP e das proteases do complexo FtsH tilacoidal e inferência sobre o processo de coevolução

3.2.1.1 – Identificação e obtenção das sequências para a análise

As sequências correspondentes aos homólogos de FIP foram encontradas através do algoritmo "blastn", utilizando-se a sequência de nucleotídeos (sequência codante) de FIP (At5g02160) como *query* (pergunta). Para se obter as sequências, FIP foi comparada as sequências depositadas no banco de dados não redundante (nr/nt) do *GENBANK* (<http://www.ncbi.nlm.nih.gov/genbank>) e no banco de dados do *Phytozome* (<http://www.phytozome.net/>). No *GENBANK* selecionou-se apenas as sequências com *e-value* $\leq e^{-05}$, enquanto que no *Phytozome* a busca é feita para grupos de sequências consenso, sendo selecionadas para o presente trabalho apenas as sequências do grupo com *e-value* $\leq e^{-05}$ (no caso o grupo possuía um *e-value* = e^{-39}). As sequências identificadas como homólogos de FIP estão listadas no Anexo A.

A obtenção das sequências homólogas a FtsH se mostrou mais complexa devido à grande identidade de sequência existente entre as diferentes proteases FtsH, principalmente entre aquelas componentes do complexo FtsH tilacoidal. Para solucionar este problema, visou-se a obtenção e classificação de todos os homólogos das proteases do complexo FtsH nos tipos A e B. Desta forma, seria possível se estudar a filogenia de cada um dos tipos independentemente e avaliar as relações destes com a proteína FIP. Assim, as sequências de cada uma das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8: AT1G06430) foram utilizadas como *query* na busca por

homólogos. Através do algoritmo “blastn”, comparando-as uma a uma, com cada organismo onde havia sido identificado um homólogo de FIP, tanto no banco de dados não redundante (nr/nt) do *GENBANK* (<http://www.ncbi.nlm.nih.gov/genbank>) como no banco de dados do *Phytozome* (<http://www.phytozome.net/>). Selecionando-se apenas sequências com e-value=0 ou a sequência de maior identidade ($\geq 70\%$). As sequências dos homólogos das proteases FtsH do complexo tilacoidal podem ser observadas na Anexo B.

De modo a conferir um maior grau de confiabilidade a classificação dos homólogos das proteases do complexo FtsH nos tipos A e B, foram incluídas na análise destes, algumas sequências homólogas das FtsHs utilizadas anteriormente nesta classificação (KATO; KOUSO; SAKAMOTO, 2012). Estas sequências adicionais podem também ser encontradas na Anexo B.

3.2.1.2 – Alinhamento múltiplo das sequências e modelo de substituição de DNA

Os alinhamentos múltiplos de sequências (MSAs, do inglês *multiple sequence alignments*) dos homólogos de FIP e das proteínas do complexo FtsH foram realizados, independentemente, na versão *online* do programa *TranslatorX* (*TranslatorX server* - <http://translatorx.co.uk/>; ABASCAL; ZARDOYA; TELFORD, 2010), pelo método de alinhamento múltiplo “MAFFT”. Os MSAs foram visualizados no programa *Jalview* v.2.8.0 (WATERHOUSE *et al.* 2009) e foram eliminadas as sequências muito curtas, redundantes e largamente incompletas, bem como as partes pobremente alinhadas do início e do final das sequências, de modo a preservar no MSA, somente as sequências dos organismos com bons alinhamentos e representantes dos homólogos de FIP e das FtsHs.

O modelo apropriado de substituição de DNA foi identificado pelo programa *jModelTest* v.2.1.4 (DARRIBA *et al.* 2012) através da comparação dos modelos disponíveis usando o *Akaike Information Criterion* (AIC). O melhor modelo de substituição identificado tanto para os homólogos de FIP, quanto para os homólogos das FtsHs foi o *General Time Reversible* com uma proporção de sítios invariáveis e taxa de substituição com distribuição *gamma* entre sítios (GTR+I+G; TAVARÉ, 1986).

3.2.1.3 – Proposição de modelo evolutivo e inferência sobre a coevolução

Os MSAs e o modelo de substituição apropriado foram submetidos a proposição filogenética bayesiana, usando o *software* MrBayes (versão 3.2.1 *mpi-enable*; RONQUIST; HUELSENBECK, 2003), de modo a inferir a topologia da árvore filogenética e o comprimento dos ramos. Foram geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas. Destas estimativas 25% foram eliminadas como *burn-in*, para assegurar a amostragem das árvores após a convergência da topologia. Finalmente utilizou-se o *software* FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) para a análise, manipulação e produção das figuras das árvores filogenéticas geradas.

Após esta etapa foi possível classificar os homólogos das proteases do complexo FtsH nos tipos A e B, portanto a partir desse ponto se tornou possível a construção de modelos evolutivos para as FtsHs do tipo A e do tipo B isoladamente, através de novas análises filogenéticas bayesianas independentes. Assim as sequências de ambos os tipos foram identificadas e selecionadas do conjunto inicial contendo todos os homólogos, de modo a conduzir todas as etapas de alinhamento e curagem das sequências independentemente. As sequências foram novamente alinhadas pelo *software* TranslatorX (*TranslatorX server* - <http://translatorx.co.uk/>; ABASCAL; ZARDOYA; TELFORD, 2010), pelo método de alinhamento múltiplo “MAFFT” e os MSAs visualizados e curados no programa *Jalview* v.2.8.0 (WATERHOUSE *et al.* 2009), exatamente como descrito anteriormente. Novamente o modelo de substituição de DNA foi determinado pelo programa *jModelTest* v.2.1.4 (DARRIBA *et al.* 2012) através da comparação dos modelos disponíveis usando o *Akaike Information Criterion* (AIC), sendo ele o modelo GTR+I+G. A proposição filogenética bayesiana foi, mais uma vez, feita com o *software* MrBayes (versão 3.2.1 *mpi-enable*; RONQUIST; HUELSENBECK, 2003), sendo geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas, com 25% das árvores eliminadas como *burn-in*. A análise, manipulação e produção das figuras das árvores filogenéticas foram realizadas no *software* FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

Em seguida as árvores filogenéticas foram utilizadas para a inferência sobre a existência de coevolução, entre os pares FIP e FtsHs tipo A, FIP e FtsHs tipo B e, por último, FtsHs tipo A e FtsHs tipo B, com o uso do *webserver* MirrorTree Server ([http:// http://csbg.cnbc.csic.es/mtserver/index.php/](http://http://csbg.cnbc.csic.es/mtserver/index.php/); OCHOA; PAZOS, 2010). O método de *MirrorTree* (homônimo ao *software* utilizado) consiste em quantificar a similaridade entre as árvores das famílias gênicas de interesse, o que é feito da seguinte forma: as matrizes de distâncias são obtidas a partir do resumo das comprimentos dos ramos que separam cada um dos pares de genes na árvore e a similaridade das árvores é calculada como uma correlação entre as matrizes de distâncias de acordo com a seguinte equação padrão:

$$r = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

Onde n é o número de elementos das matrizes, ou seja, $n = (N^2 - N)/2$ sendo N o número de organismos comuns entre as duas filogenias, R_i são os elementos da primeira matriz de distância (as distâncias entre todos os genes na primeira filogenia), S_i são os valores correspondentes a segunda matriz e R e S são as médias de R_i e S_i , respectivamente.

3.2.2 – Filogenia representativa dos homólogos das FtsH do tipo A

De modo a investigar mais profundamente e a corroborar os resultados das análises filogenéticas anteriores, visou-se a proposição de um modelo evolutivo mais completo para os homólogos das FtsHs do tipo A, no qual estariam incluídos membros do maior número de táxons possíveis, dentro de um limite razoável e que fossem representativos da filogenia completa dos homólogos das FtsHs. O objetivo era determinar se o grupo de espécies portadoras de homólogos de FIP, seria

recuperado como um grupo monofilético nesta filogenia mais completa dos homólogos de FtsH do tipo A.

Para tal utilizou-se as sequências de aminoácidos das FtsHs do tipo A de *Arabidopsis thaliana* (FtsH1: At1g50250; FtsH5: At5g42270) como *query*, através do algoritmo 'blastp', contra o banco de dados *refseq_protein* do GENBANK (<http://www.ncbi.nlm.nih.gov/genbank>). Foram selecionadas somente as sequências com e-value $\leq 1e-180$ (2073 sequências selecionadas).

O alinhamento múltiplo das sequências foi gerado no *software* MAFFT v.7 (KATO; STANDLEY, 2013) e analisado no programa *Jalview* v.2.8.0 (WATERHOUSE *et al.* 2009), onde foram eliminadas sequências muito curtas, redundantes e largamente incompletas, bem como regiões pobremente alinhadas nas extremidades. Além disso o MSA resultante foi tratado com o *software* Gblocks (com parâmetros padrão; TALAVERA; CASTRESANA, 2007; CASTRESANA, 2000), o qual elimina posições pobremente alinhadas e regiões divergentes, as quais, convenientemente, devem ser eliminadas antes de uma análise filogenética. Os blocos selecionados devem preencher alguns requisitos com respeito a ausência de grandes segmentos de posições não conservadas contíguas, ausência de *gaps* e alta conservação das posições flangeadoras, tornando o alinhamento final mais apropriado para análises filogenéticas.

O modelo de substituição de aminoácidos apropriado foi determinado pelo *software* Prottest v.3 (ABASCAL; ZARDOYA; POSADA, 2005), seguindo o parâmetro *Akaike Information Criterion* (AIC). O melhor modelo de substituição identificado foi o WAG (Matriz Whelan and Goldman; WHELAN; GOLDMAN, 2001). Utilizou-se o *software* MrBayes (versão 3.2.1 *mpi-enable*; RONQUIST; HUELSENBECK, 2003) para a proposição filogenética bayesiana, sendo que foram geradas 1×10^6 gerações de MCMC e amostradas a cada 1000 gerações, produzindo 1000 estimativas. Foram eliminadas 25% das árvores como *burn-in*, para assegurar a amostragem das árvores após a convergência da topologia. Por fim, o *software* FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) foi usado para a análise, manipulação e produção da figura da árvore filogenética.

3.3 – Detecção de sítios de pressão de seleção

De modo a inferir o tipo de pressões seletivas e forças adaptativas aos quais os homólogos de FIP e das proteases do tipo A do complexo FtsH tilacoidal, bem como detectar a presença ou ausência de sítios seletivos nas sequências destes, para tal utilizou-se o *software* CodeML do pacote PAML (Phylogenetic Analysis by Maximum Likelihood; YANG; NIELSEN, 2007) para produzir os testes evolutivos e a detecção sítio por sítio por análises de Máxima Verossimilhança. Os alinhamentos múltiplos de sequências e as árvores filogenéticas utilizados na análise foram os mesmos produzidos nas seções 4.1.1.2 e 4.1.1.3 para FIP e para as FtsHs do tipo A, respectivamente.

Uma medida de seleção de uma sequência, é a razão entre as substituições de nucleotídeos não-sinônimas (d_N), cujos resultados conduzem a uma mudança de códon e de seu aminoácido correspondente e as substituições sinônimas (d_S), as quais por sua vez são as mudanças de nucleotídeos que alteram apenas o códon e não o aminoácido e portanto se comportam como neutras, se fixando em proporção ao tempo de divergência entre as sequências. Isto porque, considerando-se que as substituições não-sinônimas (d_N) alteram a sequência de aminoácidos e que a função das proteínas depende de sua estrutura, este parâmetro está frequentemente sobre o filtro da seleção natural. Segue conseqüentemente que a razão entre as taxas de substituição não-sinônima e sinônima ($\omega = d_N/d_S$) pode ser considerada uma medida convincente da seleção (YANG; BIELAWSKI, 2000; HURST, 2002).

Portanto, a razão ω foi mensurada como uma medida da seleção natural sobre as proteínas e de acordo com tal modelo, as categorias de seleção são definidas como: 1) seleção positiva ($\omega > 1$): a variabilidade de aminoácidos na posição é mais frequente do que o esperado por deriva genética; 2) seleção negativa ($\omega < 1$): forças evolutivas tem atuado contra mudanças no aminoácido que ocupa essa posição; 3) seleção neutra ($\omega = 1$): a variabilidade de aminoácidos que ocupam essa posição é tão frequente quanto o esperado por deriva genética (YANG; NIELSEN, 2002).

Três testes de Máxima Verossimilhança foram produzidos, usando os modelos de substituição descritos na Tabela 1, de modo a possibilitar a detecção de pressão de seleção sítio por sítio. Primeiramente, o modelo de razão única de ômega (M0) foi testado e comparado ao modelo de variáveis discretas (M3 – com K=3 classes discretas), testando assim a presença de pressão de seleção variável

entre os sítios. Em seguida, outros dois pares de modelos foram utilizados para testar a probabilidade de seleção positiva nos sítios de códons (M1a X M2a; M7 X M8). O modelo aproximadamente neutro (M1a) - inclui duas categorias de sítios, são elas: seleção neutra ($\omega = 1$) e seleção negativa ($0 < \omega < 1$); e foi comparado ao modelo de seleção M2a, o qual contém uma classe adicional de sítios: seleção positiva ($\omega > 1$). Por fim, o modelo beta M7 que utiliza distribuição beta para ω (no intervalo de 0-1), foi comparado ao modelo M8, com distribuição beta para ω , mais a classe de sítios selecionados positivamente ($\omega > 1$). A comparação foi feita por meio de testes de razão de verossimilhança (LRT; do inglês *Likelihood Ratio Test*) para identificar o ajuste dos modelos mais próximos, com 4 graus de liberdade para a comparação M0/M3 e 2 para as comparações M1a/M2a e M7/M8. Quando os testes são significativos o modelo mais simples ou nulo (M0, M1a ou M7) deve ser rejeitado em favor do mais complexo (M3, M2a ou M8). Além disso, o teste Bayes Empirical Bayes (BEB) foi aplicado para calcular a probabilidade posterior de cada classe de sítio ($\omega < 1$, $\omega = 1$ e $\omega > 1$).

Tabela 1 – Parâmetros utilizados nos modelos de substituição por sítio

Modelo	Nº de sítios variáveis	Nº de parâmetros livres	Parâmetros
M0 (Razão única)	0	1	ω
M1a (neutro)	1	2	$p_0(p_1 = 1 - p_0)$, $\omega_0 < 1, \omega_1 = 1$
M2a (seleção)	2	4	$p_0, p_1 (p_2 = 1 - p_0 - p_1)$, $\omega_0 < 1, \omega_1 = 1, \omega_2 > 1$
M3 (discreto)	3	5	$p_0, p_1 (p_2 = 1 - p_0 - p_1)$ $\omega_0, \omega_1, \omega_2$
M7 (beta)	7	2	p, q
M8 (beta e ω)	8	4	$p_0, (p_1 = 1 - p_0)$, $p, q, \omega_S > 1$

*Os parâmetros nos parênteses não são livres

3.4 – Análise de conservação de sequência e estrutural do complexo FtsH

Buscou-se identificar a estrutura do complexo FtsH tilacoidal através do método de modelagem por homologia, utilizando-se para isto da sequência de aminoácidos da proteína FtsH5 de *Arabidopsis thaliana* (At5g42270), submetida ao software HHpred (BIEGERT *et al.*, 2006). Este software realiza uma busca de

sequências molde através da detecção de homologia de proteínas e predição de estrutura pelo método de comparação HMM-HMM. O HHpred parte de uma sequência *query* e constrói um alinhamento múltiplo de sequência usando a ferramenta HHblits (BIEGERT *et al.*, 2006), criando um perfil HMM. Este perfil é então comparado a um banco de dados de HMMs representando proteínas com estrutura conhecida (e.g. PDB, SCOP) ou famílias de proteínas anotadas (e.g. PFAM, SMART, CDD, COGs, KOGs). O resultado é uma lista dos homólogos mais próximos com alinhamento.

Usualmente, uma vez identificado o molde mais adequado para a modelagem da proteína (o qual é selecionado com base não somente na identidade e cobertura sobre a sequência de interesse, mas também em outros critérios como homologia de sequência secundária e ausência de gaps), pode-se passar para a etapa de modelagem da estrutura sobre o molde selecionado. Contudo, no presente trabalho devido a inviabilidade em se modelar o complexo hexamérico das proteases FtsH, no qual ainda não é conhecida a proporção exata de seus componentes (razão entre os tipos A e B de FtsH e a capacidade de substituição entre os pares dos dois tipos), decidiu-se utilizar o melhor molde como modelo para a análise, isto porque como dito anteriormente, é possível assumir uma estrutura similar entre o complexo procariótico e plastidial devido a grande similaridade das sequências de FtsHs procarióticas e plastidiais (BIENIOSSEK *et al.*, 2009). O modelo selecionado foi a estrutura cristalizada completa da região citosólica do complexo FtsH de *Thermus thermophilus* (PDB: 2DHR) com resolução de 3.9 Å e com média de 58 % de identidade.

Primeiramente, utilizou-se o Webserver ConSurf (<http://consurf2.tau.ac.il/>; Celniker *et al.*, 2013; ASHKENAZY *et al.*, 2010), para estimar a conservação evolutiva de posições de aminoácidos nos homólogos de FIP e das proteases FtsH do tipo A. No ConSurf, a taxa evolutiva é estimada com base no parentesco evolutivo entre uma proteína e seus homólogos, considerando que a similaridade entre os aminoácidos reflete as matrizes de substituição. Para tal, utilizaram-se MSAs de aminoácidos dos produtos dos genes usados nos MSAs de nucleotídeos produzidos nas seções 4.1.1.3 para as FtsHs do tipo A. Os escores de conservação de posição-específica (do inglês *position-specific*), foram calculados utilizando algoritmos bayesianos empíricos e com o modelo de substituição WAG (Matriz Whelan e Goldman; WHELAN; GOLDMAN, 2001). A distribuição contínua dos

escores de conservação é então dividida em nove classes para a visualização, começando pelas posições mais variáveis coloridos de azul turquesa, passando pelas posições conservadas intermediárias coloridas de branco, até as posições mais conservadas em castanho-avermelhado. O *software* então projeta os escores e as classes sobre as sequências de aminoácidos no MSA e gera um script para a colorização e visualização da análise de conservação na estrutura tridimensional da proteína. Neste trabalho, utilizou-se o programa PyMOL para a visualização e análise das estruturas tridimensionais de proteínas, bem como para a produção de imagens destas.

4 RESULTADOS E DISCUSSÃO

4.1 – FIP como um regulador negativo putativo da atividade do complexo FtsH tilacoidal

4.1.1 – Perfil de expressão tecido específico e ao longo do desenvolvimento

Primeiramente, buscou-se avaliar os níveis e padrões de expressão tecido específicos e ao longo do desenvolvimento de FIP e das proteases do complexo FtsH de *Arabidopsis thaliana*, de modo a investigar e buscar suporte para a hipótese de regulação da atividade do complexo FtsH por parte da proteína FIP. Para tal, foram analisados dados públicos de experimentos de microarranjo do consórcio *Bio-Array Resource for Arabidopsis Functional Genomics* (BAR) (<http://bar.utoronto.ca>) e do banco de dados do *Genevestigator* (<http://www.genevestigator.ethz.ch>), o qual integra dados dos bancos Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) e ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>), dentre outros.

Nos perfis de expressão ao longo do desenvolvimento (Figura 1) é possível observar que todos os transcritos aqui analisados apresentam uma tendência de aumento a partir dos primeiros estágios até o período de produção dos pendões e florescimento, decrescendo sistematicamente após o florescimento até a produção de sílicas e sementes. O que faz sentido, levando-se em consideração o papel do complexo FtsH no controle de qualidade das proteínas do cloroplasto, tais como a proteína D1 do PSII, a qual deve ser degradada frente a danificação desta, seja por estresse luminoso ou pelos radicais livres gerados pelo processo normal de fotossíntese (LINDAHL *et al.* 2000, BAILEY *et al.* 2002) e também alinha-se com as evidências que indicam uma possível função do complexo na formação e desenvolvimento dos cloroplastos (YU *et al.* 2005; ZALTSMAN *et al.* 2005a; KATO; SAKAMOTO, 2007), pois correlaciona um aumento da expressão dos genes do complexo FtsH com o crescimento e desenvolvimento da planta e consequente aumento do número de cloroplastos e da necessidade de lidar com os danos causados pela fotossíntese e outros estresses que atingem o cloroplasto, com os níveis de expressão somente decaindo após o florescimento (Figura 2). Além disso, a proteína FIP como regulador putativo da atividade do complexo tem a sua

expressão notadamente correlacionada àquela das FtsHs do tipo A (FtsH1 e FtsH5), inclusive em termos dos níveis de expressão dos transcritos (Figura 2).

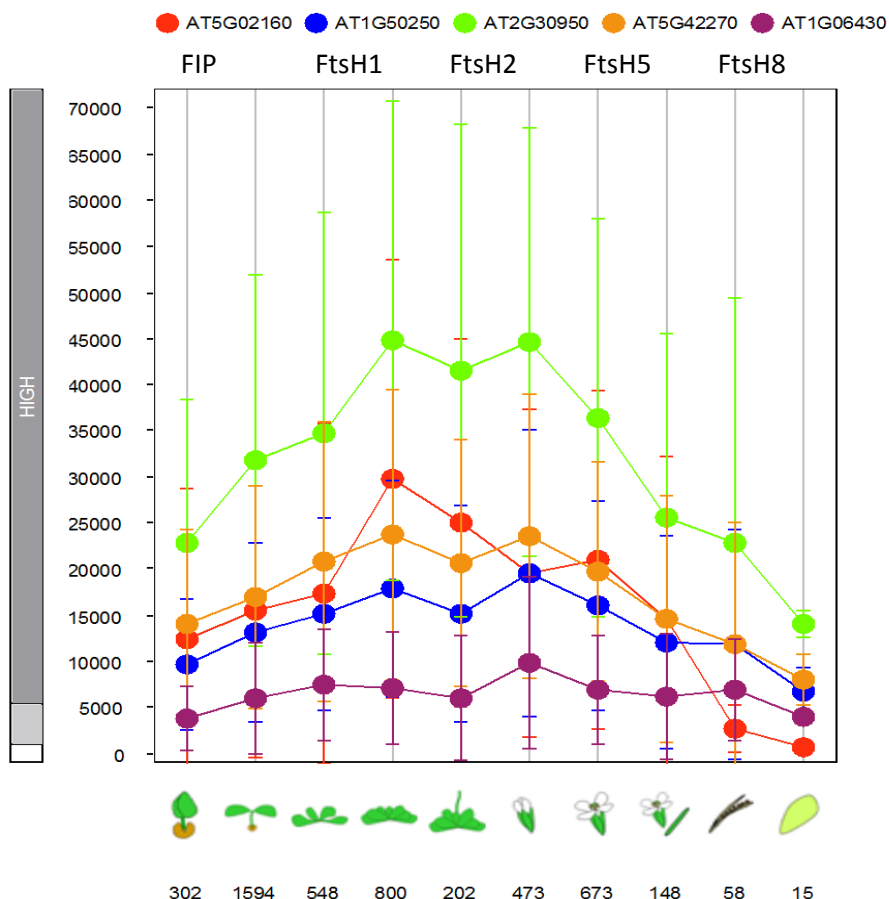


Figura 2 – Expressão de FIP e das proteases do complexo FtsH ao longo do desenvolvimento de *Arabidopsis thaliana*. Os estágios de desenvolvimento estão representados ao longo do eixo X (em ordem da esquerda para a direita: Estágio 0 (1-5.9 dias); Estágio 1 (6-13.9 dias); Estágio 1b (14-17.9 dias); Estágio 1c (18-20.9 dias); Estágio 1d (21-24.9 dias); Estágio 5 (25-28.9 dias); Estágio 6 (29-35.9 dias); Estágio 6b (36-44.9 dias); Estágio 8 (45-50 dias); Estágio 9 (>)), juntamente com o número de amostras usadas para estimar os valores de expressão de cada gene. O eixo Y representa a escala de expressão dos genes, a expressão de cada gene é informada como intensidade de sinal no arranjo 22K. Acima do gráfico a legenda determina a correspondência entre os genes e as cores utilizadas para representá-los no gráfico

Chama a atenção também o fato de que os níveis de expressão das proteases do complexo FtsH observados nos dados públicos de experimentos de microarranjo, e portanto aqui expressos em termos de acúmulo de transcritos (Figura 2), remontam os níveis de expressão de proteínas conforme acessados por experimentos de *immunoblot* por Sinvany-Villalobo e colaboradores (2004).

Os níveis e padrões de expressão de FIP e das proteases do complexo FtsH (Anexos C a G), evidenciam a predominância de expressão destes genes nos

tecidos da parte aérea da planta, notadamente nas folhas da roseta. O restante dos tecidos como sementes, sílicas e tecidos da parte floral apresentam relativamente expressões bem mais reduzidas destes genes. Novamente reproduzindo o mesmo padrão encontrado anteriormente, FtsH2 como a mais abundante de todas as proteases, FIP com níveis de expressão semelhantes aqueles encontrados em FtsH5, e FtsH8 e FtsH1 com os menores níveis de expressão (Anexos C a G). Além disso, os níveis de expressão de FtsH5 e de FIP são bastante similares e sobrepostos, corroborando o possível papel de FIP na regulação do complexo FtsH e sugerindo que esta regulação pode ser ubíqua nos tecidos da parte aérea de *Arabidopsis thaliana* (Anexos C e F).

4.1.2 – Perfil de expressão de FIP e das proteases do complexo FtsH em resposta a diversos estresses abióticos

O complexo FtsH tilacoidal sabidamente está envolvido no sistema de controle de qualidade de proteínas de cloroplastos e portanto está diretamente relacionado a resposta aos estresses ambientais, através de sua atividade proteolítica e de chaperona (chaperone-like) e principalmente pelo *turnover* de proteínas danificadas e desnaturadas, com alguns exemplos descritos na literatura (LINDAHL *et al.* 2000; BAILEY *et al.* 2002; YOSHIOKA *et al.* 2006; OSTERSETZER; ADAM 1997; MALNOE *et al.* 2011; dentre outros). Evidencia-se assim o papel multifuncional das proteases FtsH nas células. Desta forma o complexo, logicamente, deve apresentar alguma forma de modulação de sua atividade por meio da qual possa responder de maneira adequada as diferentes condições ambientais e auxiliar na manutenção da integridade dos cloroplastos frente à ocorrência de estresses.

Por esse motivo decidiu-se investigar os perfis de expressão dos genes do complexo FtsH, bem como do gene de FIP frente a diferentes estresses abióticos, de modo a inferir sobre as relações funcionais existentes entre estes e fornecer evidências para a hipótese da atuação de FIP sobre a regulação da atividade do complexo FtsH, assim como para a natureza funcional desta atuação. Nesse sentido, foram utilizados os dados públicos de experimentos de microarranjo de DNA do projeto AtGenExpress (KILIAN *et al.* 2007; <https://www.arabidopsis.org/portals/expression/microarray/ATGenExpress.jsp>), devido a

alta confiabilidade, amplitude e detalhamento do mesmo. O projeto AtGenExpress consistiu em um estudo bastante rigoroso e completo da expressão transcricional do genoma de *A. thaliana* frente a diferentes condições de estresse abiótico, além de ter sido realizado de forma a minimizar todas as possíveis fontes de variação indesejadas, através da utilização de plantas de genótipo idêntico, crescidas sobre condições idênticas, salvo as variáveis de estresse abiótico de interesse para cada condição avaliada (KILIAN *et al.* 2007).

Os perfis de expressão para as diferentes condições de estresse abiótico aqui avaliadas (Estresse a frio, à seca, a calor, osmótico, salino, oxidativo, à UV-B, por ferimento e genotóxico) podem ser observadas nas Figuras 3 a 5, bem como as análises de agrupamento hierárquico de cada perfil. De modo geral, FIP apresenta uma grande redução em seus níveis de expressão em todos os estresses, sistematicamente ao longo da série temporal (Figuras 3-5). Enquanto que as proteínas do complexo tendem a aumentar os níveis de expressão até as 6 horas de tratamento, seguido de uma redução gradativa do aumento da expressão, chegando até mesmo a ter uma expressão reduzida em alguns estresses, principalmente as 12 e 24 horas (Figuras 3-5). O que pode estar relacionado a algum mecanismo de feedback negativo da expressão destas, através da percepção dos níveis ou transcritos das proteases do complexo nos cloroplastos.

Para comparar os comportamentos da expressão dos genes ao longo da série temporal, utilizou-se a análise de agrupamento hierárquico com a correlação de Pearson. A análise agrupou FIP separadamente do restante das proteínas pertencentes ao complexo FtsH para a maioria dos estresses abióticos aqui analisados, com exceção dos estresses salino, osmótico e radiação ultravioleta (UV-B; Figuras 3-5). Porém a análise de agrupamento hierárquico levou em consideração toda a série temporal e os padrões das alterações dos valores de \log_2fc (i.e. padrões de aumento e redução da expressão) entre todos os tempos. Desta forma, é possível argumentar que FIP e as proteases do complexo FtsH, apresentam respectivamente, um comportamento de diminuição e aumento dos níveis expressão até as 6h, excetuando-se a condição de estresse por ferimento. Nesse sentido, considerando-se somente esta faixa da série temporal (0 a 6 horas), torna-se ainda mais evidente que a expressão de FIP possui um comportamento antagônico àquele apresentado pelas proteases do complexo FtsH.

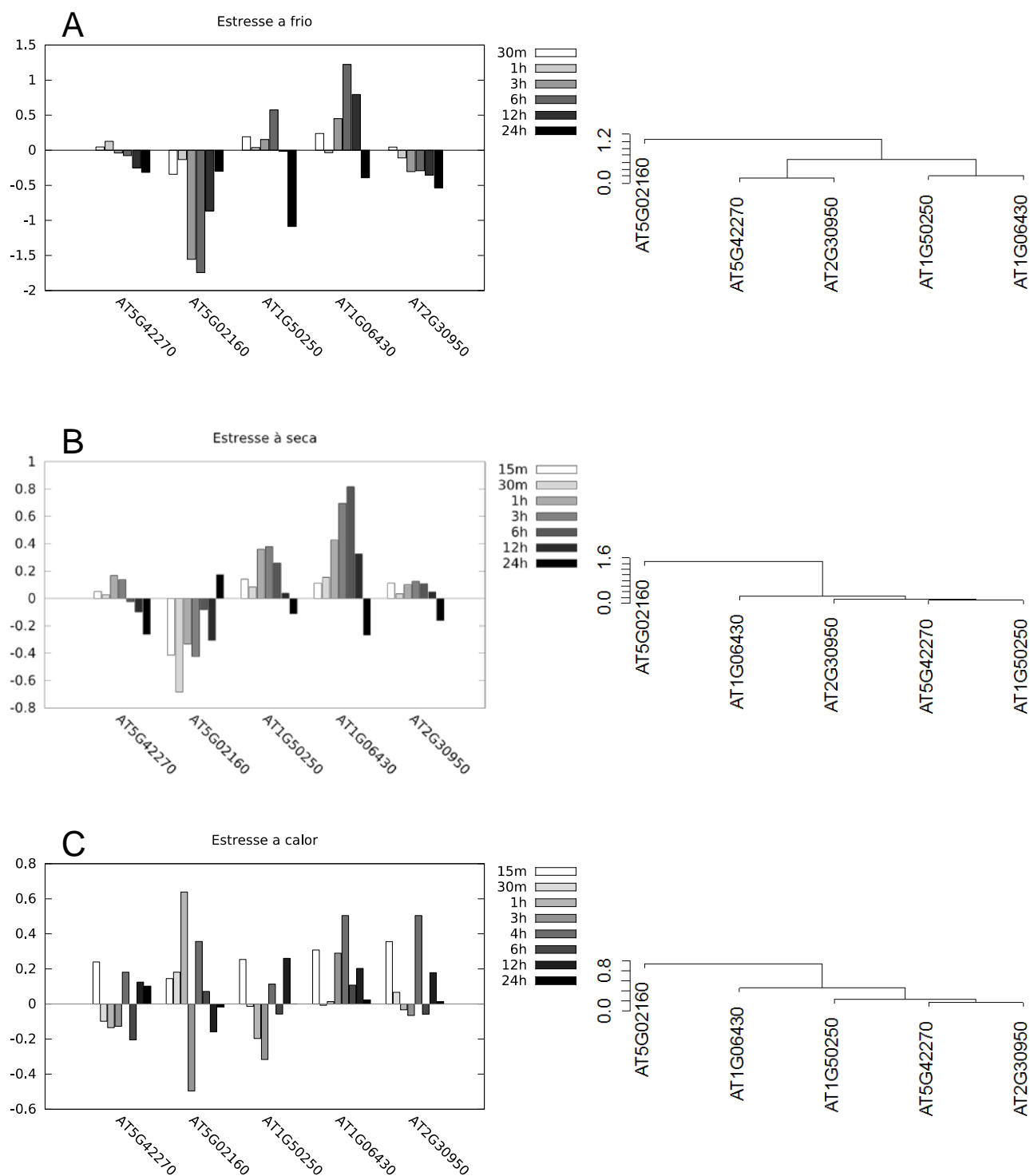


Figura 3 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. **A)** Estresse ao frio. **B)** Estresse à seca. **C)** Estresse a calor. A escala das ordenadas representa os valores de \log_2fc e legenda ao lado direito de cada gráfico representa a série temporal do experimento 15m: 15 minutos; 30m: 30 minutos; 1h: 1 hora; 3h: 3 horas; 4h: 4 horas; 6h: 6 horas; 12h: 12 horas; 24h: 24 horas

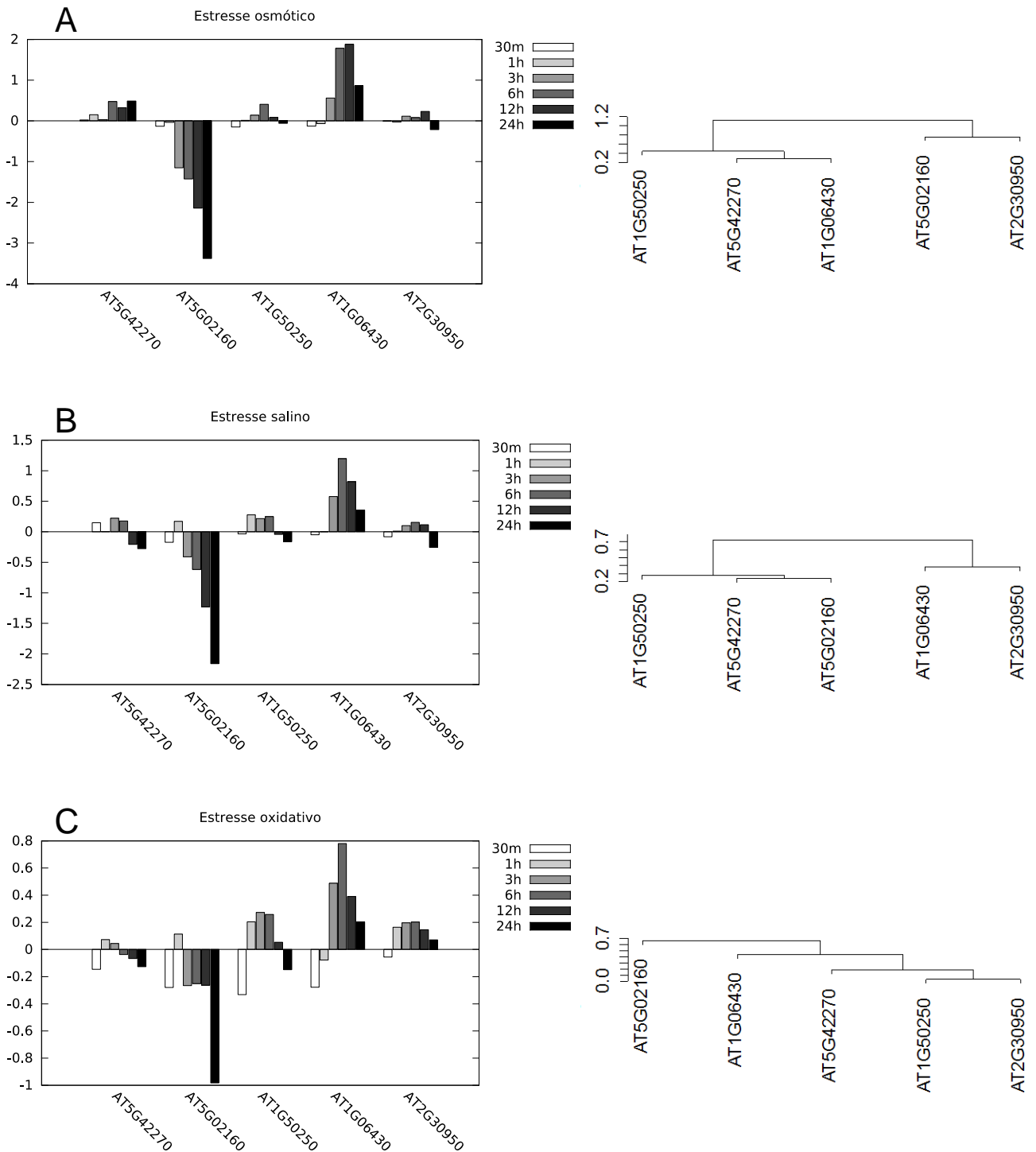


Figura 4 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. **A)** Estresse osmótico. **B)** Estresse salino. **C)** Estresse oxidativo. A escala das ordenadas representa os valores de log₂fc e legenda ao lado direito de cada gráfico representa a série temporal do experimento 15m: 15 minutos; 30m: 30 minutos; 1h: 1 hora; 3h: 3 horas; 6h: 6 horas; 12h: 12 horas; 24h: 24 horas

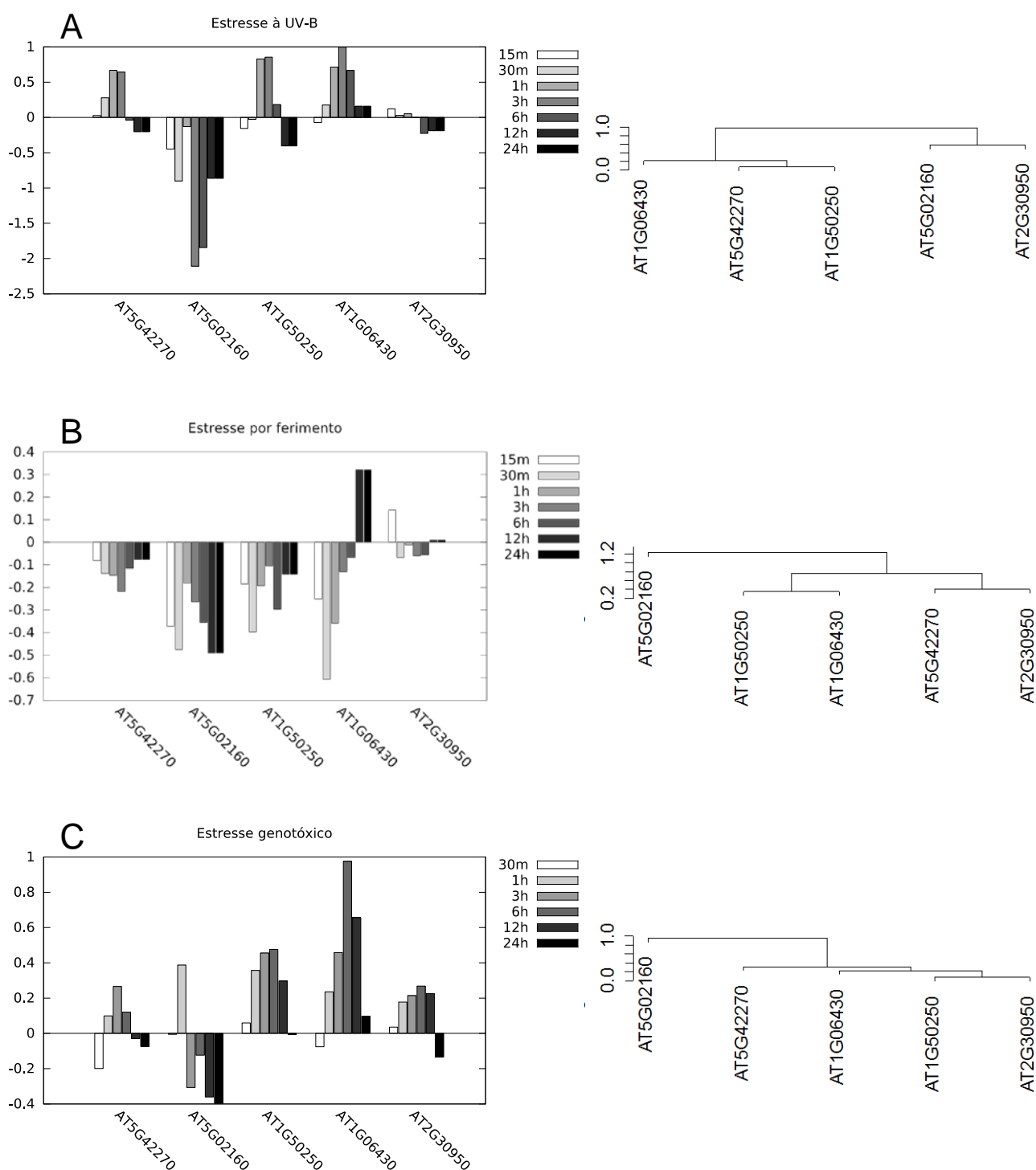


Figura 5 - Perfis de expressão de FIP (At5g02160) e das proteases do complexo FtsH (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8) sobre condições de estresse e análise de agrupamento hierárquico. **A)** Estresse à UV-B. **B)** Estresse por ferimento. **C)** Estresse genotóxico. A escala das ordenadas representa os valores de \log_2FC e legenda ao lado direito de cada gráfico representa a série temporal do experimento 15m: 15 minutos; 30m: 30 minutos; 1h: 1 hora; 3h: 3 horas; 4h: 4 horas; 6h: 6 horas; 12h: 12 horas; 24h: 24 horas

Os perfis de expressão frente as condições de estresse indicam portanto que FIP, provavelmente, atua como um regulador negativo da atividade do complexo FtsH. Corroborando os dados dos experimentos de análise de expressão obtidos anteriormente por nosso grupo de pesquisa (BRAGA, 2013). Ademais, os níveis de expressão tecido específicos e ao longo do desenvolvimento, são bastante similares e sobrepostos entre FIP e a protease FtsH5 também suportando a hipótese de FIP como regulador negativo do complexo FtsH (Figuras 1 e Anexos C e F). Conforme citado anteriormente, existem alguns exemplos descritos na literatura (auto-regulação da protease ADAM/ SMITH *et al.*, 2002; regulação das chaperonas DNA-K pelas chaperonas DNA-J/ WALSH *et al.*, 2004), de proteínas contendo domínios ricos em cisteína (*C-rich domains*) capazes de atuar na regulação da atividade de ATPases, como o domínio ATPase presente no complexo FtsH. Tal domínio é responsável por fornecer a energia necessária para a translocação e proteólise dos substratos do complexo FtsH e portanto é o candidato perfeito para a atuação na regulação da atividade do complexo através da interação com a proteína FIP.

Os resultados sugerem ainda que o complexo FtsH pode estar envolvido na resposta do cloroplasto a diferentes tipos de condições de estresse abiótico aqui avaliadas, considerando-se que as proteases do complexo tem a sua expressão aumentada de um modo geral (Figuras 2-4), excetuando-se o estresse por ferimento (Figura 4B). O que faz sentido, tendo em vista que tais estresses certamente aumentam as taxas em que as proteínas são danificadas e conseqüentemente o número de proteínas danificadas presentes nas membranas dos tilacoídes e no cloroplasto como um todo, aumentando assim a demanda pelo controle de qualidade das proteínas exercido pelo complexo FtsH.

4.2 – História evolutiva da proteína FIP e das proteases do complexo FtsH tilacoidal

4.2.1 – Inferência filogenética e classificação dos homólogos das proteases do complexo FtsH

A multiplicação dos genes das proteases FtsH se correlaciona fortemente com a evolução da fotossíntese, muito provavelmente devido ao aumento do estresse oxidativo e a conseqüente maior quantidade de danos as proteínas (ADAM *et al.*,

2005). As bactérias contêm apenas uma cópia do gene FtsH, ao passo que cianobactérias possuem quatro destes genes (MANN *et al.*, 2000) e as plantas superiores possuem ainda mais cópias, como exemplo *A. thaliana* possui uma família gênica com 17 membros, com 12 deles codificando para isoformas ativas (GARCIA-LORENZO *et al.* 2006). Esta multiplicação dos genes FtsH ocorreu depois da divergência dos cloroplastos de seus progenitores procarióticos, pela duplicação de diferentes genes de cianobactérias (SAKAMOTO *et al.*, 2003; YU *et al.*, 2004).

Vários autores buscaram estabelecer as relações filogenéticas entre os genes desta família gênica, bem como da relação destes com os genes homólogos presentes em plantas e outros taxa de organismos (MARBACH; COELHO; SILVA-FILHO, 2001; YU; PARK; RODERMEL, 2004; KATO; KOUSO; SAKAMOTO, 2012; ADAM *et al.*, 2005; CHEN *et al.*, 2000). Zaltsman, Ori e Adam (2005) demonstraram a existência de dois tipos de FtsH plastidiais, o tipo A composto pelos membros FtsH1 e FtsH5 e o tipo B com as proteases FtsH2 e FtsH8. No presente trabalho buscou-se a elucidação das relações filogenéticas entre os homólogos de FIP (Anexo A) e os homólogos das proteases do complexo FtsH (Anexo B), conforme descrito na metodologia se fez necessária a classificação das proteases FtsHs nos tipos A e B.

Como primeiro passo, identificou-se as sequências dos homólogos de FIP a serem adicionadas no respectivo modelo evolutivo e pôde-se constatar que somente musgos e plantas superiores possuem sequências homólogas (Anexo A). Nesse ponto, poderia se especular, com base nesse conhecimento, que FIP pode ter surgido como uma adaptação ao ambiente terrestre e as muitas formas de condições estressantes impostas por este novo ambiente, a partir da colonização deste, de modo a permitir um controle mais dinâmico e preciso da atividade do complexo FtsH, fornecendo assim um mecanismo para as plantas responderem mais dinamicamente aos estresses. Desta forma, o processo de colonização terrestre teria levado não somente a multiplicação e diversificação das estruturas, funções e destinos celulares das proteases FtsHs, como também ao surgimento de um regulador negativo para o controle da atividade, no caso das proteases do complexo FtsH tilacoidal. As sequências homólogas de FIP foram então utilizadas para produzir um MSA e posteriormente produziu-se a filogenia dos homólogos de FIP, a qual pode ser observada na Figura 6.

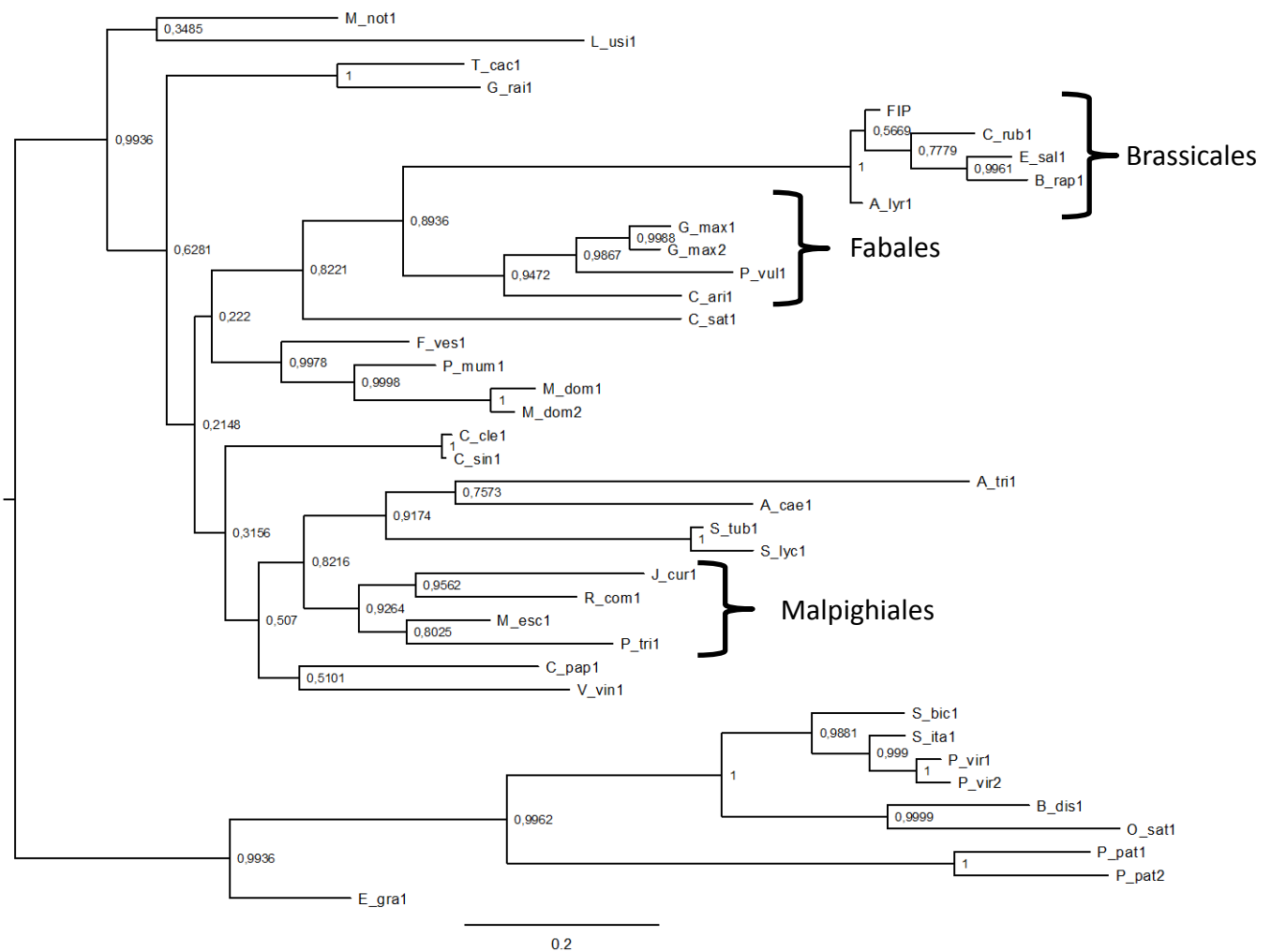


Figura 6 – Inferência filogenética bayesiana das sequências homólogas de FIP. Os nomes foram arbitrariamente escolhidos e a correspondência com os identificadores das sequências utilizadas na inferência encontram-se no Anexo A. Foram geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas. Destas estimativas 25% foram eliminadas como *burn-in*. A probabilidade posterior de cada nó encontra-se a direita de cada um dos mesmos

Dados anteriores do nosso grupo de pesquisa demonstraram que FIP interage apenas com as proteases do complexo FtsH pertencentes ao tipo A (RODRIGUES, 2011). Como as FtsHs do tipo A e B são altamente similares entre si, fez-se necessária a obtenção de todos os homólogos das proteases do complexo das espécies onde FIP havia sido encontrada, bem como a classificação correta destas nos tipos A e B.

A busca pelos homólogos de FtsH nas espécies portadores de homólogos de FIP, se alinha com a hipótese de que o número de homólogos em cada tipo A e B varia entre as espécies, conforme havia sido especulado anteriormente por Kato, Kouso & Sakamoto (2012).

A classificação dos homólogos de FtsH (Anexo B) foi feita por meio da inferência filogenética bayesiana, além disso realizou-se a inclusão na análise de sequências homólogas as proteases do complexo FtsH anteriormente utilizadas para a inferência filogenética e classificação nos tipos A e B (KATO; KOUSO; SAKAMOTO, 2012; Figura 1), as quais serviram como referência para a classificação.

Como pode se observar na Figura 7, foi possível classificar todos os homólogos das proteases do complexo FtsH em três classes: FtsHs do tipo A, FtsHs do tipo B e outras FtsHs. Todas as sequências usadas como referência foram perfeitamente classificadas nos tipos corretos e portando conferindo maior confiabilidade a classificação das sequências homólogas de interesse (Figura 7). A partir desse ponto, tornou-se possível a construção de modelos evolutivos para as FtsHs do tipo A e do tipo B isoladamente, através de novas análises filogenéticas bayesianas independentes e, conseqüentemente, a avaliação das relações destes tipos com a proteína FIP, também de forma independente.

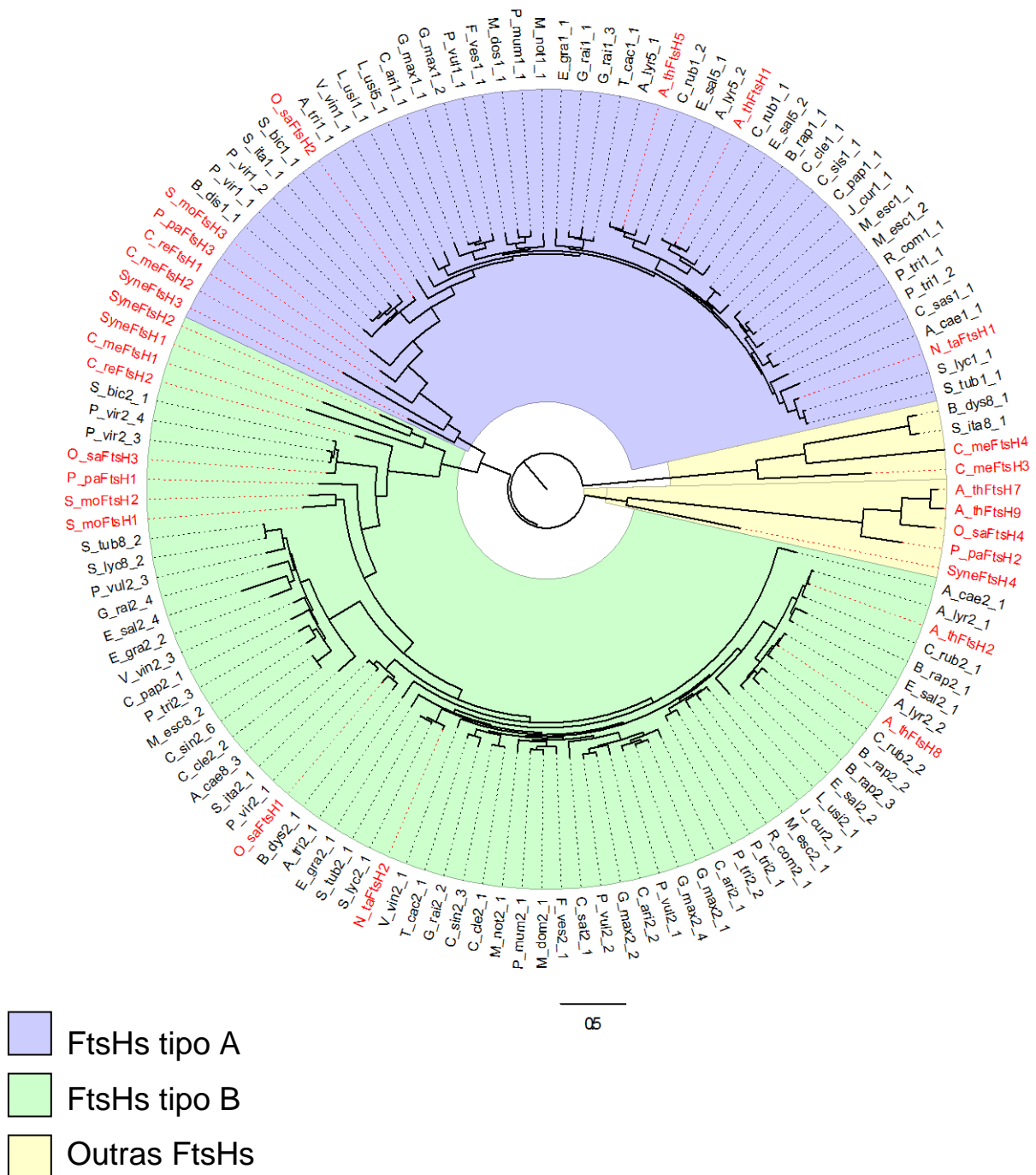


Figura 7 – Inferência filogenética bayesiana das sequências homólogas das proteases do complexo FtsH. Os nomes foram arbitrariamente escolhidos e a correspondência com os identificadores das sequências utilizadas na inferência encontram-se no Anexo B e o alinhamento destas no Anexo I. Foram geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas. Destas estimativas, 25% foram eliminadas como *burn-in*. A legenda no canto inferior esquerdo relaciona as cores e a classificação das sequências homólogas de FtsH nos tipos A, B e outras FtsHs. As sequências utilizadas como referência para a classificação estão marcadas em vermelho

4.2.2 – Evolução convergente entre as FtsHs do tipo A e FIP

Com as proteases do complexo FtsH classificadas nos tipos A e B, tornou-se possível a inferência evolutiva para os homólogos dos dois tipos de FtsH. Assim as sequências de ambos foram selecionadas e analisadas separadamente, a partir do conjunto inicial (Anexo B), sendo novamente alinhadas independentemente para a condução da inferência filogenética bayesiana. As árvores filogenéticas produzidas para os homólogos das FtsHs tipos A e B estão mostradas nas Figuras 8 e 9.

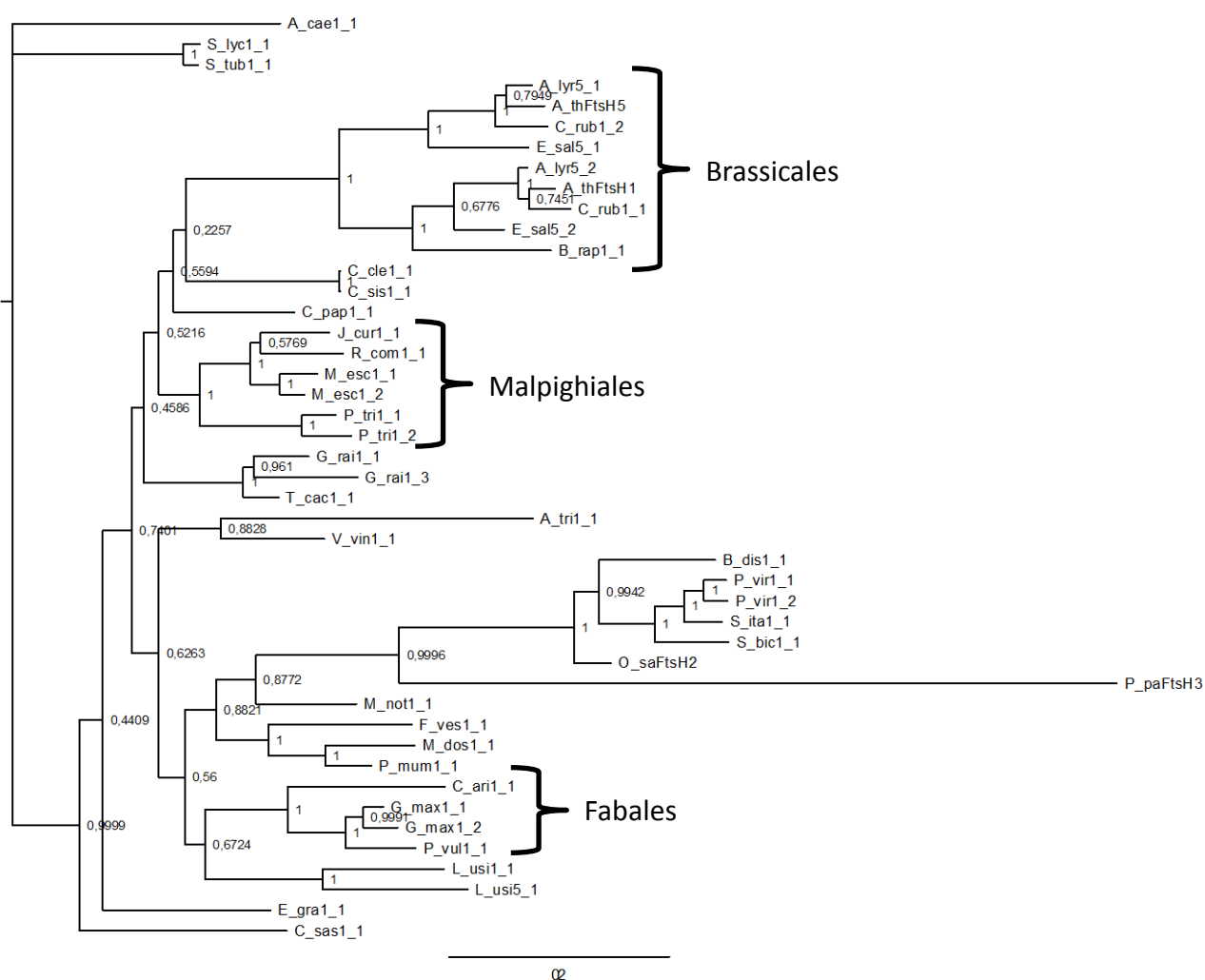


Figura 8 – Inferência filogenética bayesiana das sequências homólogas às proteases FtsHs do tipo A, do grupo de espécies portador de FIP. Os nomes foram arbitrariamente escolhidos e a correspondência com os identificadores das sequências utilizadas na inferência encontram-se no Anexo B. Foram geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas. Destas estimativas, 25% foram eliminadas como *burn-in*. A probabilidade posterior de cada nó encontra-se a direita de cada um dos mesmos

A análise visual das árvores filogenéticas de FIP (Figura 6) e das FtsHs do tipo A (Figura 8) e B (Figura 9) permite notar alguns padrões e agrupamentos de sequências recorrentes entre elas. Por exemplo, pode-se observar que as sequências provenientes das ordens *Brassicales*, *Fabales* e *Malpighiales*, assim como outras *taxa* menores e maiores, formam grupos monofiléticos em todas as três filogenias.

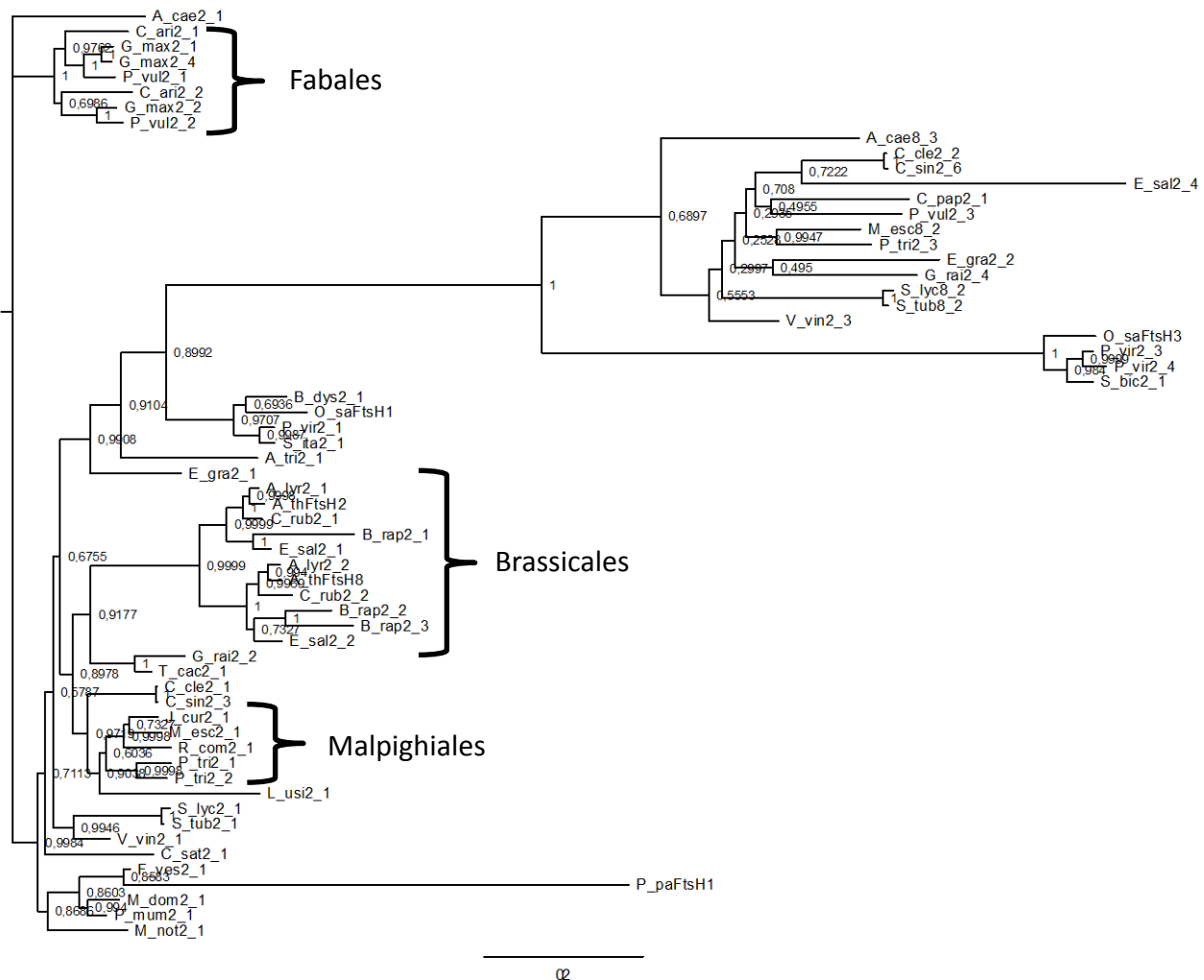


Figura 9 – Inferência filogenética bayesiana das sequências homólogas às proteases FtsHs do tipo B, do grupo de espécies portador de FIP. Os nomes foram arbitrariamente escolhidos e a correspondência como os identificadores das sequências utilizadas na inferência encontram-se no Anexo B. Foram geradas 2×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 2000 estimativas. Destas estimativas 25% foram eliminadas como *burn-in*. A probabilidade posterior de cada nó encontra-se a direita de cada um dos mesmos

Contudo, em uma simples análise visual seria impraticável determinar os padrões subjacentes as relações estabelecidas entre as sequências em cada uma destas filogenias, assim como inferir sobre qualquer tipo de relação evolutiva entre as famílias gênicas estudadas. Nesse sentido, tem sido observado que as distâncias evolutivas (distâncias entre os taxa em uma filogenia) de proteínas interagentes frequentemente apresentam um nível maior de similaridade do que as de proteínas não-interagentes. Este fato indica que proteínas interagentes estão frequentemente sujeitas a pressões e restrições evolutivas comuns e esta é a base do método conhecido como *Mirrortree*.

Tem sido proposto que proteínas interagentes devem coevoluir para manter suas interações (GOH *et al.*, 2000; PAZOS *et al.*, 1997; PAZOS; VALENCIA, 2001) e o método *Mirrortree* prediz interações proteína-proteína acessando a extensão da concordância entre as distâncias evolutivas que pode ser atribuída a coevolução. Para este propósito, matrizes de distância devem ser construídos a partir das filogenias de sequências ortólogas de um conjunto comum de espécies. O grau de coevolução entre as famílias de ortólogos é acessada calculando o coeficiente de correlação entre as matrizes de distâncias. O método mede portanto a correlação entre as distâncias evolutivas e assim, indiretamente, a correlação entre taxas evolutivas entre ramos individuais das árvores filogenéticas de duas famílias gênicas.

Nesse sentido, no presente trabalho utilizou-se o método *Mirrortree* para inferir sobre a existência de processo coevolutivo entre as famílias gênicas de FIP e das FtsHs do tipo A e B. A análise *Mirrortree* suporta a existência de coevolução entre os homólogos de FIP e das proteases FtsH do tipo A (Figura 10A), com um coeficiente de correlação de 85,3% ($r=0.853$; P valor = ≤ 0.000001). Por outro lado nenhuma correlação foi encontrada entre os homólogos de FIP e das proteases FtsH do tipo B (Figura 10B), com um coeficiente de correlação de 32.2% ($r = 0.322$; P valor = ≤ 0.000001), o que corrobora os resultados experimentais do nosso grupo de pesquisa, onde foi mostrado que FIP interage apenas com proteases FtsH do tipo A (RODRIGUES, 2011).

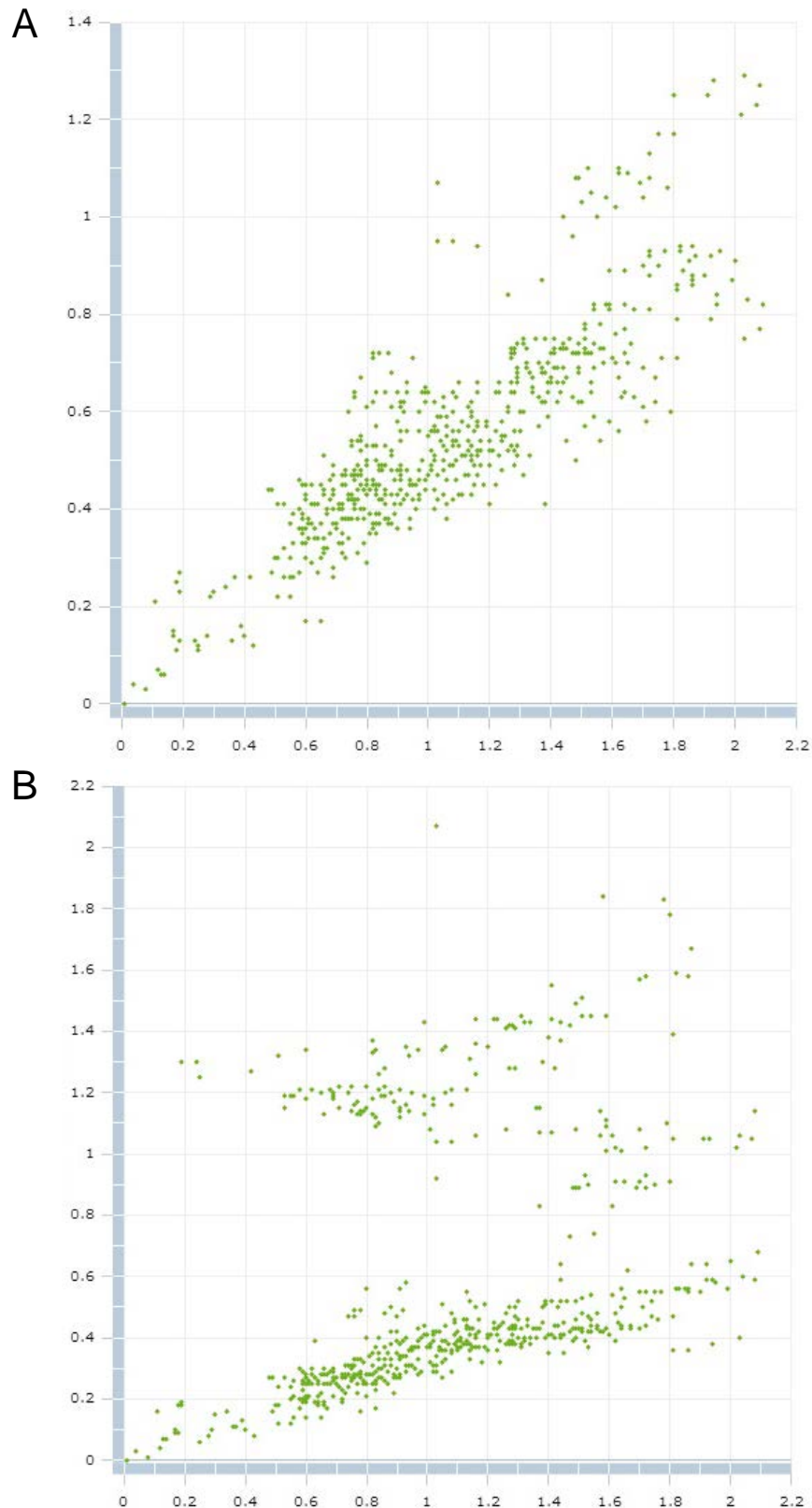


Figura 10 – Inferência sobre a existência de coevolução pelo método Mirrortree. **A)** Representação gráfica das distâncias entre espécies das famílias gênicas de FIP (eixo Y) e das proteases FtsH do tipo A (eixo X), mostrando claramente um padrão de correlação ($r=0.853$; P valor = ≤ 0.000001). **B)** Representação gráfica das distâncias entre espécies das famílias gênicas de FIP (eixo Y) e das proteases FtsH do tipo B (eixo X), com nenhuma correlação ($r = 0.322$; P valor = ≤ 0.000001). As imagens foram produzidas através do *MirrorTree Server* (<http://csbg.cnb.csic.es/mtserver/>)

Desta forma foi demonstrado tanto pelas evidências experimentais com os experimentos de duplo híbrido de lêveduras, coimunoprecipitação e *GST-Pull Down* (RODRIGUES, 2011), quanto pelas análises evolutivas, com a inferência filogenética bayesiana e pelo método de *Mirrortree*, que a proteína FIP interage com as FtsHs do tipo A do complexo tilacoidal. Além disso a análise dos perfis de expressão sugerem ainda que FIP atua como um regulador negativo da atividade do complexo. Nesse contexto, como citado anteriormente, a literatura fornece modelos para a interação e atuação putativa de FIP como um regulador negativo do complexo FtsH. A regulação muito provavelmente ocorre através da interação do domínio rico em cisteína de FIP e o domínio ATPase das FtsHs do tipo A e conseqüente regulação da atividade deste e do fornecimento de energia para a translocação e proteólise dos substratos pelo complexo.

A natureza da atuação de FIP sobre o complexo FtsH deve portanto ser comprovada através da análise de plantas mutantes nocautes e de superexpressão de FIP, de modo a investigar a relação funcional existente entre a proteína FIP e o complexo FtsH tilacoidal, bem como as conseqüências daquela para a atividade deste.

4.2.3 – A proteína FIP no contexto da história evolutiva das FtsHs

De modo a corroborar os resultados sobre o processo de coevolução entre FIP e as FtsHs do tipo A, e investigar a relação do surgimento da proteína FIP no contexto da história evolutiva completa das FtsHs, visou-se a proposição de um modelo evolutivo mais completo e representativo para estas últimas. Nesta filogenia foram incluídas seqüências homólogas as proteases FtsH do tipo A de *Arabidopsis thaliana*, provenientes da maioria dos táxons onde estão presentes FtsHs do tipo A. De forma a determinar se o grupo de espécies portadoras de homólogos de FIP, seria recuperado como um táxon monofilético nesta filogenia mais completa dos homólogos de FtsH do tipo A ou se seqüências de outras espécies não portadoras de FIP se agrupariam em conjunto com estas e o que poderia ser inferido a partir dessa análise.

A inferência filogenética bayesiana dos homólogos das FtsHs (Figura 11) demonstrou que foi possível recuperar como um táxon monofilético na filogenia com o grupo de espécies portadoras de homólogos de FIP. Porém algumas seqüências

de algas verdes (*Chlorophyta*; Tabela 2), bem como algumas sequências provenientes das espécies portadoras de homólogos de FIP (não selecionadas nas análises anteriores), se agruparam no mesmo táxon.

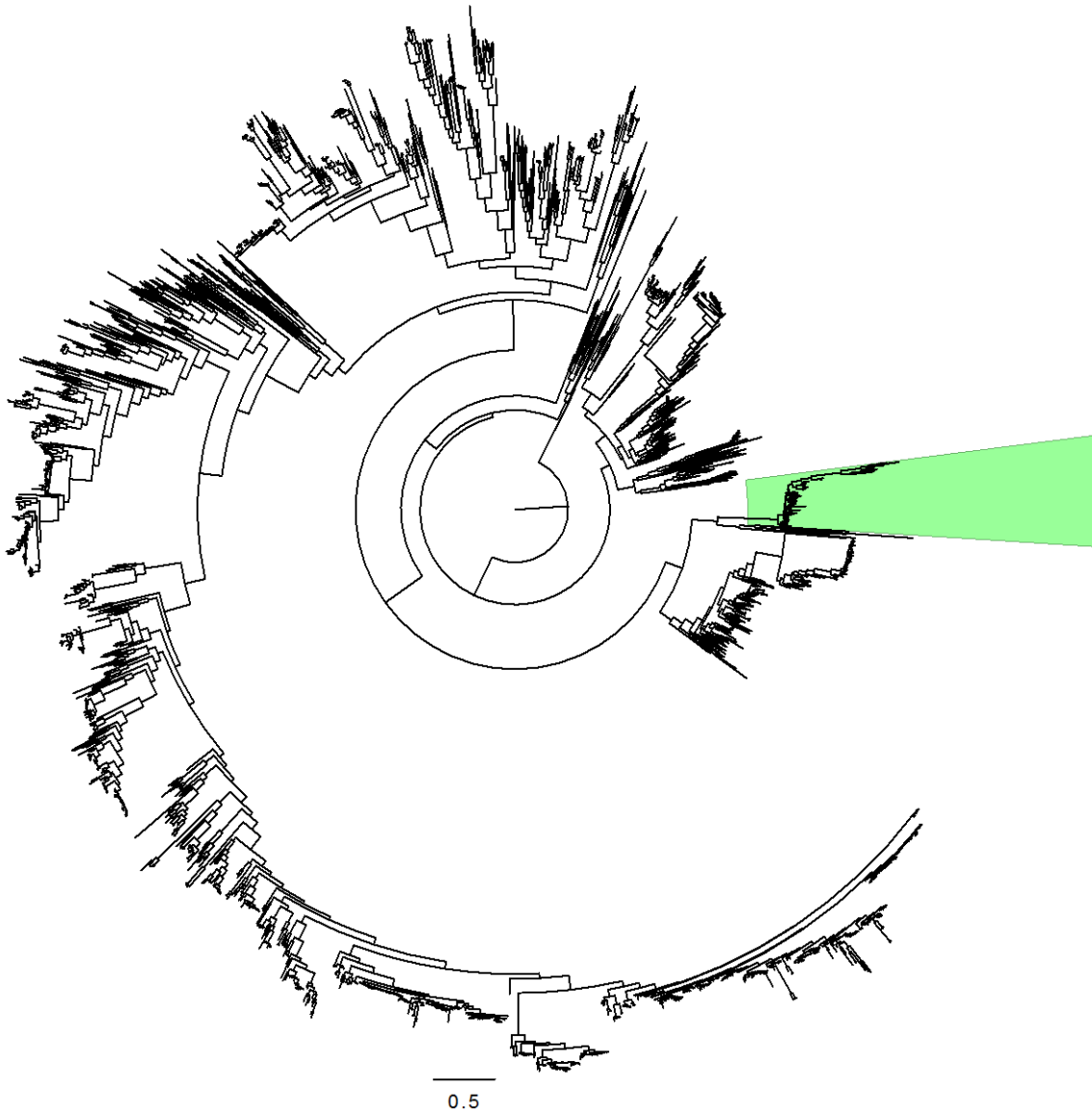


Figura 11 – Inferência filogenética bayesiana das sequências homólogas as FtsHs do tipo A. Os identificadores das sequências foram omitidos devido ao grande número de sequências utilizadas na análise (2073 sequências). Foram geradas 1×10^6 gerações de MCMC (*Markov chain Monte Carlo*) e amostradas a cada 1000 gerações, produzindo 1000 estimativas. O agrupamento de sequências provenientes de espécies portadoras de homólogos de FIP foi recuperado como um táxon monofilético na filogenia (em verde), juntamente com algumas sequências de espécies de algas verdes, não portadoras de homólogos de FIP (Tabela 2)

Tais resultados evidenciam o fato de que as sequências dos homólogos de FtsH do tipo A nas espécies portadoras de homólogos de FIP compartilham algumas características, porém algumas espécies de algas verdes (Tabela 2) também possuem estas mesmas características. Isto pode sugerir que as características desenvolvidas pelas FtsH do tipo A nos ancestrais compartilhados por algas verdes, musgos e plantas vasculares, tenham possibilitado a evolução de FIP nos dois últimos, quando ocorreu a colonização do ambiente terrestre.

Tabela 2 – Sequências homólogas a FtsHs do tipo A agrupadas no táxon monofilético das sequências provenientes de espécies portadoras de homólogos de FIP utilizadas na análise da Figura 8

Divisão	Espécie	Identificador (GI)
<i>Chlorophyta</i>	<i>Bathycoccus prasinos</i>	612390445
	<i>Chlamydomonas reinhardtii</i>	159465357
	<i>Coccomyxa subellipsoidea</i> C-169	545371675
	<i>Micromonas pusilla</i> CCMP1545	303275720
	<i>Micromonas</i> sp. RCC299	255088145
	<i>Ostreococcus lucimarinus</i> CCE9901	145350390
	<i>Ostreococcus tauri</i>	308807935
	<i>Volvox carteri</i> f. <i>nagariensis</i>	302833547
<i>Briophyta</i>	<i>Physcomitrella patens</i>	168001910
<i>Lycopodiophyta</i>	<i>Selaginella moellendorffii</i>	302753920
<i>Magnoliophyta</i>	<i>Brachypodium distachyon</i>	357123107
	<i>Cucumis melo</i>	659104635
	<i>Malus domestica</i>	657970655
	<i>Oryza brachyantha</i>	573948409
	<i>Phoenix dactylifera</i>	672113676
	<i>Prunus persica</i>	596125686
	<i>Theobroma cacao</i>	590595043
	<i>Zea mays</i>	308081381

4.3 – Pressões seletivas e conservação de sequência

Considerando-se os resultados anteriores, os quais posicionam FIP como um regulador negativo putativo da atividade do complexo FtsH, buscou-se avaliar quais forças evolutivas estão guiando a evolução da proteína FIP e das proteases FtsHs do tipo A, através da análise de pressão de seleção. Uma abordagem alternativa visou analisar a estrutura do complexo FtsH e os seus padrões de conservação de sequência no sentido de fornecer evidências para a possível região de interação de

FIP com complexo e modelos para o mecanismo de regulação da atividade do complexo.

Primeiramente, portanto, realizou-se a análise de pressão de seleção a partir dos alinhamentos múltiplos de sequências e das árvores filogenéticas utilizados anteriormente para FIP e para as FtsHs do tipo A (Figuras 6 e 8, respectivamente). Tanto para a proteína FIP quanto para as FtsHs do tipo A, a comparação do modelo de razão única de ômega (M_0) com o modelo de variáveis discretas (M_3 – com $K=3$ classes discretas), mostrou um valor significativamente maior de verossimilhança para o modelo M_3 , pelo teste de razão de verossimilhança, indicando assim a presença de pressão de seleção variável entre os sítios (Tabelas 3 e 4).

Novamente para ambos os casos, as comparações entre os modelos M_{1a} versus M_{2a} e M_7 versus M_8 , mostraram o mesmo resultado, porém nestas comparações os modelos mais simples (M_{1a} e M_7) apresentaram verossimilhanças maiores ou equivalentes pelo teste de razão de verossimilhança e portanto não se rejeitam os modelos mais simples (Tabelas 3 e 4). Assim pode-se concluir dos resultados da análise de pressão de seleção que existe uma pressão de seleção variável entre os sítios dos homólogos da proteína FIP e das FtsHs do tipo A, contudo não se pode atribuir tal variação a presença de seleção positiva ou adaptativa, devido ao resultado das comparações (M_{1a} versus M_{2a} e M_7 versus M_8) e portanto a variação é explicada pela atuação da seleção negativa ou purificadora.

Este resultado faz sentido a luz do conhecimento da interação existente entre as proteínas de interesse e da importância da função do complexo FtsH para a manutenção e desenvolvimento dos cloroplastos, bem como da regulação de sua atividade pela proteína FIP. Além disso evidencia a importância da conservação da estrutura geral tanto na proteína FIP para a sua interação com o complexo e atuação na regulação, quanto para a atividade e regulação das FtsHs do tipo A, visto que grande parte de suas sequências encontram-se sobre pressão de seleção negativa (dados não mostrados), como determinado pelo teste de probabilidade posterior *Bayes Empirical Bayes* (BEB).

Tabela 3 – Descrição dos parâmetros estimados e os escores de verossimilhança para os homólogos de FIP sob diferentes modelos de seleção com razões de ω variáveis entre os sítios

Modelo de sítio	ω	$p(\omega)$	lnL	LRT	p-value
M0 (uma razão)	$\omega=0.09990$	-	-4870.9707	367.313753	0
M3 (discreto)	$\omega_0=0.00891$ $,\omega_1=0.12032$ $,\omega_2=0.45001$	$p_0=0.55422$ $,p_1=0.25979$ $,p_2=0.18598$	-4687.313753		
M1a (proximamente neutro)	$\omega_0=0.05898$ $,\omega_1=1$	$p_0=0.82324$ $,p_1=0.17676$	-4767.9970	-	-
M2a (seleção positiva)	$\omega_0=0.05898$ $,\omega_1=1.00000$ $,\omega_2=21.22331$	$p_0=0.82324$ $,p_1=0.17676$ $,p_2=0.00000$	-4767.997097		
M7 (beta)		$p=0.28358$ $,q=2.07030$	-4686.43784	-	-
M8 (beta e ω)	$\omega=3.02483$	$p_0=0.99999$, $(p_1=0.00001)$, $p=0.28357$, $q=2.07030$	-4686.438698		

Tabela 4 – Descrição dos parâmetros estimados e os escores de verossimilhança para os homólogos de FtsH do tipo A (apenas das espécies portadoras de FIP) sob diferentes modelos de seleção com razões de ω variáveis entre os sítios

Modelo de sítio	ω	$p(\omega)$	lnL	LRT	p-value
M0 (uma razão)	$\omega=0.02355$	-	-26177.33878	1394.211135	0
M3 (discreto)	$\omega_0=0.00178$ $,\omega_1=0.05965$ $,\omega_2=0.26360$	$p_0=0.79102$ $,p_1=0.13944$ $,p_2=0.26360$	-25480.23321		
M1a (proximamente neutro)	$\omega_0=0.01537$ $,\omega_1=1$	$p_0=0.96106$ $,p_1=0.03894$	-25923.16287	0	1.00E+00
M2a (seleção positiva)	$\omega_0=0.01537$ $,\omega_1=1.00000$ $,\omega_2=100.44081$	$p_0=0.96106$ $,p_1=0.03894$ $,p_2=0.00000$	-25923.16287		
M7 (beta)		$p=0.10705$ $,q=2.54572$	-25480.84524	-	-
M8 (beta e ω)	$\omega=3.75461$	$p_0=0.99999$, $(p_1=0.00001)$, $p=0.10705$, $q=3.75461$	-25480.85111		

Na segunda abordagem, foi feita uma análise da conservação das sequências das FtsHs do tipo A, na tentativa de identificar padrões de conservação de sequência que possivelmente estariam relacionados a interação com a proteína FIP e a regulação da atividade do complexo. Os modelos foram construídos utilizando como molde a estrutura cristalizada completa da região citosólica do complexo FtsH de *Thermus thermophilus* (PDB: 2DHR) e para estimar a conservação evolutiva de posições de aminoácidos utilizou-se o MSA de aminoácidos dos produtos dos genes usados na análise da Figura 8.

De uma maneira geral as FtsHs do tipo A se mostraram bastante conservadas nas regiões mais internas da proteína, assim como nas regiões de contato entre as subunidades do complexo, com maiores níveis de variabilidade nas regiões externas (Figura 12). Isto geralmente é causado e pode ser atribuído as restrições evolutivas a mutações nas posições mais internas e cruciais para a manutenção estrutura da proteína e da função de cada subunidade, da mesma forma restrições evolutivas impostas pelas interações entre as subunidades reduzem a variabilidade nas regiões de contato entre as subunidades. Todavia, as regiões mais externas estão livres destas restrições e podem variar em maior magnitude. Tais resultados mostram que grande parte das FtsHs do tipo A apresentam pouca ou nenhuma variabilidade (Figura 12; branco até vermelho), o que corrobora os resultados da análise de pressão de seleção, os quais apontam que tais proteases encontram-se em grande parte de suas sequências sobre pressão de seleção negativa.

Porém, não foi possível identificar uma região externa claramente conservada que pudesse fornecer um modelo para a interação com a proteína FIP. Assim, estes resultados podem indicar que a interação ocorra de alguma forma nas regiões mais internas das proteases FtsHs, possivelmente na região do sítio catalítico do domínio ATPásico, conforme os exemplos descritos na literatura (SMITH *et al.*, 2002; WALSH *et al.*, 2004), tomados aqui como modelos para a interação. Um modelo tridimensional para a proteína FIP poderia ser uma solução, contudo a mesma não apresenta homologia suficiente com nenhuma estrutura proteica conhecida, o que impede a sua modelagem pelo método de homologia e por ser uma proteína transmembrana os métodos de modelagem *de novo* geram resultados muito inconclusivos, impossibilitando assim a predição da sua estrutura tridimensional e, conseqüentemente, a análise de docagem do complexo FtsH e FIP, o que poderia fornecer um modelo estrutural para a interação.

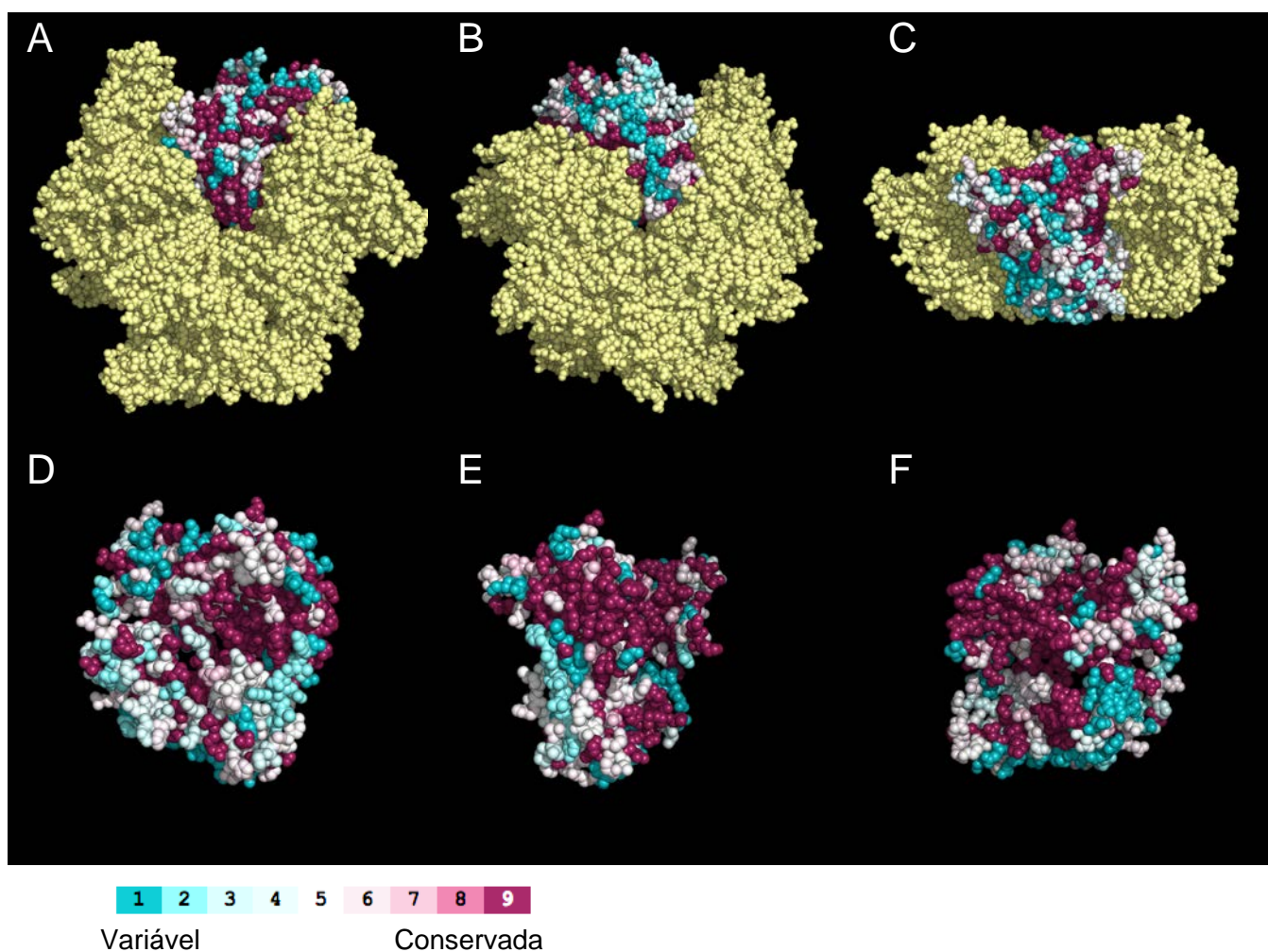


Figura 12 – Representação gráfica tridimensional das posições estruturais conservadas das FtsHs do tipo A, do grupo de espécies portador de FIP. A análise de conservação de sequências foi realizada no Webservice ConSurf (<http://consurf2.tau.ac.il/>; Celniker et al., 2013; Ashkenazy et al., 2010) e a figura gerada pelo *software* PyMOL. A sequência molde utilizada foi a estrutura cristalizada completa da região citosólica do complexo FtsH de *Thermus thermophilus* (PDB: 2DHR). A) Visão superior do complexo (região voltada para o estroma). B) Visão inferior do complexo (região voltada para a membrana). C) Visão lateral do complexo, com a região voltada para o estroma na parte superior da figura. D) Subunidade isolada (Região em contato com uma subunidade adjacente). E) Região voltada para o centro do complexo (rotação de 90 graus a direita da figura D). F) Região em contato com a outra subunidade adjacente (rotação de 90 graus a direita da figura E). Para a representação da conservação foi utilizada apenas uma cadeia polipeptídica do complexo hexamérico, com o restante das cadeias aparecendo em amarelo nas figuras A, B e C

5 CONSIDERAÇÕES FINAIS

- Os perfis de expressão tecido específicos e ao longo do desenvolvimento de FIP e das proteases do complexo FtsH de *Arabidopsis thaliana*, bem como os perfis de expressão frente a diversos estresses abióticos suportam a hipótese de FIP como um regulador negativo da atividade do complexo FtsH tilacoidal. As evidências provêm principalmente do comportamento antagônico dos perfis de expressão de FIP em relação as proteases do complexo FtsH em resposta aos diferentes estresses abióticos aqui analisados, da sobreposição e similaridade dos níveis de expressão tecido-específicos e ao longo do desenvolvimento de FIP e da protease FtsH5 e, por fim, do modelo de regulação da atividade do complexo fornecido pela literatura, onde proteínas portadoras de domínios ricos em cisteínas (CRD) tem sido mostradas capazes de regular a atividade de outras proteínas através da interação com os domínios ATPásicos destas últimas.

- A análise dos perfis de expressão também pode sugerir que o complexo FtsH pode estar envolvido na resposta a diferentes tipos de condições de estresse abiótico.

- As sequências homólogas a proteína FIP foram exclusivamente encontradas em musgos e plantas vasculares, sugerindo que a origem de FIP pode estar relacionada a colonização terrestre e as muitas formas de condições estressantes impostas por este novo ambiente. Assim, a colonização terrestre e a evolução da fotossíntese teriam conduzido não somente a multiplicação e diversificação das estruturas, funções e destinos celulares das proteases FtsHs, como também ao surgimento de um regulador negativo para o controle da atividade, no caso das proteases do complexo FtsH tilacoidal, fornecendo um mecanismo para regulação mais dinâmica e fina aos estresses. A propriedade de sensibilidade ao estado redox das cisteínas presentes em FIP, pode fornecer um mecanismo molecular para a regulação fina (*fine tuning*) da interação de FIP com o complexo FtsH tilacoidal. Através da percepção e resposta rápida a variações do estado redox dos cloroplastos provocadas pelos estresses, FIP modularia a sua interação com o complexo e consequentemente a atividade deste.

- Foi possível recuperar os grupos dos tipos A e B como táxons monofiléticos dentre os homólogos das proteases FtsH das espécies aqui utilizadas.

- A análise de *Mirrortree* suporta a existência de coevolução entre a proteína FIP e as proteases FtsH do tipo A. Por outro lado, nenhuma correlação foi encontrada entre FIP e os homólogos do tipo B das FtsH, o que corrobora os dados experimentais anteriores de nosso grupo de pesquisa, os quais mostram que FIP apenas interage com as FtsHs do tipo A. Além disso, este resultado suporta a hipótese de que os homólogos de FIP nas diferentes espécies aqui estudadas são capazes de interagir com os homólogos das proteases FtsH do tipo A e portanto a regulação da atividade do complexo FtsH por FIP pode ser ubíqua nestas espécies.

- O grupo das espécies portadoras de homólogos de FIP foi recuperado como um *taxon* monofilético em uma filogenia representativa dos homólogos das proteases FtsH do tipo A. Porém, algumas sequências de algas verdes (*Chlorophyta*) foram agrupados no mesmo *taxon*, indicando que algumas características de sequência (motivos) provavelmente são compartilhados entre o grupo de espécies portadoras de homólogos de FIP e as algas verdes. Sugerindo também que estas características, mais parcimoniosamente, foram desenvolvidas pelas FtsHs do tipo A nos ancestrais compartilhados por algas verdes, musgos e plantas vasculares, e que estas mesmas características tenham possibilitado o surgimento de FIP nos musgos e plantas devido ao evento de colonização terrestre.

- Tanto a proteína FIP quanto as proteases FtsH do tipo A estão sob pressão de seleção negativa, em grande extensão de suas sequências. Evidenciando a importância da conservação da estrutura para a função das proteases FtsH do tipo A e de FIP, bem como para a interação.

- As proteases FtsH do tipo A se mostraram bastante conservadas de um modo geral, principalmente nas regiões mais internas e de contato entre as subunidades do complexo, com a maior parte da variabilidade de sequências se encontrando nas regiões exteriores e livres de restrições evolutivas. Corroborando os resultados da análise de pressão de seleção.

REFERÊNCIAS

ABASCAL, F.; ZARDOYA, R.; POSADA, D. ProtTest: selection of best-fit models of protein evolution. **Bioinformatics**, Oxford, v.21, p.2104-2015, May 2005.

ABASCAL F.; ZARDOYA R.; TELFORD M.J. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. **Nucleic Acids Research**, Oxford, v.38, p.W7-13, April 2010.

ADAM, Z.; ZALTSMAN, A.; SINVANY-VILLALOBO, G.; SAKAMOTO, W. FtsH proteases in chloroplasts and cyanobacteria. **Physiologia Plantarum**, Lund, v. 123, p. 386-390, April 2005.

AKIYAMA, Y. Quality control of cytoplasmic membrane proteins in Escherichia coli. **Journal of Biochemistry**, Tokyo, v. 146, p. 449-454, May 2009.

AKIYAMA, Y.; KIHARA, A., MORI, H.; OGURA, T.; ITO, K. Roles of the periplasmic domain of Escherichia coli FtsH (HflB) in protein interactions and activity modulation. **The Journal of Biological Chemistry**, Baltimore, v. 273, p.22326-22333, Sept. 1998.

AKIYAMA, Y.; KIHARA, A., TOKUDA, H.; ITO, K. FtsH (HflB) is an ATP-dependent protease selectively acting on SecY and some other membrane proteins. **The Journal of Biological Chemistry**, Baltimore, v. 271, p. 31196-31201, Jan. 1996.

ARNOLD, T. LANGER, Membrane protein degradation by AAA proteases in mitochondria. **Biochimica et Biophysica Acta**, Amsterdam, v.159, p.89–96, Sept. 2002.

ASHKENAZY H.; EREZ E.; MARTZ E.; PUPKO T.; BEN-TAL N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. **Nucleic Acids Research**, Oxford, v.38, p.W529-W533, May 2010.

BANFI, S.; BASSI, M.T.; ANDOLFI, G.; MARCHITIELLO, A.; ZANOTTA, S.; BALLABIO, A.; CASARI, G.; FRANCO, B. Identification and characterization of AFG3L2, a novel paraplegin-related gene. **Genomics**, London, v.59, p.51–58, July 1999.

BAILEY, S.; THOMPSON, E.; NIXON, P.J.; HORTON, P.; MULLINEAUX, C.W.; ROBINSON, C.; MANN N.H. A critical role for the var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle in vivo. **The Journal of Biological Chemistry**, London, v.277, p.2006–2011, Nov. 2002.

BRAGA, W.G.S. **Caracterização da interação entre FIP e FtsH5: mapeamento da região de interação e análise de expressão em condições de estresse**. 2013. 67p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2013.

BIEGERT, A.; MAYER, C.; REMMERT, M.; SÖDING, J.; LUPAS, A.N. The MPI Bioinformatics Toolkit for protein sequence analysis. **Nucleic Acids Research**, Oxford, v.34, p.335-339, July 2006.

BIENIOSSEK, C.; SCHALCH, T.; BAUMANN, M.; MEISTER, M.; MEIER, R.; BAUMANN, U. The molecular architecture of the metalloprotease FtsH. **Proceedings of the National Academy of Sciences of the USA**, Whashington, v.103, p.3066–3071, Feb. 2006.

BIENIOSSEK C.; NIEDERHAUSER B.; BAUMANN U.M. The crystal structure of apo-FtsH reveals domain movements necessary for substrate unfolding and translocation. **Proceedings of the National Academy of Sciences of the USA**, Whashington, v.106, p.21579–21584, Dec. 2009.

CASTRESANA, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. **Molecular Biology and Evolution**, Chicago, v.17, p.540-552, May 2000.

CASARI G.; de FUSCO M.; CIARMATORI S.; ZIEVIANI M.; MORA M.; FERNANDEZ P.; de MICHELE G.; FILLA A.; COCOZZA S.; MARCONI R.; DÜRR, A.; FONTAINE, B.; BALLABIO, A. Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. **Cell**, Oxford, v.93, p.973–983, June 1998.

CELNIKER, G.; NIMROD, G.; ASHKENAZY, H.; GLASER, F.; MARTZ, E.; MAYROSE, I.; PUPKO, T.; BEN-TAL, N. 2013. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. **Israel Journal of Chemistry**, Jerusalem, v.53, p.199-206, Mar. 2013.

CHEN, M.; CHOI, Y.; VOYTAS, D.F.; RODERMEL, S. Mutations in the *Arabidopsis* VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. **The Plant Journal**, Oxford, v.22, p.303-313, May 2000.

COPPOLA, M.; PIZZIGONI, A.; BANFI, S.; BASSI, M.T.; CASARI, G.; INCERTI, B. Identification and characterization of Yme1L1, a novel paraplegin related gene. **Genomics**, London, v.66, p.48–54, May 2000.

DARRIBA, D.; TABOADA, G.L.; DOALLO, R.; POSADA D. jModelTest 2: more models, new heuristics and parallel computing. **Nature Methods**, London, v.9, p.772, Jan. 2012.

EISENBERG, D.; SHWARZ, E.; KOMARONY, M.; WALL, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. **Journal of Molecular Biology**, London, v.200, p.513-522, Oct.1984.

FERRO, M.; BRUGIERE, S.; SALVI, D.; SEIGNEURIN-BERNY, D.; COURT, M.; MOYET, L.; RAMUS, C.; MIRAS, S.; MELLAL, M.; LE GALL S.; KIEFFER-JAQUINOD, S.; BRULEY, C.; GARIN, J.; JOYARD, J.; MASSELON, C.; ROLLAND, N. AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. **Molecular Cell Proteomics**, Rockville, v.9, p.1063–1084, Jan. 2010.

FRISO, G.; GIACOMELLI L.; YTTERBERG A.J.; PELTIER J.B.; RUDELLA A.; SUN Q.; van WIJK K.J. In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. **The Plant Cell**, Rockville, v.16, p.478-499, Feb. 2004.

GARCIA-LORENZO M.; SJODIN A.; JANSSON S.; FUNK C. Protease gene families in *Populus* and *Arabidopsis*. **BMC Plant Biology**, London, v.6, p.30, Dec. 2006.

GOH, C.S.; BOGAN, A.A.; JOACHIMIAK, M.; WALTHER, D.; COHEN, F.E. Co-evolution of proteins with their interaction partners. **Journal of Molecular Biology**, London, v.299, p.283-293, June 2000.

HOOPER N.M. Families of zinc metalloproteases. **FEBS Letters**, Amsterdam, v.354, p.1–6, Oct. 1994.

HUGUENEY P.; BOUVIER F.; BADILLO A.; DHARLINGUE A.; KUNTZ M.; CAMARA B. Identification of a plastid protein involved in vesicle fusion and/or membrane-protein translocation. **Proceedings of the National Academy of Sciences of the USA**, Washington, v.92, p.5630–v.5634, June 1995.

HURST L.D. The Ka/Ks ratio: diagnosing the form of sequence evolution. **Trends in Genetics**, Cambridge, v.18, p.486-489, Sept. 2002.

JANSKA H. ATP-dependent proteases in plant mitochondria: what do we know about them today? **Physiologia Plantarum**, Lund, v.123, p.399–405, Dec. 2005.

JANSKA, H.; KWASNIAK, M.; SZCZEPANOWSKA J. Protein quality control in organelles – AAA/FtsH story. **Biochimica et Biophysica Acta**, Amsterdam, v.1833, p.381-387, April 2013.

ITO, K.; AKIYAMA, Y. Cellular functions, mechanism of action, and regulation of FtsH protease. **Annual Review of Microbiology**, Palo Alto, v. 59, p. 211-231, Feb. 2005.

KARIMI, M.; INZÉ, D.; DEPICKER, A. GATEWAY vectors for Agrobacterium-mediated plant transformation. **Trends in Plant Science**, Cambridge, v.7, n.5, p.193-195, May 2002.

KATO, Y.; MIURA, E.; MATSUSHIMA, R.; SAKAMOTO, W. White leaf sectors in yellow variegated2 are formed by viable cells with undifferentiated plastids. **Plant Physiology**, Lancaster, v.144, p.952–960, June 2007.

KATO, Y.; MIURA, E.; IDO, K.; IFUKU, K.; SAKAMOTO, W. The variegated mutants lacking chloroplastic FtsHs are defective in D1 degradation and accumulate reactive oxygen species. **Plant Physiology**, Lancaster, v.151, p.790–1801, Dec. 2009.

KATO Y.; KOUSO T.; SAKAMOTO W. Variegated Tobacco Leaves Generated by Chloroplast FtsH Suppression: Implication of FtsH Function in the Maintenance of Thylakoid Membranes. **Plant Cell Physiology**, London, v.53, p.391-404, Feb. 2012.

KATO, Y.; SUN, X.; ZHANG, L.; SAKAMOTO, W. Cooperative D1 degradation in the photosystem II repair mediated by chloroplastic proteases in Arabidopsis. **Plant Physiology**, Lancaster, v. 159, p. 1428-1439, Aug. 2012.

KATOH K.; STANDLEY D. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. **Molecular Biology and Evolution**, Chicago, v.30, p.772-780, Jan. 2013.

KILIAN, J.; WHITEHEAD, D.; HORAK, J.; WANKE, D.; WEINL, S.; BATISTIC, O.; D'ANGELO, C.; BORNBERG-BAUER, E.; KUDLA, J.; HARTER K. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. **The Plant Journal**, Oxford, v.50, p.347-363, Mar. 2007.

KIHARA, A.; AKIYAMA, Y.; ITO, K. Host regulation of lysogenic decision in bacteriophage lambda: transmembrane modulation of FtsH (HflB), the cII degrading

protease, by HflKC (HflA). **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 94, p. 5544-5549, May 1997.

KRZYWDA, S.; BRZOZOWSKI, A.M.; VERMA, C.; KARATA, K.; OGURA, T.; WILKINSON, A.J. The crystal structure of the AAA domain of the ATP-dependent protease FtsH of *Escherichia coli* at 1.5 Å resolution. **Structure**, London, v. 10, p. 1073-1083, Aug. 2002.

LEONHARD K.; HERRMANN J.M.; STUART R.A.; MANNHAUPT, G.; NEUPERT, W.; LANGER T. AAA proteases with catalytic sites on opposite membrane surfaces comprise a proteolytic system for the ATP-dependent degradation of inner membrane proteins in mitochondria. **The EMBO Journal**, London, v.15, p.4218–4229, Aug. 1996.

LINDAHL, M.; SPETEA, C.; HUNDAL, T.; OPPENHEIM A.B.; ADAM Z.; ANDERSSON, B. The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. **The Plant Cell**, Rockville, v.12, p.419–431, Mar. 2000.

LINDAHL, M.; TABAK, S.; CSEKE, L.; PICHERSKY, E.; ANDERSSON, B.; ADAM, Z. Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants. **The Journal Biological Chemistry**, Baltimore, v.271, p.29329–29334, Nov. 1996.

LYSKA, D.; MEIERHOFF, K.; WESTHOFF, P. How to build functional thylakoid membranes: from plastid transcription to protein complex assembly. **Planta**, Berlin, v. 237, p. 413-428, Feb. 2013.

LIU, X.; YU, F.; RODERMEL, S. Arabidopsis chloroplast FtsH, var2 and suppressor of var2 leaf variegation: a review. **Journal of Integrative Plant Biology**, Hoboken, v.52, p.750–761, Aug. 2010.

MALNOE, A.; WOLLMAN, F.A.; DE VITRY C.; RAPPAPORT, F. Photosynthetic growth despite a broken Q-cycle. **Nature Communications**, London, v.2, p.301-304, April 2011.

MARGULIS, L. **Origins of eukaryotic cells**. New Haven: Yale University Press, 1970. 349p.

MIURA, E.; KATO, Y.; MATSUSHIMA, R.; ALBRECHT, V.; LAALAMI, S.; SAKAMOTO, W. The balance between protein synthesis and degradation in chloroplasts determines leaf variegation in Arabidopsis yellow variegated mutants. **The Plant Cell**, Rockville, v.19, p.1313–1328, April 2007.

MURASHIGE, T.; SKOOG, F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. **Physiologia Plantarum**, Lund, v.5, p.473-498, April 1962.

OGURA, T.; WILKINSON, A.J. AAA(+) superfamily ATPases: common structure-diverse function. **Genes to Cells**, Malden, v.6, p.575–597, July 2001.

OGURA, T.; TOMOYASU, T.; YUKI, T.; MORIMURA, S.; BEGG, K.J. DONACHIE, W.D.; MORI, H.; NIKI, H.; HIRAGA, S. Structure and function of the ftsH gene in Escherichia coli. **Research in Microbiology**, Paris, v. 142, p. 279-282, April 1991.

OSTERSETZER, O.; ADAM, Z. Light-stimulated degradation of an unassembled Rieske FeS protein by a thylakoid-bound protease: the possible role of the FtsH protease. **The Plant Cell**, Rockville, v.9, p.957–965, June 1997.

PAZOS, F.; HELMER-CITTERICH, M.; AUSIELLO, G.; VALENCIA, A. Correlated mutations contain information about protein-protein interaction. **Journal of Molecular Biology**, London, v.271, p.511-523, Aug. 1997.

PAZOS, F.; VALENCIA, A. Similarity of phylogenetic trees as an indicator of protein-protein interaction. **Protein Engineering Design and Selection**, Madrid, v.14, p.609-614, Sept. 2001.

PIECHOTA, J.; KOLODZIEJCZAK, M.; JUSZCZAK, I.; SAKAMOTO, W.; JANSKA, H. Identification and characterization of high molecular weight complexes formed by matrix AAA proteases and prohibitins in mitochondria of Arabidopsis thaliana. **The Journal of Biological Chemistry**, Baltimore, v.285, p.12512–12521, April 2010.

RODRIGUES, R.A.O. **Estudo do direcionamento das proteases FtsH plastidiais às membranas dos tilacóides**. 2011. 123p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2011.

RODRIGUES, R.A.O.; SILVA FILHO, M.C.; CLINE, K. FtsH2 and FtsH5: two homologous subunits use different integration mechanisms leading to the same thylakoid multimeric complex. **The Plant Journal**, Oxford, v. 65, p. 600-609, Feb. 2011.

RONQUIST, F.; HUELSENBECK, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. **Bioinformatics**, Oxford, v.19, p.1572-1574, August 2003.

SAKAMOTO, W.; TAMURA, T.; HANBA-TOMITA, Y.; SODMERGEN, MURATA M. The VAR1 locus of Arabidopsis encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. **Genes to Cells**, Malden, v.7, p.769-780, Aug. 2002.

SAKAMOTO, W.; ZALTSMAN, A.; ADAM, Z.; TAKAHASHI, Y. Coordinated regulation and complex formation of YELLOW VARIEGATED1 and YELLOW VARIEGATED2, chloroplastic FtsH metalloproteases involved in the repair cycle of photosystem II in Arabidopsis thylakoid membranes. **The Plant Cell**, Rockville, v.15, p.2843-2855, Dec. 2003.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. Molecular cloning - A Laboratory Manual, Ed.2 New York: Cold Spring Harbor Laboratory Press, 1989. 1889p.

SANTOS, D.; DE ALMEIDA D.F. Isolation and characterization of a new temperature-sensitive cell division mutante of *Escherichia coli* K-12. **Journal of Bacteriology**, New York, v.124, p.1502-1507, Dec. 1975.

SEO, S.; OKAMOTO, M.; IWAI, T.; IWANO, M.; FUKUI, K.; ISOGAI, A.; NAKAJIMA, N.; OHASHI, Y. Reduced levels of chloroplast FtsH protein in tobacco mosaic virus-infected tobacco leaves accelerate the hypersensitive reaction. **The Plant Cell**, Rockville, v.12, p.917-932, June 2000.

SHI, Y.Y.; TANG, W.; HAO, S.F.; WANG, C.C. Contributions of cysteine residues in Zn2 to zinc fingers and thiol-disulfide oxidoreductase activities of chaperone DnaJ. **Biochemistry**, Washington, v. 44, p. 1683-1689, Feb. 2005.

SINVANY-VILLALOBO, G.; DAVYDOV, O.; BEN-ARI, G.; ZALTSMAN, A.; RASKIND, A.; ADAM, Z. Expression in multigene families. Analysis of chloroplast and mitochondrial proteases. **Plant Physiology**, Lancaster, v.135, p.1336-1345, July 2004.

SOKOLENKO, A.; POJIDAEVA, E.; ZINCHENKO, V.; PANICHKIN, V.; GLASER, V.M.; HERRMANN, R.G.; SHESTAKOV, S.V. The gene complement for proteolysis in the cyanobacterium *Synechocystis* sp. PCC 6803 and *Arabidopsis thaliana* chloroplasts. **Current Genetics**, New York, v.41, p.291-310, Aug. 2002.

SUMMER, E.J.; MORI, H.; SETTLES, A.M.; CLINE, K. The thylakoid delta pH-dependent pathway machinery facilitates RR-independent N-tail protein integration. **The Journal of Biological Chemistry**, Baltimore, v.275, p.23483-23490, Aug. 2000.

TAKECHI, K.; SODMERGEN, MURATA M.; MOTOYOSHI, F.; SAKAMOTO W. The YELLOW VARIEGATED (VAR2) locus encodes a homologue of FtsH, an ATP-dependent protease in Arabidopsis. **Plant Cell Physiology**, London, v.41, p.1334-1346, Dec. 2000.

TALAVERA, G.; CASTRESANA, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. **Systematic Biology**, Oxford, v.56, p.564-577, Mar. 2007.

TAVARÉ S. Some probabilistic and statistical problems in the analysis of DNA sequences. In: MIURA R.M.(Ed.). Lectures on mathematics in the life sciences. Providence (RI): **American Mathematical Society**, Washsington, v.17, p.57–86, June 1986.

TOMOYASU, T.; YUKI, T.; MORIMURA, S.; MORI, H.; YAMANAKA, K.; NIKI, H.; HIRAGA, S.; OGURA, T. The Escherichia coli FtsH protein is a prokaryotic member of a protein family of putative ATPases involved in membrane functions, cell-cycle control, and gene-expression. **Journal of Bacteriology**, New York, v.175, p.1344–1351, Mar. 1993.

URANTOWKA, A.; KNORPP, C.; OLCZAK, T.; KOLODZIEJCZAK, M.; JANSKA, H. Plant mitochondria contain at least two i-AAA-like complexes. **Plant Molecular Biology**, Boston, v.59, p.239–252, Sept. 2005.

WALSH, P.; BURSAC, D.; LAW, Y. C.; CYR, D.; LITHGOW, T. The J-protein family: modulating protein assembly, disassembly and translocation. **EMBO Reports**, London, v. 5, p. 567-571, June 2004.

WATERHOUSE, A.M.; PROCTER, J.B.; MARTIN, D.M.A.; CLAMP, M.; BARTON, G.J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, Oxford, v.25, p.1189-1191, 2009.

WHELAN, S.; GOLDMAN, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. **Molecular Biology and Evolution**, Chicago, v.18, p.691-699, May 2001.

WINTER, D.; VINEGAR, B.; NAHAL, H.; AMMAR, R.; WILSON, G.V.; PROVART, N.J. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. **PLoS One**, San Francisco, v.2, p. e718, Aug. 2007.

YANG, Z.; BIELAWSKI, J.P. Statistical methods for detecting molecular adaptation. **Trends in Ecology & Evolution**, Amsterdam, v.15, p.496-503, Dec. 2000.

YANG, Z.; NIELSEN, R. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. **Journal of Molecular Evolution**, Berlin, v.46, p.409-418, April 1998.

YANG, Z.; NIELSEN, R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. **Molecular Biology and Evolution**, Chicago, v.19, p.908-917, June 2002.

YANG, Z.; NIELSEN, R. PAML 4: phylogenetic analysis by maximum likelihood. **Molecular Biology and Evolution**, Chicago, v.24, n.8, p.1586-1591, May 2007.

YOSHIOKA, M.; NAKAYAMA, Y.; YOSHIDA, M.; OHASHI, K.; MORITA, N.; KOBAYASHI, H.; YAMAMOTO, Y. Quality control of photosystem II: FtsH hexamers are localized near photosystem II at grana for the swift repair of damage. **The Journal of Biological Chemistry**, Baltimore, v. 285, p. 41972-41981, Dec. 2010.

YOSHIOKA, M.; YAMAMOTO, Y. Quality control of photosystem II: where and how does the degradation of the D1 protein by FtsH proteases start under light stress?-- Facts and hypotheses. **Journal of Photochemistry and Photobiology**, Lausanne, v. 104, p. 229-235, Aug. 2011.

YU, F.; PARK, S.; RODERMEL, S.R. The Arabidopsis FtsH metalloprotease gene family: interchangeability of subunits in chloroplast oligomeric complexes. **The Plant Journal**, Oxford, v.37, p.864-876, Mar 2004.

YU F.; PARK, S.; RODERMEL, S.R. Functional redundancy of AtFtsH metalloproteases in thylakoid membrane complexes. **Plant Physiology**, Lancaster, v.138, p.1957-1966, Aug. 2005.

YU, F.; PARK, S.S.; LIU, X.; FODREE, A.; FU, A.; POWIKROWSKA, M.; KHROUCHTCHOVA, A.; JENSEN, PE.; KRIGER, J.N.; GRAY G.R.; RODERMEL, S.R. Suppressor of variegation4, a new var2 suppressor locus, encodes a pioneer protein that is required for chloroplast biogenesis. **Molecular Plant**, London, v.4, p.229-240, Mar. 2011.

ZALTSMAN, A.; FEDER, A.; ADAM Z. Developmental and light effects on the accumulation of FtsH protease in Arabidopsis chloroplasts — implications for thylakoid formation and photosystem II maintenance. **The Plant Journal**, Oxford, v.42, p.609-617, 2005a.

ZALTSMAN, A.; ORI, N.; ADAM Z. Two types of FtsH protease subunits are required for chloroplast biogenesis and Photosystem II repair in Arabidopsis. **The Plant Cell**, Rockville, v.17, p.2782-2790, Oct. 2005b.

ZIMMERMANN, P. HIRSCH-HOFFMANN, M.; HENNIG, L. GRUISSEM, W. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. **Plant Physiology**, Lancaster, v.136, p.2621-2632, Sept. 2004.

ZHANG, D.; KATO, Y.; ZHANG, L.; FUJIMOTO, M.; TSUTSUMI, N.; SODMERGEN, SAKAMOTO W. The FtsH protease heterocomplex in Arabidopsis: dispensability of type-B protease activity for proper chloroplast development. **The Plant Cell**, Rockville, v.22, p.3710–3725, Nov. 2010.

ANEXOS

ANEXO A – Descrição das sequências de nucleotídeos homólogas ao gene de FIP (At5g02160). Na primeira coluna encontram-se os nomes arbitrariamente atribuídos para as sequências, compostos pela primeira letra do gênero e três primeiras da espécie, recebendo um número para diferenciar entre as sequências provenientes da mesma espécie. A segunda coluna referencia a espécie de origem da sequência e a terceira coluna o identificador. Existem dois tipos de identificadores, o GI (NCBI) e *transcript name* (Phytozome), devido a presença de sequências de ambos os bancos de dados

NOME	ESPÉCIE	IDENTIFICADOR
A_tri1	<i>Amborella trichopoda</i>	586708094
A_cae1	<i>Aquilegia caerulea</i>	Aquca_009_00932.1
A_lyr1	<i>Arabidopsis lyrata</i>	297806146
FIP	<i>Arabidopsis thaliana</i>	240256250
B_dis1	<i>Brachypodium distachyon</i>	357134998
B_rap1	<i>Brassica rapa</i>	Bra009609
C_rub1	<i>Capsella rubella</i>	565461773
C_pap1	<i>Carica papaya</i>	evm.model.supercontig_169.14
C_ari1	<i>Cicer arietinum</i>	502168463
C_cle1	<i>Citrus clementina</i>	567873076
C_sin1	<i>Citrus sinensis</i>	orange1.1g032793m
C_sat1	<i>Cucumis sativus</i>	449440037
E_gra1	<i>Eucalyptus grandis</i>	Eucgr.K03108.1
E_sal1	<i>Eutrema salsugineum</i>	567169496
F_ves1	<i>Fragaria vesca</i>	470133865
G_max1	<i>Glycine max</i>	403043537
G_max2	<i>Glycine max</i>	403043669
G_rai1	<i>Gossypium raimondii</i>	Gorai.006G171200.1
J_cur1	<i>Jatropha curcas</i> cultivar GZQX0401	643726927
L_usi1	<i>Linum usitatissimum</i>	Lus10024373
M_dom1	<i>Malus domestica</i>	657956878
M_dom2	<i>Malus domestica</i>	657991126
M_esc1	<i>Manihot esculenta</i>	cassava4.1_019008m
M_not1	<i>Morus notabilis</i>	514344252
O_sat1	<i>Oryza sativa Japonica Group</i>	115461676
P_vir1	<i>Panicum virgatum</i>	Pavirv00019140m
P_vir2	<i>Panicum virgatum</i>	Pavirv00031963m
P_vul1	<i>Phaseolus vulgaris</i>	593424697
P_pat1	<i>Physcomitrella patens subsp. patens</i>	167997870
P_pat2	<i>Physcomitrella patens subsp. patens</i>	168019145
P_tri1	<i>Populus trichocarpa</i>	566175264
P_mum1	<i>Prunus mume</i>	645220536
R_com1	<i>Ricinus communis</i>	255581503
S_ita1	<i>Setaria italica</i>	514744032
S_lyc1	<i>Solanum lycopersicum</i>	461496376
S_tub1	<i>Solanum tuberosum</i>	565365602
S_bic1	<i>Sorghum bicolor</i>	242086598
T_cac1	<i>Theobroma cacao</i>	590629368
V_vin1	<i>Vitis vinifera</i>	225440409

ANEXO B – Descrição das sequências de nucleotídeos homólogas aos genes das protease do complexo FtsH tilacoidal (FtsH1: At1g50250; FtsH2: At2g30950; FtsH5: At5g42270; FtsH8: AT1G06430). Na primeira coluna encontram-se os nomes arbitrariamente atribuídos para as sequências, compostos pela primeira letra do gênero e três primeiras da espécie, recebendo um número para indicar através de qual FtsH (FtsH1, 2, 5 ou 8) identificou-se a sequência e outro para diferenciar entre as sequências de uma mesma espécie. A segunda coluna referencia a espécie de origem da sequência e a terceira coluna o identificador. Existem quatro tipos de identificadores, o GI (NCBI), *transcript name* (Phytozome), *Transcript ID* (EnsemblGenomes) e *Transcript ID* (JOINT GENOME INSTITUTE), sendo que os dois últimos correspondem apenas as sequências das espécies *Cyanidioschyzon merolae* e *Selaginella moellendorffii*, respectivamente. As sequências utilizadas como referência (Kato, Kouso & Sakamoto, 2012) estão inclusas e a nomenclatura destas diferencia-se daquela das outras sequências devido a presença da identificação real da protease, por exemplo a sequência GI 42562642 de *Arabidopsis thaliana*, a qual representa a protease FtsH1, foi nomeada A_thFtsH1

(continua)

NOME	ESPÉCIE	IDENTIFICADOR
A_tri1_1	<i>Amborella trichopoda</i>	586648540
A_tri2_1	<i>Amborella trichopoda</i>	586674718
A_cae1_1	<i>Aquilegia caerulea</i>	Aquca_015_00455.1 CDS
A_cae2_1	<i>Aquilegia caerulea</i>	Aquca_039_00040.1 CDS
A_cae8_3	<i>Aquilegia caerulea</i>	Aquca_039_00040.1 CDS
A_lyr2_1	<i>Arabidopsis lyrata subsp. lyrata</i>	297822822
A_lyr2_2	<i>Arabidopsis lyrata subsp. lyrata</i>	297848919
A_lyr5_1	<i>Arabidopsis lyrata subsp. lyrata</i>	297795318
A_lyr5_2	<i>Arabidopsis lyrata subsp. lyrata</i>	297852713
A_thFtsH1	<i>Arabidopsis thaliana</i>	42562642
A_thFtsH2	<i>Arabidopsis thaliana</i>	186504410
A_thFtsH5	<i>Arabidopsis thaliana</i>	145358751
A_thFtsH7	<i>Arabidopsis thaliana</i>	23198319
A_thFtsH8	<i>Arabidopsis thaliana</i>	186478199
A_thFtsH9	<i>Arabidopsis thaliana</i>	145359427
B_dys1_1	<i>Brachypodium distachyon</i>	357123106
B_dys2_1	<i>Brachypodium distachyon</i>	357123382
B_dys8_1	<i>Brachypodium distachyon</i>	357133400
B_rap1_1	<i>Brassica rapa</i>	Bra004247 CDS
B_rap2_1	<i>Brassica rapa</i>	Bra018288 CDS
B_rap2_2	<i>Brassica rapa</i>	Bra015499 CDS
B_rap2_3	<i>Brassica rapa</i>	Bra030652 CDS
C_rub1_1	<i>Capsella rubella</i>	565498507
C_rub1_2	<i>Capsella rubella</i>	565430212
C_rub2_1	<i>Capsella rubella</i>	565471133
C_rub2_2	<i>Capsella rubella</i>	565494629
C_pap1_1	<i>Carica papaya</i>	evm.model.supercontig_26.216 CDS
C_pap2_1	<i>Carica papaya</i>	evm.model.supercontig_140.49 CDS
C_reFtsH1	<i>Chlamydomonas reinhardtii</i>	159465356
C_reFtsH2	<i>Chlamydomonas reinhardtii</i>	159478021
C_ari1_1	<i>Cicer arietinum</i>	502131152
C_ari2_1	<i>Cicer arietinum</i>	502141847
C_ari2_2	<i>Cicer arietinum</i>	659972909
C_cle1_1	<i>Citrus clementine</i>	567885528
C_cle2_1	<i>Citrus clementine</i>	567909038

(continuação)

NOME	ESPÉCIE	IDENTIFICADOR
C_cle2_2	<i>Citrus clementine</i>	567877136
C_sin1_1	<i>Citrus sinensis</i>	568839577
C_sin2_3	<i>Citrus sinensis</i>	641844420
C_sin2_6	<i>Citrus sinensis</i>	641853434
C_sas1_1	<i>Cucumis sativus</i>	449455580
C_sat2_1	<i>Cucumis sativus</i>	449438118
C_meFtsH1	<i>Cyanidioschyzon merolae</i>	CMV129C
C_meFtsH2	<i>Cyanidioschyzon merolae</i>	CMQ295C
C_meFtsH3	<i>Cyanidioschyzon merolae</i>	CME055C
C_meFtsH4	<i>Cyanidioschyzon merolae</i>	CMF159C
E_gra1_1	<i>Eucaliptus grandis</i>	Eucgr.F01206.1 CDS
E_gra2_1	<i>Eucaliptus grandis</i>	Eucgr.A01933.1 CDS
E_gra2_2	<i>Eucaliptus grandis</i>	Eucgr.J01348.1 CDS
E_sal2_1	<i>Eutrema salsugineum</i>	567210855
E_sal2_2	<i>Eutrema salsugineum</i>	567154924
E_sal2_4	<i>Eutrema salsugineum</i>	567174135
E_sal5_1	<i>Eutrema salsugineum</i>	567207489
E_sal5_2	<i>Eutrema salsugineum</i>	567124624
F_ves1_1	<i>Fragaria vesca subsp. vesca</i>	470109892
F_ves2_1	<i>Fragaria vesca subsp. vesca</i>	470134753
G_max1_1	<i>Glycine max</i>	571448767
G_max1_2	<i>Glycine max</i>	571458510
G_max2_1	<i>Glycine max</i>	571455322
G_max2_2	<i>Glycine max</i>	571476429
G_max2_4	<i>Glycine max</i>	571470518
G_rai1_1	<i>Gossypium raimondii</i>	Gorai.002G118800.1 CDS
G_rai1_3	<i>Gossypium raimondii</i>	Gorai.009G188700.1 CDS
G_rai2_2	<i>Gossypium raimondii</i>	Gorai.002G023400.1 CDS
G_rai2_4	<i>Gossypium raimondii</i>	Gorai.012G053900.1 CDS
J_cur1_1	<i>Jatropha curcas</i> cultivar GZQX0400	643728757
J_cur2_1	<i>Jatropha curcas</i> cultivar GZQX0401	643714359
L_usi1_1	<i>Linun usitatissimum</i>	Lus10016003 CDS
L_usi2_1	<i>Linun usitatissimum</i>	Lus10001420 CDS
L_usi5_1	<i>Linun usitatissimum</i>	Lus10021824 CDS
M_dom1_1	<i>Malus domestica</i>	657970654
M_dom2_1	<i>Malus domestica</i>	658053517
M_esc1_1	<i>Manihot esculenta</i>	cassava4.1_002593m CDS
M_esc1_2	<i>Manihot esculenta</i>	cassava4.1_002590m CDS
M_esc2_1	<i>Manihot esculenta</i>	cassava4.1_002703m CDS
M_esc8_2	<i>Manihot esculenta</i>	cassava4.1_003000m CDS
M_not1_1	<i>Morus notabilis</i>	514345737
M_not2_1	<i>Morus notabilis</i>	514344960
N_taFtsH1	<i>Nicotiana tabacum</i>	5804781

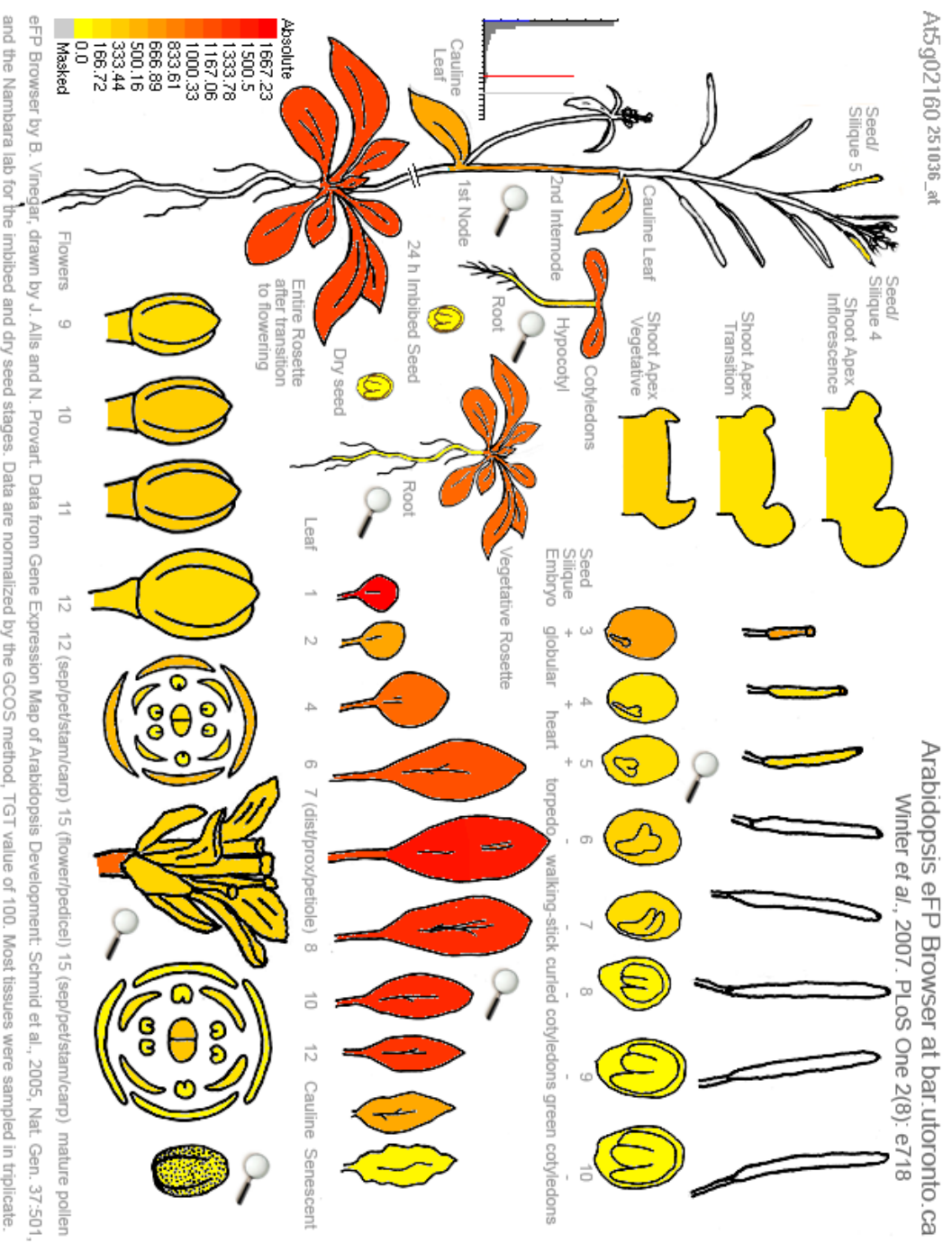
(continuação)

NOME	ESPÉCIE	IDENTIFICADOR
N_taFtsH2	<i>Nicotiana tabacum</i>	4325040
O_saFtsH2	<i>Oryza sativa Japonica</i> Group	115470051
O_saFtsH3	<i>Oryza sativa Japonica</i> Group	297724618
O_saFtsH4	<i>Oryza sativa Japonica</i> Group	115447608
P_vir1_1	<i>Panicum virgatum</i>	Pavirv00031442m CDS
P_vir1_2	<i>Panicum virgatum</i>	Pavirv00030414m CDS
P_vir2_1	<i>Panicum virgatum</i>	Pavirv00061677m CDS
P_vir2_3	<i>Panicum virgatum</i>	Pavirv00059443m CDS
P_vir2_4	<i>Panicum virgatum</i>	Pavirv00019894m CDS
P_vul1_1	<i>Phaseolus vulgaris</i>	593267926
P_vul2_1	<i>Phaseolus vulgaris</i>	593791713
P_vul2_2	<i>Phaseolus vulgaris</i>	593330812
P_vul2_3	<i>Phaseolus vulgaris</i>	593332286
P_paFtsH1	<i>Physcomitrella patens</i> subsp. <i>patens</i>	168044123
P_paFtsH2	<i>Physcomitrella patens</i> subsp. <i>patens</i>	168001965
P_paFtsH3	<i>Physcomitrella patens</i> subsp. <i>patens</i>	168016253
P_tri1_1	<i>Populus trichocarpa</i>	224065698
P_tri1_2	<i>Populus trichocarpa</i>	566173471
P_tri2_1	<i>Populus trichocarpa</i>	566204124
P_tri2_2	<i>Populus trichocarpa</i>	566159451
P_tri2_3	<i>Populus trichocarpa</i>	566212845
P_mum1_1	<i>Prunus mume</i>	645228779
P_mum2_1	<i>Prunus mume</i>	645245588
R_com1_1	<i>Ricinus communis</i>	255539167
R_com2_1	<i>Ricinus communis</i>	255558697
S_moFtsH1	<i>Selaginella moellendorffii</i>	jjgijSelmo1j152929jestExt_ Genewise1.C_390123
S_moFtsH2	<i>Selaginella moellendorffii</i>	jjgijSelmo1j149856jestExt_ Genewise1.C_240696
S_moFtsH3	<i>Selaginella moellendorffii</i>	jjgijSelmo1j144519jestExt_ Genewise1.C_80737
S_ita1_1	<i>Setaria italica</i>	514768491
S_ita2_1	<i>Setaria italica</i>	514767717
S_ita8_1	<i>Setaria italica</i>	514749446
S_lyc1_1	<i>Solanum lycopersicum</i>	350535466
S_lyc2_1	<i>Solanum lycopersicum</i>	460395389
S_lyc8_2	<i>Solanum lycopersicum</i>	460395389
S_tub1_1	<i>Solanum tuberosum</i>	565350307
S_tub2_1	<i>Solanum tuberosum</i>	565364367
S_tub8_2	<i>Solanum tuberosum</i>	565364367
S_bic1_1	<i>Sorghum bicolor</i>	Sb10g030720.1 CDS
S_bic2_1	<i>Sorghum bicolor</i>	242095231

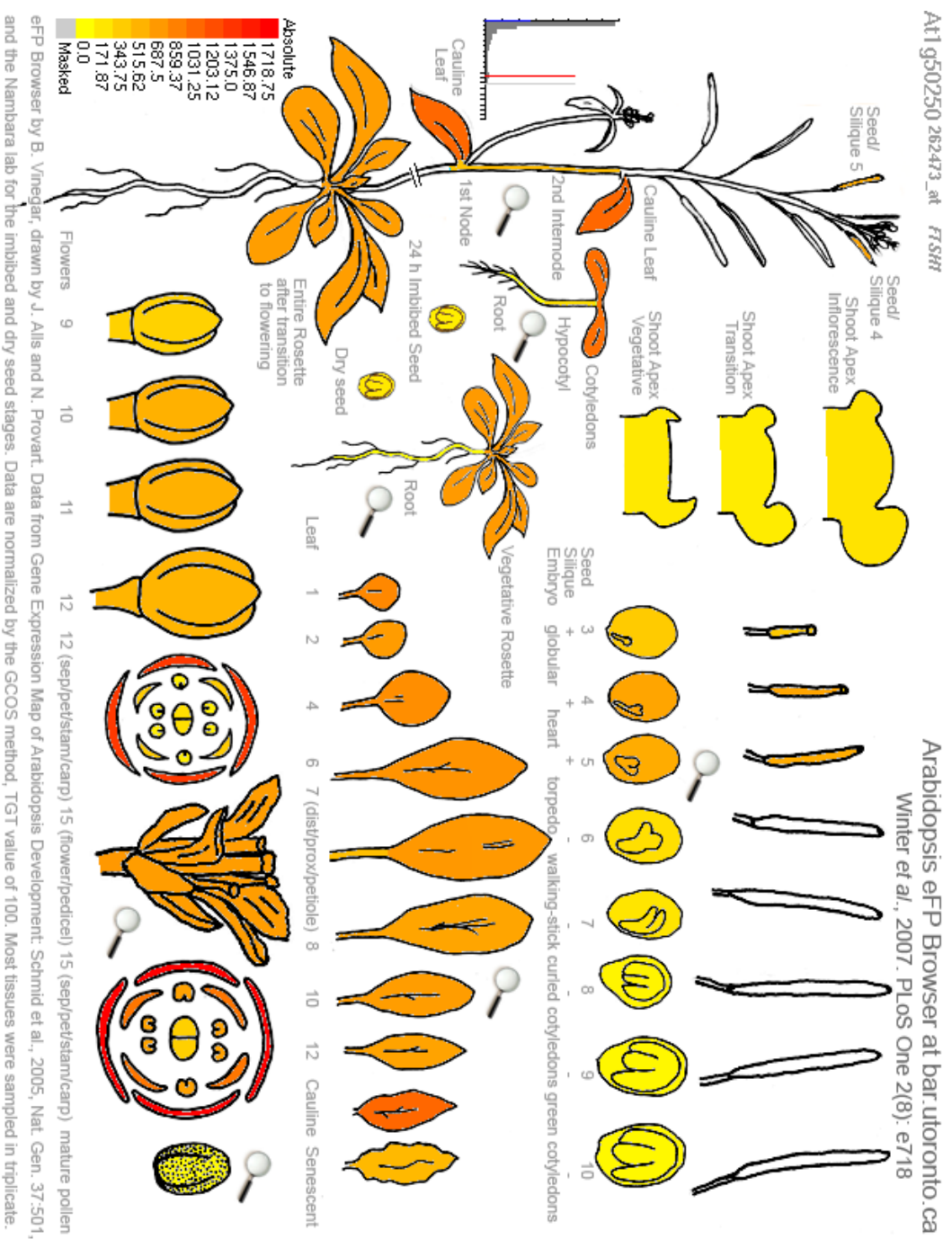
(conclusão)

NOME	ESPÉCIE	IDENTIFICADOR
SyneFtsH1	<i>Synechocystis</i> sp. PCC 6803	16329170
SyneFtsH2	<i>Synechocystis</i> sp. PCC 6803	16329170
SyneFtsH3	<i>Synechocystis</i> sp. PCC 6803	16329170
T_cac1_1	<i>Theobroma cacao</i>	590595032
T_cac2_1	<i>Theobroma cacao</i>	590692585
V_vin1_1	<i>Vitis vinifera</i>	359492543
V_vin2_1	<i>Vitis vinifera</i>	359485310
V_vin2_3	<i>Vitis vinifera</i>	147852345

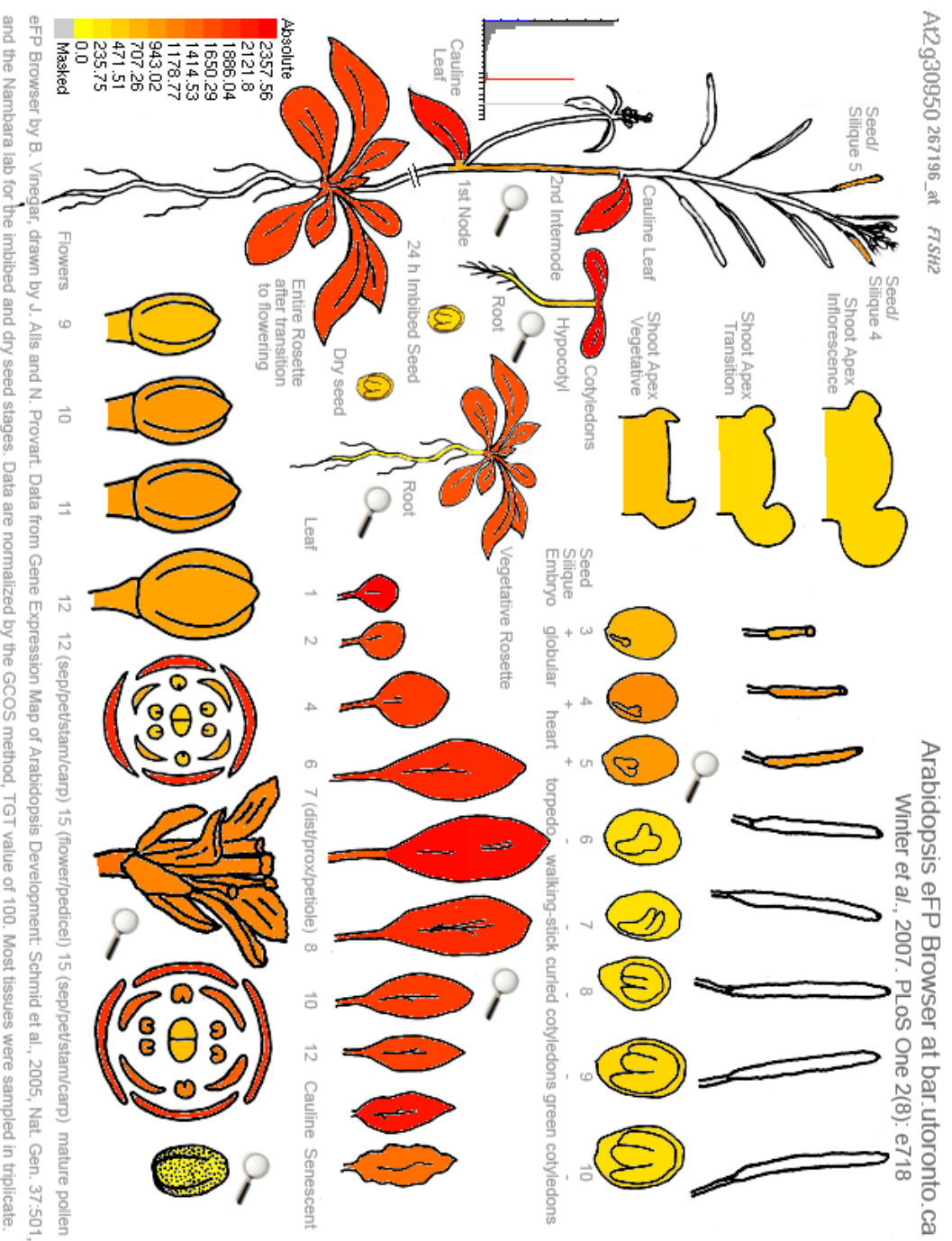
ANEXO C - Representação pictográfica dos níveis e padrões de expressão tecido-específica de FIP (At5g02160). As imagens pictográficas representando os padrões e níveis da expressão tecido-específica foram criados utilizando a ferramenta eFP browser (Winter et al., 2007; <http://bar.utoronto.ca>), a partir de dados públicos provenientes dos bancos de dados do consórcio Bio-Array Resource for Arabidopsis Functional Genomics (BAR) (<http://bar.utoronto.ca>)



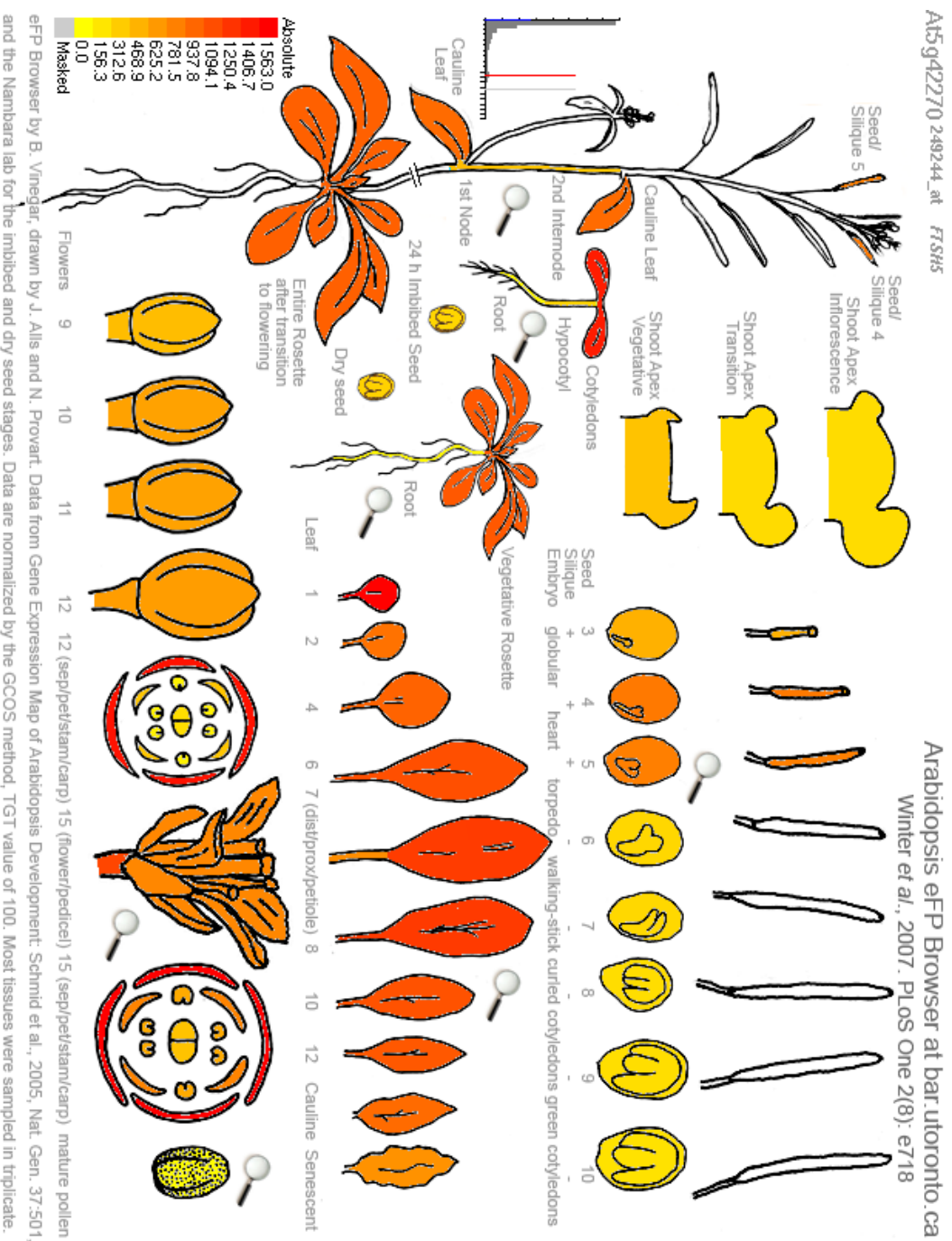
ANEXO D - Representação pictográfica dos níveis e padrões de expressão tecido-específica de FISH1 (At1g50250). As imagens pictográficas representando os padrões e níveis da expressão tecido-específica foram criados utilizando a ferramenta eFP browser (Winter et al., 2007; <http://bar.utoronto.ca>), a partir de dados públicos provenientes dos bancos de dados do consórcio Bio-Array Resource for Arabidopsis Functional Genomics (BAR) (<http://bar.utoronto.ca>)

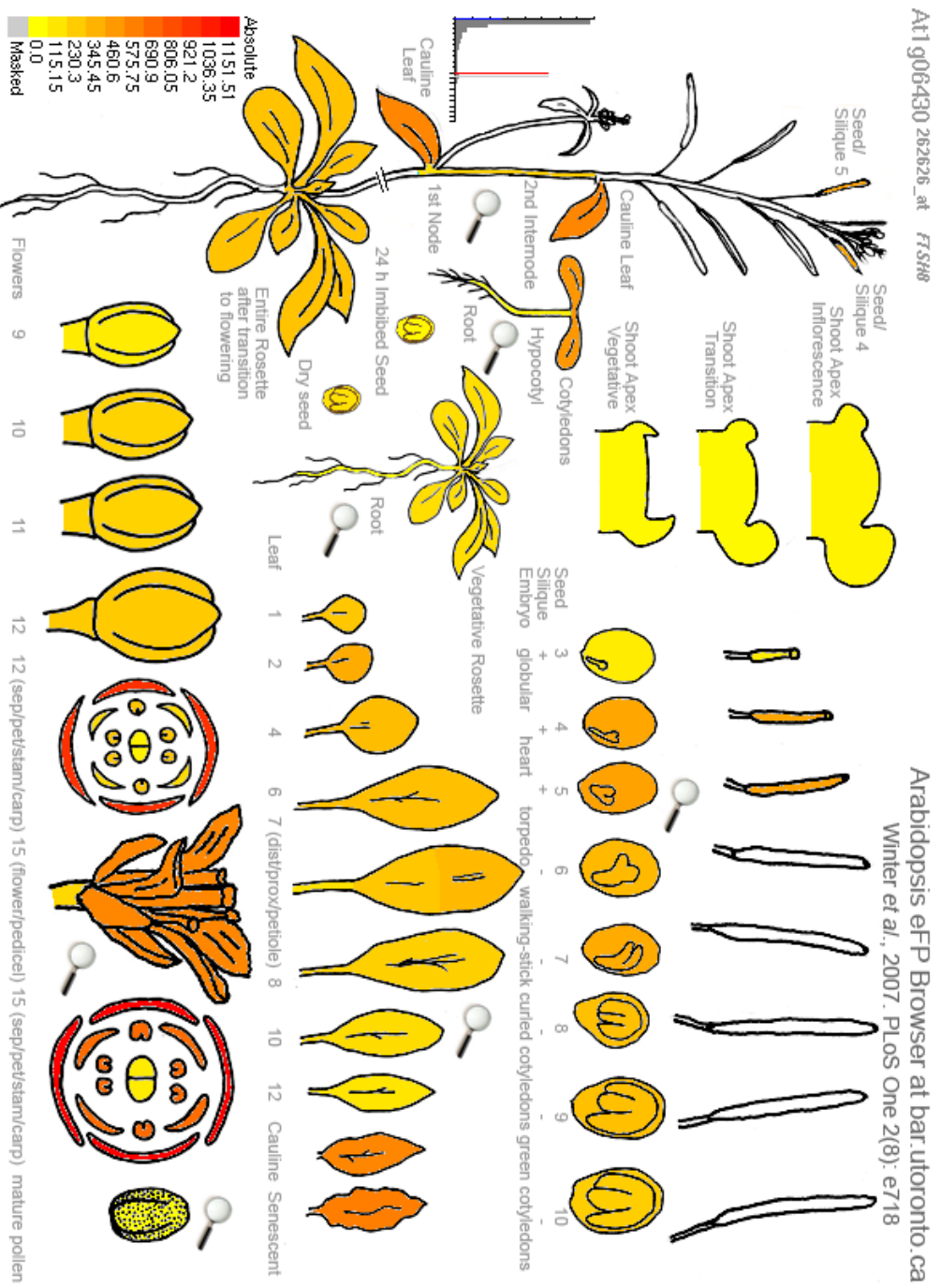


ANEXO E - Representação pictográfica dos níveis e padrões de expressão tecido-específica de FISH2 (At2g30950). As imagens pictográficas representando os padrões e níveis da expressão tecido-específica foram criados utilizando a ferramenta eFP browser (Winter et al., 2007; <http://bar.utoronto.ca>), a partir de dados públicos provenientes dos bancos de dados do consórcio *Bio-Array Resource for Arabidopsis Functional Genomics* (BAR) (<http://bar.utoronto.ca>).



ANEXO F - Representação pictográfica dos níveis e padrões de expressão tecida-específica de FISH5 (At5g42270). As imagens pictográficas representando os padrões e níveis da expressão tecido-específica foram criados utilizando a ferramenta eFP browser (Winter et al., 2007; <http://bar.utoronto.ca>), a partir de dados públicos provenientes dos bancos de dados do consórcio Bio-Array Resource for Arabidopsis Functional Genomics (BAR) (<http://bar.utoronto.ca>).





eFP Browser by B. Vinograd, drawn by J. Allis and N. Provart. Data from Gene Expression Map of Arabidopsis Development: Schmidt et al., 2005; Nat. Gen. 37:501, and the Nambara lab for the imbibed and dry seed stages. Data are normalized by the GCOS method, TGT value of 100. Most tissues were sampled in triplicate.