

HISTÓRIA EVOLUTIVA DAS METALOPROTEASES AAA

PHELLIPPE ARTHUR SANTOS MARBACH

Licenciado em Ciências Biológicas

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

Tese apresentada à Escola Superior de
Agricultura “Luiz de Queiroz”,
Universidade de São Paulo, para obtenção
do título de Doutor em Agronomia, Área de
Concentração: Genética e Melhoramento
de Plantas.

PIRACICABA
Estado de São Paulo – Brasil
Março – 2004

**Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Marbach, Phellippe Arthur Santos

História evolutiva das metaloproteases AAA / Phellippe Arthur Santos Marbach. -
Piracicaba, 2004.

83 p.

Tese (doutorado) - Escola Superior de Agricultura Luiz de Queiroz, 2004.
Bibliografia.

1. Evolução molecular 2. Filogenia 3. Proteína vegetal I. Título

CDD 574.19245

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

OFEREÇO

À minha mãe e aos verdadeiros amigos, aqueles com quem pode-se contar.
Sempre.

Faça ciência. Faça amor. A interpretação fica a seu critério

AGRADECIMENTOS

Ao Prof. Dr. Marcio de Castro Silva Filho pela oportunidade, pela orientação, e pelo exemplo profissional.

Ao Prof. Dr. Alexandre Siqueira Guedes Coelho, pelo apoio, pela coorientação, pela amizade, por me permitir ser por alguns dias um membro de sua família, a Betina, sua esposa, e Pedrão e Ana Julia, seus lindos e divertidos filhos. Não esperava encontrar nesta altura da vida mais um grande amigo, um irmão de vida.

Aos colegas do laboratório de biologia celular e molecular de plantas, ESALQ-USP, em especial a Celso Spada Fiori, pelo apoio e amizade.

Aos colegas do Laboratório de Citogenética, ESALQ - USP Dra. Janay A. Santos, Dr. Mateus Mondin, e Silvia, pela prestabilidade, pelo suporte técnico na utilização do microscópio e à Prof^a Dra Margarida L. R. de Aguiar Perecin por disponibilizar seu laboratório.

Ao Dr. Daniel Moura pelas longa e proveitosas discussões sobre ciência, pelas sugestões e correções na elaboração deste trabalho e pela amizade.

Aos os professores, alunos e funcionários do Departamento de Genética da ESALQ – USP.

Aos colegas do laboratório de cultura de tecidos de plantas e à Profa Dr. Maria Lúcia Carneiro Vieira, por disponibilizar seu laboratório

A Juliana Dantas de Almeida, a Ju, pela ajuda incondicional na elaboração deste manuscrito, pelo companherismo e amizade nos momentos de dificuldade e a Carolina Morgante Vianna, a Caróol, pela ajuda na confecção da tese.

À Fapesp, Fundação de Amparo à Pesquisa do Estado de São Paulo pelo apoio financeiro.

A Profa. Genovefa Peisino, ao Prof. Dr. Carlos Frederico Martins, ao Prof. Dr. José Maria, ao Prof. Dr. Wagner Otoni, Profa. Elizabeth Pacheco Batista Fontes, ao Prof. Dr. João Batista Teixeira e ao Prof. Dr. Marcio de Castro Silva Filho. Estes profissionais do ensino são minhas referências, são exemplos. Foram fundamentais e contribuíram a sua maneira na minha formação pessoal e profissional. Principalmente com eles aprendi a amar e a praticar uma das melhores criações humanas, a arte da ciência.

SUMÁRIO

	Página
RESUMO.....	Viii
SUMMARY.....	X
1 INTRODUÇÃO.....	1
2 REVISÃO DE LITERATURA.....	3
2.1 Proteínas AAA.....	3
2.2 Metaloproteases AAA em procariotos.....	5
2.3 Metaloproteases AAA eucarióticas.....	10
3 MATERIAL E MÉTODOS.....	14
3.1 Busca por seqüências de aminoácidos de metaloproteases AAA nos bancos de dados de seqüências.....	14
3.1.1 Estratégias empregadas.....	14
3.2 Análises filogenéticas das FtsHs procarióticas.....	15
3.2.1 Alinhamento múltiplo das seqüências de aminoácidos das FtsHs procarióticas.....	15
3.2.2 Alinhamento múltiplo das seqüências de aminoácidos das FtsHs eucarióticas.....	17
3.2.3 Método de análise filogenética empregado.....	17
3.4 Análise e caracterização estrutural das FtsHs.....	17
3.4.1 Obtenção dos perfis de hidropaticidade das FtsHs.....	17
3.4.2 Predição das estruturas secundárias das FtsHs procarióticas.....	18
3.5 Identificação e caracterização das FtsHs de Cana-de-açúcar.....	18

3.5.1 Busca nos bancos de dados do SUCEST por ESTs de genes que codificam FtsHs em cana-de-açúcar.....	18
3.5.2 Análise <i>in silico</i> da expressão das FtsHs em cana-de-açúcar.....	19
3.6 Estudo da localização subcelular das FtsHs de cana-de-açúcar.....	19
3.6.1 Construção das fusões gênicas	19
3.6.2 Obtenção das células competentes de <i>Escherichia coli</i>	21
3.6.3 Transformação de <i>Escherichia coli</i>	21
3.6.4- Reação de sequenciamento.....	23
3.7 Expressão transiente das construções gênicas em células da epiderme de cebola.....	24
4 RESULTADOS E DISCUSSÃO.....	25
4.1 História evolutiva das FtsHs bacterianas.....	25
4.1.1 Número e distribuição de Isoformas de FtsHs em genomas bacterianos.....	25
4.1.2 Relações de ancestralidade entre as isoformas de FtsHs bacterianas.....	29
4.1.3 Análise estrutural das FtsHs bacterianas.....	40
4.2 História evolutiva das FtsHs eucarióticas.....	45
4.2.1 Filogenia das FtsHs eucarióticas.....	45
4.2.2 FtsHs em <i>Arabidopsis thaliana</i> e Cana-de-açúcar.....	49
4.2.3 Análise da expressão <i>in silico</i> dos grupos filogenéticos de FtsHs nas bibliotecas de cDNA do SUCEST.....	54
4.2.4 Relações filogenéticas entre FtsHs procarióticas e eucarióticas.....	55
4.3 Localização subcelular das FtsHs p1, p2 e m1 de cana-de-açúcar.....	59
5 CONCLUSÕES.....	65
REFERÊNCIAS BIBLIOGRÁFICAS.....	66
APÊNDICE.....	82

HISTÓRIA EVOLUTIVA DAS METALOPROTEASES AAA

Autor: PHELLIPPE ARTHUR SANTOS MARBACH

Orientador: Prof. Dr. MARCIO DE CASTRO SILVA FILHO

RESUMO

As metaloproteases AAA são ubíquas nos Domínios Bacteria e Eukarya, contudo não são encontradas no domínio Archaea. Estas enzimas são proteínas integrais de membrana e estão envolvidas em processos celulares centrais como respiração celular, fotossíntese e biogênese de organelas. A metaloprotease AAA de *Escherichia coli* denominada FtsH é o membro mais estudado deste grupo de proteínas. A análise da distribuição dos genes que codificam esta enzima em genomas procarióticos revelou que a grande maioria das espécies bacterianas possuem apenas uma isoforma deste gene no genoma. A existência de mais de uma isoforma de FtsH por genoma parece estar limitada a alguns grupos bacterianos. A análise filogenética das metaloproteases AAA bacterianas mostrou a existência de três isoformas distintas resultantes de antigos eventos de duplicação. Estas isoformas possuem grande similaridade em nível de estrutura primária e secundária. A análise filogenética das metaloproteases eucarióticas mostrou que elas podem ser agrupadas em cinco clados. Membros de um mesmo clado foram preditos terem a mesma localização subcellular, plastídeo or mitocôndria. Os clados plastidiais foram denominados FtsHp1, FtsHp2 e FtsHp3, e os

mitocondriais, FtsHm1 e FtsHm2. A distribuição das metaloproteases AAA nos domínios da vida, bem como sua localização subcelular, sugerem uma origem endossimbiótica para os representantes eucarióticos. De fato, a análise filogenética das metaloproteases AAA eucarióticas e bacterianas mostrou uma relação de ancestralidade entre metaloproteases AAA de cianobactérias e plastidiais. O mesmo foi observado entre as metaloproteases eucarióticas do grupo FtsHm1 e o ancestral das atuais proteobactérias. Surpreendentemente, esta análise filogenética sugere que as FtsHm2s originaram de um evento de transferência horizontal entre uma bactéria da divisão CFB e a célula eucariota primitiva. A localização subcelular das FtsHp1, FtsHp2 e FtsHm1 de cana-de-açúcar foi estudada para testar hipótese de que o perfil filogenético das metaloproteases AAA de plantas pode ser usado para determinar a sua localização subcelular. Fusões gênicas envolvendo a seqüência de direcionamento destas proteínas e GFP (*Green Fluorescent Protein*) foram expressas em células do epitélio de cebola e a localização da GFP foi feita em microscópio de fluorescência. A GFP fusionada à seqüência de direcionamento das FtsHp1 e FtsHp2 foi direcionada para os plastídeos enquanto que a GFP fusionada à seqüência de direcionamento da FtsH-m1 foi enviada para as mitocôndrias. Os resultados confirmam a hipótese de que localização subcelular das metaloproteases AAA vegetais pode ser predita com base no seu perfil filogenético.

EVOLUTIVE HISTORY OF METALLOPROTEASES AAA

Author: PHELLIPPE ARTHUR SANTOS MARBACH

Adviser: Prof. Dr. MARCIO DE CASTRO SILVA FILHO

SUMMARY

Metalloproteases AAA are ubiquitous in Bacteria and Eukarya domains, however they are not found in Archaeas. These enzymes are integral proteins involved in essential cellular processes such as oxidative phosphorylation, photosynthesis and organellar biogenesis. The *Escherichia coli* metalloprotease AAA named FtsH, is the most studied member of this protein. Analysis of the distribution of the AAA metalloprotease encoded genes in bacterial genomes revealed that the majority of the bacterial genomes contains only one copy of the metalloprotease gene. The existence of more than one FtsH isoform per genome seems to be limited to some bacterial groups. Phylogenetic analysis of AAA metalloproteases showed the existence of three distinct isoforms, resulting from old duplication events. These isoforms show high similarity at the primary and secondary structural level. Phylogenetic analysis of eukaryotic AAA metalloproteases showed that they can be clustered into five clades. Members of the same clade were predicted to have the same subcellular localization, plastid or mitochondria. The plastidial clades were named FtsHp1, FtsHp2 e FtsHp3, and the mitochondrial clades FtsHm1 e FtsHm2. The distribution of metalloproteases AAA in the domains of life as well

as their predicted subcellular localization, suggested an endosymbiotic origin of the eukaryotic members. In fact, the phylogenetic analysis including both bacterial and eukaryotic AAA metalloproteases, showed an ancient relationship among cyanobacterial e plastidial-derived AAA metalloproteases. The same observation was shown for the FtsHm1 metalloproteases clade and the ancestor of actual proteobacterias. Surprisingly, this phylogenetic analysis suggest that the FtsHm2 clade ancestor was originated from a horizontal gene transfer event between the primitive eukaryotic cell and a species from CFB bacterial division. The subcellular localization of sugarcane FtsHp1, FtsHp2 e FtsHm1 were studied to test the hypothesis that the phylogenetic profile of the plant metalloproteases AAA can be used to predict their subcellular localization. Gene fusion involving the targeting sequence of these proteins and GFP (Green Fluorescent Protein) were expressed in onion epidermal cells and the subcellular localization of GFP was monitored with a fluorescent microscope. The GFP fused to the FtsHp1 e FtsHp2 targeting sequences was targeted to plastids while the GFP fused to the FtsH-m1 targeting sequence of was exclusively targeted to mitochondria. These results confirm the hypothesis that the subcellular localization of the plant AAA metalloproteases can be predicted based on their phylogenetic profile.

1 INTRODUÇÃO

A conversão da energia química em atividade biológica certamente foi um processo central na origem da vida, independente de onde ela tenha se originado. Atualmente este processo é feito por proteínas e ocorre, geralmente, de uma maneira indireta, sendo o GTP e o ATP preferencialmente utilizados pelas células, procariontas e eucariontas como fonte de energia, em vários processos celulares tais como divisão celular, síntese proteica e sinalização.

A genômica tem produzido uma grande quantidade de informação provocando mudanças conceituais radicais nos estudos evolutivos, permitindo a caracterização do padrão de evolução de famílias de proteínas inteiras de uma maneira sem precedentes na história da ciência (Jia et al., 2004). O estudo evolutivo das famílias de proteínas permite elucidar os mecanismos moleculares envolvidos na sua origem e divergência funcional, fornecendo bases mais sólidas para o desenvolvimento de estudos funcionais de seus membros (Leipe et al., 2002).

Vários autores têm proposto o uso do perfil filogenético como uma maneira rápida de fazer a predição da localização subcelular de proteínas, uma informação fundamental na caracterização funcional de uma proteína (Marcotte et al., 2000; Hannenhalli & Russell, 2000; Marbach et al., 2001).

As proteínas AAA (**A**TPases **A**ssociadas a diferentes **A**tividades celulares) constituem uma família particular de proteínas capazes de converter energia em atividade biológica. Seus membros estão envolvidos em diversos processos celulares tais como o controle do ciclo celular, a regulação da expressão gênica e a biogênese de organelas.

As AAA são divididas em várias subfamílias de acordo com sua similaridade e função (Beyer, 1997), dentre elas destacam-se as metaloproteases AAA. Estas proteínas possuem características evolutivas e funcionais que as distinguem das demais proteínas AAA. São as únicas proteínas AAA cuja dispersão está limitada aos domínios Bacteria e Eukarya (Swaffield & Purugganan, 1997). Além disso, possuem acopladas no mesmo peptídeo as funções ATPásica e proteolítica.

Contudo, a característica mais marcante das metaloproteases AAA é sua multifuncionalidade em virtude de sua ação proteolítica sobre uma ampla gama de substratos distintos. Estas proteases estão envolvidas em processos celulares variados que vão desde a montagem de complexos protéicos de membrana até a degradação de proteínas citoplasmáticas e/ou de membranas (Begg et al., 1992; Herman et al., 1993; Nilsson et al., 1994; Tomoyasu *et al.*, 1995; Akiyama *et al.*, 1996).

As metaloproteases AAA eucarióticas conhecidas possuem altos níveis de similaridade com alguns ortólogos procarióticos. As proteínas estão localizadas nas mitocôndrias ou plastídeos onde participam da fosforilação oxidativa e/ou fotossíntese. Estas proteínas são codificadas por genes nucleares e pós-traducionalmente importadas para suas respectivas organelas por meio de seqüências de direcionamento específicas (Hugueney *et al.*, 1995; Leonhard *et al.*, 1996; Chen *et al.*, 2000).

A dispersão das metaloproteases AAA nos domínios da vida e sua localização subcelular em eucariotos sugerem que os ortólogos eucarióticos destas proteínas originaram-se a partir de eventos de migração dos genes dos endossimbiontes ancestrais das mitocôndrias e plastídeos para o núcleo eucariótico primitivo (Swaffield & Purugganan, 1997; Margulis, 1970).

Neste contexto, este trabalho teve como principal objetivo estudar a filogenia das metaloproteases AAA com o propósito de estabelecer as relações de ancestralidade entre os membros procarióticos e eucarióticos.

2 REVISÃO DE LITERATURA

2.1 Proteínas AAA

A atividade catalítica das NTPases, enzimas que hidrolisam nucleosídeos trifosfatados, é utilizada pela célula virtualmente em todos os processos celulares. Estas enzimas são agrupadas, de acordo com sua estrutura terciária, em várias classes. Dentre estas enzimas, as que possuem o motivo *P-loop*, *P-loop* NTPases, representam de 10 a 18% de todos os produtos gênicos conhecidos (Leipe et al., 2002).

As *P-loop* NTPases caracterizam-se por possuírem o motivo Walker A N-terminal posicionado entre uma fita β pregueada e uma α -hélice, o motivo *P-loop*, seguido do motivo Walker B. A função do motivo Walker A é posicionar corretamente a porção trifosfato do dNTP enquanto o motivo Walker B possui um resíduo de ácido aspártico conservado situado no final de uma fita β pregueada que se liga a um íon Mg^+ por intermédio de uma molécula de água (Leipe et al., 2002).

As *P-loop* NTPases são divididas em dois grupos estruturalmente distintos. Nos membros do primeiro grupo, que inclui as nucleotídeo kinases e GTPases, as fitas β pregueadas do *P-loop* e do Motivo Walker B são vizinhas (Figura 1A). Já nos membros do segundo grupo existe uma fita β adicional entre as fitas β do *P-loop* e do Motivo Walker B (Figura 1B), a família das proteínas AAA faz parte do segundo grupo (Lupas & Martin, 2002; Leipe et al., 2002).

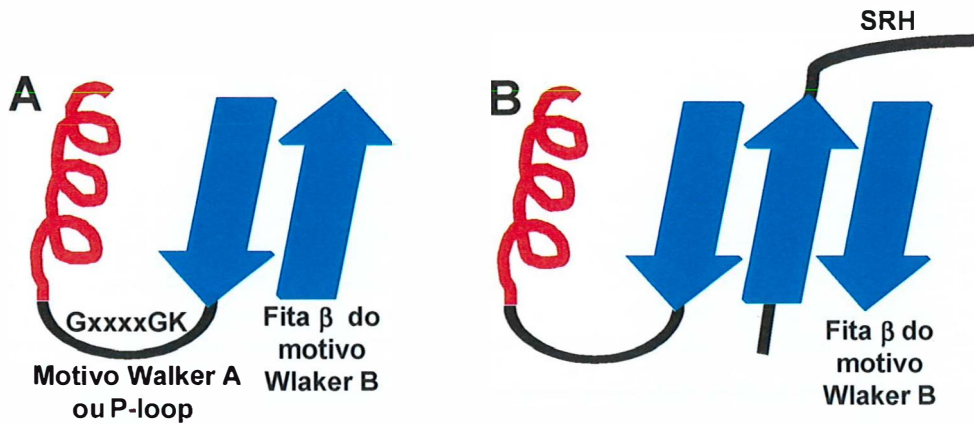


Figura 1 – Arranjos estruturais encontrados nas P-loop NTPases

A família das proteínas AAA é um grupo distinto dentro da superfamília das proteínas AAA⁺ (Neuwald et al., 1999). Seus membros são caracterizados pela presença de uma ou duas cópias do módulo AAA. Este módulo é um domínio ATPásico altamente conservado que é definido por uma seqüência conservada de 220-250 aminoácidos (Figura 2). O módulo AAA possui os motivos Walker A e B e o motivo SRH (*Second Region of Homology*) envolvido na hidrólise do ATP. Este último distingue as proteínas AAA das demais AAA⁺ (Karata et al., 1999; Neuwald et al., 1999; Lupas & Martin, 2002).

Proteínas AAA já foram descritas em espécies dos três domínios da vida: Archaea, Bacteria e Eukarya (Swaffield & Purugganan, 1997). Isto sugere que a proteína AAA ancestral já estava presente no cenancestor, o mais recente ancestral comum dos organismos. Além disso, indica que a divergência funcional entre as AAAs foi baseada em uma das mais antigas invenções da natureza, a conversão da energia química armazenada no ATP em atividade biológica.

Estas proteínas estão envolvidas em diversos processos celulares como controle do ciclo celular, regulação da expressão gênica, biogênese de organelas, transporte, associação e degradação de proteínas (Patel & Latterich, 1998).

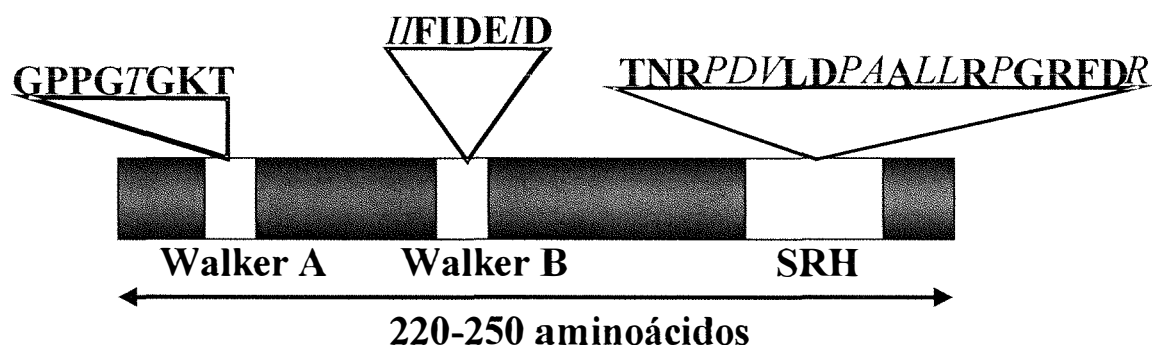


Figura 2 – Diagrama do módulo AAA mostrando o arranjo dos motivos Walker A, Walker B e SRH. As seqüências dos motivos é o consenso das metaloproteases AAA

A dispersão das proteínas AAA nos três domínios da vida sugere que formação das ATPases proteassomais, cdc48 e metaloproteases foram os primeiros eventos de diversificação funcional entre as proteínas AAA. Esta divergência funcional inicial foi subsequente à separação das linhagens Bacteria e Archaea/Eukarya. As demais divergências funcionais entre as proteínas AAA ocorreram principalmente na linhagem Eukarya (Swaffield & Purugganan 1997; Patel & Latterich, 1998; Lupas & Martin, 2002).

2.2 Metaloproteases AAA em procariotos

As metaloproteases AAA são ubíquas nos domínios Bacteria e Eukarya, mas ainda não foram descritas em Archaea (Langer, 2000). Um membro típico desta subfamília possui uma cópia do módulo AAA e se distingue das demais proteínas AAA pela presença na região C-terminal do motivo conservado HEXGH para ligação a zinco envolvido na sua atividade proteolítica (Figura 3A) e uma provável estrutura *coiled coil*. Esta protease está fortemente associada à membrana por um ou dois segmentos transmembranas situados na região N-

terminal (Figura 3A-B), antes do módulo AAA (Tomoyasu et al., 1995; Shotland et al., 2000).

O membro desta subfamília mais estudado é a metaloprotease AAA de *E. coli* denominada FtsH. Esta protease foi isolada a partir da caracterização do mutante *ftsH* Y16 de *E. coli* (Santos & Almeida, 1975; Ogura et al., 1991), que pertence a um grupo de mutantes denominados *fts*, *filamentation temperature sensitive*. Estes mutantes quando cultivados em altas temperaturas (42 °C), não são capazes de formar o septo que separa as células durante a divisão celular.

A filamentação termosensível do mutante *ftsH* Y16 é causada pelo decréscimo da proteína PBP3 codificada pelo gene *ftsI* (Ferreira et al., 1987). A transformação deste mutante com um plasmídeo carregando o gene *ftsI* reverte a filamentação termosensível, contudo, o transformante mantém a termosensibilidade para o crescimento (Ferreira et al., 1987; Ogura et al., 1991).

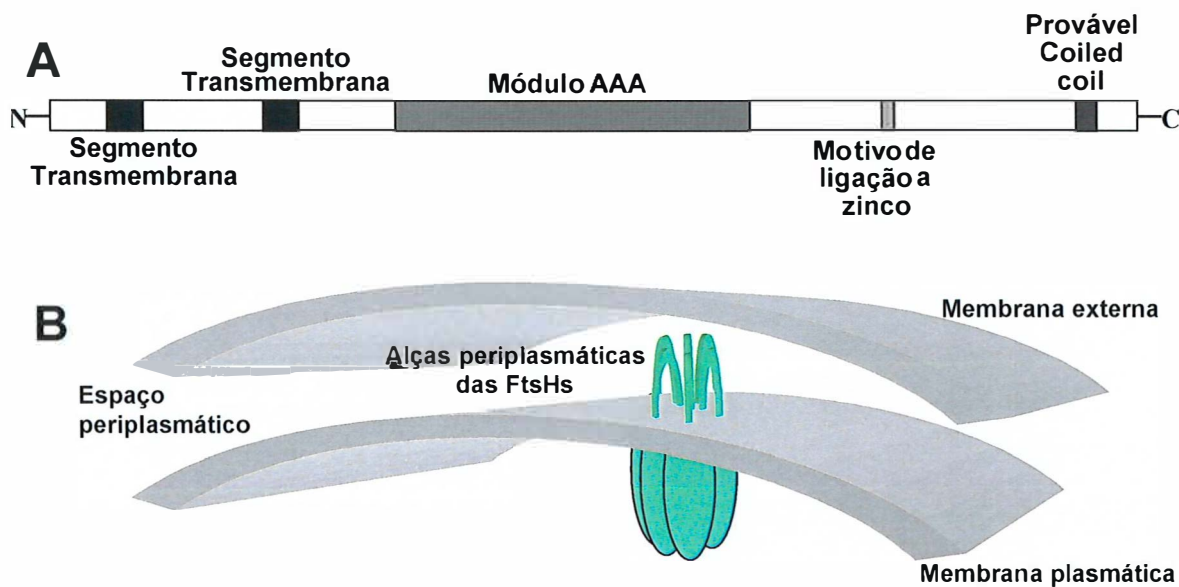


Figura 3 – Diagramas de uma típica metaloprotease AAA procariótica mostrando seus domínios e motivos característicos (A) e seu ancoramento na membrana plasmática (B). O módulo AAA contém os motivos Walker A, Walker B e SRH

Posteriormente, Begg et al. (1992) mostraram que o mutante *ftsH* Y16 possuía na verdade mutações em dois genes, uma no gene *ftsI*, responsável pela filamentação termosensível, e outra no gene *ftsH*, responsável pelo crescimento termosensível, denominada mutação *ftsH1*. Em seguida, verificou-se que o gene *ftsH* codifica uma proteína integral de membrana com atividade ATPásica que foi denominada FtsH (Tomoyasu et al., 1993a; Tomoyasu et al., 1993b)

A FtsH está envolvida em vários processos celulares importantes como exportação de proteínas, associação de proteínas à membrana plasmática e degradação de proteínas de membrana e solúveis, refletindo a característica mais marcante desta proteína, seu amplo espectro de ação (Begg et al., 1992; Herman et al., 1993; Nilsson et al., 1994; Tomoyasu et al., 1995; Akiyama et al., 1996). Esta característica resultou em pelo menos quatro designações diferentes para o gene que codifica esta enzima em *E. coli*: *ftsH*, *hflB*, *tolZ* e *mrsC*. Estes nomes foram dados por diferentes grupos que faziam seleção para diferentes fenótipos mutantes (Schumann, 1999).

O mutante *hflB* (*high frequency of lysogenization*) foi a primeira descrição de um fenótipo associado à FtsH (Gaustsch e Wulff, 1974). Estes autores estavam interessados em genes de *E. coli* que estivessem relacionados com o modo de infecção (lise *versus* lisogênese) após a infecção da *E. coli* com o Bacteriófago λ .

A proteína λ CII é estabilizada no mutante *hflB29*. Esta proteína é um fator transcricional que regula positivamente a expressão dos genes *λ int* e *λ cl*. O primeiro gene codifica uma integrase, enzima necessária para a integração do DNA do fago no cromossomo bacteriano, enquanto que o segundo gene codifica a proteína repressora do fago, que o mantém no estado lisogênico uma vez estabelecido.

Herman et al. (1993) verificaram que a superexpressão da FtsH no mutante *hflB29* suprimia a alta frequência de lisogênese, estimulando a degradação da proteína λ CII. Mais tarde, os autores demonstraram que a

proteína FtsH era na verdade uma protease responsável pela degradação da proteína λ CII e σ^{32} , principal determinante da resposta a choque térmico em *E. coli* (Herman et al., 1995). A ação proteolítica da FtsH sobre σ^{32} foi também demonstrada por Tomoyasu et al. (1995). Estas foram as primeiras evidências experimentais da atividade proteolítica da protease FtsH.

Por outro lado, o mutante *tolZ21* foi isolado em uma seleção para mutantes com tolerância a colicinas, fenótipo Tol⁻. Outros fenótipos mutantes observados em *tolZ* foram: crescimento termossensível, fenótipo Ts; incapacidade de crescer em meio contendo apenas fonte de carbono não fermentável, fenótipo Nfc⁻; alta frequência de lisogenização, fenótipo Hfl⁻ e , fenótipo Std⁻. Curiosamente, os mutantes *ftsH1*, *hflB29* e *ftsH101* não possuem os fenótipos mutantes Tol⁻ e Nfc⁻ (Qu et al., 1996).

Granger et al. (1998) identificaram um gene envolvido na degradação de mRNA que foi denominado *mrsC* (**mRNA stability**). Estudos posteriores revelaram que o gene em questão é o gene *ftsH* (Wang et al., 1998).

O mutante ER437 foi isolado originalmente como um mutante Ts tolerante a colicina, Tol⁻ (Nomura e Witten, 1967). Mais tarde, verificou-se que ele possui o fenótipo Hfl⁻ e três mutações na região 3' do gene *ftsH* (Oppenheim et al., 1974; Teff et al., 2000).

A FtsH forma uma estrutura homoligomérica semelhante às estruturas em forma de anel também observadas em outras proteases AAA⁺ ATP-dependentes, como as Clp/ClpP e HslU/HslV (Figura 3B). A região N-terminal inclui os dois segmentos transmembrana e um pequeno domínio periplasmático possuindo importante função na homoligomerização e interação com o complexo periplasmático HflKC formado pela associação das proteínas integrais HflK e HflC (Kihara et al., 1996; Akiyama & Ito, 2000).

Este complexo protéico interage com a alça periplasmática da FtsH modulando negativamente sua atividade proteolítica sobre as proteínas de membrana SeY, YccA e subunidade *a* da ATPase F₀ (Kihara et al., 1996). Entretanto, mutações sem sentido nos genes integrais *hflK* e *hflC* não

estabilizam as proteínas citosólicas λ CII e σ^{32} , sugerindo que a modulação da atividade proteolítica da FtsH sobre proteínas integrais e citosólicas é distinta (Kihara et al., 1998). De fato, a degradação de λ CII e σ^{32} requer a participação de fatores adicionais, HflD e DnaK, respectivamente, para uma eficiente ação proteolítica (Tomoyasu, et al., 1995; Kihara et al., 2001).

Homólogos da protease FtsH já foram encontrados em outras bactérias gram negativas. Na α -proteobactéria *Caulobacter crescentus*, a FtsH não é produto de um gene essencial. Os mutantes deste gene possuem uma elevada concentração de σ^{32} indicando que esta FtsH está envolvida na resposta a choque térmico. Estas células são viáveis em condições normais de crescimento, mas são sensíveis a temperaturas elevadas e estresse salino. Além disso, a FtsH está envolvida no controle do ciclo celular (Fischer et al., 2002).

Em *Bradyrhizobium japonicum*, uma outra α -proteobactéria, a FtsH tem um comportamento diferente do observado em *E. coli* e *Caulobacter crescentus* (Narberhaus et al., 1999). Trata-se de uma proteína essencial, como em *E. coli*, contudo é expressa constitutivamente não sendo induzida por choque térmico. Curiosamente, esta protease degrada, *in vivo*, o fator σ^{32} e seus homólogos em *Bradyrhizobium*, ou seja, RpoH₁, RpoH₂ (Urech et al., 2000). Além disso, é capaz de restaurar a viabilidade da cepa de *E. coli* AR3291 cujo gene *ftsH* foi nocauteado, indicando que pode compensar a perda de algumas funções importantes da FtsH de *E. coli* (Narberhaus et al., 1999).

A condição de gene essencial também foi descrita para a FtsH da ϵ -proteobactéria *Helicobacter pylori*. Este gene faz parte do operon *sro* que é induzido por choque térmico e cobre (Beier et al., 1997; Ge & Taylor, 1997). Melchers et al. (1998) mostraram em um ensaio de complementação que a FtsH de *Helicobacter felis* foi capaz de restaurar a viabilidade em *E. coli* com o gene *ftsH* inativado.

Em bactérias gram positivas o cenário não é diferente. A FtsH de *Bacillus subtilis* não é produto de um gene essencial, mas seu mutante possui fenótipo pleiotrópico. Este apresenta um crescimento anormal, não esporula e é sensível a estresse salino e a choque térmico. Contudo, apesar da FtsH ser necessária para a adaptação a temperaturas elevadas, ela não é essencial para a regulação da resposta a choque térmico (Deuerling et al., 1997).

Assim como em *Bacillus subtilis*, a FtsH de *Lactococcus lactis* não é produto de um gene essencial, entretanto, o mutante com o gene rompido é sensível a sal, calor e frio. Além disso, a FtsH de *Lactococcus lactis* foi capaz de complementar *E. coli* com o gene *ftsH* mutado (Nilsson et al., 1994).

Synechocystis é a única cianobactéria cuja FtsH foi estudada. Na verdade, *Synechocystis* possui quatro genes que codificam proteases homólogas a FtsH a saber: *slr0228*, *sll1463*, *slr1390* e *slr1604*. Mann et al. (2000) conseguiram obter mutantes apenas para *slr0228* e *sll1463*. Isto sugere que *slr1390* e *slr1604* possam ser genes essenciais. O rompimento de *slr0228* causou uma redução de 60% na abundância do fotossistema I funcional indicando seu envolvimento na fotossíntese. Posteriormente, Silva et al. (2002) mostraram que a FtsH *slr0228* estava envolvida na degradação da proteína D1.

2.3 Metaloproteases AAA eucarióticas

Vários ortólogos eucarióticos da FtsH já foram identificados e aparentemente estão exclusivamente localizados ou em mitocôndrias ou em plastídeos. Geralmente, são codificados por genes nucleares e pós-traducionalmente importadas para suas respectivas organelas por meio de seqüência de direcionamento específica (Hugueney et al., 1995; Leonhard et al., 1996; Chen et al., 2000).

Em levedura, foram identificadas três proteínas homólogas a FtsH, todas localizadas na mitocôndria: YME1 (YTA11p), YTA10 (AFG3), e YTA12 (Rca1) (Thorsness et al., 1993; Schnall et al., 1994). YME1 (Yeast mitochondrial

escape) forma um complexo homoligomérico que expõe seus sítios catalíticos no espaço intermembrana e por isso é denominado protease *i*-AAA.

O módulo AAA da Yme1 possui atividade chaperona, ou seja, interage com substratos desnaturados, o que garante a especificidade da sua atividade proteolítica. A função chaperona também foi descrita em outras proteínas AAA (Leonhard *et al.*, 1999).

A inativação da YME1 causa pelo menos quatro fenótipos recessivos em células de levedura: elevada taxa de escape do mtDNA, atividade respiratória deficiente em temperaturas elevadas, crescimento sensível ao frio (14 °C) em meio rico em glicose, e letalidade em células *rho*⁻ (Thorsness *et al.*, 1993). Além disso, a morfologia das mitocôndrias deste mutante é severamente alterada, sugerindo o envolvimento da YME1 na manutenção da morfologia das mitocôndrias. (Campbell *et al.*, 1994).

As bases moleculares dos fenótipos apresentados pelo mutante *yme1* ainda não foram elucidadas. A subunidade Cox2 da citocromo oxidase *c* é o único substrato conhecido da protease *i*-AAA (Nakai *et al.*, 1995). Contudo, em função dos múltiplos fenótipos apresentados pelo mutante *yme1*, acredita-se que esta protease esteja envolvida em outros processos celulares além da degradação de Cox2.

A busca por supressores dos fenótipos do mutante *yme1* levou à identificação de uma proteína cuja atividade está relacionada com a proteína YME1 (Hanekamp & Thorsness, 1999; Kominsky *et al.*, 2002). A inativação do gene *YNT20* causa a redução na taxa de escape de DNA mitocondrial no mutante *yme1*, sugerindo que *YNT20* está envolvido na via de escape de DNA mitocondrial mediada por *yme1* (Hanekamp & Thorsness, 1999).

A complementação do mutante *yme1* com ortólogos de *Neurospora crassa* e *Homo sapiens* indica que estas proteínas ainda desempenham funções filogeneticamente relacionadas nestas espécies (Shah *et al.*, 2000, Klanner *et al.*, 2001). Entretanto, o fato do ortólogo de *Neurospora crassa* ter

complementado parcialmente o mutante *yme1* indica a existência de especificidade funcional, mesmo entre linhagens próximas.

YTA10 e YTA12 (*Yta*: Yeast tat-binding analogs), ao contrário de YME1, têm seus sítios catalíticos expostos na matriz mitocondrial além de formarem um complexo heteroligomérico denominado protease *m*-AAA (Leonhard *et al.*, 1996). A *m*-protease usa sua atividade chaperona para montar complexos protéicos da cadeia respiratória. O controle de qualidade da membrana interna, degradando cadeias polipeptídicas não incorporadas nos complexos protéicos, também é feito por esta protease (Paul & Tzagoloff, 1995; Arlt *et al.*, 1996; Arlt *et al.*, 1998). Além disso, a *m*-protease controla a expressão de Cox1 e Cob, que são codificadas por genes mitocondriais que possuem introns. Este controle ocorre durante o processamento do transcrito primário, uma vez que os níveis dos mRNAs de Cox1 e Cob são drasticamente reduzidos em mutantes *Δyta10*, e *Δyta12*, enquanto a transcrição e tradução ocorrem normalmente (Arlt *et al.*, 1998).

Ortólogos de Yme1p, Yta10p, e Yta12p foram encontrados em outros eucariotos. Em humanos, o ortólogo de Yta10p e Yta12p está relacionado com uma forma de paraplegia hereditária, uma doença neurodegenerativa (Itoh *et al.*, 1999; Casari *et al.*, 1998). Por sua vez, a FtsH mitocondrial identificada em *Pisum sativum* é capaz de complementar seus ortólogos de levedura YTA10, e YTA12. O fato dela ter complementado o duplo mutante *Δyta10/Δyta12* indica que ela forma um complexo homoligomérico em *Pisum sativum*, o contrário do observado para seus ortólogos YTA10 e YTA12 em leveduras (Kolodziejczak *et al.*, 2002).

Pro-hibitinas são proteínas altamente conservadas que estão localizadas em mitocôndrias em eucariotos (Takahashi *et al.*, 2003; Steglich *et al.*, 1999). Esta proteína foi originalmente descrita como um fator supressor de tumor devido a sua atividade celular antiproliferativa (McClung *et al.*, 1989). Células de levedura deficientes em pro-hibitinas possuem um ciclo celular mais curto. Além disso elas parecem desempenhar alguma função relacionada com a herança

mitocondrial (Berger et al., 1998). Em *Caenorhabditis elegans* estas proteínas são essenciais durante o desenvolvimento embrionário (Sanz et al., 2003). Por outro lado, estas proteínas parecem estar envolvidas na reação de hipersensibilidade em resposta a ataques de patógenos em arroz (Takahashi et al., 2003).

Curiosamente as pro-hibitinas possuem similaridade significativa com as proteínas moduladoras da FtsH, notadamente as HflK e HflC, sugerindo que estas possam desempenhar funções semelhantes em eucariotos. De fato, Steglich et al., (1999) demonstraram que as pro-hibitinas formam um complexo com a protease *m*-AAA modulando negativamente sua atividade. Estes resultados mostram uma conservação no modo como as FtsHs são reguladas em procariotos e eucariotos.

Pftf (Plastid fusion and/or translocation factor) de *Capsicum annuum* foi a primeira FtsH identificada em plantas (Huguency et al., 1995). Em *Arabidopsis thaliana* dois homólogos da FtsH, VAR1 e VAR2 foram identificados nos mutantes para variegação *var1* e *var2*. Estas proteases estão localizadas nos tilacóides e participam da biogênese dos plastídeos (Chen et al., 2000; Sakamoto et al., 2002).

Uma terceira FtsH tilacoidal encontrada em *Arabidopsis thaliana* participa da degradação da proteína D1 do fotossistema II (Lindahl et al., 2000). A proteína D1 é uma das subunidades do fotossistema II, e está envolvida no transporte de elétrons, sofrendo constantes danos fotooxidativos. Com a reposição da proteína D1 danificada, a planta consegue sobreviver ao estresse oxidativo. (Spetea et al., 1999). A FtsH identificada em *Nicotiana tabacum* também está localizada nos tilacóides com os sítios catalíticos expostos no estroma. A superexpressão desta proteína retarda a resposta de hipersensibilidade enquanto que a redução nos nível de expressão acelera a resposta de hipersensibilidade (Seo et al., 2000).

3 MATERIAL E MÉTODOS

3.1 Busca por seqüências de aminoácidos de metaloproteases AAA nos bancos de dados de seqüências

3.1.1 Estratégias empregadas

Os programas BLASTP e TBLASTN (Altschul et al., 1997) foram utilizados para fazer as buscas por seqüências de aminoácidos similares à FtsH de *Escherichia coli* nos bancos de dados do NCBI - National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>), do Tigr - The institute for genomic research (<http://tigrblast.tigr.org/>) e do Instituto Sanger (<http://www.sanger.ac.uk/>). O filtro para as regiões das seqüências com baixa complexidade foi desabilitado em todas as buscas. Os critérios utilizados para estabelecer a relação de homologia das seqüências recuperadas com a seqüência de aminoácido da FtsH de *Escherichia coli* foram a presença e arranjo dos principais motivos que caracterizam as FtsHs (Figura 3A).

Inicialmente, foi feita uma busca usando o programa BLASTP nos bancos de seqüências de proteínas conectados ao NCBI. Estas buscas foram feitas separadamente nos bancos de dados de bactérias e de eucariotos. Após o alinhamento das seqüências obtidas nesta primeira busca verificou-se que as regiões N- e C-terminal de algumas seqüências não foram corretamente identificadas ou tinham sido parcialmente anotadas.

Assim, foi feita uma busca adicional usando o programa TBLASTN em todos os genomas com seqüenciamento completo ou em andamento,

disponíveis no NCBI, individualmente. Neste caso, a seqüência de uma espécie filogeneticamente próxima foi usada como “*query*” (seqüência referência). A seqüência da FtsH de *Bradyrhizobium japonicum*, uma α -proteobactéria foi usada como seqüência *query* na busca feita no genoma de *Caulobacter crescentus*, também uma α -proteobactéria, e assim por diante. Esta estratégia de busca permitiu corrigir a anotação de várias FtsHs. Ao final, foram recuperadas 161 seqüências de FtsHs procarióticas e 35 seqüências eucarióticas.

3.2 Análises filogenéticas das FtsHs procarióticas

3.2.1 Alinhamento múltiplo das seqüências de aminoácidos das FtsHs procarióticas

Os programas ClustalX (Thompson et al., 1997) e o SAM-T99 Alignment Tuneup (Karplus et al., 1998) foram utilizados para fazer o alinhamento múltiplo das seqüências de aminoácidos das FtsHs. O ClustalX e o SAM-T99 Alignment Tuneup estão disponíveis nos sítios <ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/> e <http://www.cse.ucsc.edu/research/compbio/HMM-apps/HMM-applications.html>, respectivamente. Os valores de abertura e extensão de *gap* utilizados no ClustalX foram respectivamente, 35 e 0,75 para o alinhamento par a par, e 15 e 0,3 para o alinhamento múltiplo.

Inicialmente, foi feito um alinhamento com todas as seqüências, sendo as as regiões N e C-terminais excluídas devido ao fato de serem altamente variáveis. Este alinhamento foi utilizado para construir uma árvore filogenética pelo método do Agrupamento de vizinhos, *Neighbor-joining*. Apesar dos valores de *bootstrap*, suporte estatístico, não serem significativos todos os grupos bacterianos reconhecidamente monofiléticos, foram recuperados nesta árvore: α , β , γ , δ e ϵ -proteobactérias, Clamídias, Espiroquetas, Cianobactérias, CFB, *Deinococcus/thermus*, planctomicetales e gram positivas.

Assim, com o objetivo de refinar o alinhamento final e eliminar ao máximo possíveis homoplasias, optou-se por uma estratégia de alinhamento múltiplo mais refinada. Nesta etapa foram utilizados os programas ClustalX e o SAM-T99 Alignment Tuneup. Estes programas usam métodos diferentes para construir os alinhamentos múltiplos. O primeiro método utiliza o alinhamento progressivo par a par enquanto que a segunda metodologia usa uma abordagem probabilística.

Primeiramente, as seqüências provenientes de regiões conservadas de cada grupo monofilético foram alinhadas com ambos os programas. Esta região estende-se do início do módulo AAA até o provável motivo *coiled coil* localizado na região C-terminal (Figura 3A). As regiões N- e C-terminais altamente variáveis, que flanqueiam a região conservada, foram excluídas a priori. Posteriormente, os alinhamentos resultantes foram comparados e analisados visualmente. As melhores hipóteses de homologia de cada alinhamento bem como a posição de abertura de indels (**in**serção ou **de**leção) foram escolhidas e mantidas no alinhamento final.

A estrutura secundária predita das FtsHs foi utilizada para auxiliar na escolha da posição de abertura dos indels, pois, estes eventos mutacionais ocorrem preferencialmente em regiões menos conservadas como em alças que ficam expostas na superfície da proteína. A comparação e edição dos alinhamentos foram realizadas com o auxílio de programa BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) (Hall, 1999).

No passo seguinte, os alinhamentos múltiplos de grupos filogenéticos relacionados foram efetuados usando a opção de alinhamento de perfis do ClustalX. O resultado obtido foi por sua vez comparado com o alinhamento do grupo filogenético mais próximo dos anteriores, e assim por diante, até obter-se um alinhamento múltiplo com todos os grupos filogenéticos. Os indels inseridos durante a construção dos alinhamentos foram preferencialmente mantidos.

3.2.2 Alinhamento múltiplo das seqüências de aminoácidos das FtsHs eucarióticas

A estratégia para construir o alinhamento das FtsHs eucarióticas foi essencialmente a mesma descrita para as FtsHs procarióticas, exceto pela exclusão prévia da seqüência de direcionamento situada na região N-terminal das proteínas. As seqüências de direcionamento de proteínas pertencentes a diferentes grupos filogenéticos eucarióticos não derivaram de uma seqüência ancestral, ou seja, não são homólogas, portanto, não foram utilizadas.

3.2.3 Método de análise filogenética empregado

As reconstruções filogenéticas foram realizadas utilizando o método de Agrupamento de vizinhos, *Neighbor-joining* (Saitou & Nei, 1987).

A análise filogenética pelo método Agrupamento de vizinhos foi realizada usando o programa ClustaX. A matriz de Gonnet foi usada para gerar a matriz de distância a partir dos alinhamentos. Esta matriz de distância foi posteriormente usada para construir a árvore filogenética. O suporte estatístico dos clados foi estimado por *bootstrap*, usando 1000 reamostragens dos dados.

3.4 Análise e caracterização estrutural das FtsHs

3.4.1 Obtenção dos perfis de hidropaticidade das FtsHs

Inicialmente o programa TMpred foi utilizado para gerar os perfis de hidropaticidade das FtsHs com uma janela de 17 aminoácidos. Este programa está disponível no sítio http://www.ch.embnet.org/software/TMPRED_form.html (Hofmann & Stoffel, 1993).

Os índices de hidropaticidade de proteínas que pertenciam a um mesmo grupo filogenético, definido a priori na análise filogenética, foram plotados

manualmente com o programa Excel num mesmo gráfico, de maneira sobreposta, de acordo com o alinhamento das respectivas proteínas no ClustalX.

3.4.2 Predição das estruturas secundárias das FtsHs procarióticas

O programa Jpred (<http://www.compbio.dundee.ac.uk/~www-jpred/>) foi utilizado para fazer a predição da estrutura secundária de uma seqüência representativa de cada grupo monofilético obtido na primeira reconstrução filogenética pelo método do vizinho mais próximo (Item 3.2 e 3.3). Estas seqüências foram alinhadas e as regiões de cada seqüência onde foram preditas α -hélices ou fitas β -pregueadas foram destacadas no alinhamento com as cores vermelho e azul, respectivamente. A região usada nesta análise estende-se do início do módulo AAA até o provável motivo "*coiled coil*" localizado na região C-terminal, a mesma utilizada no alinhamento múltiplo usado nas análises filogenéticas.

3.5 Identificação e caracterização das FtsHs de Cana-de-açúcar

3.5.1 Busca nos bancos de dados do SUCEST por ESTs de genes que codificam FtsHs em cana-de-açúcar

A base de dados do SUCEST utilizada nas buscas foi a *Cluster consensi*. Esta base de dados é constituída de seqüências consenso de grupos de ESTs (**Expressed Sequence Tags**) que foram agrupadas pelo programa CAP3. O programa utilizado nas buscas foi o TBLASTN. Foram usadas como seqüências *query* seqüências de aminoácidos de FtsH eucarióticas de todos os grupos filogenéticos obtidos na análise filogenética das FtsHs eucarióticas.

Cento e cinquenta e três grupos de ESTs foram identificados usando um valor de corte de $E < 1e-5$. Cada grupo de ESTs identificado foi usado como

seqüência *query* em uma busca no banco de dados de proteínas conectados ao NCBI. Apenas aqueles grupos de ESTs que alinhavam primeiramente com as FtsHs foram considerados grupos de ESTs de FtsHs.

Estes grupos de ESTs foram posteriormente discriminados com base no seu nível de similaridade com as FtsHs dos grupos filogenéticos e, pela presença ou ausência de assinaturas de seqüências típicas de cada grupo filogenético.

3.5.2 Análise *in silico* da expressão das FtsHs em cana-de-açúcar

A análise da expressão *in silico* das FtsHs nas bibliotecas de cDNA de cana-de-açúcar do SUCEST foi feita para cada grupo filogenético das FtsHs. Inicialmente, foi determinada a origem de cada EST integrante dos grupos de ESTs de FtsHs, isto é, a biblioteca de cDNA em que foi clonada. Estes grupos de ESTs foram discriminados a priori por grupo filogenético de FtsHs (item 3.5.1).

Deste modo, foi feita uma razão do número de ESTs de um dado grupo filogenético encontrados em uma determinada biblioteca de cDNA, pelo número total de ESTs desta biblioteca. Esta razão, reflete o nível de expressão do grupo filogenético das FtsHs na biblioteca de cDNA. Esta análise foi feita para cada grupo filogenético nas bibliotecas não normalizadas de cDNA de cana-de-açúcar agrupadas por tecido.

3.6 Estudo da localização subcelular das FtsHs de cana-de-açúcar

3.6.1 Construção das fusões gênicas

A estratégia utilizada em todas as construções foi fusionar a região 5' dos cDNAs das FtsHs p1, p2 e m1 em fase de leitura com a mGFP5* (Green

Fluorescent Protein) nos sítios *Nco*I e *Bgl*II do vetor binário pCAMBIA 1302 (Figura 4).

A região 5' das FtsHs foi amplificada via PCR (Polymerase Chain Reaction) por meio de iniciadores específicos 5' e 3', contendo na extremidade 5' os sítios *Nco*I e *Bgl*II, respectivamente (Figura 5). As extremidades do produto dos PCRs das FtsHs p1 e p2 foram digeridas simultaneamente com *Nco*I e *Bgl*II e ligado ao vetor pCAMBIA 1302 previamente digerido com as mesmas enzimas (Figura 5A). A confirmação da inserção dos fragmentos no sítio de clonagem foi feita via PCR com iniciadores complementares às seqüências do promotor 35S e do gene da GFP (Figura 5B).

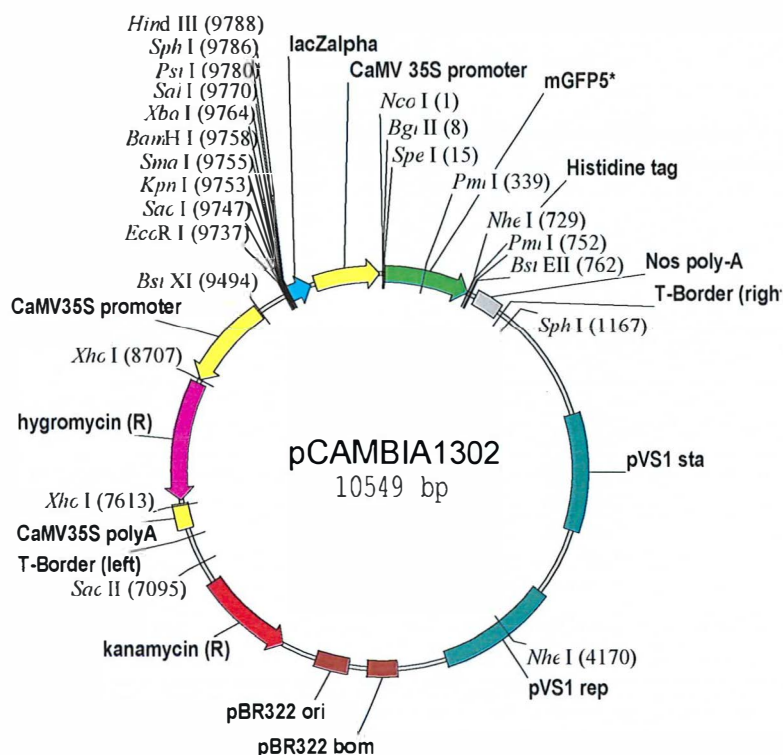


Figura 4 - Plasmídeo pCAMBIA 1302 utilizado nas fusões gênicas da seqüência de direcionamento das FtsHs p1, p2 e m1 com GFP

Como o fragmento amplificado da FtsH m1 possui um sítio interno para a enzima *Bgl*II foi necessário modificar a estratégia para obter o fragmento inteiro

com as extremidades 5' e 3' digeridas com as enzimas *NcoI* e *BglII*, respectivamente. Primeiramente, o fragmento amplificado foi clonado no sítio da enzima *SmaI* do vetor pUC 18, que não possui sítios de restrição *NcoI* e *BglII* (Figura 5C). Em seguida, o vetor foi linearizado com a enzima *NcoI* e submetido à digestão parcial com enzima *BglII* (Figura 5C). O fragmento oriundo da digestão com *BglII* apenas no sítio localizado na extremidade do inserto foi purificado do gel e clonado no pCAMBIA 1302 (Figura 5C e 5A). Todas as construções foram confirmadas por PCR (Figura 3B) e seqüenciamento.

3.6.2 Obtenção das células competentes de *Escherichia coli*

Células de *Escherichia coli*, cepa Jm109, foram crescidas por 16 horas em 3 mL de meio LBA a 37°C. 500 µL da cultura foram inoculados em 50 mL de meio LB em Erlenmeyer de 250 mL estéril e crescidos sob agitação até atingir a fase de crescimento exponencial ($OD \sim 0,5 \rightarrow \lambda = 600\text{nm}$). Neste ponto as células foram transferidas para o gelo (para cessar o crescimento) e depois para tubo de polipropileno estéril. Após centrifugação a 5000 rpm por 1 minuto a 4°C, o sobrenadante foi descartado e o precipitado ressuspendido em solução gelada de cloreto de cálcio 50 mM estéril. Em seguida, as células foram incubadas por 30 minutos no gelo e submetidas a nova centrifugação a 5000 rpm por 1 minuto a 4°C. O sobrenadante foi novamente descartado e o precipitado ressuspendido em 2,5 mL de cloreto de cálcio 50 mM.

3.6.3 Transformação de *Escherichia coli*

O volume total da reação de ligação, 10 µL, foi misturado com 200 µL de células competentes e a mistura incubada por 30 minutos no gelo. Em seguida as células foram submetidas a um choque térmico: 2 minutos a 42 °C seguidos de 2 minutos no gelo. Adicionou-se 1 mL de meio LB e a células foram incubadas por 1 hora a 37°C a 180 rpm. Após centrifugação a 6000 rpm por 6

minutos o sobrenadante foi parcialmente descartado, deixando no tubo aproximadamente 300 μL .

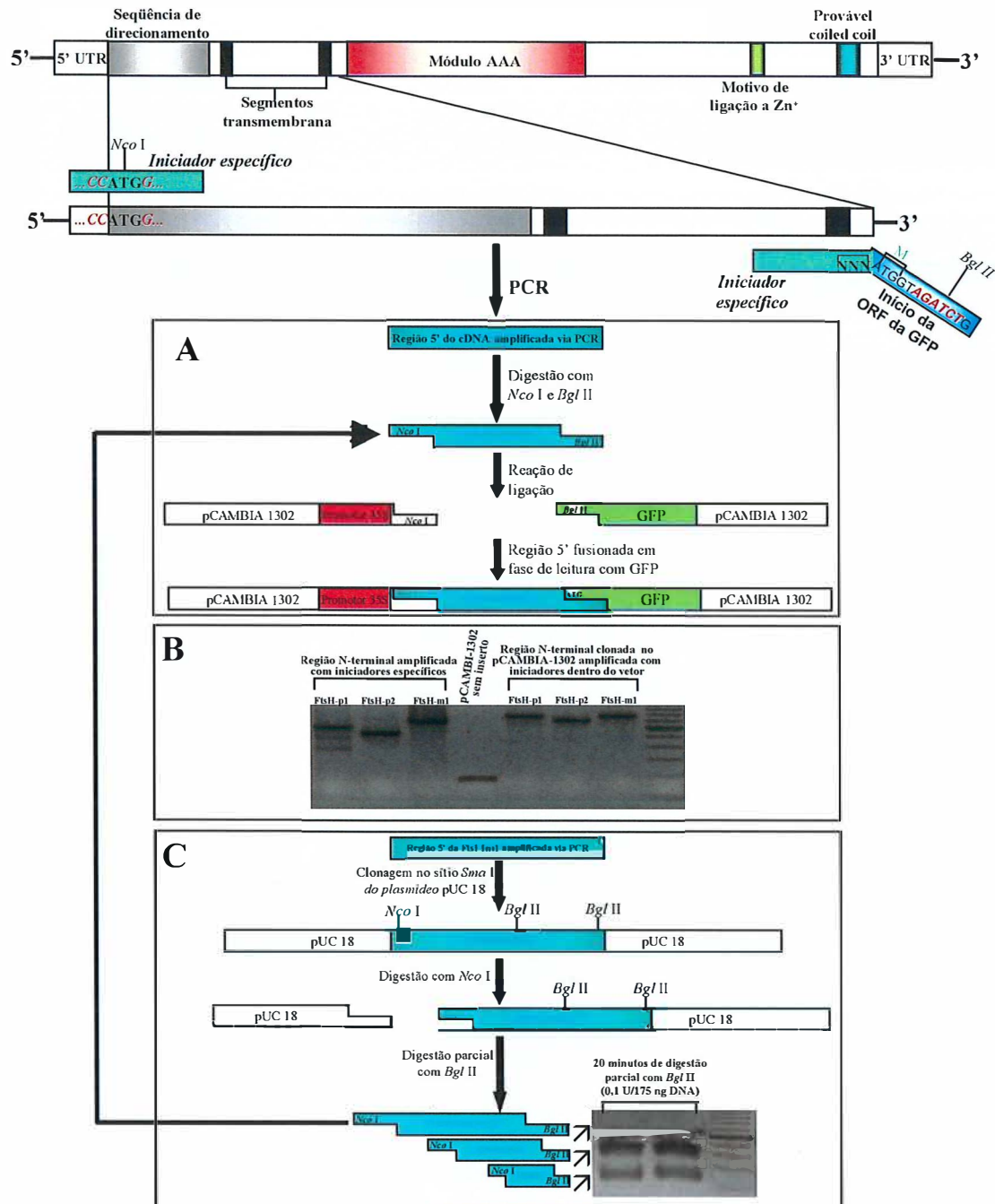


Figura 5 - Estratégias utilizadas na construção das fusões gênicas envolvendo GFP e a região N-terminal das FtsHs p1, p2 e m1

Finalmente este volume foi espalhado em placas de Petri (50 μ L por placa) contendo meio LB suplementado com 100 μ g/mL de ampicilina (LBA100), X-gal e IPTG. As células foram incubadas por 12-16 horas a 37°C. Foram selecionados 15 clones desta minibiblioteca que foram novamente plaqueados em meio LB suplementado com 100 μ g/mL de ampicilina (LBA100), X-gal e IPTG.

3.6.4- Reação de sequenciamento

O DNA plasmidial dos clones selecionados foi extraído, utilizando o método da lise alcalina. A composição da reação de seqüenciamento está descrita abaixo:

- 2 μ L Kit mix
- 1 μ L iniciador Reverso (3pmol/ μ L)
- 6 μ L Água deionizada
- 6 μ L tampão

Após a adição do DNA (200-500ng), o volume da reação foi ajustado para 20 μ L com água deionizada. A reação de seqüenciamento foi realizada nas seguintes condições:

- 96 °C por 2 minutos
 - 96 °C por 20 segundos
 - 50 °C por 10 segundos
 - 60 °C por 4 minutos
 - 72 °C por 7 minutos
- } 25 ciclos

O seqüenciamento foi realizado no seqüenciador ABI PRISM 377.

3.7 Expressão transiente das construções gênicas em células da epiderme de cebola

Células da epiderme adaxial (2cm x 2cm) de cebola foram transformadas via biobalística com as construções gênicas envolvendo a seqüência de direcionamento das FtsHs p1, p2 e m1 e GFP. A preparação das partículas de tungstênio para o bombardeamento foi feita essencialmente conforme Brasileiro e Carneiro (1999).

As partículas cobertas com 8-10 μ g de DNA das construções gênicas foram aceleradas no bombardeador da *Bio-Rad* modelo PDS 1000/He e os parâmetros de bombardeamento utilizados foram: vácuo, 27 inches Hg; distância do alvo, 6 cm; pressão do hélio, 900 psi (Giglione *et al.*, 2000). Após o bombardeamento as epidermes de cebola foram incubadas a 22°C no escuro por 19 horas (Giglione *et al.*, 2000) e examinadas em microscópio de fluorescência Axioplan 2 (Zeiss).

Os vetores pCAMBIA-GFP (35S), RecA-GFP (Akashi *et al.*, 1998) e β -GFP (cedida pelo Prof. Marc Boutry) foram utilizados como controles citoplasmático, plastidial e mitocondrial (35S-GFP), respectivamente.

4 RESULTADOS E DISCUSSÃO

4.1 História evolutiva das FtsHs bacterianas

4.1.1 Número e distribuição de Isoformas de FtsHs em genomas bacterianos

A busca por isoformas de FtsH foi realizada nos genomas de 115 espécies bacterianas distribuídas em 21 grupos monofiléticos (Tabela 1). Esta análise revelou que apenas 29 espécies possuem mais de uma isoforma por genoma, aproximadamente 25% do total (Tabela 1). A existência de mais de uma isoforma de FtsH parece ser uma cocaracterística de alguns grupos bacterianos.

Das seis espécies de cianobactérias analisadas cinco possuem quatro isoformas de FtsH e uma, a *Trichodesmium erythraerum*, possui 3 isoformas (Tabela 1). O sequenciamento do genoma de *Trichodesmium erythraerum* ainda está em progresso, isto significa que é possível que a quarta isoforma ainda não tenha sido anotada. Estas espécies estão distribuídas entre as principais classes deste Filo: Chroococcales Oscillatoriales e Nostocales, sugerindo que a existência de múltiplas formas de FtsH por genoma é uma característica do grupo.

Situação semelhante também foi observada entre as Firmicutes da classe *Clostridia*. As espécies estudadas pertencem às principais ordens desta classe, Clostridiales e Thermoanaerobacteriales. Quatro espécies possuem três isoformas por genoma e uma espécie, *Clostridium perfringens*, possui apenas

Tabela 1 Número e distribuição das isoformas de FtsHs nos genomas bacterianos

Organismos	Sequenciamento e anotação	Nº de isoformas	Nº de acesso das isoformas
Proteobacteria			
γ-proteobactérias (GM-)			
<i>Escherichia coli</i>	Completo	1	1789568
<i>Haemophilus influenzae</i>	Completo	1	16273245
<i>Pasteurella multocida</i>	Completo	1	15602303
<i>Pseudomonas aeruginosa</i>	Completo	1	15599945
<i>Pseudomonas putida</i>	Completo	1	26991401
<i>Buchnera sp.</i>	Completo	1	15616986
<i>Salmonella enterica</i>	Completo	1	16762058
<i>Vibrio cholerae</i>	Completo	1	15640657 – Cr. I
<i>Vibrio vulnificus</i>	Completo	1	27365051 – Cr. I
<i>Wigglesworthia brevipalpis</i>	Completo	1	24324149
<i>Xanthomonas campestris</i>	Completo	1	21231166
<i>Xylella fastidiosa</i>	Completo	1	15836698
<i>Yersinia pestis</i>	Completo	1	16123648
<i>Shewanella oneidensis</i>	Completo	1	24372780
<i>Azotobacter vinelandii</i>	Em progresso	2	23104777, 23104135
β-proteobactérias (BT-)			
<i>Neisseria meningitidis</i>	Completo	1	15676696
<i>Ralstonia solanacearum</i>	Completo	1	17546245
<i>Ralstonia metallidurans</i>	Em progresso	1	22976335
<i>Chromobacterium violaceum</i>	Completo	1	34499252
<i>Nitrosomonas europaea</i>	Em progresso	2	22954515, 22954858
<i>Burkholderia fungorum</i>	Em progresso	3	22984076, 22984177, 22986544
<i>Burkholderia cepacia</i>	Em progresso	2	#
<i>Burkholderia pseudomallei</i>	Em progresso	2	#
<i>Bordetella bronchiseptica</i>	Em progresso	1	#
<i>Bordetella avium</i>	Em progresso	1	#
α-proteobactérias (AF-)			
<i>Rhodospseudomonas palustris</i>	Completo	1	22963104
<i>Rhodospirillum rubrum</i>	Em progresso	1	22966736
<i>Agrobacterium tumefaciens</i>	Completo	1	15891238 – Cr. linear
<i>Bradyrhizobium japonicum</i>	Completo	1	27382257
<i>Brucella suis</i>	Completo	1	23502549 – Cr. I
<i>Mesorhizobium loti</i>	Completo	1	13473326
<i>Caulobacter crescentus</i>	Completo	1	16127456
<i>Rickettsia conorii</i>	Completo	1	15891991
<i>Rickettsia prowazekii</i>	Completo	1	15603922 – Cr. I
<i>Sinorhizobium meliloti</i>	Completo	1	15966488
<i>Magnetospirillum magnetotacticum</i>	Em progresso	2	23015574, 23007846
δ-proteobactérias (DT-)			
<i>Geobacter sulfurreducens</i>	Completo	2	39996282, 39996908
<i>Desulfovibrio desulfuricans</i>	Em progresso	2	23475650, 23473478
<i>Geobacter metallireducens</i>	Em progresso	2	23056302, 23056372
<i>Myxococcus xanthus</i>	Em progresso	2	#
<i>Bacteriovorax marinus</i>	Em progresso	1	#
ε-proteobactérias (EP-)			
<i>Helicobacter hepaticus</i>	Completo	2	32262371, 32266378
<i>Campylobacter jejuni</i>	Completo	2	15792441, 15792332
<i>Helicobacter pylori</i>	Completo	2	15611424, 15611341
<i>Wolinella succinogenes</i>	Em progresso	2	34557037, 34557847
<i>Helicobacter felis</i>	–	1	3122116
Fibrobacter (FBR-)			
<i>Acidobacterium capsulatum</i>	Em progresso	1	#
<i>Fibrobacter succinogenes</i>	Em progresso	1	#
Clamídiæ (CL-)			
<i>Chlamydia muridarum</i>	Completo	1	15834849
<i>Chlamydia trachomatis</i>	Completo	1	15605576
<i>Chlamydomydia pneumoniae</i>	Completo	1	16752029
<i>Chlamydomydia caviae</i>	Completo	1	29840520
<i>Chlamydomydia abortus</i>	Em progresso	1	#
Espiroqueta (SP-)			
<i>Borrelia burgdorferi</i>	Completo	1	15595134
<i>Leptospira interrogans</i>	Completo	1	24216518
<i>Treponema pallidum</i>	Completo	1	15639752
<i>Treponema denticola</i>	Em progresso	1	#
<i>Treponema denticola</i>	Em progresso	1	#
Thermotogae (TH-)			
<i>Thermotoga maritima</i>	Completo	1	15643346

Tabela 1 Número e distribuição das isoformas de FtsHs nos genomas bacterianos

Organismos	Sequenciamento e anotação	Nº de isoformas	Nº de acesso das isoformas
Aquificae (AQ-)			
<i>Aquifex aeolicus</i>	Completo	1	15606261
Deinococcus-Thermus (D/T-)			
<i>Deinococcus radiodurans</i>	Completo	3	Cr. I, 15805610 e 15806043; Cr. II, 15807950
<i>Thermus thermophilus</i>	Em progresso	1	8051696
Divisão Cytophaga, Flexibacter, Bacterioides (CFB-)			
<i>Bacteroides thetaiotaomicron</i>	Completo	1	29349415
<i>Chlorobium tepidum</i>	Completo	2	21673136, 21672968
<i>Porphyromonas gingivalis</i>	Completo	1	34539920
<i>Cytophaga hutchinsonii</i>	Em progresso	1	23137721
<i>Bacteroides forsythus</i>	Em progresso	1	#
<i>Bacteroides fragilis</i>	Em progresso	1	#
<i>Prevotella intermedia</i>	Em progresso	1	#
Cyanobacteria (CB-)			
<i>Thermosynechococcus elongatus</i>	Completo	4	22297675, 22298276, 22298071, 22299375
<i>Nostoc sp.</i>	Completo	4	17228756, 17231134, 17232268, 17232428
<i>Synechocystis sp.</i>	Completo	4	16329602, 16330069, 16329797, 16331432
<i>Prochlorococcus marinus</i>	Completo	4	33862651, 33862949, 33863028, 33864065
<i>Synechococcus sp.</i>	Em progresso	4	33864841, 33865745, 33865834, 33866121
<i>Trichodesmium erythraeum</i>	Em progresso	3	23040279, 23041054, 23043520
Chloroflexi (CF-)			
<i>Chloroflexus aurantiacus</i>	Em progresso	2	22969877, 22974948
<i>Dehalococcoides ethenogenes</i>	Em progresso	1	#
Actinobacteria (AA+)			
<i>Streptomyces coelicolor</i>	Completo	2	21221831, 21223942
<i>Tropheryma whippelii</i>	Completo	1	28493555
<i>Corynebacterium glutamicum</i>	Completo	1	19553890
<i>Mycobacterium leprae</i>	Completo	1	15827019
<i>Mycobacterium tuberculosis</i>	Completo	1	15610746
<i>Bifidobacterium longum</i>	Completo	1	23466230
<i>Mycobacterium smegmatis</i>	Em progresso	1	3426264
<i>Clavibacter michiganensis</i>	Em progresso	1	#
<i>Thermobifida fusca</i>	Em progresso	1	23019087
Firmicutes/Clostridia (FC+)			
<i>Desulfotobacterium hafniense</i>	Em progresso	3	23112207, 23112299, 23114744
<i>Thermoanaerobacter tengcongensis</i>	Completo	3	20808757, 20808395, 20806651
<i>Clostridium tetani</i>	Completo	3	28209972, 28212045, 28210271
<i>Clostridium acetobutylicum</i>	Completo	3	15896449, 15893891, 15894242
<i>Clostridium perfringens</i>	Completo	2	18311452, 18309580
<i>Helibacillus mobilis</i>	-	1	7521916
Firmicutes/Bacilli (FB+)			
<i>Listeria innocua</i>	Completo	1	16799329
<i>Bacillus subtilis</i>	Completo	1	16077137
<i>Bacillus anthracis</i>	Completo	1	21398029
<i>Bacillus halodurans</i>	Completo	1	15612648
<i>Oceanobacillus iheyensis</i>	Completo	1	23097534
<i>Staphylococcus epidermidis</i>	Completo	1	27469190
<i>Staphylococcus aureus</i>	Completo	1	15926188
<i>Bacillus firmus</i>	-	1	3122117
Firmicutes/Lactobacillales (FL+)			
<i>Streptococcus agalactiae</i>	Completo	1	22536201
<i>Streptococcus mutans</i>	Completo	1	24378545
<i>Streptococcus pneumoniae</i>	Completo	1	15902056
<i>Enterococcus faecalis</i>	Completo	1	29374906
<i>Lactobacillus plantarum</i>	Completo	1	28270263
<i>Lactococcus lactis</i>	Completo	1	15672003
<i>Oenococcus oeni</i>	Em progresso	1	23036826
Firmicutes/Mollicutes (FM+)			
<i>Mycoplasma genitalium</i>	Completo	1	12045316
<i>Mycoplasma penetrans</i>	Completo	1	26554472
<i>Mycoplasma pneumoniae</i>	Completo	1	13508410
<i>Mycoplasma pulmonis</i>	Completo	1	15829250
<i>Ureaplasma urealyticum</i>	Completo	1	13357662
<i>Mycoplasma gallisepticum</i>	Completo	1	31544230
Planctomycetes (PLC-)			
<i>Pirellula sp.</i>	Completo	1	32472252
<i>Gemmata obscuriglobus</i>	Em progresso	2	#
Fubobacteria (FU+)			
<i>Fusobacterium nucleatum</i>	Completo	1	19705274

Cr., cromossomo; -, Isoforma não é oriunda de sequenciamento de genoma; #, Isoforma ainda não anotada

duas isoformas. Dentre as quatro espécies que possuem três isoformas duas são do gênero *Clostridium*, sugerindo que *Clostridium perfringens* tenha perdido uma de suas isoformas (Tabela 1).

Todas as espécies da classe ϵ -proteobactéria, cujos genomas foram sequenciados, possuem duas isoformas. A mesma situação foi observada em 4 das cinco espécies da classe δ -proteobactéria cujos genomas foram analisados (Tabela 1). Todavia, o sequenciamento e anotação do genoma da espécie em que foi encontrada apenas uma isoforma, *Bacteriovorax marinus*, ainda está em andamento e a outra isoforma pode não ter sido ainda anotada.

Das dez espécies da classe β -proteobacteira seis possuem apenas uma isoforma por genoma. Dentre as quatro espécies com mais de uma isoforma de FtsH três pertencem ao gênero *Burkholderia*. Isto sugere que a presença de mais de uma isoforma seja uma característica apenas das espécies deste gênero (Tabela 1).

Nas proteobactérias *Escherichia coli* e *Bradyrhizobium japonicum*, a FtsH é um gene essencial. Estas espécies possuem apenas uma isoforma no genoma (Begg et al., 1992; Narberhaus et al., 1999). Contudo, o gene que codifica a FtsH não é essencial em *Caulobacter crescentus*, uma outra proteobactéria com apenas uma isoforma no genoma (Fischer et al., 2002). Mann et al. (2000) conseguiu obter mutantes apenas para duas das quatro isoformas existentes no genoma de *Synechocystis sp*, indicando que as outras isoformas são essenciais.

Estas observações indicam que não existe uma relação entre a condição de gene essencial e número de isoformas no genoma. A condição de ser ou não gene essencial deve estar relacionada com funções cruciais para a viabilidade celular que foram adquiridas ou perdidas pela FtsH após a divergência das espécies. Em *Escherichia coli* a FtsH é uma proteína essencial e uma peça chave na regulação da resposta a choque térmico (Herman et al.,

1995). O contrário do observado em *Bacillus subtilis*, uma espécie onde a FtsH não é uma proteína essencial e não está envolvida na regulação da resposta a choque térmico (Deuerling et al., 1997).

A existência de apenas uma isoforma por genoma parece ser característica da maioria dos grupos monofiléticos amostrados. Entretanto, mesmo em alguns destes grupos foi encontrado pelo menos uma espécie com mais de uma isoforma de FtsH. A relação de ancestralidade entre as diferentes isoformas de FtsH encontrada no genomas bacterianos será discutida a seguir.

4.1.2 Relações de ancestralidade entre as isoformas de FtsHs bacterianas

A existência de mais de uma isoforma de um gene em um genoma pressupõe a ocorrência de transferência lateral de genes (TLG) e/ou duplicação gênica. Para elucidar a origem das isoformas de FtsHs encontradas nos genomas bacterianos foram realizadas análises filogenéticas onde foram usadas seqüências de aminoácidos recuperadas nas buscas feitas nos bancos de dados públicos.

A árvore filogenética das isoformas de FtsH obtida pelo método de Agrupamento de vizinhos possui três grandes clados (Figura 6). No clado 1 estão agrupadas 3 das 4 FtsHs de cianobactérias, a FtsH das espécies do filo Chloroflexi, ambas Gram negativas, e FtsHs de todas as espécies de bactérias gram positivas, exceto as pertencentes à classe Mollicutes (Figura 7).

No caso das espécies de bactérias Gram positivas com mais de uma isoforma por genoma, *Streptomyces coelicolor* e as Firmicutes/Clostridia (Tabela 1), apenas uma delas das isoformas foi agrupada neste clado, indicando que estas isoformas não são oriundas de um evento de duplicação que ocorreu na linhagem da espécie. A mesma inferência pode ser feita para as FtsHs de cianobactérias que não foram agrupadas com as outras 3 isoformas do clado 1.

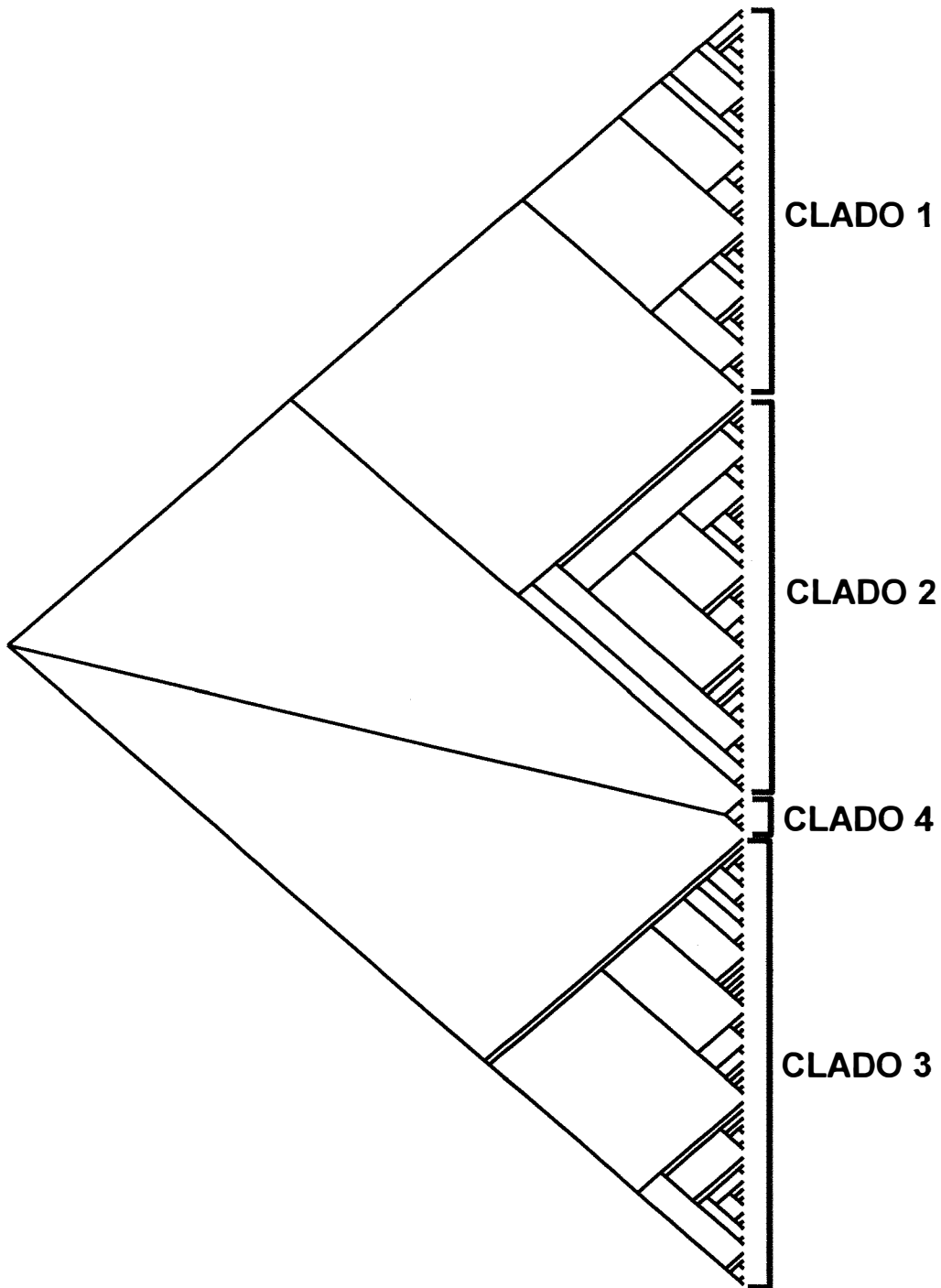


Figura 6 - Árvore filogenética não enraizada das isoformas de FtSHs bacterianas obtida pelo método de vizinho mais próximo

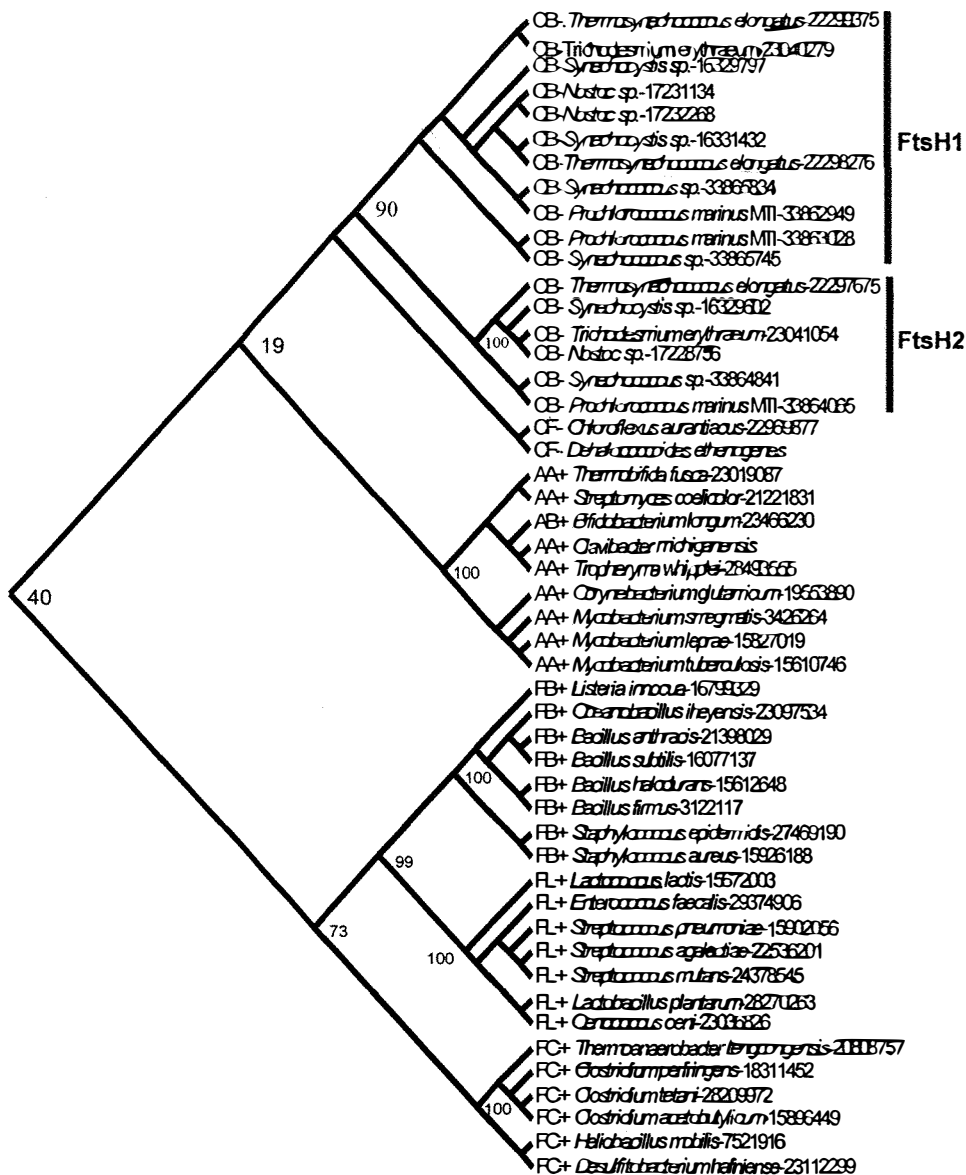


Figura 7 - Detalhe do clado 1 da filogenética das isoformas de FTsH bacterianas. As siglas dos filos bacterianos antecedem os nomes das espécies e estão definidas na tabela 1

O clado 2 é constituído exclusivamente por FtsHs de espécies de bactérias Gram negativas: CFBs, Espiroquetas, Clamídias, Fibrobacteres e proteobactérias, exceto ϵ -proteobactérias (Figura 8). Todos os filós agrupados neste clado possuem pelo menos uma espécie com mais de uma isoforma por genoma. Entretanto, apenas uma isoforma foi agrupada neste clado. O mesmo observado no clado 1.

A topologia dos cladós 1 e 2 revela um aspecto importante sobre a filogenia das FtsHs que pertencem a estes grupos. Ela recupera grupos bacterianos reconhecidamente monofiléticos (Figuras 7 e 8). Além disso, a ordem de emergência grupos bacterianos monofiléticos é congruente com o consenso das filogenias de bactérias obtidas a partir de outros conjuntos de dados como RNA ribossomal e outras proteínas (Brochier et al., 2002; Daubin et al., 2001; Woese, 1987).

O terceiro clado é um verdadeiro mosaico bacteriano. Nele estão agrupadas as duas isoformas presentes nos genomas das ϵ -proteobactérias e Firmicutes/Clostridia (Figura 6 e 9). Também fazem parte deste clado as FtsHs de grupos monofiléticos de bactérias Gram positivas e Gram negativas que possuem apenas uma isoforma como Firmicutes/Mollicutes e CFBs, além dos parálogos das isoformas dos cladós 1 e 2 (Figura 9).

O clado 3 possui uma bifurcação interna originando dois cladós menores. As duas isoformas das ϵ -proteobactérias ficam separadas nestes subcladós, uma em cada ramo. O mesmo é observado para as duas isoformas das Firmicutes/Clostridia. Isto sugere que estas isoformas já estavam presentes no ancestral das bactérias Gram positivas e Gram negativas.

A transferência lateral de genes (TLG) é um fenômeno comum em procariotos sendo considerado uma das principais forças que dirige a evolução nestes organismos (Doolittle, 1999). Entretanto, diante do cenário exposto acima é mais parcimonioso inferir que pelo menos dois antigos eventos de duplicações gênicas, e não TLG, originaram as diferentes isoformas de FtsHs encontradas nos genomas bacterianos.

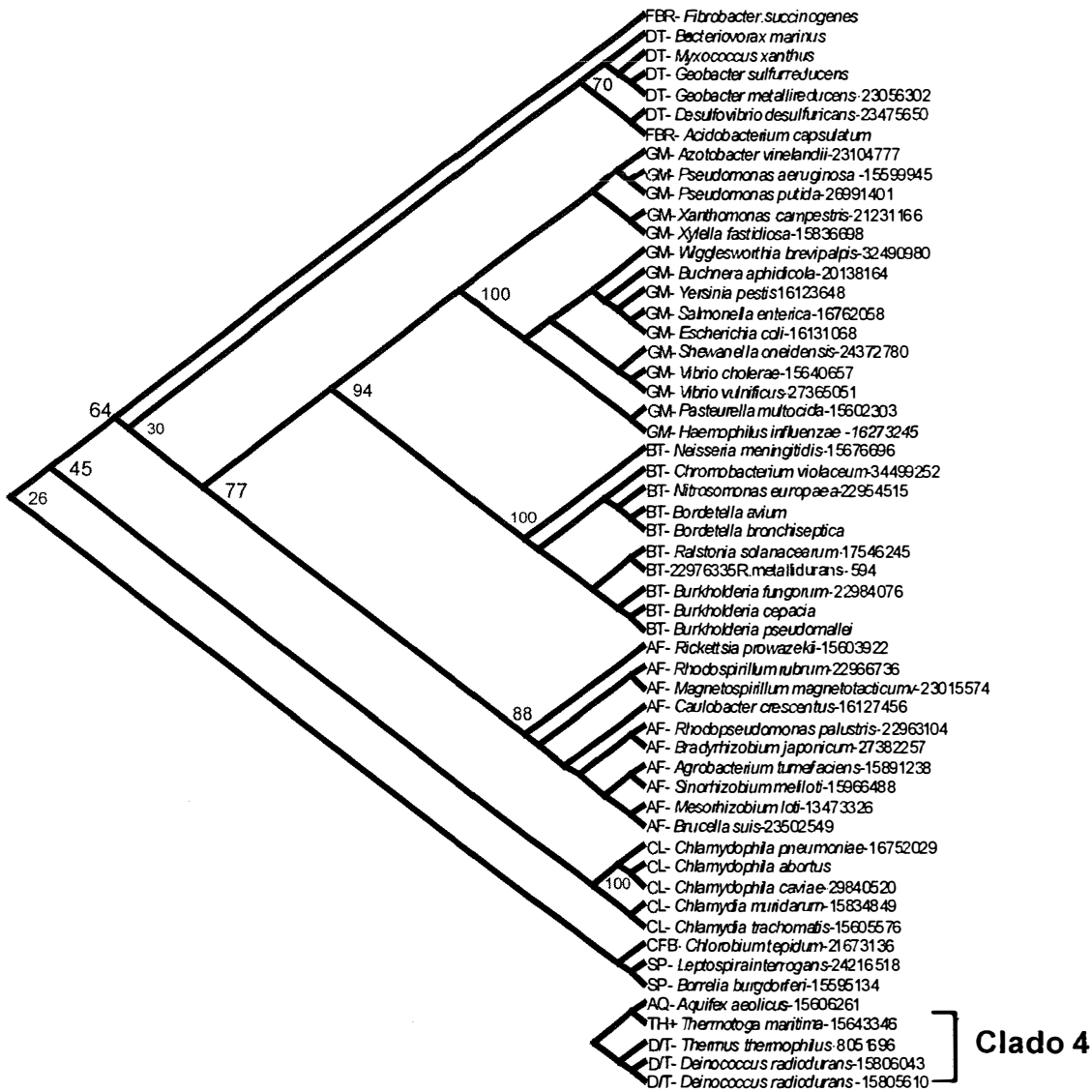


Figura 8 - Detalhe do clado 2 e o clado 4 obtido na reconstrução filogenética das isoformas de FtsH bacterianas. As siglas dos filos bacterianos antecedem os nomes das espécies e estão definidas na tabela 1



Figura 9 - Detalhe do clado 3 da reconstrução filogenética das isoformas de FtsH bacterianas. As siglas dos filos bacterianos antecedem os nomes das espécies e estão definidas na tabela 1

A primeira duplicação originou a FtsH ancestral das isoformas dos clados 1 e 2 e a ancestral das isoformas do clado 3. Uma nova duplicação da FtsH ancestral das isoformas do clado 3 originou as FtsHs ancestrais dos dois subclados deste grupo. A partir de agora as FtsHs dos clados 1 e 2 serão

referidas como isoformas 1 e as isoformas dos subclados do clado 3 como Isoformas 2 e 2'.

A existência das isoformas 1, 2 e 2' em bactérias Gram positivas e Gram negativas posiciona os eventos de duplicação que originaram estas isoformas pelo menos antes da divergência destes dois grande grupos bacterianos. Qualquer tentativa de posicionar estes eventos antes da divergência entre Gram positivas e Gram negativas, é meramente especulativa, pois a filogenia bacteriana ainda não está bem resolvida.

Existe atualmente uma grande controvérsia a respeito de quais são os grupos monofiléticos basais do Domínio Bacteria, isto é, os grupos mais primitivos. Vários autores usando diferentes metodologias já propuseram diferentes grupos candidatos, dentre eles, os Filos Thermotogae e Aquificae são o mais recorrentes (Woese, 1987; Gupta, 2002; Brochier & Philippe, 2002; Di Giulio, 2003). As FtsHs destas espécies estão agrupadas no clado 4 fora dos clados 1, 2 e 3 (Figura 8). Isto poderia sugerir que ele seria um grupo basal. Entretanto, dependendo do modo como a reconstrução filogenética é feita, eliminando os indels ou usando apenas o alinhamento da região do módulo AAA, este clado fica fora dos clados 1, 2 e 3 ou dentro de um deles. Isto significa que sua posição não está bem resolvida, talvez por ser grupo é eqüidistante dos clados 1, 2 e 3.

O envolvimento de TLG na história evolutiva destas isoformas não é descartado, entretanto, se ocorreram foram pouco freqüentes, muito antigos e de difícil detecção. Como discutido anteriormente, a ordem de emergência alguns Filos bacterianos, bem como o grupo basal, ainda não estão bem resolvidos. Isto inviabiliza um enraizamento da árvore filogenética das FtsHs bacterianas que possibilitaria detectar eventuais TLGs.

Alguns autores postulam que proteínas envolvidas em múltiplas interações com outras proteínas são menos sujeitas a serem transferidas lateralmente. Elas representam um grupo de proteínas que pode ser usado para recuperar a filogenia bacteriana (Brochier et al., 2002). A FtsH possui este perfil

e, como já exposto anteriormente, sua filogenia possui um elevado nível de congruência com a filogenia bacteriana (Figuras 6, 7 e 8).

Independentemente de qual seja este grupo basal, a distribuição das isoformas 1, 2 e 2' sugere que os eventos de duplicação que originaram estas isoformas são muito antigos. Um cenário provável é que estas duplicações tenham ocorrido no organismo ancestral das bactérias ou logo após a divergência de poucos grupos basais. Estas isoformas foram sendo mantidas, perdidas e/ou eventualmente transferidas lateralmente nas linhagens de bactérias que emergiram posteriormente. Esta hipótese explica com parcimônia a distribuição das isoformas de FtsH entre os grupos monofiléticos bacterianos, além de ser congruente com a topologia da árvore filogenética principalmente nos clados 1 e 2.

A isoforma 2 presente em algumas proteobactérias também é um bom exemplo de congruência da topologia da árvore filogenética com a filogenia bacteriana. A maioria das espécies de α , β e γ -proteobactéria possuem apenas a isoforma 1 enquanto que as δ -proteobactérias possuem uma isoforma 1 e uma isoforma 2. Contudo, existe pelo menos uma espécie dentro das classes α , β e γ -proteobactérias que possui as isoformas 1 e 2 (Tabela 1 e Figura 8). A topologia da árvore filogenética indica que a isoforma 1 estava presente no ancestral das proteobactérias e, que a isoforma 2 é na verdade um parálogo da isoforma 1.

Estes dados poderiam, a princípio, sugerir que a isoforma 2 encontrada nas α , β e γ -proteobactérias foram adquiridas por TLG. Mas, a topologia da árvore filogenética não corrobora esta hipótese. Estas isoformas 2 formam um clado distinto que inclui também a isoforma 2 de δ -proteobactérias, indicando que elas derivam de um único ancestral (Figura 9). Como as proteobactérias são um grupo monofilético é razoável inferir que a isoforma 2 ancestral estava presente no ancestral das proteobactérias. Posteriormente ela foi perdida em algumas linhagens e mantida em outras. A isoforma 2 da α proteobactéria *A.*

vinelandii foi excluída da análise filogenética. Ela está parcialmente anotada e não Possui a maior parte da região C-terminal. Contudo, a região que foi anotada possui maior identidade com os membros do clado da isoforma 2 de proteobactérias referido acima.

Os valores de *bootstrap* dos clados 1, 2 e 3 da árvore filogenética obtida pelo método de Agrupamento de vizinhos foram baixos independentemente de como foi feita a análise, incluindo ou excluindo posições com indels ou fazendo a correção para múltiplas substituições.

Valores de *bootstrap* baixos podem ser devido à presença de pouco sinal filogenético nos dados resultantes da saturação das substituições nos sítios ou do baixo nível de divergência das seqüências usadas nas análises. O nível médio de identidade entre todas as isoformas usadas nas análises é de 54,6% (Figura 10). Por sua vez o nível médio de identidade das seqüências dentro dos grupos monofiléticos é de 74,1%, 19,5% maior que nível médio de identidade de todas as isoformas (Figura 11). Este resultado mostra que o nível de divergência das seqüências dentro dos grupos monofiléticos não difere muito do nível de divergência das seqüências entre os grupos monofiléticos.

O módulo AAA é a região mais conservada das FtsHs. Ele representa a maior parte do alinhamento utilizado para fazer as reconstruções filogenéticas, 250 posições de um total de 460 posições. Ou seja, uma parte considerável do alinhamento usado para recuperar a história evolutiva das FtsHs é uma região funcionalmente conservada nas FtsHs. Talvez este seja o fator que esteja contribuindo para a obtenção de níveis baixos de *Bootstrap*.

Entretanto, a topologia da árvore recupera todos grupos bacterianos reconhecidamente monofiléticos bem como sua a ordem de emergência (Figuras 6, 7 e 8). Isto significa que o sinal filogenético existente é suficiente para recuperar a história evolutiva das FtsHs mas, insuficiente para dar suporte estatístico aos clados formados. Além disso, a relação entre os clados 1, 2 e 3 pode ser recuperada usando outra abordagem. A comparação do nível médio de similaridade entre as seqüências dos 3 clados revela que as seqüências dos

clados 1 e 2 são mais similares entre si do que quando comparadas com as seqüências do clado 3 (Figura 10). Nestas circunstâncias faz mais sentido o argumento biológico do que o argumento estatístico.

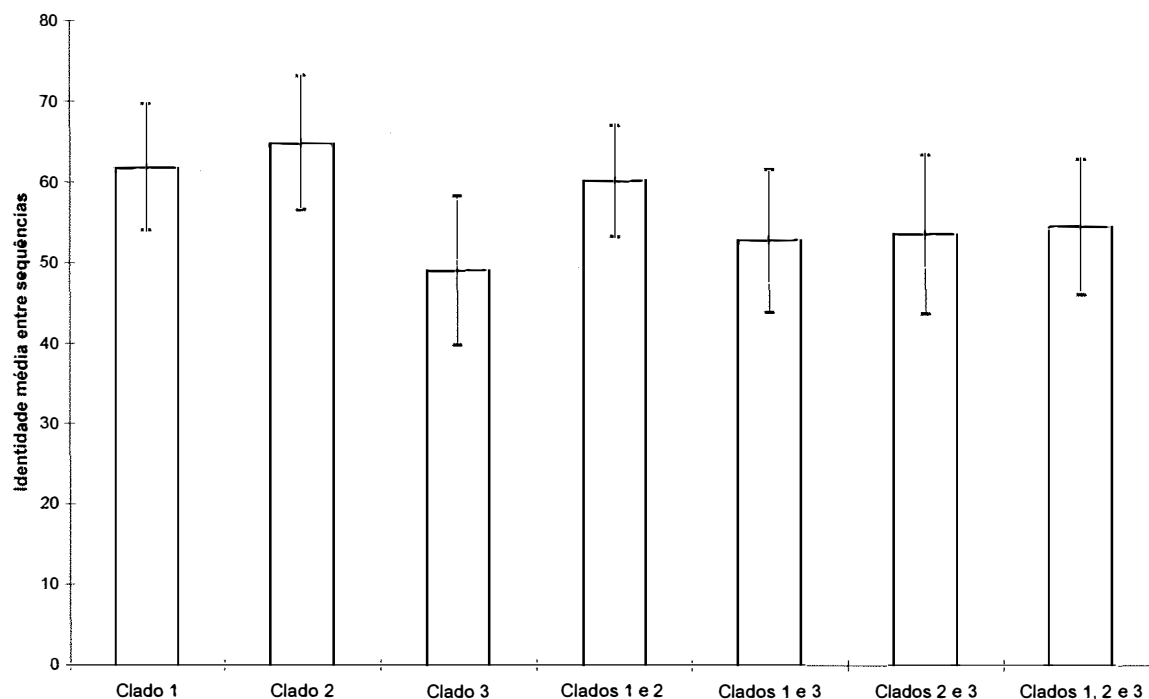


Figura 10 - Relação das médias de identidade entre as seqüências de FtsHs dentro dos clados e entre os clados. Os valores de identidade média foram obtidos a partir da matriz de distância genética das seqüências gerada pelo ClustalX

Na verdade, estes níveis de identidade altos podem estar refletindo o nível de conservação funcional destas proteínas. A FtsH de *Lactococcus lactis*, uma bactéria Gram positiva, é capaz de restaurar a viabilidade celular de *Escherichia coli*, uma bactéria Gram negativa, cuja FtsH foi inativada (Nilsson et al., 1994). Este resultado indica que a FtsH destas espécies de bactérias ainda possuem algumas funções em comum. O fato desta complementação ser

parcial sugere que estas enzimas passaram a desempenhar função(ões) específica(s) dentro destas linhagens ou que esta incapacidade é devido a coevolução enzima/moduladores/substrato dentro da linhagem da espécie.

Em *Escherichia coli* o fator transcricional σ^{32} é degradado pela FtsH. *Bradyrhizobium japonicum* possui 3 homólogos de σ^{32} , RpoH₁, RpoH₂ e RpoH₃. Urech et al., (2000) expressaram estes homólogos em *Escherichia coli* e verificaram que a FtsH nativa degradou *in vivo* apenas RpoH₁, RpoH₂. Entretanto, ela foi capaz de degradar todos os 3 homólogos em ensaios *in vitro*. Estes resultados sugerem que a degradação de RpoH₃ requer algum tipo de modulador e/ou interação distintos dos necessários para a degradação de σ^{32} , RpoH₁ e RpoH₂ em *Escherichia coli*.

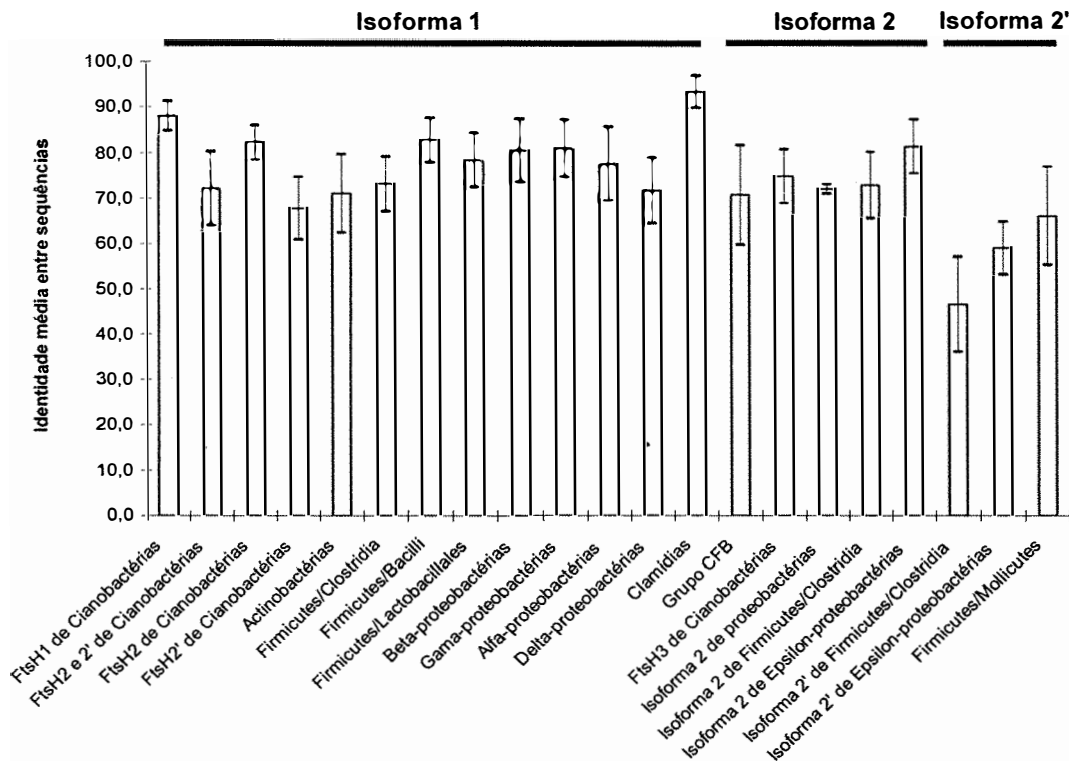


Figura 11 - Identidade média entre as seqüências de FtsH dentro dos grupos monofiléticos bacterianos. Os valores de identidade média foram obtidos a partir da matriz de distância genética das seqüências gerada pelo ClustalX

4.1.3 Análise estrutural das FtsHs bacterianas

Proteínas funcionalmente relacionadas geralmente possuem estrutura secundária similar, o que faz com que a classificação de proteínas baseada na estrutura secundária seja comparável à sua classificação estrutural (Przytycka et al., 1999). Apesar de possuírem um alto nível de identidade na sua seqüência de aminoácidos é possível distinguir as diferentes isoformas das FtsHs com base nas diferenças nas suas estruturas primárias (Figura 6).

Contudo, altos níveis de similaridades também são observados entre proteínas filogeneticamente relacionadas mas que são funcionalmente distintas (Beyer, 1997). A similaridade média entre o módulo das proteínas AAA do proteassomo e o módulo das FtsHs é de aproximadamente 60%.

Testes de complementação mostram que FtsHs de bactérias que pertencem a Filos muito divergentes ainda possuem conservação funcional (Nilsson et al., 1994; Narberhaus et al., 1999; Fischer et al., 2002). Embora, estes testes de complementação envolverem sempre a mesma isoforma de FtsH, à exceção das FtsHs de *Helicobacter pylori* e *Escherichia coli*. Assim sendo, a estrutura secundária das diferentes isoformas de FtsHs foi analisada comparativamente com o objetivo de investigar se elas podem também ser distinguidas neste nível estrutural.

Inicialmente a estrutura secundária de uma seqüência representativa de cada clado monofilético bacteriano foi predita com o auxílio do programa Jpred (Cuff et al. 1998). As três isoformas foram representadas. Em seguida estas seqüências foram alinhadas e identificadas as regiões onde foram preditas estruturas secundárias.

Esta análise revelou que o alto nível de similaridade entre as 3 isoformas também é observada em nível de estrutura secundária (Figura 12). O nível de similaridade da estrutura primária das isoformas é variável, sendo maior na região N-terminal (região onde localiza-se o módulo AAA) do que na região C-terminal (região do domínio proteolítico). Entretanto, a similaridade de

estruturas secundárias ao longo da proteína entre sequências é próximo de 100%.

A estrutura secundária predita pelo programa Jpred é um consenso da predição de 5 programas e tem um percentual médio de acerto de 73% (Andreas & Ouellette, 2001). Para avaliar o nível de confiança do Jpred sua predição foi comparada com a estrutura secundária do módulo AAA da FtsH de *Escherichia coli* obtida por cristalografia (Krzywda, et al., 2002). Cerca de 81% das estruturas secundárias preditas pelo Jpred estão em concordância com a estrutura secundária do módulo AAA da FtsH de *Escherichia coli* resolvida por cristalografia (Figura 12).

As seqüências apresentaram variação na predição de duas estruturas secundárias dentro do módulo AAA, a $\beta 1$ e a $\alpha 6$. Estas estruturas estão presentes na FtsH de *Escherichia coli*. Contudo, esta variação não está correlacionada com a filogenia das FtsHs, nem tão pouco com a filogenia bacteriana de forma que não puderam ser usadas para discriminar as isoformas das FtsHs.

Vários sítios invariáveis podem ser observados em posições onde não foram preditas estruturas secundárias (Figura 12) embora seqüências de todas as isoformas, incluindo todos os grupos monofiléticos bacterianos, tenham sido incluídas neste alinhamento. Isto indica uma forte pressão de seleção nestes sítios, o que revela uma grande importância funcional dos aminoácidos conservados nestas posições. As funções de apenas duas destas regiões já foram bem caracterizadas funcionalmente, o motivo Walker A e o SRH, as demais ainda estão por ser elucidadas.

Atividade chaperona já foi descrita para a região N-terminal da FtsH mitocondrial YME1 (Leonhard et al., 1999). Esta região é homóloga à região N-terminal anterior à fita β pregueada $\beta 2$ da FtsH de *Escherichia coli* (Figura 12). Vários sítios invariáveis estão presentes na alça entre a α -hélice $\alpha 1$ e a fita β pregueada $\beta 2$. Estes sítios são fortes candidatos a estarem funcionalmente relacionados com a atividade chaperona das FtsHs. Como apenas a região do

módulo AAA da FtsH *Escherichia coli* foi resolvida por cristalografia não foi possível avaliar a confiabilidade das estruturas secundárias preditas na região C-terminal das isoformas de FtsH.

Uma alternativa para avaliar esta confiabilidade é verificar a existência de associação entre as estruturas secundárias preditas e algum fenômeno biológico. Várias mutações capazes de inativar a atividade proteolítica da FtsH de *Escherichia coli* estão localizadas na região C-terminal FtsHs (Qu et al., 1996; Wang et al., 1998; Gaustsch e Wulff, 1974). Portanto, a alternativa escolhida foi mapear estas mutações e verificar se existe alguma correlação com as estruturas secundárias preditas (Figura 12). 5 das 6 mutações identificadas ficam dentro de regiões onde foram preditas estruturas secundárias, o que permite associar uma função biológica a estas regiões.

Uma característica marcante da região C-terminal é o baixo nível de conservação da sua estrutura primária quando comparada com a região do módulo AAA (Figura 12), embora seja observada uma alta similaridade quanto à estrutura secundária. O nível de conservação da estrutura primária na região C-terminal é maior quando seqüências de FtsHs de um mesmo grupo bacteriano são comparadas. Entretanto, o nível de conservação nesta região é variável entre os grupos monofiléticos bacterianos, assim como as regiões que apresentam uma alta similaridade (Figura 13).

Uma hipótese atraente é que estas variações específicas estejam relacionadas com a diversificação funcional destas proteínas dentro das linhagens bacterianas. As mutações mapeadas na região C-terminal estão muitas vezes relacionadas com perda da atividade proteolítica causando diferentes fenótipos mutantes, provavelmente por afetarem de maneira específica a degradação dos diferentes substratos das FtsHs (Qu et al., 1996; Wang et al., 1998; Gaustsch e Wulff, 1974).

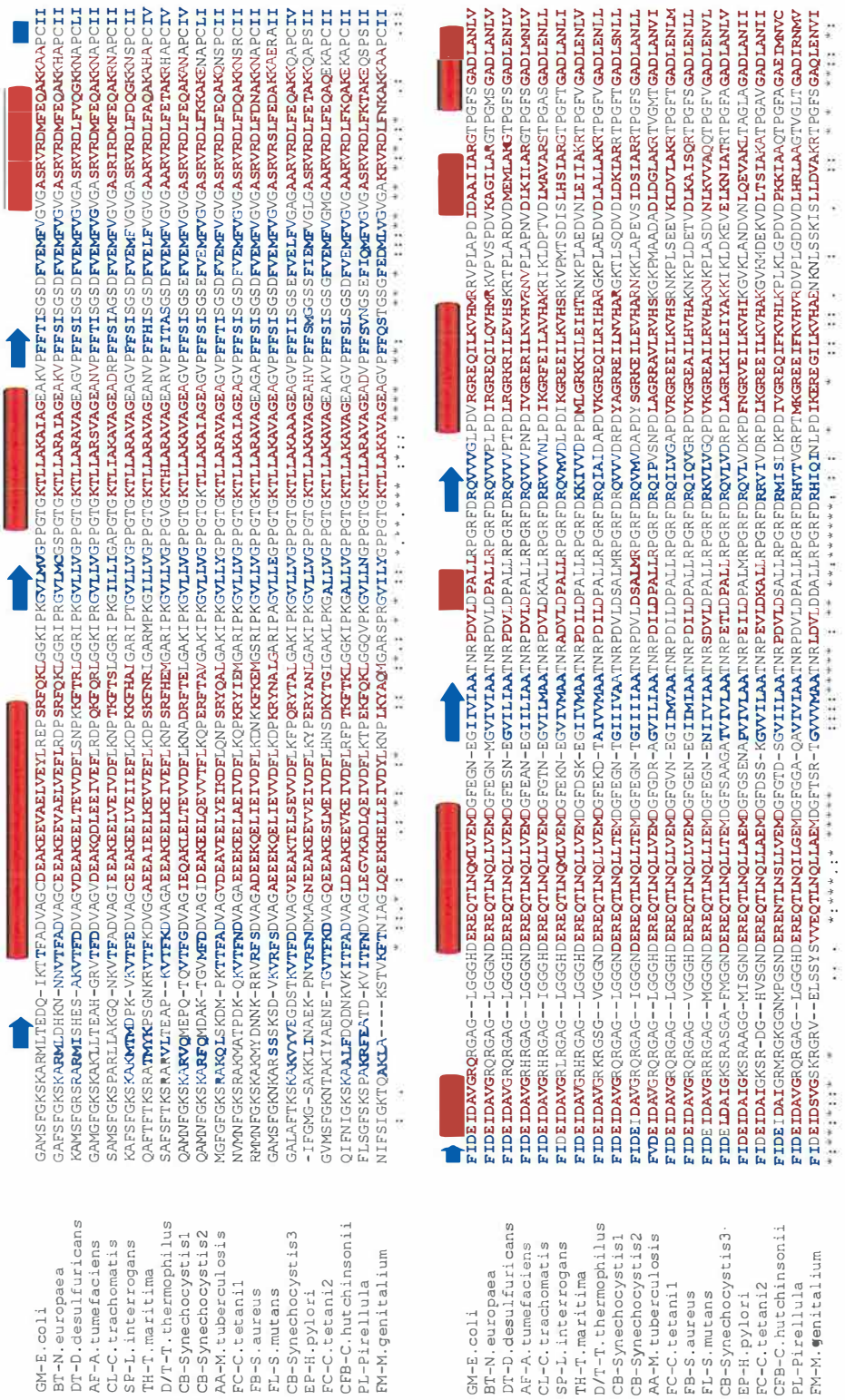


Figura 12 - Alinhamento de seqüências representativas de todos os principais grupos monofiléticos de bactérias. A região alinhada estende-se do início do módulo AAA até o provável motivo "coiled coil" localizado na região C-terminal. As estruturas secundárias preditas pelo programa Jpred estão destacadas, vermelho, α -hélice; Folha β -pregueada, azul. A estrutura secundária do módulo da FisH de *Escherichia coli* esta representada na forma de cilindro, α -hélices, e setas, Folha β -pregueada. As setas indicam as posições onde mutações causam perda de atividade proteolítica da FisH de *Escherichia coli*

mitocondriais FtsHm1 e FtsHm2. Todos os grupos foram suportados por consistentes valores de bootstrap (Figura 14).

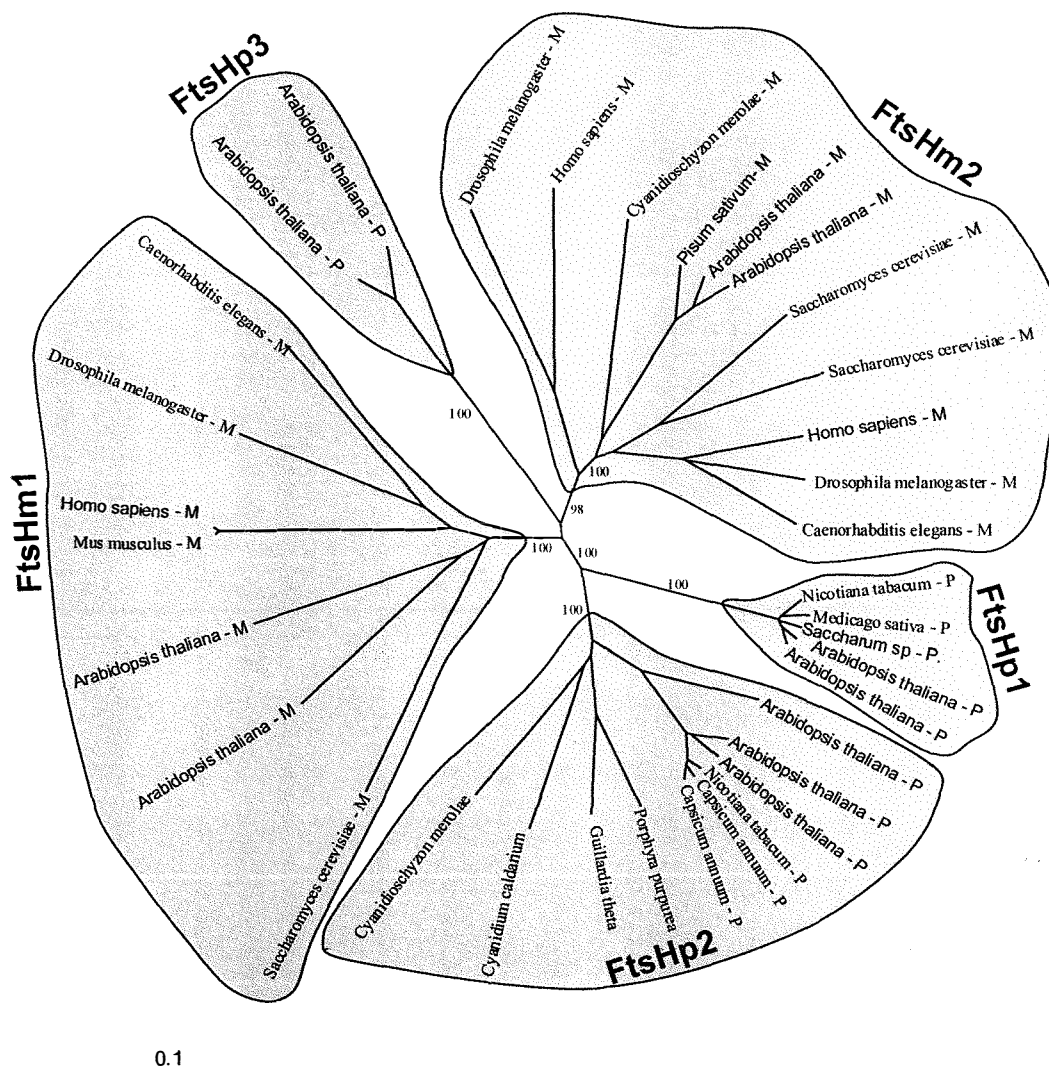


Figura 14 - Árvore filogenética não enraizada das FtsHs eucarióticas obtida pelo método de Agrupamento de vizinhos. Os números indicam os valores de *bootstrap*. P, Plastídeo; M, Mitocôndria

Evidências adicionais que suportam estes grupos filogenéticos foram obtidas a partir da análise do perfil de hidropaticidade das FtsHs eucarióticas (Figura 15). Esta análise mostra a distribuição de regiões hidrofóbicas e hidrofílicas ao longo da estrutura primária de uma proteína. Ela é

comumente utilizada na localização dos possíveis segmentos transmembrana da proteína. Esta abordagem também foi utilizada por Lolkema e Slotboom (2003) para caracterizar famílias de proteínas de membrana.

A observação de que proteínas pertencentes a um mesmo grupo filogenético possuem perfis de hidropaticidade altamente conservados sugere que estes possam ser utilizados como assinatura do grupo (Figura 15). Isto indica também que estas proteínas sejam funcionalmente relacionadas pois o perfil de hidropaticidade determina, em última análise, a estrutura terciária das proteínas (Lolkema e Slotboom, 2003).

Foram incluídas na análise filogenética a FtsH-p2 das algas vermelhas e de *Guillardia theta*, uma cryptophyta. Nestes organismos a FtsHp2 está codificada no genoma plastidial. Estas algas divergiram antes do ancestral das plantas superiores. Assim sendo, pode-se inferir que o evento de transferência do gene que codifica a FtsHp2 em plantas para o genoma nuclear ocorreu após a divergência entre o ancestral das atuais plantas superiores e as linhagens ancestrais destas algas.

Os clados monofiléticos das FtsHs mitocondriais, FtsHm1 e FtsHm2, possuem representantes em plantas, fungos e metazoários. A topologia destes clados mostra que seus membros derivaram de um único ancestral que foi adquirido pela célula eucariótica primitiva. O que está de acordo com a origem endossimbiótica destes genes.

A estrutura da árvore filogenética das FtsHs eucarióticas mostra que os clados possuem um histórico particular de eventos duplicação. Animais e fungos possuem apenas uma isoforma da FtsHm1 enquanto que *Arabidopsis*, uma dicotiledônea, possui duas isoformas denominadas FtsHm1A e FtsHm1B (Figuras 14 e 16).

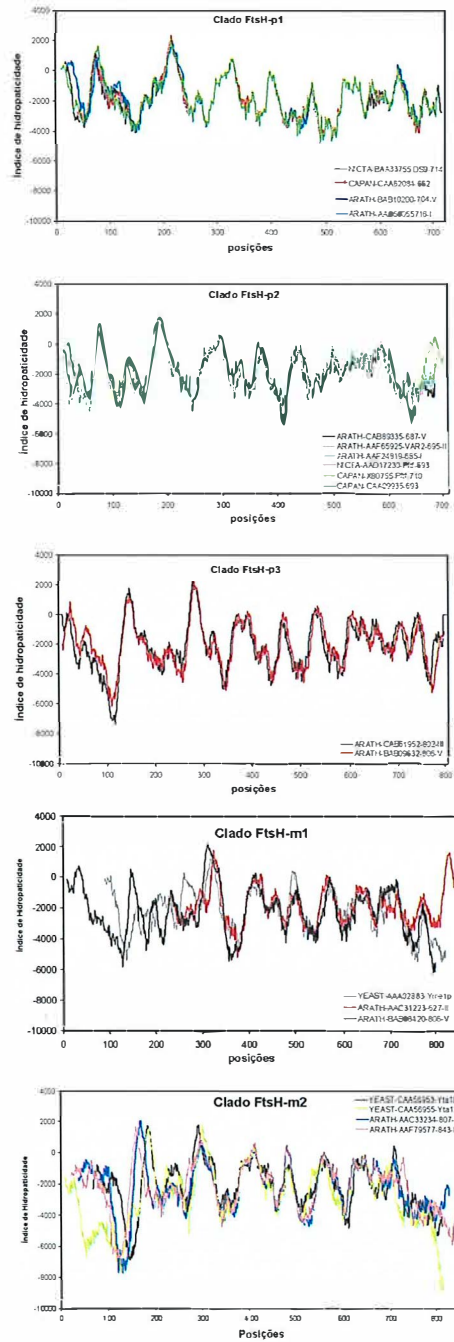


Figura 15 - Sobreposição dos perfis de hidropaticidade de FtsHs pertencentes a um mesmo grupo filogenético

Como representantes de monocotiledôneas não foram incluídos nesta análise, não é possível definir se este evento de duplicação ocorreu dentro da linhagem de *Arabidopsis* ou é mais antiga. Ou seja, ocorreu antes da divergência entre monocotiledôneas e dicotiledôneas. O mesmo pode ser dito para os demais clados.

Por sua vez, a topologia do clado das FtsHm2 sugere um evento de duplicação antigo, no ancestral das plantas animais e fungos. Posteriormente, uma das isoformas foi perdida no ancestral das plantas e fungos. Uma nova duplicação desta isoforma ocorreu em *Arabidopsis* e levedura. Como representantes de monocotiledôneas e outras espécies de fungos não foram incluídos na análise não é possível inferir quando estas duplicações ocorreram.

4.2.2 FtsHs em *Arabidopsis thaliana* e Cana-de-açúcar

As FtsHs são mais extensivamente estudadas em procariotos e levedura. Existe pouca informação a respeito da distribuição e função destas proteínas em plantas, apesar do seu papel central na fotossíntese e respiração celular. Entretanto, a disponibilidade da seqüência completa do genoma de *Arabidopsis thaliana* e do transcriptoma de cana-de-açúcar foi possível analisar comparativamente, a ocorrência e distribuição destas proteínas em dicotiledôneas e monocotiledôneas e assim ter uma visão mais completa desta família de proteínas nas plantas superiores.

Arabidopsis thaliana possui 12 genes que codificam proteínas FtsHs no seu genoma, com pelo menos dois representantes em cada grupo filogenético das FtsH eucarióticas (Figura 14) dispersos nos cromossomos I, II, III e V (Figura 16). A nomenclatura usada para estes genes neste trabalho leva em consideração seu perfil filogenético e sua localização cromossomal (Figura 16). Por exemplo, o gene localizado no cromossomo 5 que codifica a FtsH pertencente ao grupo filogenético FtsHp1 foi denominado FtsHp1-V.

Acredita-se que o atual genoma de *Arabidopsis thaliana* é o resultado de pelo menos 3 eventos de duplicação gênica em larga escala. Simillion et al.

(2002) identificaram as regiões do genoma de *Arabidopsis* oriundas destes eventos de duplicação. Estas regiões foram devidamente mapeadas e estão disponíveis no site www.psb.rug.ac.be/.

Os membros dos grupos filogenéticos FtsHp1, FtsHp2, FtsHp3 e FtsHm2 estão localizados dentro de regiões envolvidas em eventos de duplicação. Isto é uma evidência de que as isoformas destes grupos se originaram de eventos de duplicação dentro da linhagem de *Arabidopsis*, ou seja, são parálogos. No caso das FtsHp2 e FtsHm2 esta evidência é visual. A disposição e orientação dos genes que codificam as FtsHp2-I/FtsHm2-I é a mesma que dos genes que codificam FtsHp2-II/FtsHm2-II, embora seja invertida (Figura 16).

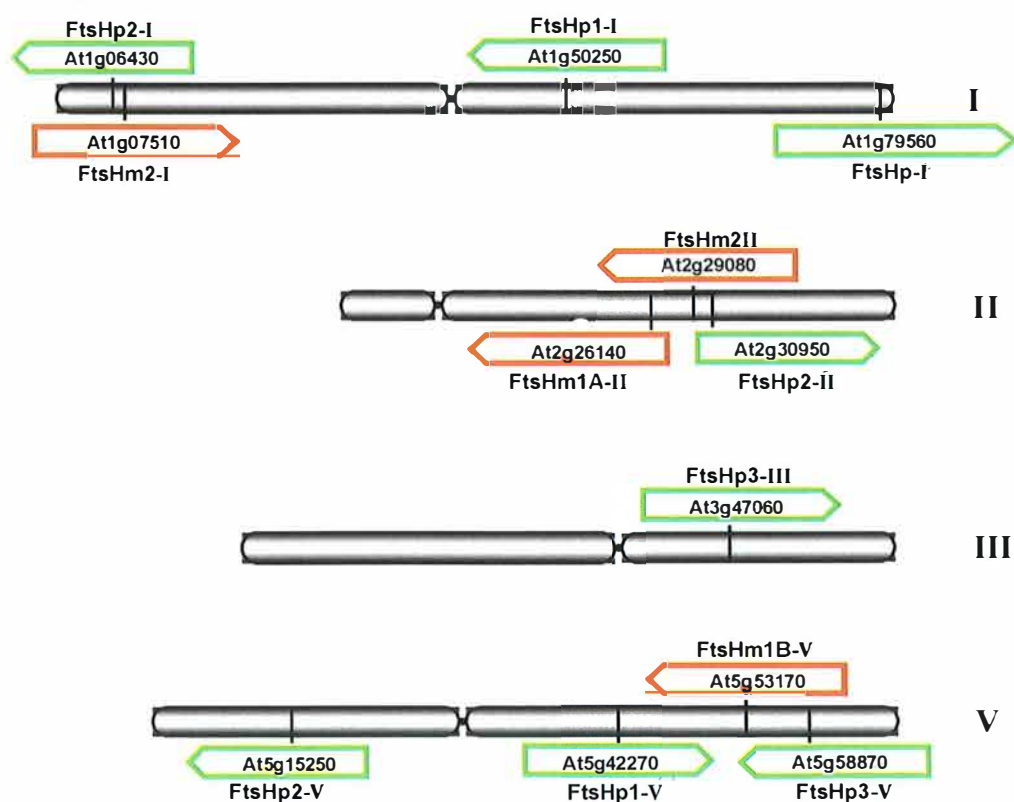


Figura 16 - Distribuição das FtsHs nos cromossomos de *Arabidopsis thaliana*. As setas indicam a orientação dos genes. A cor verde indica os genes que codificam FtsHs plastidiais enquanto que a cor vermelha indica genes que codificam FtsHs mitocondriais

Os genes que codificam as FtsHm1A e FtsHm1B não foram mapeados dentro de regiões derivadas de um evento de duplicação. Portanto não foi possível estabelecer elas se originaram de um evento de duplicação dentro da linhagem de *Arabidopsis thaliana* ou anterior a divergência desta linhagem.

O perfil filogenético das FtsH eucarióticas foi usado para procurar grupos de ESTs de FtsH no banco de dados de cDNA de cana-de-açúcar, o SUCEST. Inicialmente, 153 grupos de ESTs apresentando similaridade com FtsHs foram encontrados, sendo 23 confirmados por meio de comparação via BLASTX nos bancos de dados do GenBank e de *Arabidopsis thaliana* (Figura 17).

Foram encontrados grupos de ESTs homólogos a todos os grupos filogenéticos de FtsHs eucarióticas definido a priori. O perfil de hidropaticidade dos grupos filogenéticos foi usado para confirmar a relação de homologia dos grupos de ESTs com os grupos filogenéticos de FtsHs eucarióticas. Apenas o grupo SCVPRT2083G10.g um apresentou um perfil de hidropaticidade incompatível com o perfil de hidropaticidade do grupo filogenético com o qual ele apresentou maior similaridade (Figura 18). Este grupo de EST codifica para 135 aminoácidos e apresentou maior similaridade com a FtsHp2 de *Capsicum annuum*. O *e-value* do alinhamento par a par foi de $5e^{-16}$, isto mostra o poder do perfil de hidropaticidade na detecção de homologia.

A existência de grupos de ESTs com maior similaridade a FtsHm1A ou a FtsHm1B indica que estas isoformas também existem em cana-de-açúcar. Isto posiciona o evento de duplicação que deu origem as duas isoformas antes da divergência entre monocotiledôneas e dicotiledôneas (Figura 17 e Figura 18).

A função de membros dos grupos plastidiais FtsH-p1 e FtsHp-2 foi recentemente descrita. Em *Arabidopsis* estas proteínas estão localizadas nas membranas do tilacóide e estão aparentemente envolvidas na reposição da proteína D1 do fotossistema II e biogênese das membrana do tilacóide, respectivamente (Lindahl *et al.*, 2000; Chen *et al.*, 2000).

Em tabaco um membro do grupo FtsH-p1 parece estar associado com a reação de hipersensibilidade (Seo *et al.*, 2000). Estas proteínas apresentam alta identidade/similaridade com grupos de de ESTs de cana-de-açúcar (figura 17 e Figura 18). Considerando que estes eventos são cruciais em monocotiledôneas e dicotiledôneas, é provável que estas proteínas possuam funções filogeneticamente conservadas nos plastídeos de ambas linhagens.

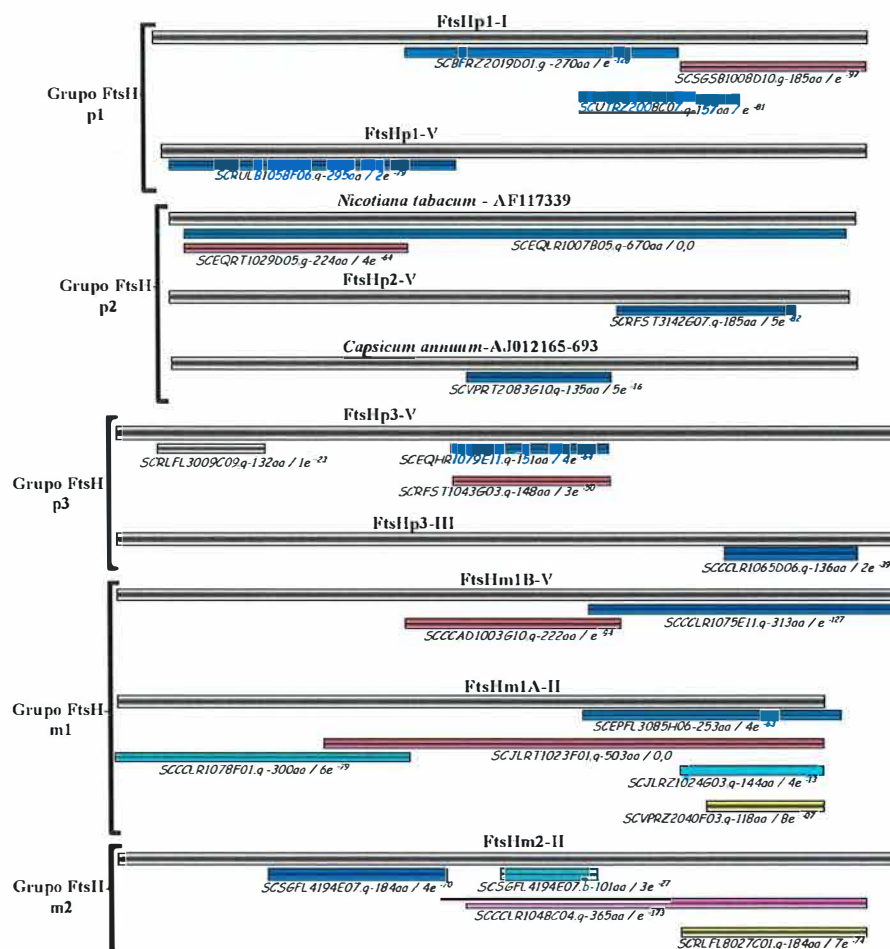


Figura 17 - Representação diagramática do alinhamento dos grupos de ESTs de cana-de-açúcar (barras coloridas) em relação às seqüências de proteínas depositadas do banco de dados do NCBI com a qual apresentam maior identidade (barras cinzas). O nome do grupo de ESTs, o número de aminoácidos codificado pela seqüência consenso do grupo e o *e-value* do alinhamento da seqüência consenso de cada grupo de ESTs estão indicados embaixo de suas respectivas representações diagramáticas

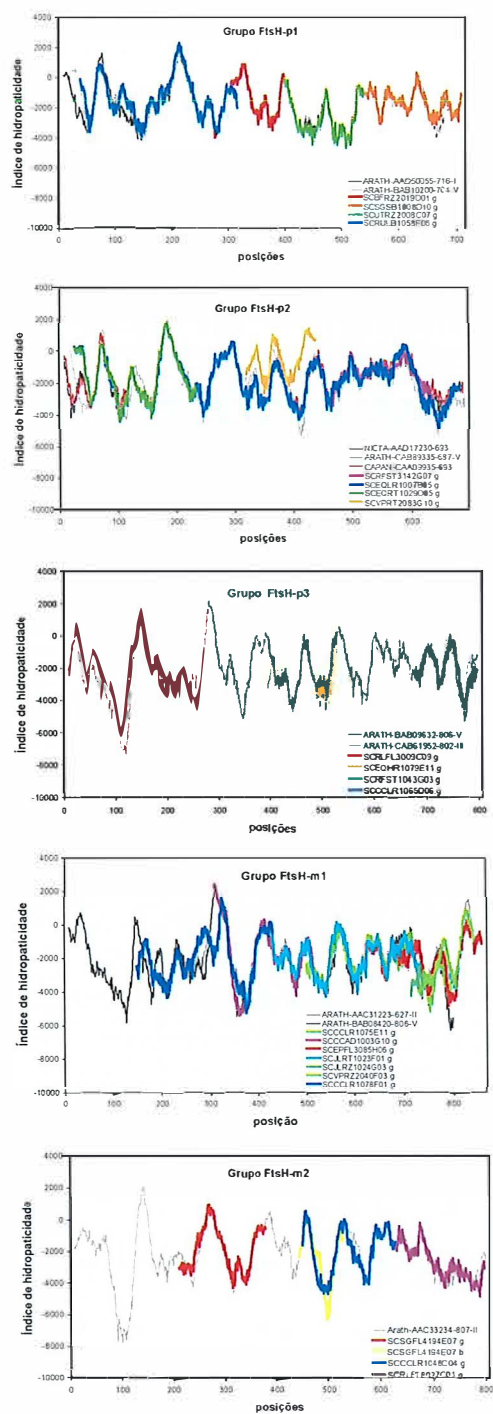


Figura 18 - Sobreposição dos perfis de hidropaticidade dos grupos de ESTs e das respectivas FtsHs homólogas depositadas no banco de dados do NCBI

Em levedura as FtsHs estão envolvidas na montagem de complexos da cadeia respiratória e na degradação das subunidades não incorporadas nestes complexos. Estas funções parecem ser conservadas entre eucariotos visto que o ortólogo humano de *Yme1* complementa o mutante de levedura com o gene *yme1* nocauteado (Shah *et al.*2000). Recentemente, Kolodziejczak *et al.* (2002) mostraram que a FtsHm2 de *Pisum sativum* é capaz de complementar seus ortólogos de levedura YTA10, e YTA12. Isto sugere que as FtsHs eucarióticas são filogeneticamente e funcionalmente relacionadas assim como em procariotos.

A duplicação que originou a FtsHm1B ocorreu depois da divergência do reino Plantae dos reinos Fungi/Metazoa e antes da divergência entre Monocotiledôneas e dicotiledôneas. Ou seja, ela é uma FtsH tipicamente vegetal. O fato dela ter sido mantida em Monocotiledôneas e dicotiledôneas sugere uma importância funcional, provavelmente distinta dos seus homólogos dos reinos Fungi e Metazoa.

4.2.3 Análise da expressão *in silico* dos grupos filogenéticos de FtsHs nas bibliotecas de cDNA do SUCEST

O grupo FtsH-p2 é qualitativa e quantitativamente mais expresso que FtsH-p1 e FtsH-p3. Sua expressão é maior em folhas, meristema apical, casca do colmo e raiz (figura 19A). A expressão dos grupos plastidiais na gema lateral não foi detectada, sugerindo que estas proteínas não são expressas ou são expressas em níveis muito baixos nestes órgãos. Por outro lado, todos os grupos plastidiais são expressos em sementes e em plantas infectadas com *Herbaspirillum rubrisubalbicans*. De um modo geral, exceto para gemas laterais, as FtsHs são expressas diferencialmente na maioria dos tecidos de cana-de-açúcar (figura 19A).

Um fato interessante observado foi a expressão de todos os grupos plastidiais na zona de transição colmo-raiz bem como nas raízes. O estudo da

regulação da expressão destas proteínas em tecidos de cana-de-açúcar não expostos a luz torna-se interessante visto que, seus ortólogos em *Arabidopsis* são regulados pela luz (Lindahl, *et al.*, 1996; Chen *et al.*, 2000; Sakamoto *et al.*, 2002).

Os grupos de FtsHs mitocondriais também possuem um grande espectro de expressão gênica e são aparentemente diferencialmente regulados. Entretanto, o grupo FtsH-m1 possui maiores níveis de expressão em tecido foliar que o grupo FtsH-m2 (Figura 19B).

A multifuncionalidade é uma importante característica das metaloproteases AAA como a FtsH de *Escherichia coli* e YME1, YTA10 e YTA12 de levedura. Além disso a atividade proteolítica das metaloproteases AAA parece estar desacoplada da sua atividade chaperona. Uma intrigante questão é se as metaloproteases plastidiais comportam-se como as formas mitocondriais, tanto no aspecto funcional como estrutural. De acordo com as observações sobre a expressão das FtsHs em tecidos não fotossintéticos pode-se sugerir uma nova função celular. Embora especulativo, isto pode trazer mais esclarecimentos ao atual entendimento da família das proteínas FtsH-like eucarióticas.

4.2.4 Relações filogenéticas entre FtsHs procarióticas e eucarióticas

As evidências de uma origem endossimbiótica para as FtsH eucarióticas são muitas, conforme discutido anterior. As seqüências de FtsHs disponíveis nos bancos de dados representam bem o universo eucariótico destas proteínas, e razoavelmente bem o procariótico. Isto permitiu fazer uma análise filogenética de forma a tentar recuperar as relações de ancestralidade entre as FtsHs eucarióticas e procarióticas. Nesta análise foram incluídas seqüências procarióticas usada nos estudo de filogenia da FtsHs bacterianas e dois representantes de cada grupo monofilético eucariótico.

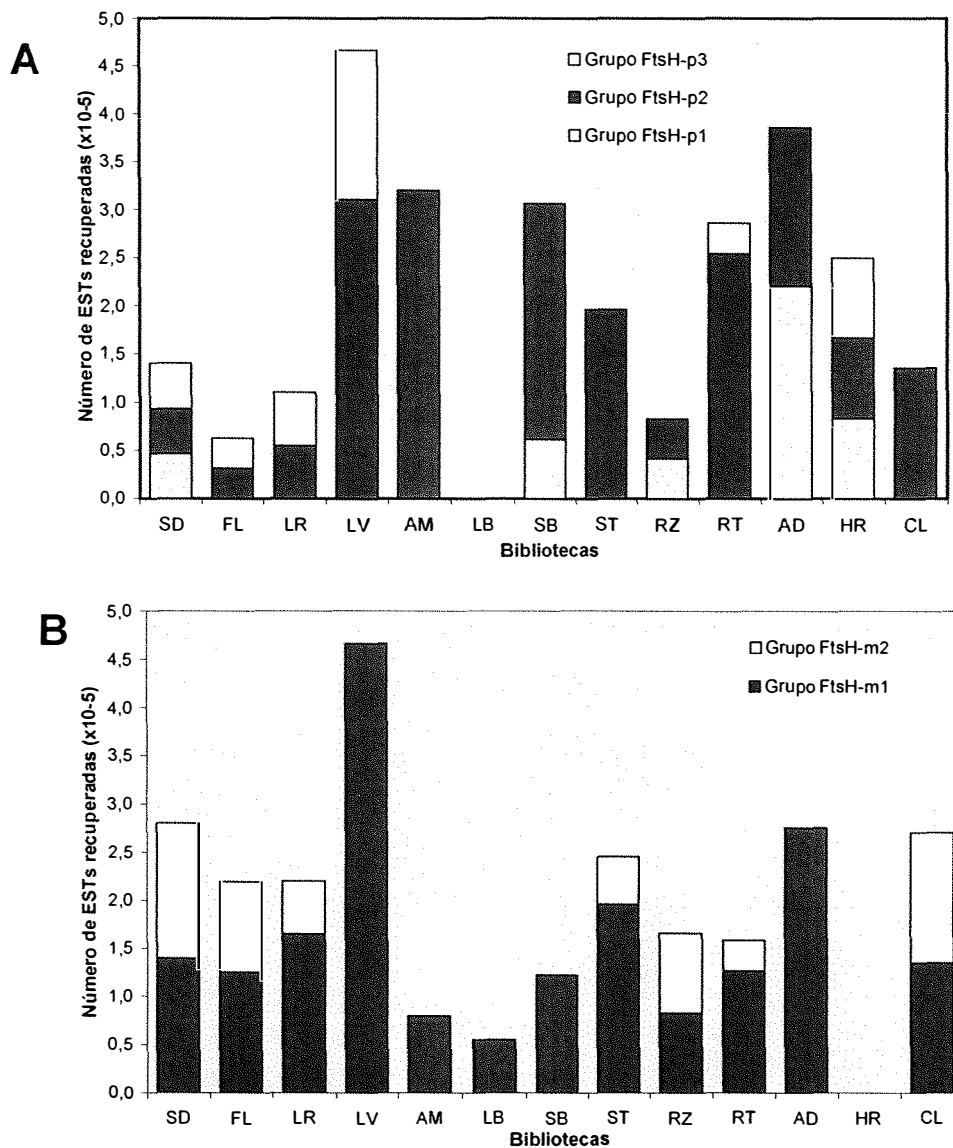


Figura 19 - análise da expressão *in silico* dos grupos de ESTs de FtsHs em diferentes bibliotecas de cDNA de cana-de-açúcar. Abreviações: SD, sementes; FL, flores; LR, cartucho; LV, folhas estioladas; AM, meristema apical; LB, gemas laterais; SB, casca da colmo; ST, medula; RZ, zona de transição folha-raiz; RT, raiz; AD, plântulas crescidas *in vitro* infectadas com *Gluconacetobacter Diazotrophicans*; HR, plântulas crescidas *in vitro* infectadas com *Herbaspirillum rubrisubalbicans*

Taxas divergências elevadas são comuns em proteínas de organelas endossimbióticas, principalmente mitocôndrias. Provavelmente isto é consequência do evento de endossimbiose (Lopez-Garcia & Moreira, 1999). Assim, para prevenir um eventual problema de atração dos ramos longos seqüências de FtsHs bacterianas mais divergentes não foram incluídas nesta análise (Gribaldo & Philippe, 2002; Philippe 1999).

A análise filogenética mostrou que as FtsHs eucarióticas p1, p2 e p3 derivaram do mesmo ancestral que as FtsH1, FtsH2 e FtsH3 de cianobactérias, respectivamente (Figura 20). Este resultado corrobora a teoria endossimbionte para a origem dos plastídeos (Margulis, 1970). Além disso, mostra que estas três isoformas já estavam presentes no endossimbionte que originou os plastídeos. Resultado semelhante foi obtido para as FtsHm1, entretanto com menor suporte estatístico. Elas emergem junto as proteobactérias, atuais descendentes do endossimbionte que originou as mitocôndrias (Figura 20).

O resultado surpreendente nesta análise foi a relação de ancestralidade entre as FtsHm2 e as FtsHs procarióticas. As FtsHm2 emergiram dentro do grupo monofilético das CFBs mostrando uma relação direta com as FtsHs dos organismos desta divisão bacteriana. Não há relatos na literatura de relações simbióticas entre a célula eucariota primitiva e bactérias desta divisão.

Doolittle (1998) propôs um mecanismo para explicar a presença de gene bacterianos cuja origem não está relacionada com os eventos endossimbióticos que originaram as mitocôndrias e plastídeos. De acordo com este modelo, a origem destes genes seria de bactérias que serviriam de alimento para a célula eucariota primitiva. Eventualmente o DNA destas bactérias contendo um determinado gene se integraria no genoma nuclear da célula eucariota primitiva. A eventual a perda do gene nativo, homólogo do gene bacteriano, levaria a fixação do gene bacteriano na população. Este é um interessante mecanismo para explicar a origem da FtsHm2.

A transferência de genes dos endossimbiontes que originaram as mitocôndrias e plastídeos para o genoma nuclear da célula eucariota primitiva é explicado por um mecanismo semelhante (Doolittle, 1998).

Em princípio estes modelos parecem apenas especulativos pois carecem de evidências experimentais que possam validá-los. Entretanto, a evolução é um processo contínuo e certos mecanismos são recorrentes. Stegemann et al. (2003) e Huang et al. (2003) mostraram elegantemente que a transferência do genoma plastidial para o núcleo é um processo contínuo e ocorre com uma frequência relativamente alta. De acordo com Huang et al. (2003) 16 em cada 250000 gametas de tabaco contém uma nova inserção de DNA plastidial no genoma nuclear. Os autores acreditam que este número é subestimado.

4.3 Localização subcelular das FtsHs p1, p2 e m1 de cana-de-açúcar

De acordo com a teoria endossimbionte, plastídeos e mitocôndria derivam, respectivamente, de antigos eventos simbióticos entre o ancestral das α -proteobactérias e cianobactérias e a célula eucariota primitiva.

As FtsHs estão distribuídas apenas nos Domínios Bacteria e Eukarya. Além disso, a predição de direcionamento feita pelo programa TargetP (Emanuelsson *et al.*, 2000) sugere a localização subcelular mitocondrial ou plastidial para membros de um mesmo grupo filogenético de FtsHs eucarióticas. Por último, a análise das relações filogenéticas entre FtsHs eucarióticas e procarióticas corrobora a origem endossimbiótica para a maioria da FtsHs eucarióticas.

Com base nestes dados é intuitiva a hipótese de que pode-se determinar a localização subcelular de uma FtsHs eucariótica com base no seu perfil filogenético. A confirmação desta hipótese teria uma grande importância prática, pois a localização subcelular é um dos primeiros passos na caracterização funcional de uma proteína.

Para testar esta hipótese foi determinada a localização subcelular de 3 FtsHs de cana-de-açúcar que pertencem ao grupos filogenéticos FtsHp1, FtsHp2 e FtsHm1. Primeiramente a região do cDNA que codifica as seqüências de direcionamento destas proteínas foi fusionadas em fase de leitura com o gene que codifica a GFP (Green Fluorescent Protein). Estas construções gênicas foram expressas transientemente em células da epiderme de cebola e a localização subcelular da expressão transiente da GFP determinada via microscopia de fluorescência.

Os vetores de expressão pCAMBIA-GFP (35S-GFP), β -GFP, GFP e RecA-GFP foram usados como controles citosólico, mitocondrial e plastidial, respectivamente (Figura 21)

A GFP fusionada à seqüência de direcionamento das FtsHs p1 e p2 foi enviada para plastídeos enquanto que a GFP fusionada com a seqüência de direcionamento da FtsH-m1 foi enviada para as mitocôndrias (Figura 22). Esses resultados estão de acordo com a predição de direcionamento destas proteínas baseada no seu perfil filogenético.

Os resultados aqui apresentados indicam que a localização subcelular das FtsHs eucarióticas pode ser determinada com base no seu perfil filogenético. Isto torna-se relevante quando considera-se que após o seqüenciamento do genoma de um organismo aproximadamente de 20 a 30% dos genes descobertos são anotados como gene prováveis ou gene hipotéticos. Ou seja, naquela região do genoma pode estar codificada a informação para a síntese uma proteína ainda desconhecida em todos os aspectos, estrutural, funcional. Este foi o caso de 9 dos 11 genes que codificam FtsHs no genoma de *Arabidopsis thaliana*.

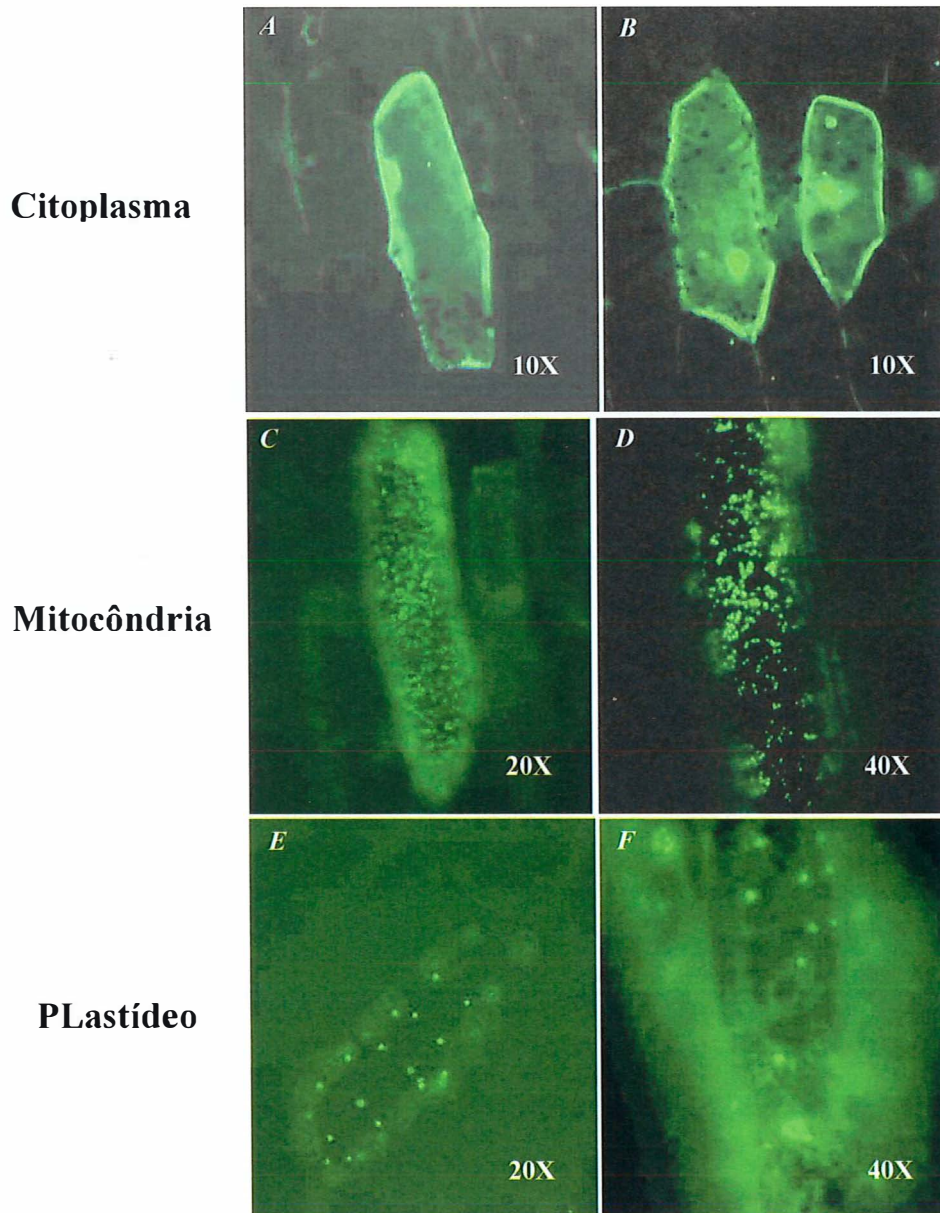


Figura 21 - Expressão transiente em células do epitélio adaxial de cebola das fusões de GFP utilizadas como controle. A e B) Citoplasma: pCAMBIA-GFP (35S-GFP). C e D) Mitocôndria: β -GFP - gentilmente cedida pelo Prof. Marc Boutry. E e F) Cloroplasto: RecA-GFP (Akashi et al., 1998). A fluorescência da GFP foi observada em filtro FITC (excitação 470 nm e emissão 520 nm)

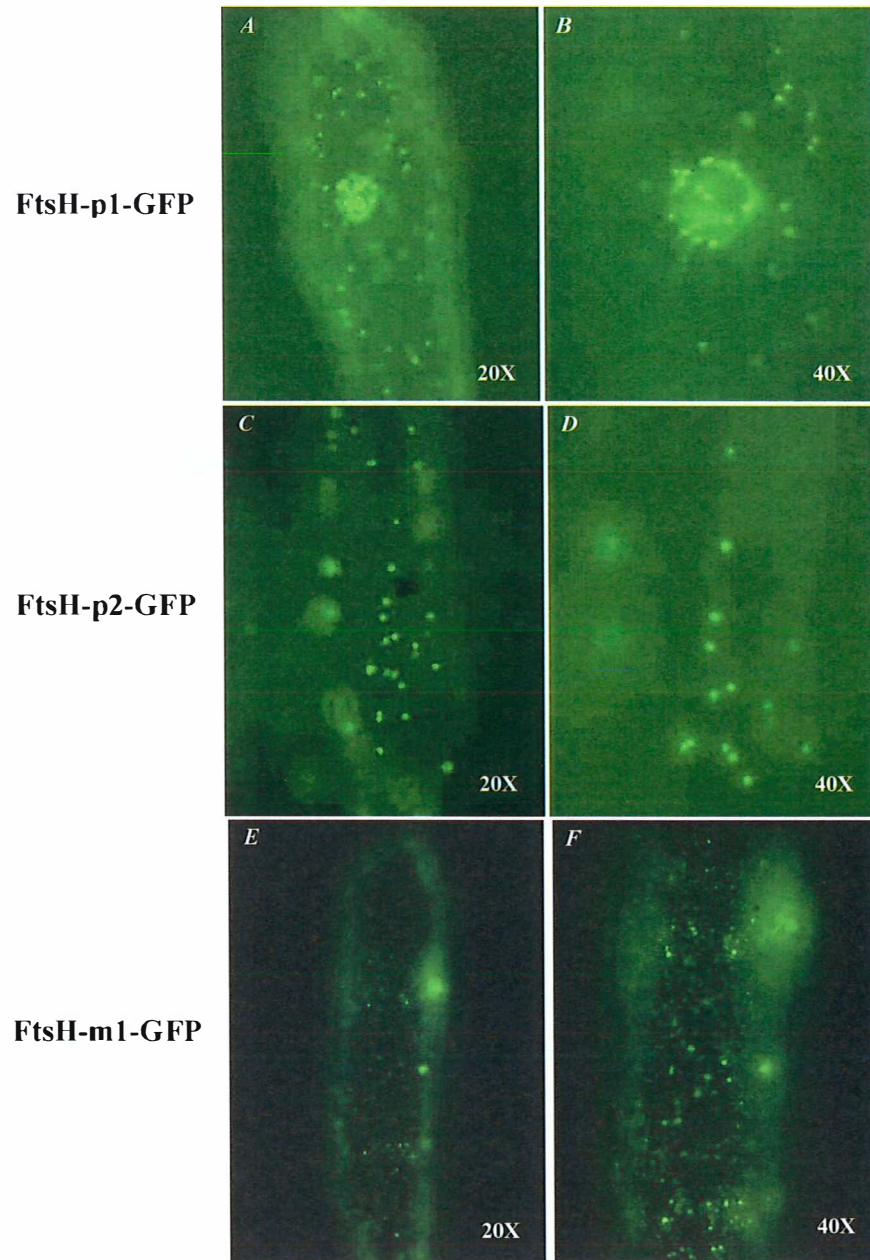


Figura 22 - Expressão transiente em células do epitélio adaxial de cebola das construções gênica envolvendo a seqüência de direcionamento das FtsHs p1, p2 e m1 e GFP . A e B) FtsH-p1-GFP. C e D) FtsH-p2-GFP. E e F) FtsH-m1-GFP. A fluorescência da GFP foi observada em filtro FITC (excitação 470 nm e emissão 520 nm)

Os resultados da localização subcelular das FtsHs indicam que a predição da localização subcelular de FtsH eucarióticas pode ser feita com base no seu perfil filogenético. Recentemente, Sakamoto et al. (2003) determinaram a localização subcelular das FtsHs de *Arabidopsis thaliana*. Das 11 FtsHs cuja localização subcelular foi determinada 10 estão de acordo com o seu perfil filogenético. A única exceção é a FtsHm1B que foi direcionada para os plastídeos. Este resultado pode ser devido ao uso de um sistema heterólogo de expressão gênica para a determinação da localização subcelular.

Como discutido anteriormente, esta FtsH é típica de plantas. A duplicação que originou esta isoforma ocorreu na linhagem Plantae, antes da divergência entre monocotiledôneas e dicotiledôneas. A confirmação do direcionamento da FtsHm1B para os plastídeos será uma grande novidade, tanto do ponto de vista evolutivo como funcional. A atividade proteolítica das FtsHs geralmente está associada a funções regulatórias bem como o controle de qualidade de complexos protéicos de membranas. Em *Arabidopsis thaliana* 8 FtsHs estão localizadas nos plastídeos. O Entendimento das interações destas isoformas será um dos grandes desafios no estudo funcional destas proteínas em plantas superiores.

A análise do perfil funcional das FtsHs procarióticas e eucarióticas revela aspectos importantes. Estas proteínas estão usualmente relacionadas com o controle de qualidade e resposta a estresses. A complementação parcial de *E. coli* com a FtsH de *Lactococcus lactis* e de levedura com a FtsH de *Homo sapiens* e *pisum sativum* revela a existência de relações funcionais entre FtsHs de bactérias gram negativas e gram positivas bem como entre as FtsHs de levedura e *Homo sapiens* e *Pisum sativum* (Nilsson et al., 1994; Kolodziejczak et al., 2002).

Estas complementações revelam a existência de funções constitutivas que ainda são desempenhadas pela FtsH nas espécies usadas no experimentos. Isto sugere a existência de mecanismos de ação, substratos ou mesmo fatores moduladores conservados para a FtsH nestas linhagens

celulares distintas. Estas linhagens divergiram há centenas de milhões de anos e são morfológica e fisiologicamente distintas. O fato das complementações serem parcial sugere que as FtsHs podem estar envolvidas em processos celulares específicos para cada linhagem, não essenciais para a viabilidade da célula. O que não é surpreendente, em razão do tempo de divergência entre as linhagens em questão.

Pode-se especular que estas funções constitutivas das FtsH sejam as mais primitivas, isto é, eram desempenhadas pela proteína ancestral das FtsHs atuais. A atividade chaperona observada nas FtsHs já foi descrita em outras proteínas AAA. Além disso, ela pode ser desacoplada da atividade proteolítica (Leonhard et al., 1999). Isto sugere que estas funções constitutivas primitivas possam estar relacionadas com sua atividade chaperona. Elas poderiam ainda ser comum entre FtsHs procarióticas e eucarióticas

5 CONCLUSÕES

- A existência de mais de uma isoforma de FtsH por genoma é uma característica de alguns grupos bacterianos;
- Existem três isoformas de FtsHs bacterianas oriundas de eventos de duplicação que antecederam a divergência entre bactérias Gram positivas e Gram negativas;
- As três isoformas de FtsHs bacterianas possuem alto nível de similaridade quanto à estrutura primária e secundária;
- As FtsHs eucarióticas podem ser agrupadas em cinco clados filogenéticos sendo três clados compostos por isoformas plastidiais e dois clados composto por isoformas mitocondriais;
- A origem das FtsHs eucarióticas dos clados FtsHp1, FtsHp2, FtsHp3 e FtsHm1 ocorreu a partir da transferência lateral de genes dos genomas dos endossimbiontes para o núcleo da célula eucariota primitiva;
- As FtsHm2s originaram-se de um evento de transferência horizontal entre uma bactéria da divisão CFB e a célula eucariota primitiva;
- A localização subcelular das FtsHs eucarióticas pode ser predita com base no seu perfil filogenético.

REFERÊNCIAS BIBLIOGRÁFICAS

AKASHI, K.; GRANDJEAN, O.; SMALL, I. Potential dual targeting of an *Arabidopsis* archaeobacterial-like histidyl-tRNA synthase to mitochondria and chloroplasts. **FEBS Letters**, v.431, n.1, p.39-44, 1998.

AKIYAMA, Y.; ITO, K. Roles of multimerization and membrane association in the proteolytic functions of FtsH (HflB). **EMBO Journal**, v. 19, n. 15, p. 3888-3895, 2000.

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**, v. 271, n. 49, p. 31196-31201, 1996.

ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n.17, p. 3389-3402, 1997.

ANDREAS, D.B.; OUELLETTE, B.F.F. **Bioinformatics**: a practical guide to the analysis of genes and proteins. 2. ed. New York: Wiley interscience, 2001. 1.v.

- ARLT, H.; TAUER, R.; FELDMANN, H.; NEUPERT, W.; LANGER, T. The YTA10-12 complex, an AAA protease with chaperone-like activity in the inner membrane of mitochondria. **Cell**, v. 85, n. 6, p. 875-885, 1996.
- ARLT, H.; STEGLICH, G.; PERRYMAN, R.; GUIARD, B.; NEUPERT, W.; LANGER, T. The formation of respiratory chain complexes in mitochondria is under the proteolytic control of the m-AAA protease. **EMBO Journal**, v. 17, n. 16, p. 4837-4847, 1998.
- BEGG, K.J.; TOMOYASU, T.; DONACHIE, W.D.; KHATTAR, M.; NIKI, H.; YAMANAKA, K.; HIRAGA, S.; OGURA, T. *Escherichia coli* mutant Y16 is a double mutant carrying thermosensitive *ftsH* and *ftsI* mutations. **Journal of Bacteriology**, v. 174, n. 7, p. 2416-2417, 1992.
- BEIER D, SPOHN G, RAPPUOLI R, SCARLATO V. Identification and characterization of an operon of *Helicobacter pylori* that is involved in motility and stress adaptation. **Journal of Bacteriology**, v. 179, n. 15, p. 4676-4683, 1997
- BERGER, K.H.; YAFFE, M.P. Prohibitin family members interact genetically with mitochondrial inheritance components in *Saccharomyces cerevisiae*. **Molecular & Cell Biology**, v. 18, n. 7, p. 4043-4052, 1998.
- BEYER, A. Sequence analysis of the AAA protein family. **Protein Science**, v. 6, n. 10, p. 2043-2058, 1997
- BROCHIER, C.; BAPTESTE, E.; MOREIRA, D.; PHILIPPE, H. Eubacterial phylogeny based on translational apparatus proteins. **Trends in Genetics**, v. 18, n. 1, p. 1-5, 2002.

- CAMPBELL, C.L.; TANAKA, N.; WHITE, K.; THORSNESS, P.E. Mitochondrial morphological and functional defects in yeast caused by *yme1* are suppressed by mutation of a 26S protease subunit homologue. **Molecular Biology of the Cell**, v. 5, n. 8, p. 899-905, 1994.
- CASARI, G.; DE FUSCO, M.; CIARMATORI, S.; ZEVIANI, M.; MORA, M.; FERNANDEZ, P.; DE MICHELE, G.; FILLA, A.; COCOZZA, S.; MARCONI, R.; DURR, A.; FONTAINE, B.; BALLABIO, A. Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. **Cell**, v. 93, n. 6, p. 973-983, 1998.
- 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. **Plant Journal**, v. 22, n. 4, p. 303-313, 2000.
- CUFF, J.A.; CLAMP, M.E.; SIDDIQUI, A.S.; FINLAY, M.; BARTON, G.J. Jpred: A Consensus Secondary Structure Prediction Server. **Bioinformatics**, v. 14, n. 10, p. 892-893, 1998.
- DAUBIN, V.; GOUY, M.; PERRIERE, G. Bacterial molecular phylogeny using supertree approach. **Genome Informatics**, v. 12, n. 5, p. 155-164, 2001.
- DEUERLING, E.; MOGK, A.; RICHTER, C.; PURUCKER, M.; SCHUMANN, W. The *ftsH* gene of *Bacillus subtilis* is involved in major cellular processes such as sporulation, stress adaptation and secretion. **Molecular microbiology**, v. 23, n. 5, p. 921-933, 1997
- DI GIULIO, M. The universal ancestor and the ancestor of bacteria were hyperthermophiles. **Journal of Molecular Evolution**, v. 57, n. 6, p. 721-730, 2003.

- DOOLITTLE, W.F. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. **Trends in Genetics**, v. 14, n. 8, p. 307-311, 1998.
- DOOLITTLE, W.F. Phylogenetic classification and the universal tree. **Science**, v. 284, n. 5423, p. 2124-2129, 1999.
- EMANUELSSON, O.; NIELSEN, H.; BRUNAK, S.; HEIJNE, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. **Journal of Molecular Biology**, v. 300, n. 4, p. 1005-1016, 2000.
- FERREIRA, L.C.; KECK, W.; BETZNER, A.; SCHWARZ, U. In vivo cell division gene product interactions in *Escherichia coli* K-12. **Journal of Bacteriology**, v. 169, n. 12, p. 5776-5781, 1987.
- FISCHER, B.; RUMMEL, G.; ALDRIDGE, P.; JENAL, U. The FtsH protease is involved in development, stress response and heat shock control in *Caulobacter crescentus*. **Molecular microbiology**, v. 44, n. 2, p. 461-478, 2002.
- GAUTSCH, J.W.; WULFF, D.L. Fine structure mapping, complementation, and physiology of *Escherichia coli* hfl mutants. **Genetics**, v. 77, n. 3, p. 435-448, 1974.
- GE, Z.; TAYLOR, D.E. Sequencing, expression, and genetic characterization of the *Helicobacter pylori* ftsH gene encoding a protein homologous to members of a novel putative ATPase family. **Journal of Bacteriology**, v. 178, n. 21, p. 6151-6157, 1996.

GIGLIONE, C.; SERERO, A.; PIERRE, M.; BOISSON, B.; MEINNEL, T. Identification of eukaryotic peptide deformylases reveals universality of N-terminal protein processing mechanisms. **EMBO Journal**, v. 19, n. 21, p. 5916-5929, 2000.

GRANGER, L.L.; O'HARA, E.B.; WANG, R.F.; MEFFEN, F.V.; ARMSTRONG, K.; YANCEY, S.D.; BABITZKE, P.; KUSHNER, S.R. The *Escherichia coli* mrsC gene is required for cell growth and mRNA decay. **Journal of Bacteriology**, v. 180, n. 7, p. 1920-1928.

GRIBALDO, S.; PHILIPPE, H. Ancient phylogenetic relationships. **Theoretical population biology**, v. 61, n. 4, p. 391-408, 2002

GUPTA, R.S.; GRIFFITHS, E. Critical issues in bacterial phylogeny. **Theoretical population biology**, v. 61, n. 4, p. 423-434, 2002

HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999.

HANEKAMP, T.; THORSNESS, P.E. YNT20, a bypass suppressor of yme1 yme2, encodes a putative 3'-5' exonuclease localized in mitochondria of *Saccharomyces cerevisiae*. **Current Genetics**, v. 34, n. 6, p. 438-448, 1999

HANNENHALLI, S.S.; RUSSELL, R.B. Analysis and prediction of functional subtypes from protein sequence alignments. **Journal of Molecular Biology**, v. 303, n. 1, p. 61-76, 2000.

- HERMAN, C.; THEVENET, D.; D'ARI, R.; BOULOC, P. Degradation of sigma 32, the heat shock regulator in *Escherichia coli*, is governed by HflB. **Proceedings of the National Academy Science of the USA**, v. 92, n. 8, p. 3516-3520, 1995
- HERMAN, C.; OGURA, T.; TOMOYASU, T.; HIGARA, S.; AKIYAMA, Y.; ITO, K.; THOMAS, R.; D'ARI, R.; BOULOC, P. Cell growth and λ phage development controlled by the same essential *Escherichia coli* gene, *fts/hflB*. **Proceedings of the National Academy Science of the USA**, v. 90, n. 22, p. 10861-10865, 1993.
- HOFMANN, K.; STOFFEL, W. TMbase - A database of membrane spanning proteins segments. **Biological Chemistry Hoppe-Seyler**, v. 374, p. 166-172, 1993.
- HUANG, C.Y.; AYLIFFE, M.A.; TIMMIS, J.N. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. **Nature**, v. 422, n. 6927, p. 72-76, 2003.
- HUGUENEY, P.; BOUVIER, F.; BADILLO, A.; D'HARLINGUE, A.; KUNTZ, M.; CAMARA, B. Identification of a plastid protein involved in vesicle fusion and/or membrane protein translocation. **Proceedings of the National Academy Science of the USA**, v. 92, n. 12, p. 5630-5634, 1995.
- ITOH, R.; TAKANO, H.; OHTA, N.; MIYAGISHIMA, S.; KUROIWA, H.; KUROIWA, T. Two *ftsH*-family genes encoded in the nuclear and chloroplast genomes of the primitive red alga *Cyanidioschyzon merolae*. **Plant Molecular Biology**, v. 41, n. 3, p. 321-337, 1999.

JIA, L.; CLEGG, M.T.; JIANG, T. Evolutionary Dynamics of the DNA-Binding Domains in Putative R2R3-MYB Genes Identified from Rice Subspecies *indica* and *japonica* Genomes. **Plant Physiology**, v. 134, n. 2, p. 575-585, 2004.

KARATA, K.; INAGAWA, T.; WILKINSON, A.J.; TATSUTA, T.; OGURA T. Dissenting the role of a conserved motif (the second region of homology) in the AAA family of ATPases. **The Journal of Biological Chemistry**, v. 274, n. 37, p. 26225-26232, 1999.

KARPLUS, K.; BARRETT, C.; HUGHEY, R. Hidden Markov models for detecting remote protein homologies. **Bioinformatics**, v. 14, n. 10, p. 846-856, 1998

KIHARA, A.; AKIYAMA, Y.; ITO, K. A protease complex in the *Escherichia coli* plasma membrane: HfIKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. **EMBO Journal**, v. 15, n. 22, p. 6122-6131, 1996.

KIHARA, A.; AKIYAMA Y.; ITO K. Revisiting the lysogenization control of bacteriophage lambda. Identification and characterization of a new host component, HflD. **The Journal of Biological Chemistry**, v. 276, n. 17, p. 13695-13700, 2001.

KLANNER, C.; PROKISCH, H.; LANGER, T. MAP-1 and IAP-1, two novel AAA proteases with catalytic sites on opposite membrane surfaces in mitochondrial inner membrane of *Neurospora crassa*, **Molecular Biology of the Cell**, v. 12, n. 9, p. 2858-2869, 2001

KOŁODZIEJCZAK, M.; KOLACZKOWSKA, A.; SZCZESNY, B.; URANTOWKA, A.; KNORPP, C.; KIELECZAWA, J.; JANSKA, H. A higher plant mitochondrial homologue of the yeast m-AAA protease. Molecular cloning, localization, and putative function. **The Journal of Biological Chemistry**, v. 277, n. 46, p. 43792-43798, 2002

KOMINSKY, D.J.; BROWNSON, M.P.; UPDIKE, D.L.; THORSNESS, P.E. Genetic and biochemical basis for viability of yeast lacking mitochondrial genomes. **Genetics**, v. 162, n. 4, p. 1595-1604, 2002

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**, v. 10, n. 8, p. 1073-83, 2002.

LANGER, T. AAA proteases: Cellular machines for degrading membrane protein. **Trends in Biochemical Sciences**, v. 25, n. 5, p. 247-251, 2000.

LEIPE, D.D.; WOLF, Y.I.; KOONIN, E.V.; ARAVIND, L. Classification and evolution of P-loop GTPases and related ATPases. **Journal of Molecular Biology**, v. 317, n. 1, p. 41-72, 2002.

LEONHARD, K.; STIEGLER, A.; NEUPERT, W.; LANGER T. Chaperone-like activity of the AAA domain of the yeast Yme1 AAA protease. **Nature**, v. 398, n. 6725, p. 348-351, 1999.

LEONHARD, K.; HERRMANN, 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. **EMBO Journal**, v. 15, n. 16, p. 4218-4229, 1996.

- 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 of Biological Chemistry**, v. 271, n. 46, p. 29329-29334, 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**, v. 12, n. 3, p. 419-431, 2000.
- LOLKEMA, J.S.; SLOTBOOM, D.J. Classification of 29 families of secondary transport proteins into a single structural class using hydrophathy profile analysis. **Journal of Molecular Biology**, v. 327, n. 5, p. 901-909, 2003.
- LOPEZ-GARCIA, P.; MOREIRA, D. Metabolic symbiosis at the origin of eukaryotes. **Trends in Biochemical Science**, v. 24, n. 3, p. 88-93, 1999.
- LUPAS, A. N.; MARTIN, J. AAA proteins. **Current Opinion in Structural Biology**, v. 12, n. 6, p. 746-753, 2002
- MCCLUNG, J.K.; DANNER, D.B.; STEWART, D.A.; SMITH, J.R.; SCHNEIDER, E.L.; LUMPKIN, C.K.; DELL'ORCO, R.T.; NUELLE, M.J. Isolation of a cDNA that hybrid selects antiproliferative mRNA from rat liver. **Biochemical and Biophysical Research Communications**, v. 312, n. 2, p. 355-359, 2003.
- MANN, N.H.; NOVAC, N.; MULLINEAUX, C.W.; NEWMAN, J.; BAILEY, S.; ROBINSON, C. Involvement of an FtsH homologue in the assembly of functional photosystem I in the cyanobacterium *Synechocystis* sp. PCC 6803. **FEBS Letters**, v. 479, n. 1, p. 72-77, 2000

MARBACH, P.A.S.; COELHO, A.S.G.; SILVA-FILHO, M.C. Mitochondrial and chloroplast localization of FtsH-like proteins in sugarcane based on their phylogenetic profile. **Genetics and Molecular Biology**, v. 24, n. 1, p. 183-190, 2001.

MARCOTTE, E.M.; XENARIOS, I.; BLIEK, A.M; EISENBERG, D. Localizing proteins in the cell from their phylogenetic profiles. **Proceedings of the National Academy Science of the USA**, v. 97, n. 22, p. 12115-12120, 2000.

MARGULIS, L. **Origin of eukaryotic cells**. New Haven: Yale University Press, 1970. v 1.

MELCHERS, K.; WIEGERT, T.; BUHMANN, A.; POSTIUS, S.; SCHAFER, K.P.; SCHUMANN, W. The *Helicobacter felis* ftsH gene encoding an ATP-dependent metalloprotease can replace the *Escherichia coli* homologue for growth and phage lambda lysogenization. **Archives of Microbiology**, v. 169, n. 5, p. 393-396, 1998

NAKAI, T.; YASUHARA, T.; FUJIK, Y.; OHASHI, A. Multiple genes including a member of the AAA family, are essential for degradation of unassembled subunit 2 of Cytochrome c oxidase in yeast mitochondria. **Molecular and Cell Biology**, v.15, n. 8, p. 4441-4452, 1995.

NARBERHAUS, F.; URECH, C.; HENNECKE, H. Characterization of the *Bradyrhizobium japonicum* ftsH gene and its product. **Journal of Bacteriology**, v. 181, n. 23, p. 7394-7397, 1999.

NEUWALD, A.F.; ARAVIND, L.; SPOUGE, J.L.; KOONIN, E.V. AAA⁺: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. **Genome Research**, v. 9, n. 1, p. 27-43, 1999.

NILSSON, D.; LAURIDSEN, A.A.; TOMAYASU, T.; OGURA, T. A *Lactococcus lactis* gene encodes a membrane protein with putative ATPase activity that is homologous to the essential *Escherichia coli* *ftsH* gene product. **Microbiology**, v. 140, n. 10, p.2601-2610, 1994.

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**, v. 142, n. 2, p. 279-282, 1991.

OPPENHEIM, A.; HONIGMAN, A.; OPPENHEIM, A.B. Interference with phage lambda cro gene function by a colicin-tolerant *Escherichia coli* mutant. **Virology**, v. 61, n. 1, p. 1-10, 1974.

PAGE, R.D.M. TreeView: An application to display phylogenetic trees on personal computers. **Bioinformatics**, v. 12, n. 4, p. 357-358, 1996.

PATEL, S; LATTERICH, M. The AAA team: related ATPases with diverse functions. **Trends in Cell Biology**, v. 8, n. 2, p. 65-71, 1998.

PAUL, M.F.; TZAGOLOFF, A. Mutations in *RCA1* and *AFG3* inhibit F₁-ATPase assembly in *Saccharomyces cerevisiae*. **FEBS Letters**, v. 373, n. 1, p. 66-70, 1995.

PHILIPPE, H.; FORTERRE, P. The rooting of the universal tree of life is not reliable. **Journal of Molecular Evolution**, v. 49, n. 4, p. 509-523, 1999

PRZYTYCKA, T.; AURORA, R.; ROSE, G.D. A protein taxonomy based on secondary structure. **Nature Structural & Molecular Biology**, v. 6, n. 7, p. 672-682, 1999.

QU, J.N.; MAKINO, S.I.; ADACHI, H.; KOYAMA, Y.; AKIYAMA, Y.; ITO, K.; TOMOYASU, T.; OGURA, T.; MATSUZAWA, H. The *tolZ* gene of *Escherichia coli* is identified as the *ftsH* gene. **Journal of Bacteriology**, v. 178, n. 12, p. 3457-3461, 1996.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**. v. 4, n. 4, p. 406-425, 1987.

SALI, A.; BLUNDELL, T.L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v. 234, n. 3, p. 779-815, 1993.

SANTOS, D.; DE ALMEIDA, D.F. Isolation and characterization of a new temperature-sensitive cell division mutant of *Escherichia coli* K-12. **Journal of Bacteriology**, v. 124, n. 3, p. 1502-1507, 1975.

SANZ, M.A.; MADAN, V.; CARRASCO, L.; NIEVA, J.L. Interfacial domains in Sindbis virus 6K protein. Detection and functional characterization. **The Journal of Biological Chemistry**, v. 278, n. 3, p. 2051-2057, 2003

- SAKAMOTO, W.; TAMURA, T.; HANBA-TOMITA, Y.; MURATA, M.; SODMERGEN. The VAR1 locus of Arabidopsis encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. **Genes to Cells**, v. 7, n. 8, p. 769-780, 2002.
- SCHUMANN, W. FtsH-a single-chain charonin? **FEMS Microbiology Reviews**, v. 23, n. 1, p. 1-11, 1999.
- SCHNALL, R.; MANNHAUPT, G.; STUCKA, R.; TAUER, R.; EHNLE, S.; SCHWARZLOSE, C.; VETTER, I.; FELDMANN, H. Identification of a set of yeast genes coding for a novel family of putative ATPases with high similarity to constituents of the 26S protease complex. **Yeast**, v. 10, n. 9, p. 1141-1155, 1994
- SEO, S.; OKAMOTO, M.; IWAI, T.; IWANO, M.; FUKUI, K.; ISOGAI, A. NAKAJIMA, N.; OHASHI, Y. Reduced level of chloroplast FtsH protein in tobacco mosaic virus-infected tobacco leaves accelerate the hypersensitive reaction. **The Plant Cell**, v. 12, n. 6, p. 917-932, 2000.
- SHAH, Z.H.; HAKKAART, G.A.J.; ARKU, B.; DE JONG, L.; SPEK, H.V.D.; GRIVELL, L.A.; JACOBS, H.T. The human homologue of the yeast mitochondrial AAA metalloprotease Yme1p complements a yeast *yme1* disruptant. **FEBS Letters**, v. 478, n. 3, p. 267-270, 2000.
- SHOTLAND, Y.; TEFF, D.; KOBY, S.; KOBILER, O.; OPPENHEIM, A.B.. Characterization of a conserved α -helical, coiled-coil motif at the C-terminal domain of the ATP-dependent FtsH (HfIB) protease of *Escherichia coli*. **Journal of Molecular Biology**, v. 299, n. 4, p. 953-964, 2000.

- SILVA, P.; CHOI, Y.J.; HASSAN, H.A.; NIXON, P.J. Involvement of the HtrA family of proteases in the protection of the cyanobacterium *Synechocystis* PCC 6803 from light stress and in the repair of photosystem II. **Philosophical Transactions: Biological Sciences**, v. 357, n. 1426, p. 1461-1467, 2002
- SIMILLION, C.; VANDEPOELE, K.; VAN MONTAGU, M.C.; ZABEAU, M.; VAN DE PEER, Y. The hidden duplication past of *Arabidopsis thaliana*. **Proceedings of the National Academy Science of the USA**, v. 99, n. 21, p. 13627-13632, 2002
- SPETEA, C.; HUNDAL, T.; LOHMANN, F.; ANDERSSON, B. GTP bound to chloroplast thylakoid membranes is required for light-induced multienzyme degradation of the photosystem II D1 protein. **Proceedings of the National Academy Science of the USA**, v. 96, n. 11, p. 6547-6552, 1999.
- STEGEMANN, S.; HARTMANN, S.; RUF, S.; BOCK, R. High-frequency gene transfer from the chloroplast genome to the nucleus. **Proceedings of the National Academy Science of the USA**, v. 100, n. 15, p. 8828-8833, 2003.
- SWAFFIELD, J.C.; PURUGGANAN, M.D. The evolution of the conserved ATPase domain (CAD): reconstructing the history of an ancient protein module. **Journal of Molecular Evolution**, v. 45, n. 5, p. 549-563, 1997.
- TAKAHASHI, A.; KAWASAKI, T.; WONG, H.L.; SUHARSONO, U.; HIRANO, H.; SHIMAMOTO, K. Hyperphosphorylation of a mitochondrial protein, prohibitin, is induced by calyculin A in a rice lesion-mimic mutant *cdr1*. **Plant Physiology**, v. 132, n. 4, p. 1861-1869, 2003.

- TEFF, D.; KOBAYASHI, S.; SHOTLAND, Y.; OGURA, T.; OPPENHEIM, A.B. A colicin-tolerant *Escherichia coli* mutant that confers Hfl phenotype carries two mutations in the region coding for the C-terminal domain of FtsH (HflB). **FEMS Microbiology Letters**, v. 183, p. 115-117, 2000.
- THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.G. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, v. 25, n. 24, p. 4876-4882, 1997.
- THORSNESS, P.E.; WHITE, K.H.; FOX, T.D. Inactivation of *YME1*, a member of the ftsH-sEC18-PAS1-CDC48 family of putative ATPase-encoding genes, causes increased escape of DNA from mitochondria in *Saccharomyces cerevisiae*. **Molecular and Cellular Biology**, v. 13, n. 9, p. 5418-5426, 1993.
- 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**, v. 175, n. 5, p. 1344-1351, 1993a.
- TOMOYASU, T.; YAMANAKA, K.; MURATA, K.; SUZAKI, T.; BOULOC, P.; KATO, A.; NIKI, H.; HIRAGA, S.; OGURA, T. Topology and subcellular localization of FtsH protein in *Escherichia coli*. **Journal of Bacteriology**, v. 175, n. 5, p. 1352-1357, 1993b.

TOMOYASU, T.; GAMER, J.; BUKAU, B.; KANEMORI, M.; MORI, H.; RUTMAN, A.J.; OPPENHEIM, A.B.; YURA, T.; YAMANAKA, K.; NIKI, H.; HIRAGA, S.; OGURA, T. *Eschechia coli* FtsH is a membrane-bound ATP-dependent protease which degrades the heat-shock transcription factor σ^{32} . **EMBO Journal**, v. 14, n.11, p.2251-2260, 1995.

URECH, C.; KOPY, S.; OPPENHEIM, A.B.; MUNCHBACH, M.; HENNECKE, H.; NARBERHAUS, F. Differential degradation of *Escherichia coli* sigma32 and *Bradyrhizobium japonicum* RpoH factors by the FtsH protease. **European Journal of Biochemistry**, v. 267, n. 15, p. 4831-4839, 2000.

WANG, R.F.; O'HARA, E.B.; ALDEA, M.; BARGMANN, C.I.; GROMLEY, H.; KUSHNER, S.R. *Escherichia coli* mrsC is an allele of hflB, encoding a membrane-associated ATPase and protease that is required for mRNA decay. **Journal of Bacteriology**, v. 180, n. 7, p. 1929-1938, 1998.

WOESE, C.R. Bacterial evolution. **Microbiology Reviews**, v. 51, n. 2, p. 221-271, 1987.

APÉNDICE 1

Tabela 2 Lista de seqüências de FtsHs eucarióticas usadas nas análises filogenéticas

Espécies	Número de acesso	Predição da localização subcelular pelo programa TargeP
<i>Arabidopsis thaliana</i>	AAD50055	Plastídeos
<i>Arabidopsis thaliana</i>	BAB10200	Plastídeos
<i>Arabidopsis thaliana</i>	CAB89335	Plastídeos
<i>Arabidopsis thaliana</i>	AAF24819	Plastídeos
<i>Arabidopsis thaliana</i>	AAF65925	Plastídeos
<i>Arabidopsis thaliana</i>	CAB61952	Plastídeos
<i>Arabidopsis thaliana</i>	AAD30220	Plastídeos
<i>Arabidopsis thaliana</i>	BAB09632	Plastídeos
<i>Arabidopsis thaliana</i>	AAC31223	Mitocôndria
<i>Arabidopsis thaliana</i>	AAC33234	Mitocôndria
<i>Arabidopsis thaliana</i>	AAF79577	Mitocôndria
<i>Arabidopsis thaliana</i>	BAB08420	Mitocôndria
<i>Capsicum annuum</i>	CAA62084	Plastídeos
<i>Capsicum annuum</i>	CAA09935	Plastídeos
<i>Capsicum annuum</i>	X80755	Plastídeos
<i>Nicotiana tabacum</i>	BAA33755	Plastídeos
<i>Nicotiana tabacum</i>	AAD17230	Plastídeos
<i>Porphyra purpurea</i>	AAC08213	<i>Codificado no genoma plastidial</i>
<i>Guillardia theta</i>	AAC35738	<i>Codificado no genoma plastidial</i>
<i>Cyanidioschyzon merolae</i>	BAA88165	<i>Codificado no genoma plastidial</i>
<i>Cyanidium caldarium</i>	AAB82667	<i>Codificado no genoma plastidial</i>
<i>Mesostigma viride</i>	AAF43852	<i>Codificado no genoma plastidial</i>
<i>Cyanidioschyzon merolae</i>	BAA88164	Mitocôndria
<i>Saccharomyces cerevisiae</i>	CAA56953	Mitocôndria
<i>Saccharomyces cerevisiae</i>	CAA56955	Mitocôndria
<i>Saccharomyces cerevisiae</i>	AAA02883	Mitocôndria
<i>Caenorhabditis elegans</i>	AAF60660	Mitocôndria
<i>Caenorhabditis elegans</i>	CAA88955	Mitocôndria
<i>Drosophila melanogaster</i>	AAF46922	Mitocôndria
<i>Drosophila melanogaster</i>	AAF45806	Mitocôndria
<i>Drosophila melanogaster</i>	AAF49365	Mitocôndria
<i>Mus musculus</i>	AAC35558	Mitocôndria
<i>Homo sapiens</i>	CAB48398	Mitocôndria
<i>Homo sapiens</i>	CAB51858	Mitocôndria
<i>Homo sapiens</i>	CAA76314	Mitocôndria