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

**Caracterização do papel da glutamil-tRNA sintetase na
localização subcelular de proteínas**

Luíza Lane de Barros Dantas

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

**Piracicaba
2010**

Luíza Lane de Barros Dantas
Bacharel em Ciências Biológicas

Caracterização do papel da glutamil-tRNA sintetase na localização subcelular de proteínas

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

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

Piracicaba
2010

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

Dantas, Luíza Lane de Barros

Caracterização do papel da glutamil-tRNA sintetase na localização subcelular de proteínas /
Luíza Lane de Barros Dantas. - - Piracicaba, 2010.
112 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2010.
Bibliografia.

1. Aminoacil tRNA sintetase 2. Glutamina 3. Leveduras 4. Proteínas - Localização I. Título

CDD 581.19245
D192c

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

Dedico a *mainha* e *painho*,
pelo apoio incondicional e
por esse amor que não
conhece fronteiras

AGRADECIMENTOS

Agradeço ao meu orientador, Marcio de Castro Silva Filho, pela oportunidade, pela confiança, pela paciência budista, pela liberdade de horários, pela porta sempre aberta e por ter me mostrado o que é *high science*. Termino esta etapa com mais maturidade científica graças à sua influência. É impossível conviver com você e não se empolgar com a ciência. Meus mais sinceros agradecimento, respeito e admiração.

Ao professor Daniel Scherer de Moura pelas consultas, pela disponibilidade em ajudar, por todos os conselhos e pelas críticas (até as mais severas). Você não tem idéia do impacto (positivo, claro) de cada uma das nossas conversas e reuniões sobre mim. Foi um privilégio poder contar com você.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, pelo suporte financeiro que permitiu dedicação exclusiva a este trabalho.

À querida amiga e companheira de trabalho Marcela Scarso. Obrigada pela ajuda, pela troca tão valiosa de informações, pela cumplicidade, pelo ombro amigo, pelo carinho, pela disposição e por me fazer acreditar nos piores momentos que tudo iria funcionar. Você foi essencial para este trabalho. Obrigada, Marcela!

A Marcelo Brandão, pela parceira, pelo incentivo, pela paciência e por ter me mostrado o poder da bioinformática. Agradeço também pela ajuda valiosa com as análises filogenéticas e pela leitura crítica deste trabalho. Sem isto, o trabalho não teria alcançado o mesmo êxito. Obrigada por tudo isso e pela amizade, Marcelo. Conte comigo sempre.

À maravilhosa e amiga equipe do Laboratório de Biologia Molecular de Plantas, Ligia, Fabiana, Ricardo, Larissa, Eveline, Isabela, Susan, e aos amigos da equipe de pesquisa de peptídios hormonais, Wellington, Juliana e Celso, pela troca de informações e experiência, pela ajuda e por compartilhar os melhores e os piores momentos comigo. Vocês são incríveis!!! Ao querido Rafael Colombi, pela dedicação, ajuda e disponibilidade dentro do laboratório.

A Juliano Guerreiro, pela leitura crítica deste trabalho, que muito ajudou a melhorá-lo, e pela disponibilidade em ouvir e aconselhar.

A Christine Stock, pela ajuda fundamental nos ensaios de duplo-híbrido, pela acolhida, pela amizade, por toda a disponibilidade em ensinar e ajudar.

A Mariana Barros, minha irmã, amiga, cúmplice e a melhor técnica em imagens de que se tem notícia. Obrigada por todo o apoio e pela ajuda em obter o melhor de cada uma das imagens que fazem parte deste trabalho. Te amo e você sabe disso!

A *painho* e *mainha*, sem dúvida as melhores pessoas do mundo. Obrigada pelo apoio incondicional, pelas palavras certas de estímulo, por vibrar comigo a cada pequena conquista, pelo suporte financeiro, pela coragem (a herdada e a ensinada) e por nunca, jamais, em hipótese alguma ter soltado a minha mão. A vocês, que mesmo longe foram absolutamente essenciais para este trabalho, meu agradecimento e todo o meu amor.

Às minhas amigas, divas e parceiras, Virgínia, Paulinha, pelo amparo e pelo resgate à distância, e Thais, pelo resgate imediato. Vocês são aquela família que eu escolhi e foram fundamentais para meu equilíbrio, minha diversão e para que eu tivesse forças. Não podemos nos perder!

SUMÁRIO

RESUMO.....	09
ABSTRACT.....	11
1 INTRODUÇÃO.....	13
2 REVISÃO BIBLIOGRÁFICA.....	17
2.1 Localização subcelular e duplo direcionamento.....	17
2.2 As Aminoacil-tRNA sintetases.....	25
2.3 A Glutamil-tRNA sintetase (GluRS).....	31
3 MATERIAL E MÉTODOS.....	35
3.1 Experimentos para as construções utilizadas.....	35
3.1.1 Material Biológico.....	35
3.1.2 Extração de RNA total e síntese de cDNA.....	35
3.1.3 Amplificação das sequências codificadoras.....	35
3.1.4 Clonagem nos vetores pGADT7 e pGBKT7 de duplo-híbrido.....	37
3.1.5 Clonagem nos vetores pSPYNE-35S e pSPYCE-35S de BiFC.....	39
3.2 Ensaios de duplo-híbrido em levedura.....	40
3.2.1 Material Biológico.....	40
3.2.2 Construção da biblioteca de cDNA de <i>Arabidopsis thaliana</i>	40
3.2.3 Transformação de levedura para triagem (<i>screening</i>) em ensaios de duplo-híbrido.....	41
3.2.4 Extração de DNA plasmidial de leveduras positivas para a triagem (<i>screening</i>) em duplo-híbrido.....	43
3.2.5 Análise das sequências codificadoras dos genes para as proteínas interagentes na triagem (<i>screening</i>) em duplo-híbrido.....	44
3.2.6 Transformação de levedura para confirmação de interação proteína-proteína detectada pela triagem (<i>screening</i>) em duplo-híbrido.....	44
3.2.7 Transformação de levedura para mapeamento da(s) região(ões) de interação proteína-proteína nos ensaios de duplo-híbrido.....	45
3.3 Confirmação das interações <i>in planta</i> por BiFC.....	45
3.3.1 Transformação de <i>Agrobacterium tumefaciens</i>	45

3.3.2 Material biológico vegetal.....	46
3.3.3 Transformação de plantas de tabaco por agroinfiltração.....	46
3.3.4 Visualização das plantas transformadas.....	47
3.4 Análises filogenéticas da GluRS.....	47
4 RESULTADOS.....	51
4.1 Interações proteína-proteína identificadas na triagem (screening) em ensaios de duplo-híbrido em levedura.....	51
4.2 Confirmação da interação das proteínas em duplo-híbrido em levedura.....	52
4.3 Mapeamento da(s) região(ões) de interação proteína-proteína nos ensaios de duplo-híbrido em levedura.....	54
4.4 Confirmação das interações <i>in planta</i> por BiFC.....	55
4.5 Análises filogenéticas da região N-terminal de interação da GluRS.....	58
5 DISCUSSÃO.....	63
6 CONSIDERAÇÕES FINAIS	69
REFERÊNCIAS.....	71
APÊNDICES.....	91
ANEXOS.....	103

RESUMO

Caracterização do papel da glutamil-tRNA sintetase na localização subcelular de proteínas

Nos organismos eucariotos, aproximadamente 50% das proteínas traduzidas no citoplasma são transportadas para as organelas, onde irão desempenhar suas funções. Com isso, surgiu um intrincado sistema de transporte intracelular de proteínas. Nas plantas, a presença de uma segunda organela endossimbionte, o plastídio, tornou este sistema mais complexo e gerou demanda adicional por transporte. Ainda, grande maioria das proteínas mitocondriais e plastidiais são codificadas por genes nucleares e importadas do citosol. O dogma uma proteína-uma localização foi associado ao conceito de um gene-uma proteína na biologia celular. Entretanto, proteínas individuais podem ter mais de uma função, e mais recentemente, proteínas codificadas por um único gene foram identificadas em mais de um compartimento subcelular, o que deu origem ao conceito de duplo direcionamento (DD). Um exemplo bem estudado de DD vem das proteínas da família das aminoacil-tRNA sintetases (aaRS), que participam da síntese protéica ao acoplar o aminoácido ao seu tRNA cognato. Dentre as aaRSs, a glutamil-tRNA sintetase citosólica (GluRS), através de sua extensão N-terminal, parece estar envolvida com outras funções além da tradução. Em *Arabidopsis thaliana*, há dois genes nucleares que codificam a GluRS, um para uma proteína de duplo direcionamento (DD) e outro para uma proteína citosólica. Resultados recentes em nosso laboratório mostraram que a GluRS citosólica pode estar relacionada ao controle da localização subcelular de proteínas organelares em *Arabidopsis*. Para verificar um eventual papel desta proteína na localização subcelular de outras proteínas, foram realizados ensaios de duplo-híbrido em levedura, os quais mostraram interação entre a GluRS e a glutamina sintetase (GS) de *Arabidopsis thaliana*, proteína de DD para mitocôndrias e cloroplastos. Esta interação foi confirmada *in planta*, sendo a sequência da GluRS responsável pela interação localizada na região N-terminal, do resíduo 207 ao 316. Análises filogenéticas apontam que esta região encontra-se ausente nas bactérias e que originou-se provavelmente em Archea, entre 2,6 e 1,8 bilhões de anos. Além disso, observa-se que esta sequência é conservada em fungos, musgos e plantas vasculares, tendo originado-se em *Arabidopsis* há cerca de 2 bilhões de anos.

Palavras-chave: Localização subcelular; Duplo direcionamento; Glutamil-tRNA sintetase; Glutamina sintetase

ABSTRACT

Characterization of the role of glutamyl-tRNA synthetase in the protein subcellular localization

In eukaryotic organisms, about 50% of cytoplasmic translated proteins are transported to the organelles, where they can play their roles. Thus, a complex system for intracellular transport was established. In plants, the presence of a second endosymbiont organelle, the plastid, turned this system still more intricate and required an additional transport mechanism. Besides, most of organellar proteins are coded by nuclear genes and imported from the cytosol. The one protein-one localization was associated to the idea of one gene-one protein, which has long been established in molecular biology. However, individual proteins can show more than one function, and recently, proteins coded by one single gene were identified in more than one subcellular compartment, which has originated the concept of dual targeting. One of the most studied example of dual targeted proteins is the aminoacyl-tRNA synthetase (aaRS) family, which are related to protein synthesis by attaching the correct amino acid onto the cognate tRNA molecule. Among the aaRSs, cytosolic glutamyl-tRNA synthetase (GluRS), through its N-terminal extension, seems to be involved in other cellular role beyond translation. In *Arabidopsis thaliana*, there are two genes encoding GluRS, one for a dual-targeted protein and other for a cytosolic protein. Recent results in our laboratory showed that GluRS interacts with proteins destined to other organelles, which suggest that this protein might have a role in interfering on protein localization in *Arabidopsis*. In order to gain some information on the role of this protein in subcellular localization, yeast two-hybrid assays were performed. These studies showed the interaction between GluRS and glutamine synthetase (GS), a mitochondrial and chloroplastic dual-targeted protein. This interaction was confirmed *in planta*. In addition, the GluRS sequence associated to protein interaction was localized at its N-terminal portion, between the residues 207 – 316. Phylogenetic analysis revealed that this region is absent in bacteria and it probably arose from Archea between 2.6 and 1.8 billion years ago. Also, this sequence is conserved in fungi, moss and all the green plants investigated. Finally, datation analysis showed that this sequence arose in *Arabidopsis* between 2 and 1.7 billion years ago.

Keywords: Subcellular localization; Dual targeting; Glutamyl-tRNA synthetase; Glutamine synthetase

1 INTRODUÇÃO

O papel desempenhado por uma proteína está relacionado com sua localização celular. Ao longo do tempo, os organismos eucariotos desenvolveram vias complexas de direcionamento de proteínas para seus diversos compartimentos celulares. Em plantas, as estimativas apontam para milhares de proteínas codificadas por genes nucleares que são sintetizadas nos ribossomos e transportadas para mitocôndrias e cloroplastos, onde são ativas (HEAZLEWOOD et al., 2005). Algumas destas proteínas atuam em vias metabólicas específicas a cada uma destas organelas, como fotossíntese e fosforilação oxidativa, de modo que é necessário um sistema de direcionamento preciso. Por outro lado, outras proteínas que participam de processos genéricos, como expressão de genes organelares, replicação do DNA, transcrição e tradução, podem atuar nas duas organelas (MACKENZIE, 2005) e serem duplamente direcionadas a ambas, num mecanismo de transporte conhecido como duplo direcionamento (DD). A importação das proteínas por mitocôndrias e cloroplastos se dá através de sequências de direcionamento reconhecidas por complexos de proteínas receptoras localizados na membrana externa destes compartimentos, respectivamente complexos Tom (*Translocon of the Outer Membrane*), SAM (*Sorting and Assembly Machinery*) e Tim (*Translocon of the Inner Membrane*) e os complexos Toc (*Translocon of the Outer Membrane of Chloroplasts*) e Tic (*Translocon of the Inner Membrane of Chloroplasts*) do cloroplasto são responsáveis pelo transporte da maior parte das proteínas direcionadas a estas organelas (NEUPERT; HERMANN, 2007; HÖRMAN et al., 2007; MILLAR; WHELAN; SMALL, 2006). As proteínas que possuem DD são importadas por vários mecanismos: múltiplos transcritos, processamento (*splicing*) alternativo, início da tradução alternativo, redirecionamento, sequências de direcionamento ambíguas (KARNIELY; PINES, 2005; PEETERS; SMALL, 2001; SILVA-FILHO, 2003).

Um exemplo bem estudado de proteínas com DD vem da família das aminoacil-tRNA sintetases (aaRS) (DUCHÊNE et al., 2005). Estas proteínas têm como função canônica acoplar o tRNA com seu aminoácido cognato durante a síntese protéica. Entretanto, vários estudos atribuem a estas proteínas outras funções, como transcrição, tradução, processamento, apoptose e até inflamação (PARK; EWALT; KIM, 2005).

Estas outras funções são desempenhadas pela associação entre aaRS e outras proteínas, formando complexos conhecidos como complexos multi-aminoacil-tRNA sintesases (MSC). Em um dos casos mais bem estudados de MSCs eucarióticos, as proteínas glutamyl-tRNA sintetase (GluRS), metionil-tRNA sintetase (MetRS) e aminoacil-tRNA sintetase cofator 1 (Arc1p) interagem num complexo que, além de melhorar a ligação do tRNA a cada uma das aaRS (SIMOS et al., 1998; GALANI et al., 2001; SIMADER et al., 2006; GOLINELLI-COHEN; MIRANDE, 2007), também é responsável por regular a localização subcelular de cada um dos componentes associados (GALANI et al., 2001; GOLINELLI-COHEN; MIRANDE, 2007). As proteínas deste complexo interagem através de seus domínios N-terminais (GALANI et al., 2001).

Em *Arabidopsis*, há dois genes nucleares para a GluRS, um deles (At5g64050) codifica uma GluRS de DD de 570 aminoácidos (DUCHÊNE et al., 2005) e o outro (At5g26710) codifica uma GluRS citosólica de 719 aminoácidos (DAY; GOLOVKIN; REDDY, 1998). A GluRS citosólica possui uma extensão N-terminal de 119 aminoácidos não observada na proteína de DD. Resultados obtidos em nosso laboratório com ensaios de duplo-híbrido em levedura revelaram recentemente que a região correspondente a 316 aminoácidos N-terminais da GluRS citosólica de *Arabidopsis* interage com a proteína tiamina monofosfato (TH1), que participa da biossíntese de tiamina e é importada para o cloroplasto (AJJAWI; TSEGAYE; SHINTANE, 2006).

Baseando-se nestas informações, este estudo teve como objetivo principal verificar um possível papel da GluRS de *Arabidopsis thaliana* na regulação da localização subcelular de proteínas, sobretudo aquelas de DD. Mais especificamente, este estudo visa identificar e caracterizar interações entre a GluRS citosólica (At5g26710) e proteínas de DD através de bibliotecas de cDNA de *Arabidopsis thaliana* inseridas no sistema duplo-híbrido de levedura; mapear a(s) região(ões) onde ocorre a interação entre as proteínas através de deleções; confirmar a interação entre as proteínas *in vivo in planta* por microscopia usando complementação fluorescente bimolecular (sistema BiFC); analisar filogeneticamente a região N-terminal da GluRS, esclarecendo sua origem e datação. A análise das interações encontradas entre a GluRS e proteínas de DD juntamente com as análises filogenéticas ajudarão a

compreender o mecanismo que regula a localização subcelular desta classe de proteínas e o possível papel da GluRS neste processo.

2 REVISÃO BIBLIOGRÁFICA

2.1 Localização subcelular e duplo direcionamento

A função de uma proteína e sua localização celular estão intimamente relacionados. Há proteínas cujos genes são nucleares, sintetizadas em ribossomos livres no citoplasma ou associados ao retículo endoplasmático e permanecem localizadas no citoplasma desempenhando suas funções. Outras proteínas podem ser codificadas no núcleo, sintetizadas no citoplasma e transportadas para os diferentes compartimentos celulares. Ainda, há uma classe de proteínas codificadas por genes únicos nucleares capazes de serem transportadas para mais de um compartimento subcelular (SILVA-FILHO, 2003).

Ao longo da evolução, a célula eucariótica dividiu espacialmente em diferentes compartimentos seus principais processos bioquímicos. A célula eucariótica vegetal, por sua vez, compartimentalizou alguns de seus principais eventos metabólicos, fosforilação oxidativa e fotossíntese, após os eventos endossimbiontes. Esta compartimentalização trouxe a vantagem de manter os gradientes de metabólitos, íons, pH e o estado redox em níveis toleráveis para as demais funções celulares, mas também trouxe um problema. A maioria dos genes codificadores de proteínas que atuam nas organelas responsáveis pela fotossíntese (cloroplasto) e fosforilação oxidativa (mitocôndria) ao longo de milhões e milhões de anos foi perdida ou incorporada ao genoma nuclear (algumas poucas proteínas são codificadas por genes organelares). Logo, tornou-se necessário desenvolver um mecanismo de transporte destas proteínas para as organelas (MILLAR; WHELAN; SMALL, 2006).

Figura 1 – Visão geral do duplo direcionamento (DD) de proteínas para mitocôndria e cloroplasto em uma célula de planta. As proteínas (em laranja) são codificadas por genes nucleares, traduzidas no citoplasma e direcionadas às organelas através de sua sequência de direcionamento (em rosa). Proteínas direcionadas apenas ao cloroplasto têm o peptídeo de trânsito (em verde); proteínas direcionadas apenas para mitocôndrias são reconhecidas através de sua pré-sequência (em vermelho). Todas as sequências das proteínas importadas, de DD ou não, são reconhecidas por maquinarias encontradas nas superfícies organelares (modificado de CARRIE; GIRAUD; WHELAN, 2009)

Para os processos metabólicos específicos a cada uma das organelas, tornou-se imprescindível para a célula assegurar grande precisão no direcionamento das proteínas. Assim, há proteínas codificadas no núcleo e direcionadas exclusivamente a cada uma das organelas e há também proteínas codificadas no núcleo e duplamente direcionadas tanto para as mitocôndrias quanto para os cloroplastos, num evento

denominado duplo direcionamento (DD) (Figura 1). Uma alternativa à sobreposição de reações entre mitocôndrias e cloroplastos como explicação para a presença de proteínas de DD é o fato de que, sendo as mitocôndrias presentes na célula vegetal há mais tempo do que os cloroplastos, já havia uma via de importação implementada, de modo que algumas proteínas cloroplastidiais fizeram uso da maquinaria já existente (PEETERS; SMALL, 2001; SILVA-FILHO, 2003; MACKENZIE, 2005).

Em *Arabidopsis*, mais de 1000 proteínas são direcionadas aos cloroplastos, enquanto que mais de 500 proteínas são transportadas para as mitocôndrias (MILLAR; WHELAN; SMALL, 2006). As proteínas de DD para mitocôndrias e cloroplastos compreendem cerca de 50 em *Arabidopsis* (CARRIE; GIRAUD; WHELAN, 2009; MORGANTE et al., 2009). Também há casos de DD de proteínas para cloroplastos e núcleo (SCHWACKE et al., 2007), cloroplastos e peroxissomos (SAPIR-MIR et al., 2008; REUMANN et al., 2007) e, na alga *Chlamydomonas reinhardtii*, para cloroplastos e retículo endoplasmático (LEVITAN et al., 2005). Em plantas, o primeiro caso de uma proteína de DD para mitocôndrias e cloroplastos observado a partir de um sistema *in vivo* foi a glutatona redutase de *Pisium sativum* (ervilha) (CREISSEN et al., 1995).

A informação necessária ao DD encontra-se, em sua maioria, na região N-terminal das proteínas. Como a proteína é transportada tanto para as mitocôndrias quanto para os cloroplastos, é de se esperar que a sequência de DD possua similaridade com aquelas responsáveis pela importação de proteínas para mitocôndrias e cloroplastos (PUJOL; MARECHAL-DROUARD; DUCHÊNE, 2007). No caso de proteínas direcionadas às mitocôndrias, a sequência N-terminal responsável pela importação é chamada de pré-sequência, ao passo que nos cloroplastos estas sequências de direcionamento são conhecidas como peptídios de trânsito. Em todos os casos, estas sequências são pouco conservadas, variam muito em comprimento e interagem com os complexos protéicos ancorados às membranas da organelas para só então as proteínas serem transportadas (BHUSHAN et al., 2006; PUJOL; MARECHAL-DROUARD; DUCHÊNE, 2007).

Geralmente, as pré-sequências variam de 18 a 107 aminoácidos, sendo ricas em serina, arginina e alanina. A sua estrutura pode formar α -hélices anfifílicas positivamente carregadas e ter resíduos de aminoácidos hidrofóbicos interagindo com os receptores

de membranas organelares (MOBERG et al., 2004). Por sua vez, o peptídeo de trânsito não tem uma estrutura precisa em solução aquosa, mas apresenta a habilidade de formar α -hélice num ambiente que mimetiza a membrana plastidial (AMBROGGIO et al., 2007); pode exibir de 13 a 146 aminoácidos e está estruturado em resíduos de aminoácidos hidroxilados (serina e treonina) que podem sofrer fosforilação, de modo a facilitar a interação com proteínas citosólicas para o transporte (SMALL et al., 2004). Em comum, pré-sequências e peptídios de trânsito possuem sua composição estrutural, rica em aminoácidos hidrofóbicos (alanina, leucina, fenilalanina ou valina), aminoácidos hidroxilados (serina e treonina) e aminoácidos carregados positivamente (arginina e lisina) e baixo conteúdo de aminoácidos negativamente carregados (VON HEIJNE; NISHIKAWA, 1989; CLAROS; VICENS, 1997). As diferenças estruturais compreendem a distribuição de cargas positivas e a arginina encontradas na região N-terminal mitocondrial (BHUSHAN et al., 2006; PUJOL; MARECHAL-DROUARD; DUCHÊNE, 2007), ao passo que no peptídeo de trânsito os aminoácidos positivos são abundantes na sua região central (SMALL et al., 2004).

A sequência N-terminal que permite o DD de uma proteína para mitocôndrias e cloroplastos necessita, como já foi dito previamente, guardar semelhança com as pré-sequências e com os peptídios de trânsito. Ela pode ser do tipo conhecido por pré-sequência gêmea ou do tipo pré-sequência ambígua (PEETERS; SAMLL, 2001). A chamada pré-sequência gêmea pode ser definida estruturalmente como um arranjo em *tandem* das sequências de direcionamento mitocondrial e cloroplastidial. Este arranjo geralmente é observado nos transcritos de genes que possuem dois sítios de início de transcrição, dois sítios alternativos de início de tradução, ou sítios de processamento alternativos para permitir a produção de peptídios sinais para mitocôndria e cloroplasto. Por outro lado, a presença simultânea de sequências de direcionamento mitocondriais ou plastidiais fusionadas à proteínas quiméricas mostrou que aquela que é localizada na extremidade N-terminal confere a especificidade de endereçamento da proteína às organelas (de CASTRO SILVA-FILHO et al., 1996).

A pré-sequência ambígua é uma sequência de direcionamento única reconhecida por mitocôndrias e cloroplastos, com diferentes regiões de reconhecimento para as diferentes maquinarias de importação das duas organelas. Ela apresenta resíduos de

aminoácidos carregados positivamente (arginina, mas não tanto quanto é observado na pré-sequência), hidroxilados (serina), poucos aminoácidos ácidos e glicina, além de poder conter alguns resíduos hidrofóbicos (PUJOL; MARECHAL-DROUARD; DUCHÊNE, 2007). Foi desenvolvido recentemente um programa computacional capaz de efetuar a predição de proteínas de DD ambíguas (MITSCHKE et al., 2009). As predições reveladas neste trabalho foram experimentalmente comprovadas e, baseados em ampla varredura nos genomas de plantas, estimaram a existência de cerca de 400 proteínas de DD transportadas para mitocôndrias e cloroplastos.

Proteínas localizadas na membrana externa de mitocôndrias e cloroplastos são desprovidas de sequências de direcionamento. Há ainda situações em que proteínas transportadas para mitocôndrias, importação e tradução ocorrem de forma coordenada e simultânea na superfície mitocondrial. Em levedura, alguns transcritos de proteínas mitocondriais codificados no núcleo vão para a membrana externa da organela através de polissomos ligados à mitocôndria (MACKENZIE, 2005). Estudos realizados em mamíferos sugerem que as mitocôndrias possuem receptores onde os ribossomos podem se ligar (MACKENZIE; PAYNE, 2004). Embora esta segregação de mRNA como uma rota alternativa co-traducional para importação de proteínas organelares tenha sido descrita apenas em fungos e mamíferos, é possível que também exista em células vegetais. Os genes de origem procariótica estão entre aqueles que seguem a rota alternativa de segregação de mRNAs para a mitocôndria, ao passo que os genes de origem eucariótica possuem, em sua grande maioria, a pré-sequência e parecem ser traduzidos nos polissomos citoplasmáticos (MARC et al., 2002).

A função da sequência N-terminal de direcionamento para as organelas é permitir a interação da proteína com os complexos receptores de importação localizados na membrana de mitocôndrias e cloroplastos, de modo a permitir que a proteína seja transportada. Os complexos Tom (*Translocon of the Outer Membrane*), Sam (*Sorting and Assembly Machinery*) e Tim (*Translocon of the Inner Membrane*) das mitocôndrias e os complexos Toc (*Translocon of the Outer Membrane of Chloroplasts*) e Tic (*Translocon of the Inner Membrane of Chloroplasts*) do cloroplasto são responsáveis pelo transporte da maior parte das proteínas direcionadas a estas organelas.

(NEUPERT; HERMANN, 2007; HÖRMAN et al., 2007; MILLAR; WHELAN; SMALL, 2006).

O complexo Tom existe em todos os organismos, entretanto, estudos mostram que, ao passo que alguns componentes do complexo são conservados entre os eucariotos, várias proteínas são específicas (MACASEV et al., 2000; MACASEV et al., 2004). Sendo assim, alguns componentes protéicos do complexo encontrados em animais e fungos não são observados em plantas e vice-versa (PERRY et al., 2008). Nas plantas este complexo compreende sete proteínas (MACASEV et al., 2000; MACASEV et al., 2004; CHAN et al., 2006; CHEW et al., 2004). O complexo Sam compreende três proteínas integrais, que atuam na montagem de proteínas do complexo Tom (CHACINSKA et al., 2010). O complexo Tim de plantas apresenta homologia com o de levedura e consiste em quatro proteínas de membrana integrais (DEKKER et al., 1997; MILISAV et al., 2001; TRUSCOTT et al., 2001). Proteínas dos complexos Tom e Tim podem vir a interagir durante o transporte de proteínas (CHACINSKA et al. 2005; MEINNECKE et al. 2006; MOKRANJAC et al. 2009; TAMURA et al. 2009). Nos cloroplastos, o complexo Toc consiste em três tipos de compartimentos, cada um pertencendo a uma família protéica (ANDRÈS; AGNE; KESSLER, 2010). No complexo Tic, oito subunidades foram descritas (KOVÁCS-BOGDÁN; SOLL; BÖLTER, 2010).

O controle da localização subcelular de uma proteína para mais de um destino celular, o que inclui o DD, envolve diferentes mecanismos: múltiplos transcritos, início de tradução alternativo, processamento alternativo ou modificações pós-traducionais (PEETERS; SMALL, 2001; SILVA-FILHO, 2003; CARRIE; GIRAUD; WHELAN, 2009). O controle transcricional é uma das principais maneiras de controlar a importação, baseando-se na produção de múltiplos transcritos com diferentes regiões 5' ou no processamento alternativo (OBARA; SUMI; FUKUDA, 2002; HAMADA et al., 2002). Outro mecanismo importante envolve eventos pós-transcricionais, sobretudo envolvendo sítios alternativos de início de transcrição de um único transcrito. Neste caso, um único mRNA que contém dois códons de início é traduzido a partir do reconhecimento da maquinaria de tradução de apenas um dos AUG (CHABREGAS et al., 2001, 2003; HASHIMOTO et al., 2002).

Fatores ambientais ou mesmo sinais internos podem interferir na localização da proteína, podendo ser considerado como DD ou redirecionamento (SILVA-FILHO, 2003). Por exemplo, sabe-se que a luz pode regular a localização de proteínas dentro da célula vegetal. A partição nucleocitoplásmica dos fotorreceptores é influenciada pelo ciclo diurno de luz. A localização subcelular dos fitocromos A, B, C, D e E, assim como a fototropina 1, possui uma relação direta com a luz e com o ritmo diurno. Todos os fitocromos são importados para o núcleo e formam estruturas mosqueadas de acordo com as diferentes condições de luz (KIRCHER et al., 2002). Por sua vez, a fototropina 1 localiza-se na membrana plasmática de plântulas etioladas. Na presença de luz azul, uma fração da fototropina 1 é lançada no citoplasma (SAKAMOTO; BRIGGS, 2002). Um exemplo de influência metabólica sobre a localização de proteínas ocorre em leveduras, precisamente com a catalase A (Cct1p). Quando células de levedura crescem na presença de oleatos, a catalase A é direcionada para o peroxissomo. Entretanto, mesmo na ausência de uma pré-sequência, a proteína pode localizar-se na matriz mitocondrial quando as células de levedura crescem em fonte de carbono não-fermentável (PETROVA et al., 2004).

As adaptações biológicas acerca da evolução e a história do próprio processo evolutivo em si do DD só agora começam a emergir. No início dos estudos sobre proteínas com duplo direcionamento, pensou-se ser um evento raro, devido às funções especializadas de ambas as organelas, à especificidade do direcionamento e ao pequeno número de proteínas de DD. Entretanto, nos últimos anos o número de proteínas que apresenta duplo direcionamento aumentou, graças ao crescente interesse no estudo dos processos compartilhados pelas duas organelas, como replicação e reparo, transcrição, tradução e proteólise (MACKENZIE, 2005).

Alguns estudos especulam que as proteínas de DD evoluíram a partir de proteínas direcionadas para um único destino (UEDA et al., 2008). Neste contexto, o ganho do DD seria um mecanismo restritivo comparado ao direcionamento a uma única organela para evitar a perda do direcionamento para uma das organelas (CHEW; WHELAN, 2003; MICHALECKA et al., 2003). O DD de proteínas não aumenta o proteoma celular, pelo contrário, mas ele permite a expansão da(s) função(ões) de uma única proteína, dado que a proteína duplamente direcionada pode atuar em diferentes

processos metabólicos dependendo da organela onde está localizada (CARRIE; GIRAUD; WHELAN, 2009). Estudos mais recentes apontam para o DD como um ganho de função associado a duplicações gênicas ocorridas muito cedo na história evolutiva dos seres vivos (BRANDÃO; SILVA-FILHO, submetido). O DD parece ter surgido antes da divergência que deu origem aos grupos de plantas monocotiledôneas e eudicotiledôneas (UEDA et al., 2008). De fato, proteínas de duplo direcionamento são conservadas em arroz e *Arabidopsis*, sugerindo um padrão evolutivo semelhante para as famílias gênicas que codifiquem proteínas de DD (MORGANTE et al., 2009).

As proteínas de DD até agora encontradas são incluídas na categoria de proteínas de manutenção (*housekeepers*), isto é, relacionadas a processos compartilhados pelas organelas, tais como ciclo celular, síntese de DNA e proteínas (MACKENZIE, 2005). Isto pode ser entendido pelo fato de que foram descritas poucas proteínas com DD e, dentro do grupo que reconhecidamente possui DD, muitas têm função ainda desconhecida quando comparadas àquelas codificadas pelo genoma nuclear (CARRIE; GIRAUD; WHELAN, 2009). Outros estudos dedicados à evolução do DD especulam que este mecanismo pode ser uma maneira de comunicação inter-organelar. Ao direcionar as mesmas proteínas para as duas organelas simultaneamente, é possível que mitocôndrias e cloroplastos possam executar as funções desempenhadas por estas proteínas de maneira coordenada (CARRIE; GIRAUD; WHELAN, 2009; LISTER et al., 2008).

Pouco se sabe também sobre o mecanismo regulatório do DD. Não há evidências experimentais que demonstrem a presença de regulação no recrutamento de proteínas de DD entre as organelas. Existem apenas evidências de que este recrutamento não é um processo passivo. A maioria das proteínas de DD contém na sua sequência de direcionamento um motivo de fosforilação. Este motivo é reconhecido por uma proteína quinase citosólica, que por sua vez atua como um guia e a direciona ao complexo receptor Toc cloroplastidial (WAEGEMANN; SOLL, 1996; MARTIN et al., 2006).

Sem dúvida, uma das maiores questões sobre o estudo do DD de proteínas organelares em plantas é como ocorre a regulação desta rota de repartição. É possível que o mRNA de proteínas de DD seja direcionado para a superfície organelar,

resultando em mudanças na distribuição de proteínas (RUDHE et al., 2004). Sem dúvida desvendar o papel dos fatores citosólicos envolvidos neste mecanismo, sejam eles de natureza protéica ou não, pode explicar como mitocôndrias e cloroplastos conseguem importar a quantidade necessária de proteínas de DD.

2.2 As Aminoacil-tRNA sintetases

Um dos mais interessantes exemplos de DD vem do grupos de proteínas das aminoacil-tRNA sintetases (aaRSs). Nas plantas, a síntese protéica ocorre em três compartimentos subcelulares: citoplasma, mitocôndria e cloroplasto, o que gera demanda pelo transporte subcelular de aaRS para as organelas, já que todos os genes para estas proteínas estão no núcleo. Aparentemente, o duplo direcionamento é bastante comum nesta família protéica em *Arabidopsis*, já que de 23 proteínas deste grupo importadas para as organelas, 16 têm DD (DUCHÊNE et al., 2005). Muito provavelmente, o mesmo acontece em outras espécies vegetais (CREISSEN et al., 1995; ROKOV-PLAVEC et al., 2008).

Evolutivamente, as aaRSs surgiram durante a transição da era do RNA enquanto molécula-chave da vida para a era das proteínas, a fim de estabelecer o código genético ao atuarem num dos processos celulares mais fundamentais, a síntese de proteínas (IBBA; SÖLL, 2004). A tradução do mRNA em proteínas requer a presença de aaRSs, que fornece à cadeia polipeptídica em formação cada um dos 20 aminoácidos na forma de aminoacil-tRNAs específicos (há também aaRSs específicas para cada aminoácido), que nada mais são do que os tRNAs devidamente carregados com seus aminoácidos cognatos.

Cabe a estas proteínas agir como catalisadores na reação de ligação covalente dos aminoácidos aos seus respectivos tRNAs. Esta função é desempenhada através de uma reação de aminoacilação que tem espaço nos domínios de ligação específicos a cada um de seus tRNAs correspondentes (IBBA; SÖLL, 2000) (Figura 2).

Figura 2 – Esquema simplificado de duas reações atribuídas às aaRSs. No primeiro passo da reação de aminoacilação, o aminoácido é carregado com ATP, produzindo aminoaciladenilato. Depois, o aminoácido ativado é transferido para o tRNA, formando aminoacil-tRNA e liberando AMP. As aaRSs também catalizam uma reação química que produz Ap4A (ou diadenosina tetrafosfatase) e AMP ao utilizar ATP ao invés de tRNA (adaptado de PARK; EWALT; KIM, 2005)

Inicialmente, o aminoácido é carregado com ATP, formando o intermediário aminoaciladenilato. Uma vez ativado, o aminoácido é então transferido para a adenosina da extremidade 3' do seu tRNA correspondente. As aaRSs também catalisam uma reação química secundária que produz diadenosina oligofosfatases (ApnA) utilizando ATP ou ADP ao invés de tRNA. Foi sugerido que Ap4A funciona como mensageiro secundário na regulação de diversas funções celulares essenciais, como apoptose e ciclo celular (KISSELEV et al., 1998).

Embora o mecanismo de aminoacilação seja comum a todas as aaRSs, elas estão divididas em duas classes (classe I e classe II), de acordo com sua estrutura e com a sequência de seus sítios catalíticos (CUSACK et al., 1990; ERIANI et al., 1990). Dentro de cada classe existe apenas um domínio, chamado de domínio catalítico, que é conservado entre todos os membros de cada classe. Os outros domínios apresentados por estas proteínas estão envolvidos na ligação ao anticódon do tRNA, estabilização do complexo aaRS-tRNA e na desacetilação de tRNAs carregados incorretamente. Além da divisão em classes, as aaRSs são subdivididas em três

subgrupos (a; b; c), baseado na natureza química dos seus aminoácidos substratos (CUSACK, 1997).

A classe I reúne 11 aaRSs, incluindo MetRS, ValRS, IleRS, CysRS, GluRS, GlnRS, LysRS, ArgRS, TrpRS e TyrRS. Esta classe compreende as aaRSs geralmente monoméricas e formadas por um domínio de ligação ao nucleotídeo do tipo Rossmann com três camadas de folhas- β e α -hélices alternadas ($\alpha/\beta/\alpha$) e um núcleo interior de cerca de cinco folhas- β . Além disso, a classe I contém os motivos altamente conservados KMSKS (Lys-Met-Ser-Lys-Ser) e HIGH (His-Ile-Gly-His). As aaRSs desta classe reconhecem o tRNA a partir do sulco menor do seu braço aceptor, via aminoacilação da adenosina terminal da molécula de tRNA no carbono de posição 2'-OH.

A classe II de aaRSs reúne 10 proteínas: AlaRS, HisRS, ProRS, ThrRS, SerRS, GlyRS, PheRS, AspRS, AsnRS e LysRS, esta última presente nas duas classes. A LysRS da classe I está presente na maioria dos Archea e em alguns representantes de Eubacteria, enquanto a LysRS da classe II está presente nos eucariontes, num pequeno número de Archea e na maioria de Eubacteria (AMBROGELLY; KORENCIC; IBBA, 2002; WOESE et al., 2000). As aaRSs da classe II são geralmente multiméricas e se estruturam numa mistura de folhas- β e α -hélices com um núcleo central de folhas- β antiparalelas flanqueadas por α -hélices. Há três pequenos motivos conservados na classe II de aaRSs, motivos 1, 2 e 3 (ERIANI et al., 1990). O motivo 1 está envolvido na formação de homodímeros, enquanto que os motivos 2 e 3 formam os componentes do sítio ativo. As aaRSs da classe II reconhecem o tRNA a partir do sulco maior do tRNA receptor, carregando no carbono de posição 3'-OH da adenosina terminal.

As aaRSs são proteínas com multidomínios. Dentro de cada classe há apenas um domínio característico, o domínio catalítico, que é conservado entre todos os membros de uma mesma classe (O'DONOGHUE; LUTHEY-SCHULTEN, 2003). Esse domínio contém o sítio ativo que interage com o braço aceptor do tRNA. Outros domínios estão envolvidos na ligação ao anticódon do tRNA, estabilização do complexo aaRS-tRNA e desacetilação dos tRNAs carregados com aminoácidos incorretos (BUECHTER; SCHIMMEL, 1993; SCHIMMEL et al., 1993).

Estas proteínas são encontradas em todos os domínios da vida e foram conservadas entre as espécies ao longo da evolução (WOESE et al., 2000). Mais precisamente, é certo afirmar que a estrutura básica das aaRSs é conservada, até quando há menos de 15% de identidade de sequência verificada. As aaRSs não necessariamente representam o resultado de uma evolução por convergência porque têm a mesma função bioquímica básica e carregam o tRNA com seu aminoácido cognato num processo comum a todas estas proteínas.

As aaRSs encontradas presentes nos organismos vivos são produtos de duplicações gênicas muito antigas. As especificidades de cada aminoacil-tRNA sintetase, como parte do processo de tradução, surgiram antes da raiz da árvore filogenética universal. De fato, estudos evolutivos sobre as aaRSs especulam que estas proteínas podem ter sido os primeiros pioneiros moleculares a cruzar o limiar darwiniano, que consiste numa transição muito antiga do fluxo gênico horizontal, que dominava a dinâmica evolutiva, para uma fase em que a herança vertical tornou-se a força evolutiva dominante, levando à especiação inicial da raiz da árvore filogenética universal (O'DONOGHUE; LUTHEY-SCHULTEN, 2003).

Um novo estudo sobre as aaRSs organelares de *Arabidopsis* mostra que elas são fruto de eventos de transferência horizontal de genes de bactérias, embora não tenha sido indicado um único ancestral para esta transferência, se cianobactérias ou proteobactérias. Ainda, este estudo reforça a idéia de que o DD é resultado de duplicações gênicas e representa um verdadeiro ganho de função (BRANDÃO; SILVA-FILHO, submetido).

É cada vez mais comum encontrar estudos que mostram associações entre as aaRSs e outras proteínas, formando complexos multienzimáticos (HAUSMANN; IBBA, 2008). Embora a maioria destas associações tenha sido descrita inicialmente em organismos eucariotos, vários complexos multienzimáticos (na verdade, complexos bimoleculares em sua maioria) contendo aaRSs foram recentemente identificados em Eubacteria e Archea.

Em Eubacteria, existem complexos multienzimáticos ProRS•Ybak, AspRS•GatCAB amidotransferase e TrpRS•óxido nítrico sintetase voltados para, respectivamente, hidrólise da ligação Cys-tRNA^{Pro} (AN; MUSIER-FORSYTH, 2004;

RUAN; SÖLL, 2005); conversão de Asp-tRNA_{Asn} para Asn-tRNA_{Asn} (BAILLY et al., 2007); nitração de triptofano e produção de 4-nitro-triptofano (BUDDHA ; CRANE, 2005). Já em Archea, um complexo envolvendo ProRS•Mj1338, LysRS e AspRS foi observado, embora sua função ainda não tenha sido desvendada (LIPMAN et al., 2003); outro complexo envolvendo LysRS•LeuRS•ProRS (PRAETORIUS-IBBA et al., 2005; PRAETORIUS-IBBA et al., 2007; HAUSMANN et al., 2007) atua melhorando a aminoacilação destas proteínas.

Nos grupos de eucariotos, entretanto, é que são observados os mais bem estudados casos de complexos multienzimáticos de aaRSs voltados para funções não-canônicas. Nestes organismos, os complexos tendem a ser maiores do que aqueles encontrados em procariotos e também atuam num espectro de funções bem mais abrangente, de aminoacilação a apoptose (HAUSMANN; IBBA, 2008).

Em levedura, o complexo Arc1p•MetRS•GluRS atua definindo a localização subcelular destas proteínas e também melhora a afinidade do tRNA aos seus aminoácidos cognatos, Met e Glu (SIMOS et al., 1996; GALANI et al., 2001; KARANASIOS et al., 2007); o complexo SerRS•Pex21p melhora a afinidade do tRNA à serina (ROCAK et al., 2002; GODINIC et al., 2007); o complexo TyrRS•Knr4 tem função ainda desconhecida, embora estudos sugiram que há uma relação entre o complexo e formação de ditirosina durante a esporulação (DAGKESSAMANSAIA et al., 2001; IVAKHNO; KORNELYUK, 2005).

Em mamíferos, o complexo ValRS•EF-1A melhora a reação de aminoacilação e restringe proteínas associadas no citoplasma (MOTORIN et al., 1988; VENEMA et al., 1991; BEC; KERJAN; WALLER, 1994; NEGRUTSKII et al., 1999). O complexo formado por AspRS•EF-1A atua na liberação de Asp-tRNA_{Asp} (REED; WASTNEY; YANG, 1994; REED; YANG, 1994). O complexo LysRS•AspRS•EF-1A•p38 melhora a atividade de aminoacilação das aaRSs envolvidas (GUZZO; YANG, 2008). O complexo GluRS•ProRS•L13a•GAPDH•NSAP1 atua no silenciamento da tradução (MAZUMDER; SESHADRI; FOX, 2003; SAMPATH et al., 2004; RAY; ARIF; FOX, 2007). Um outro complexo mais recente foi observado entre LysRS e a proteína Gag do vírus HIV em humanos. Este complexo favorece a incorporação de tRNA_{Lys} pelo vírus, assim como

sua montagem (JAVANBAKHT et al., 2003; GUO et al., 2005; KOVALENSKI et al., 2006).

Existe, ainda, em mamíferos, um grande complexo envolvendo nove aaRSs (LysRS, LeuRS, IleRS, GluRS, ProRS, MetRS, GlnRS, ArgRS e AspRS) e três fatores protéicos (p18, p38 e p43) cuja função, embora não totalmente desvendada, aponte para a melhoria da aminoacilação, localização subcelular das proteínas envolvidas e controle de atividades não-canônicas desempenhadas por estas aaRSs (ROBINSON; KERJAN; MIRANDE, 2000).

Embora a atividade catalítica das aaRSs seja sua função celular essencial, vários estudos mostram que estas proteínas são versáteis e multifuncionais. Esta flexibilidade funcional das aaRSs parece estar condicionada às suas interações protéicas, que promovem uma expansão de função para além da síntese de proteínas. Estas interações entre aaRSs e outras aaRSs ou outras proteínas formam os complexos multienzimáticos que estão associados às funções não-canônicas atribuídas às aaRSs, isto é, quando estas proteínas estão interagindo dentro dos complexos, elas atuam na reação de aminoacilação, ao passo que quando elas se encontram livres, podem exercer outras funções que não necessariamente estejam relacionadas à síntese de proteínas. Embora a maioria destas associações tenha sido descrita inicialmente em organismos eucariotos, vários complexos multienzimáticos (na verdade, complexos bimoleculares em sua maioria) contendo aaRSs foram recentemente identificados em Eubacteria e Archea (HAUSMANN; IBBA, 2008).

Assim, as aaRSs, em associação a outras proteínas, também participam de vários outros processos celulares, tais como transcrição de rRNA (MetRS em humanos); regulação da transcrição (AlaRS em *E. Coli* e LysRS em humanos); processamento de íntrons (LeuRS e TyrRS em fungos); controle da tradução (ThrRS em *E. Coli*); silenciamento da tradução (GluRS e ProRS em humanos); apoptose (GlnRS em humanos), angiogênese (TrpRS e TyrRS em humanos); inflamação (HisRS e LysRS em humanos); montagem do capsídeo viral e incorporação de tRNA_{Lys} (LysRS em humanos infectados pelo vírus HIV); transporte extranuclear de tRNA em leveduras e anfíbios; localização subcelular de proteínas (GluRS e MetRS em levedura) (LUND; DAHLBERG, 1998; SIMOS et al., 1998; SARKAR; AZAD; HOPPER, 1999;

GROSSHANS; HURNT; SIMOS, 2000; LEE et al., 2004; PARK; EWALT; KIM, 2005; RYCKELYNK; GIEGE; FRUGIER, 2005; HAUSMANN; IBBA, 2008).

As aaRS eucarióticas se distinguem das procarióticas pela presença de extensões na suas regiões C- ou N-terminais, que não são observadas nas proteínas correspondentes em Archea e Eubactéria (MIRANDE, 1991). Estes apêndices anexos ao centro catalítico de várias aaRSs não representam ou funcionam como domínios catalíticos voltados para a sua função canônica na síntese de proteínas. Ao invés disso, eles atuam precisamente nas interações proteína-proteína ou como domínios de ligação ao RNA (CAHUZAC et al., 2000; GUIGOU; SHALAK; MIRANDE, 2004; ROBINSON; KERJAN; MIRANDE, 2000), os quais são responsáveis pela expansão funcional das aaRSs. É possível afirmar, do ponto de vista evolutivos, que estes apêndices representam um ganho de função para as aaRSs eucarióticas.

Evolutivamente, a associação das aaRSs em complexos multienzimáticos voltados para desempenhar funções não-canônicas destas proteínas pode ser vista como uma estratégia de economia para maximizar a diversidade funcional das proteínas com uma fonte genética limitada. A expansão funcional das aaRSs provavelmente está relacionada a uma necessidade persistente de novas funções no desenvolvimento de sistemas complexos e a distribuição de proteínas em todo e qualquer tipo celular, onde estão expostas e precisam adaptar-se a um arranjo de novas pressões seletivas (PARK; EWALT; KIM, 2005).

2.3 A Glutamil-tRNA sintetase (GluRS)

Dentre as aaRSs, uma delas chama atenção pelas funções não-canônicas, isto é, não relacionadas à síntese de proteínas a ela atribuídas: a glutamil-tRNA sintetase (GluRS). Em humanos, ela atua no silenciamento da tradução (MAZUMDER; SESHADRI; FOX, 2003; SAMPATH et al., 2004; RAY; ARIF; FOX, 2007) e num complexo cujas funções compreendem o controle de localização das aaRSs que fazem parte do complexo (ROBINSON; KERJAN; MIRANDE, 2000). Em leveduras, a GluRS está associada ao controle da localização subcelular das proteínas que fazem parte do complexo (GluRS, MetRS, Arc1p) (GOLINELLI-COHEN; MIRANDE, 2007).

A GluRS é uma aaRS da classe I. Sendo assim, ela apresenta em seu domínio catalítico os motivos consenso da classe I HIGH e KMSKS (ERIANI et al., 1990). A GluRS transfere ácido glutâmico para o tRNA_{Glu} num processo de dois passos. Inicialmente, o aminoácido ácido glutâmico e a molécula de ATP ligam-se à GluRS, no seu sítio ativo, onde eles ficam covalentemente unidos por ligações do tipo éster entre o α -carboxilato do aminoácido e o α -fosfato do ATP. Os produtos desta reação são pirofosfato e o aminoacil-adenilato. No segundo passo, o aminoácido é transferido para o terminal 3'-hidroxil (típico da classe I) do braço acceptor do tRNA_{Glu}, para formar o tRNA aminoacilado Glu-tRNA_{Glu} (FREIST et al., 1997). Em termos evolutivos, sabe-se é que a GluRS de Archea compartilha similaridades estruturais com a GluRS eucariótica e que houve, ao longo dos anos, duplicações gênicas e transferência horizontal (BROWN; DOOLITTLE, 1999).

Além de seu papel na síntese de proteínas e de suas funções não-canônicas associadas aos complexos multienzimáticos dos quais faz parte, a GluRS também atua indiretamente na biossíntese de compostos tetrapirrólicos como clorofila, heme e derivados através da via do C5 encontrada em plantas, Archea e na maior parte das bactérias (BEALE; GOUGH; GRANIC, 1975; SCHÖN et al., 1986; JAHN; VERKAMP; SÖLL, 1992), além de controlar o nível celular de heme (LEVICÁN et al., 2006).

Em Arabidopsis, há dois genes que codificam a GluRS, ambos nucleares. Um deles, At5g64050, codifica uma GluRS que tem duplo direcionamento (DD) para mitocôndria e cloroplasto e apresenta similaridade com a sequência de GluRS de cianobactéria, o que sugere sua origem evolutiva no ancestral plastidial. Já foi observado, inclusive experimentalmente, que o DD da GluRS organelar é conservado entre monocotiledôneas e eudicotiledôneas (MORGANTE et al., 2009)

O outro gene, At5g26710, codifica uma GluRS que é citosólica e apresenta similaridade de sequência com outras enzimas citosólicas. Este gene possui 2928 pares de base e está organizado em 6 éxons separados por 5 íntrons. O seu cDNA possui 2513 pares de bases e a sequência codificadora possui 2160 pares de bases. A proteína tem 719 aminoácidos e peso molecular de aproximadamente 81kDa (DAY; GOLOVKIN; REDDY, 1998). A GluRS citosólica difere da organelar (DUCHÊNE et al., 2005), sobretudo na sua região N-terminal, onde há um ganho de 119 aminoácidos. Em

outras espécies, as extensões N-terminais da GluRS estão associadas a uma expansão funcional da proteína (HAUSMANN; IBBA, 2008).

Há fortes indícios de que há uma relação entre a GluRS citoplasmática e sua região N-terminal e a localização subcelular de outras proteínas. Esta proteína estaria envolvida no endereçamento de outras proteínas através de interações proteína-proteína (SIMOS et al., 1996). Como acontece em levedura, a extensão N-terminal da GluRS forma, junto com a região N-terminal da MetRS e a região N-terminal de Arc1p, o complexo multienzimático Arc1p•MetRS•GluRS (Figura 3). Neste complexo, Arc1p atua como uma molécula que atrai o tRNA, de modo a aumentar a disponibilidade de tRNA para as aaRSs (GALANI et al., 2001; FRECHIN et al., 2010).

Figura 3 – Esquema da estrutura do complexo Arc1p•MetRS•GluRS. O complexo é mantido por interações entre as regiões N-terminais de cada uma das proteínas envolvidas. No centro em vermelho está Arc1p; à esquerda em azul está MetRS (MRS); à direita em amarelo está a GluRS (ERS). As aaRSs foram esquematizadas na reação de aminacilação, transferindo aminoácidos para os tRNAs cognatos (adaptado de FRECHIN et al., 2010)

Mais do que isso, Arc1p, que é um homólogo de levedura da proteína p43, funciona como uma espécie de plataforma de ancoramento tanto para MetRS quanto

para GluRS, isto é, a presença de Arc1p está relacionada com a localização subcelular destas proteínas. Quando o complexo está formado, as três proteínas ficam confinadas ao citoplasma, entretanto, quando complexo encontra-se dissociado, estas proteínas podem ser importadas para diferentes compartimentos celulares e neles desempenhar outras funções (GALANI et al., 2001; GOLINELLI-COHEN; MIRANDE, 2007).

Em leveduras que crescem sob condições normais, quando a fermentação é utilizada para produzir energia, níveis considerados basais de atividade mitocondrial são observados e, assim, a maior parte da GluRS encontra-se no complexo Arc1p•MetRS•GluRS. Entretanto, quando a levedura passa a crescer sob condições de respiração aeróbica, a expressão de Arc1p é reduzida drasticamente, levando à liberação de GluRs para a mitocôndria, onde participa da reação de transaminação para formar glutaminil-tRNGln (FRECHIN et al., 2010).

Além disso, resultados recentes obtidos em nosso laboratório através de ensaios de duplo-híbrido em levedura que utilizaram como proteína “isca” a proteína tiamina monofosfato (TH1) de Arabidopsis, a qual participa da via de biossíntese de tiamina e é importada para o cloroplasto (AJJAWI; TSEGAYE; SHINTANE, 2006), revelaram uma interação com a GluRS citosólica de Arabidopsis (MARCELA E. SCARSO, comunicação pessoal). A análise da região da GluRS que interagiu com a proteína TH1 revelou que trata-se da sequência dos primeiros 950 pares de bases da sequência codificadora do gene da GluRS, o que corresponde à sequência de 316 aminoácidos da região N-terminal desta. Este achado reforça o indício de que não só a GluRS citoplasmática de planta poderia também estar envolvida na localização subcelular de outras proteínas, como também sugere que pode haver a formação de complexos multienzimáticos envolvendo aaRSs em plantas.

Do ponto de vista evolutivo, pouco se sabe sobre a origem da GluRS citosólica de plantas. Não se sabe ao certo sobre quando surgiu e qual a sua ancestralidade, principalmente no que diz respeito à extensão N-terminal, que em outras espécies de eucariotos é a grande responsável pelas interações proteína-proteína que formam os complexos multienzimáticos e que está por trás da expansão funcional adquirida por esta proteína.

3 MATERIAIS E MÉTODOS

3.1 Experimentos para as construções utilizadas

3.1.1 Material Biológico

Sementes de *Arabidopsis thaliana*, ecótipo Columbia, foram semeadas em vasos plásticos contendo iguais proporções de substrato organomineral (Plantamax) e vermiculita e mantidas em câmara de crescimento (Conviron ATC26) sob condições controladas de fotoperíodo (16 horas claro e 8 horas escuro), aproximadamente 300 μ E de intensidade luminosa e temperatura constante (25°C). As plantas foram irrigadas em dias alternados, com solução nutritiva [$\text{Ca}(\text{NO}_3)_2$ 5 mM, KNO_3 5 mM, MgSO_4 2 mM, KH_2PO_4 1 mM, FeEDTA e micronutrientes] e água.

3.1.2 Extração de RNA total e síntese de cDNA

Os tecidos congelados em nitrogênio líquido de *Arabidopsis thaliana* foram macerados e a extração de RNA total foi realizada utilizando o reagente Trizol® (Invitrogen), seguindo instruções do fabricante. As amostras foram tratadas com 1 unidade de DNase (Fermentas, EN0521), para remoção de DNA, por 20 minutos a 37°C e o RNA total foi reextraído com Trizol® para remover traços de DNA e DNase. O RNA total foi quantificado por leitura no espectrofotômetro a 260nm.

A síntese de cDNA de primeira fita foi feita em volume de 2 μ L, contendo 2 μ g de RNA total, 2,5 μ M de poly-(dT) primer e 1 μ L de transcriptase reversa, seguindo as instruções do fabricante (Impron-II™ Reverse Transcriptase, Promega, Madison, WI).

3.1.3 Amplificação das sequências codificadoras

As sequências codificadoras para a glutamil-tRNA sintetase, GluRS, (At5g26710) e para glutamina sintetase, GS, (At5g35630) foram obtidas com o auxílio do banco de dados TAIR (The Arabidopsis Information Resource – www.arabidopsis.org) e então utilizadas para a confecção de iniciadores específicos. Nas extremidades 5' dos iniciadores foram adicionados sítios para enzimas de restrição. Para as clonagens nos vetores de duplo-híbrido, foram utilizadas ClaI e BamHI para a GluRS, tanto sequência completa quanto os 950 pares de base iniciais (a ser clonada em ambos os vetores,

pGADT7 e pGBKT7) e *NdeI* e *EcoRI* para GS (a ser clonada em ambos os vetores, pGADT7 e pGBKT7). Para as clonagens nos vetores de BiFC, foram utilizadas *BamHI* e *Clal* para a GluRS (a ser clonada no vetor pSPYCE-35S) e *Clal* e *XhoI* para a GS (a ser clonada no vetor pSPYNE-35S).

Para todas as clonagens das quatro deleções de GluRS no vetor pGBKT7 de duplo-híbrido, foram utilizadas *BamHI* e *Sall*. Estas deleções estão todas localizadas na sequência que corresponde à região N-terminal da GluRS (Figura 4).

Os iniciadores específicos para as sequências codificadoras destas proteínas foram desenhados manualmente e checados com o auxílio do programa Oligo Calc, (<http://www.basic.northwestern.edu/biotools/oligocalc.html>). As sequências dos iniciadores encontram-se no APÊNDICE 1, para GluRS e GS, e no APÊNDICE 2, para as deleções na região N-terminal de GluRS. A amplificação das sequências codificadoras foi realizada por PCR. Para estas reações utilizou-se a enzima Platinum® Hig Fidelity Taq DNA Polymerase (Invitrogen), seguindo instruções do fornecedor. A reação de PCR continha 5 µL de tampão para PCR 10X; 2 µL de MgSO₄ 50 mM; 1 µL de cada um dos iniciadores (10 µM cada); 2 µL de dNTPs 10 µM; aproximadamente 50 ng de cDNA, 0,5 µL de Platinum® Hig Fidelity Taq DNA Polymerase e água mili-Q até completar o volume final de 50 µL. As condições da PCR foram as seguintes: 94°C por 3 minutos; 94°C por 30 segundos, 50°C por 30 segundos, 72°C por 2,5 minutos, sendo estes três últimos passos repetidos 32 vezes; 72°C por 8 minutos e 8°C até retirar as amostras. As reações de PCR foram cheçadas através de eletroforese em gel de agarose 1% em tampão TBE [tris-fosfato 0,08 M, EDTA 0,002 M].

Figura 4 – Esquema das deleções em cima da região N-terminal da GluRS citosólica (cor preta). As sequências codificadoras para cada uma delas foram clonadas no vetor de duplo-híbrido pGBKT7 obtidas a partir de iniciadores específicos cujas sequências encontram-se disponíveis no APÊNDICE 2

3.1.4 Clonagem nos vetores pGADT7 e pGBKT7 de duplo-híbrido

Os produtos de amplificação de PCR contendo os sítios para enzimas de restrição para clonagem nos vetores de duplo-híbrido pGADT7 e pGBKT7 (Figura 5). Os próprios vetores, assim como os produtos de amplificação das sequências codificadoras, foram digeridos com as enzimas de restrição (*Clal* e *BamHI* para a GluRS completa e para a sequência de 950 pares de base iniciais e *NdeI* e *EcoRI* para GS; *BamHI* e *Sall* para as deleções de GluRS) durante 2 horas a 37°C para adquirirem extremidades coesivas. Posteriormente, foi realizada a ligação com o Rapid DNA Ligation Kit (Fermentas), conforme instruções do fabricante. As reações foram incubadas durante 1 hora a 25°C, depois 12 horas a 16°C e 5 horas a 4°C.

Figura 5 – Esquema dos vetores utilizados no sistema duplo-híbrido de levedura, pGADT7 (esquerda) e pGBKT7 (direita) (adaptado de www.clonetech.com)

Após este período em diferentes temperaturas, 2 μL das reações de ligação foram usados para transformar 70 μL de células de *E. coli* eletrocompetentes, cepa DH5α. As bactérias transformadas foram incubadas a 37°C por 1 hora em meio SOB [20 g/L de triptona, 5 g/L de extrato de levedura, 0,5 g/L de NaCl em pH 7.5] e, em seguida, plaqueadas em LB sólido [10 g/L de triptona, 5 g/L de extrato de levedura, 10 g/L de NaCl, 15 g/L de ágar bacteriológico em pH 7.5] contendo 200 μg/mL para as reações de ligação no pGADT7 e 50 μg/mL de canamicina para as reações de ligação no pGBKT7 e incubadas durante 16 horas a 37°C. As cepas recombinantes e transformadas foram selecionadas e transferidas para o meio LB líquido contendo 200 μg/mL para as reações de ligação no pGADT7 e 50 μg/mL de canamicina para as reações de ligação no pGBKT7 e incubadas por 16 horas a 37°C sob agitação de 200 rpm. A extração do DNA plasmidial foi realizada de acordo com protocolo clássico de minipreparação plasmidial descrito por Sambrook e Russell (SAMBROOK; RUSSELL, 2001). A confirmação das transformações foi feita através de PCR sob as condições previamente descritas, utilizando os iniciadores específicos para cada sequência (APÊNDICE 1 e APÊNDICE 2), e também por digestão com as enzimas de restrição utilizadas para as

reações de ligação previamente descritas. Os plasmídios extraídos foram sequenciados, seguidos por verificação no programa BLAST.

3.1.5 Clonagem nos vetores pSPYNE-35S e pSPYCE-35S de BiFC

Os produtos de amplificação de PCR contendo os sítios para enzimas de restrição para clonagem nos vetores de BiFC pSPYNE-35S e pSPYCE-35S (Figura 6) e os próprios vetores foram digeridos com as enzimas de restrição (*BamHI* e *Clal* para a GluRS e *Clal* e *XhoI* para a GS) durante 2 horas a 37°C para adquirirem extremidades coesivas. Posteriormente, foi realizada a ligação com o Rapid DNA Ligation Kit (Fermentas), conforme instruções do fabricante. As reações foram incubadas durante 1 hora a 25°C, depois 12 horas a 16°C e 5 horas a 4°C.

Após este período em diferentes temperaturas, 2 µL das reações de ligação foram usados para transformar 70 µL de células de *E. coli* eletrocompetentes, cepa DH5α. As bactérias transformadas foram incubadas a 37°C por 1 hora em meio SOC e, em seguida, plaqueadas em LB sólido contendo 50 µg/mL de canamicina e incubadas durante 16 horas a 37°C. As cepas recombinantes e transformadas foram selecionadas e transferidas para o meio LB líquido contendo 50 µg/mL de canamicina e incubadas por 16 horas a 37°C sob agitação de 200 rpm. A extração do DNA plasmidial foi realizada de acordo com protocolo clássico de minipreparação plasmidial descrito em Sambrook e Russell (SAMBROOK; RUSSELL, 2001). A confirmação das transformações foi feita através de PCR sob as condições previamente descritas, utilizando os iniciadores específicos para cada sequência (APÊNDICE 1), e também por digestão com as enzimas de restrição utilizadas para as reações de ligação previamente descritas. Os plasmídios extraídos foram sequenciados, seguidos por verificação no programa BLAST.

Figura 6 – Esquema simplificado dos vetores utilizados no BiFC. Acima, pSPYNE-35S; abaixo, pSPYCE-35S (adaptado de WALTER et al. 2004)

3.2 Ensaios de duplo-híbrido em levedura

3.2.1 Material Biológico

Células de levedura *Saccharomyces cerevisiae* da cepa AH109 estocadas a -80°C em 1 mL de meio YPDA [10 g/L de extrato de levedura, 20 g/L de peptona, 20 g/L de glicose, 100 mg/L de hemissulfato de adenina] e glicerol 10% foram plaqueadas em YPDA sólido [10 g/L de extrato de levedura, 20 g/L de peptona, 20 g/L de glicose, 100 mg/L de hemissulfato de adenina, 20 g/L de ágar bacteriológico] e incubadas 24 horas a 30°C sob agitação de 200 rpm. Após este período de incubação, algumas colônias resultantes foram aleatoriamente escolhidas para transformação.

3.2.2 Construção da biblioteca de cDNA de *Arabidopsis thaliana*

A biblioteca de cDNA de *Arabidopsis thaliana*, ecótipo Columbia, foi obtida através do banco de dados TAIR (The Arabidopsis Information Resource – www.arabidopsis.org) e clonada no vetor pACT (Figura 7). A biblioteca utilizada, CD4-10, foi produzida a partir de mRNA extraído de todos os tecidos da planta e transcrito reversamente com iniciadores aleatórios. O método para a construção da biblioteca seguiu a descrição de Durfee e colaboradores (DURFEE et al., 1993).

3.2.3 Transformação de levedura para triagem (*screening*) em ensaios de duplo-híbrido

O protocolo utilizado para transformação de levedura para triagem de interações proteína-proteína em ensaios de duplo-híbrido com biblioteca de cDNA de *Arabidopsis thaliana* e com a construção contendo os primeiros 950 pb da sequência codificadora do gene da glutamil-tRNA sintetase citosólica (At5g26710) segue aquele proposto por Agatep e colaboradores (AGATEP et al., 1998).

Colônias de levedura *Saccharomyces cerevisiae* da cepa AH109 previamente incubadas durante 24 horas a 30°C em meio YPDA sólido foram aleatoriamente escolhidas, inoculadas em 25 mL de meio YPDA líquido e mantidas sob agitação de 200 rpm por 16 horas a 30°C. Após este período, foi determinada a titulação celular e calculado o volume celular que continha $2,5 \times 10^8$ células para cada 50 mL de cultura necessária. O volume determinado de 1,7 mL de cultura foi então despejado em microtubo estéril e as células foram coletadas por centrifugação por 5 minutos a 3000 x g. O precipitado resultante foi ressuscitado no volume de 50 mL de YPDA líquido pré-aquecido a 30°C e devidamente armazenado em frasco de vidro estéril. Esta nova cultura foi incubada a 30°C durante aproximadamente 4 horas sob agitação de 200 rpm, quando foi alcançada a titulação celular de 2×10^7 células/mL.

As células foram então coletadas por centrifugação por 5 minutos a 3000 x g. O meio sobrenadante foi descartado e o precipitado de células foi ressuscitado com 25 mL de água mili-Q estéril, lavado e novamente as células foram centrifugadas por 5 minutos a 3000 x g. O precipitado celular foi ressuscitado em 3 mL de acetado de lítio 100mM estéril, transferido para um tubo estéril adequado e incubado por 15 minutos a 30°C. As células foram então novamente coletadas por centrifugação por 5 minutos a 3000 x g e o sobrenadante foi removido.

Figura 7 – Esquema dos vetor pACT utilizado para clonagem na clonagem da biblioteca de cDNA de *Arabidopsis thaliana* no sistema duplo-híbrido de levedura (adaptado de www.clonetech.com)

Ao precipitado celular foram adicionados, necessariamente nesta ordem, 2,4 mL de PEG (polietilenoglicol) a 50%, 360 μ L de 1M de acetato de lítio, 500 μ L de sal de DNA de salmão (Sigma® Life Sciences), 100 μ L de biblioteca de cDNA de *Arabidopsis thaliana* clonada no vetor pACT, 100 μ L de extração plasmidial purificada da construção contendo a sequência codificadora GluRS clonada no vetor pGBKT7, chamado de “isca” e 140 μ L de água mili-Q estéril. A mistura foi agitada em aparelho de vortex até que o precipitado celular estivesse completamente ressuspendido. A mistura foi incubada a 30°C durante 30 minutos. Posteriormente, as mistura contendo as células foi submetida a choque térmico de 42°C durante 30 minutos, com mistura por inversão a cada 5 minutos durante 15 segundos.

As células foram coletadas por centrifugação por 5 minutos a 3000 x g, o sobrenadane foi descartado e as células foram cuidadosamente ressuspendidas em 10 mL de água mili-Q estéril e plaqueadas em meio de cultura sólido contendo marcadores seletivos (meio TRAF0) na ausência dos aminoácidos leucina e triptofano

(Leu- Trp-) [20 g/L de glicose, 6,7 g/L de nitrogênio básico, 2 g/L de mistura sintética de nutrientes Drop Out Powder (USBiological®) sem os aminoácidos leucina (L/Leu), triptofano (W/Trp), histidina (H/His), sem adenina (A/Ade) e sem nitrogênio básico, 20 g/L de ágar bacteriológico, 20 mg/L de histidina e 40 mg/L de adenina].

As placas foram incubadas a 30°C durante 4 dias e a seleção de clones positivos foi feita através do sistema Matchmaker Two-Hybrid System 3 (Clontech) com meio de cultura contendo marcadores seletivos. Os clones transformantes foram inoculados nos meios sólidos TRAFO repórteres Leu- Trp- His-, Leu- Trp- Ade- e Leu- Trp- His- Ade-.

3.2.4 Extração de DNA plasmidial de leveduras positivas para a triagem (*screening*) em duplo-híbrido

A extração de DNA plasmidial de leveduras positivas para as triagens feitas em ensaios de duplo-híbrido foi realizada pelo método descrito por Hoffman e Winston (HOFFMAN; WINSTON, 1987).

5 ml de meio líquido seletivo TRAFO ausente dos aminoácidos leucina e triptofano foram inoculados com colônias de levedura positivas para o sistema de seleção Matchmaker Two-Hybrid System 3 (Clontech) durante 16 horas a 30°C. Posteriormente, as células foram coletadas por centrifugação por 15 segundos a 3000 x g. O sobrenadante foi descartado, o precipitado celular foi ressuspensionado em 400 µL do tampão de lise [2% de triton, 1% de SDS, 100 mM de NaCl, 10 mM de Tris-HCl (pH 8), 1mM de Na₂EDTA] e agitada em vortex. A esta mistura foram adicionados 400 µL de solução 25% de fenol : 24% de clorofórmio : 1% de álcool isoamílico. Novamente a mistura foi agitada em vortex e depois foi adicionado 0,20 g de pérola de vidro lavada em ácido. A mistura foi agitada por 2 minutos em vortex.

A fase aquosa foi coletada por centrifugação durante 2 minutos a 14000 rpm. A esta fase foram adicionados 20 µL de 3M de acetato de potássio. Após mistura por inversão, foram adicionados 2 volumes de etanol 100%. O precipitado resultante foi coletado por centrifugação durante 30 minutos a 4°C e 14000 rpm. Posteriormente, foi lavado com 400 µL de etanol 80%. O precipitado foi coletado por centrifugação durante

5 minutos a 14000 rpm e foi seco naturalmente durante cerca de 20 minutos. Finalmente, o precipitado seco foi ressuspensionado em 50 µL de água mili-Q estéril.

3.2.5 Análise das sequências codificadoras dos genes para as proteínas interagentes na triagem (*screening*) em duplo-híbrido

As colônias de levedura positivas para os testes de interação proteína-proteína da triagem em duplo-híbrido tiveram seu DNA plasmidial extraído e o material resultante foi analisado através de PCR em condições previamente descritas com iniciadores específicos para o vetor pACT. O material plasmidial extraído também foi sequenciado utilizando os iniciadores específicos para o vetor pACT e analisado pelo programa BLAST (<http://blast.ncbi.nlm.nih.gov/>).

3.2.6 Transformação de levedura para confirmação de interação proteína-proteína detectada pela triagem (*screening*) em duplo-híbrido

O protocolo utilizado para transformação de levedura para confirmar as interações detectadas pelos ensaios de duplo-híbrido segue aquele proposto por Gietz e Woods (GIETZ; WOODS, 2002). Foram feitas trocas, de modo que a proteína utilizada como “isca” no vetor pGBKT7 e a proteína “presa” no vetor pGADT7 tiveram sua interação testada através da transformação com os vetores expressando as proteínas inteiras, além de uma troca na qual a proteína “isca” foi clonada no vetor pGADT7, passando a ser “presa” e a proteína “presa” foi clonada do vetor pGBKT7, passando a ser “isca”.

As placas resultantes da transformação foram incubadas a 30°C por 4 dias em meio sólido TRAF0 sólido ausente dos aminoácidos leucina e triptofano. A seleção de clones positivos foi feita através do sistema Matchmaker Two-Hybrid System 3 (Clontech) com meio de cultura contendo marcadores seletivos. Os clones transformantes foram inoculados no meio sólido TRAF0 repórter Leu– Trp– His– Ade–. Foi utilizado como controle positivo a interação positiva detectada nas triagens e como controle negativo foi feita transformação utilizando um dos vetores vazios e o outro vetor com a sequência codificadora do gene a ele clonado, além dos dois vetores,

pGADT7 e pGBKT7 vazios. Todas estas transformações foram repetidas duas vezes no intervalo de duas semanas.

3.2.7 Transformação de levedura para mapeamento da(s) região(ões) de interação proteína-proteína nos ensaios de duplo-híbrido

O protocolo utilizado para transformação de levedura para ensaios de duplo-híbrido para mapear a(s) região(ões) onde ocorre interação proteína-proteína foi realizado através da transformação de levedura com quatro deleções diferentes obtidas dentro dos primeiros 950 pb da sequência codificadora do gene da glutamil-tRNA sintetase citosólica (At5g26710) segue aquele proposto por Gietz e Woods (GIETZ; WOODS, 2002). A região correspondente a cada deleção encontra-se detalhada no APÊNDICE 3.

As placas resultantes da transformação foram incubadas a 30°C por 4 dias em meio sólido TRAF0 sólido ausente dos aminoácidos leucina e triptofano. A seleção de clones positivos foi feita através do sistema Matchmaker Two-Hybrid System 3 (Clontech) com meio de cultura contendo marcadores seletivos. Os clones transformantes foram inoculados no meio sólido TRAF0 repórter Leu– Trp– His– Ade–. Foi utilizado como controle positivo a interação positiva detectada nas triagens e como controle negativo foi feita transformação utilizando o vetor pGADT7 vazio e o vetor pGBKT7 com cada uma das deleções a ele clonado, além dos dois vetores, pGADT7 e pGBKT7 vazios. Todas as transformações foram repetidas duas vezes no intervalo de duas semanas.

3.3 Confirmação das interações *in planta* por BiFC

3.3.1 Transformação de *Agrobacterium tumefaciens*

Para confirmar possíveis interações entre GluRS e GS foi utilizado o sistema BiFC (Complementação de Fluorescência Bimolecular) voltado para expressão transiente em plantas de *Nicotiana tabacum*. Para tanto, no plasmídio que codifica a região N-terminal da YFP, pSPYNE-35S, foi clonada a sequência codificadora para a proteína GS, enquanto que no plasmídio que codifica a região C-terminal da YFP,

sPYCE-35S, foi clonada a sequência codificadora da GluRS. As reações de ligação foram clonadas inicialmente em *E. coli*, cepa DH5 α , conforme previamente descrito.

Posteriormente à clonagem, foi realizada a transformação de *Agrobacterium tumefaciens* GV3101 por eletroporação. Após 2 horas incubadas a 28°C em meio SOC, as bactérias transformadas foram plaqueadas em LB sólido contendo 50 μ g/mL de canamicina, 50 μ g/mL de rifampicina e 50 μ g/mL de gentamicina e incubadas durante 48 horas a 28°C. Em seguida, as cepas recombinantes transformadas foram selecionadas para confirmação por PCR sob as condições previamente descritas utilizando os iniciadores específicos para cada sequência codificadora.

3.3.2 Material biológico vegetal

Sementes de tabaco selvagem da espécie *Nicotiana tabacum* foram semeadas em vasos plásticos contendo iguais proporções de substrato organomineral (Plantamax) e mantidas em câmara de crescimento (Convion ATC26) sob condições controladas de fotoperíodo (12 horas claro e 12 horas escuro), aproximadamente 300 μ E de intensidade luminosa e temperatura constante (25°C). As plantas foram irrigadas diariamente com a mesma solução nutritiva utilizada no cultivo de *Arabidopsis thaliana* e água.

3.3.3 Transformação de plantas de tabaco por agroinfiltração

Folhas de plantas de tabaco com aproximadamente 4 semanas foram utilizadas para transformação transiente mediada por *Agrobacterium tumefaciens*. Duas colônias de bactérias positivas, uma para cada construção, uma colônia de bactéria transformada com o vetor p19, que ajuda na expressão das proteínas expressas de forma transiente (VOINNET et al., 2003), duas colônias, sendo uma de cada controle negativo e duas colônias, sendo uma de cada controle positivo, foram transferidas para 5 mL de meio MGL líquido [2,5 g/L de extrato de levedura, 5 g/L de triptona, 5 g/L de NaCl, 5 g/L de manitol, 1,16 g/L de glutamato monossódico, 0,25 g/L de KH₂PO₄, 0,1 g/L de MgSO₄.7H₂O e 0,1 g/L de biotina a pH 7] contendo 50 μ g/mL de canamicina, 50 μ g/mL de rifampicina e 50 μ g/mL de gentamicina e incubadas durante 16 horas a 28°C

sob agitação de 200 rpm. Posteriormente, 1 mL de cada cultura foi coletado por centrifugação por 5 minutos a 5000 rpm à temperatura ambiente e ressuspendidas em 1 mL do tampão de infiltração [500 mM de MES (pH 5.6), 2 mM de Na_3PO_4 , 0,5% de glicose e 100 μL de acetosiringona]. Em seguida, as células foram novamente coletadas por centrifugação por 5 minutos a 5000 rpm, novamente ressuspendidas no tampão de infiltração e coletadas por centrifugação nas mesmas condições. A suspensão bacteriana foi diluída com tampão de infiltração para ajustar o inóculo à concentração que fornece a densidade ótica DO 600 de aproximadamente 0,1.

Antes de proceder à transformação, as plantas de tabaco foram expostas por 2 horas à luz solar direta e suas folhas foram regadas com água abundante, a fim de manter os estômatos abertos para facilitar a infiltração. Em seguida, utilizando uma lâmina de um bisturi, foram feitas 5 pequenas injúrias na face abaxial de folhas de tabaco, totalizando 4 folhas por planta, e, com o auxílio de uma seringa de 1 mL, foi infiltrada com ligeira pressão as culturas de agrobactérias diluídas em tampão de infiltração. As plantas foram incubadas sob as mesmas condições de cultivo de tabaco selvagem acima descritas durante três dias, quando seguiram para observação.

Os controles positivos para este experimento consistiram na transformação de tabaco de acordo com o protocolo acima descrito utilizando a proteína bZIP63 homodimerizada através de interação entre seus monômeros, um fusionado à porção N-terminal da YFP no vetor pSPYNE-35s e outro fusionado à porção C-terminal da YFP no vetor pSPYCE-35S. Os controles negativos consistiram na transformação de tabaco também de acordo com o protocolo acima com os vetores pSPYNE-35S e pSPYCE-35S vazios. Este experimento foi repetido três vezes ao longo de um mês.

3.3.4 Visualização das plantas transformadas

Para a visualização das folhas infiltradas com *Agrobacterium tumefaciens*, foram feitas lâminas com pequenos cortes de folhas transformadas três dias antes. Os cortes de folhas com a epiderme abaxial voltada para cima foram hidratados com água mili-Q. Estas lâminas foram observadas em microscópio confocal de varredura a laser

Olympus® FV1000 em comprimento de onda 488 nm. As imagens foram analisadas no programa Olympus Fluoview FV10 – ASW.

3.4 Análises filogenéticas da GluRS

Inicialmente, foram feitas buscas em bancos de dados para a GluRS citosólica e foi feito um alinhamento das sequências encontradas em espécies de Eubacteria, Archea, fungos e organismos fotossintetizantes, sendo este último grupo compreendendo musgos, algas e plantas vasculares. A este alinhamento foi incluída a região N-terminal da GluRS de 333 aminoácidos (codificada pelos primeiros 1000 pares de base, chamada de A1B1 na Figura 4).

A hipótese filogenética para a GluRS foi proposta baseada no método da análise Bayesiana com o auxílio do programa MrBayes CVS versão (HUELSENBECK; RONQUIST, 2001) configurado para computação em paralelo. Os modelos de substituição de aminoácidos foram escolhidos usando o sistema PROTTest 2.4 (ABASCAL; ZARDOYA; POSADA, 2005). A análise Bayesiana foi desenvolvida sob o modelo WAG (WHELAN; GOLDMAN, 2001), com distribuição gama (Γ) e com proporção de sítios invariáveis (I). Seis cadeias simultâneas de Markov e Monte Carlo (MCMC) foram conduzidas para 5×10^6 gerações, com amostragem de árvores a cada 500 ciclos. As primeiras 2500 árvores foram descartadas como “burn in” para construção da árvore consenso. A consistência de cada ramo foi determinada usando-se probabilidade posterior. Para todas as análises, sequência da GluRS de algumas bactérias foram utilizadas como grupo externo.

A divergência temporal entre fungos e vegetais (HEDGES et al. 2004), eubactérias e eucariotos (DOOLITTLE et al, 1996), musgos e plantas vasculares (HEDGES et al. 2004; HECKMAN et al. 2001) e fungos ascomicetos e basidiomicetos (HECKMAN et al. 2001; HEDGES et al. 2004) foi utilizada para datação da GluRS em *Arabidopsis thaliana*. A filogenia no nível de espécie proposta anteriormente, foi analisada utilizando-se o modelo temporal relaxado incorrelato lognormal Bayesiano para obtenção das prováveis datas de divergência dos ramos constantes na árvore original. Para tanto utilizou-se o sistema Beast versão 1.4.8 (DRUMMOND; RAMBAUT, 2007), seguindo-se o tutorial padrão para datação, resumido a seguir. O conjunto de

dados foi analisado seguindo o modelo de substituição proposto anteriormente (WAG + I + G) implementado no Beast, duas cadeias Monte Carlo Markov foram corridas por 90.000.000 gerações, utilizando o modelo de especiação de Yule, amostrando as árvores a cada 10.000 gerações. Para geração da árvore final levou-se em consideração 10% de “*burn in*”.

4 RESULTADOS

4.1 Interações proteína-proteína identificadas na triagem (screening) em ensaios de duplo-híbrido em levedura

O sistema duplo-híbrido em levedura é uma das técnicas indicadas para identificar proteínas que interagem num contexto *in vivo* (Gietz et al., 1997). Ao utilizar dois plasmídios, um que expressa um domínio de ligação ao DNA fusionada à proteína de sequência conhecida cujas possíveis interações se quer investigar (chamada de “isca”) e outro que expressa um domínio de ativação fusionado a uma biblioteca de cDNA (chamada genericamente de “presa”), a interação entre as duas proteínas “isca” e “presa” causa interação dos domínios de ligação e de ativação, levando à transcrição de um gene repórter quando a levedura transformada é cultivada em meio seletivo.

Neste estudo, a possível função não-canônica voltada para a localização subcelular de proteínas em plantas associada à glutamil-tRNA sintetase (GluRS) foi investigada através de ensaios de duplo-híbrido que revelassem interações entre a GluRS de *Arabidopsis thaliana* e outras proteínas desta espécie, que fossem importadas para os diversos compartimentos celulares.

As triagens dos ensaios de duplo-híbrido, em que a sequência contendo os 950 pares de base iniciais, o que corresponde aos 316 primeiros aminoácidos da GluRS, funcionou como “isca” contra uma biblioteca (Figura 4) de cDNA de *Arabidopsis* cujas proteínas codificadas funcionaram como “presas”, revelaram a interação entre a GluRS e a glutamina sintetase (GS) (At5g35630), que atua catalizando a conversão de amônia para a atóxica Gln (Figura 8). O mais interessante de se observar é que esta proteína tem duplo direcionamento (DD) para mitocôndrias e cloroplastos (TAIRA et al., 2004). Os testes feitos no meio seletivo Leu– Trp– His– Ade– nas leveduras positivas para a interação foram todos positivos.

Figura 8 – Resultado obtido nas triagens (*screenings*) de duplo-híbrido em levedura. A proteína glutamyl-tRNA sintetase citosólica (GluRS), aqui utilizada com “isca” nos experimentos, interagiu com a proteína glutamina sintetase (GS), que tem duplo direcionamento. A interação foi confirmada por dois genes repórteres, um para síntese de His e outro para síntese de Ade. O controle utilizado foi a proteína GluRS fusionada ao domínio de ligação ao DNA codificado no vetor pGBKT7, servindo de “presa”, e o domínio de ativação codificado no vetor pGADT7 sem nenhuma proteína fusionada servido de “isca”. Não houve interação entre as proteínas no meio seletivo com genes repórteres de síntese de His e Ade

4.2 Confirmação da interação das proteínas em duplo-híbrido em levedura

Novos ensaios de duplo-híbrido, desta vez envolvendo as proteína GluRS (inteira) e GS, foram feitos para dar maior confiabilidade à triagem em levedura. O objetivo deste experimento era o de verificar se havia interação entre as duas proteínas quando, ao invés de apenas a região N-terminal, toda a proteína estava expressa. Para tanto, um primeiro ensaio foi realizado com a sequência codificadora da GluRS inteira clonada no vetor pGBKT7, de modo que a proteína GluRS serviria como “isca” e a sequência codificadora da GS clonada no vetor pGADT7, sendo a GS “presa”. Os resultados do ensaio revelaram que a interação ainda existia, com testes feitos em meio seletivo Leu– Trp– His– Ade– (Figura 9).

Figura 9 – Confirmação da interação entre as proteínas GluRS e GS em duplo-híbrido em levedura. A proteína GluRS serviu de “isca” (fusionada ao domínio de ligação), enquanto a proteína GS serviu de “presa” (fusionada ao domínio de ativação), mostrando interação obtida nas triagens. A troca de fusão, com GluRS de “presa” e GS de “isca” também demonstrou que a interação persiste. A série de controles negativos utilizados consistia em uma das proteínas fusionada ao domínio de ligação ou ativação, servido, respectivamente, como “isca” ou “presa” transformada apenas com um dos dois domínios, sem proteína fusionada. A ausência de interações demonstra que estes dois domínios não estão interagindo sozinhos. O controle positivo foi a própria interação GluRS-GS obtida na triagem. Todas as interações foram testadas com dois genes repórteres, um para síntese de His e outro para síntese de Ade

Noutro ensaio de confirmação, foi realizada uma troca de vetores na qual a GluRS inteira, com sequência codificadora clonada no vetor pGADT7, agora seria a “presa”, enquanto que a GS, com sequência clonada no vetor pGBKT7, serviria como

“isca” para os ensaios de duplo-híbrido. Novamente, a interação ainda existia, com testes feitos em meio seletivo Leu– Trp– His– Ade– (Figura 9). O controle positivo utilizado também funcionou em meio seletivo Leu– Trp– His– Ade–. Não foi observada nenhuma interação para os controles negativos utilizados.

4.3 Mapeamento da(s) região(ões) de interação proteína-proteína nos ensaios de duplo-híbrido em levedura

Uma vez encontrada interação entre nos ensaios de duplo-híbrido GluRS e GS e confirmada, o passo seguinte foi mapear a região desta interação. Com o embasamento de estudos anteriores em levedura (SIMOS et al., 1996; GALANI et al., 2001; KARANASIOS et al., 2007) e em *Arabidopsis*, além dos próprios resultados obtidos com as triagens em duplo híbrido, ficou evidente que a região de interação estava na porção N-terminal da proteína, possivelmente nos seus primeiros 300 aminoácidos.

Para mapear esta região, foram construídas deleções variadas (Figura 6) com o intuito de realizar ensaios de duplo-híbrido utilizando cada uma das deleções como proteínas “presa” e a GS, inteira, como proteína “isca”. O objetivo foi precisar em qual região ocorria, de fato, a interação. Este resultado também seria fonte de estudos filogenéticos em busca de domínios e da origem desta região.

Os resultados revelaram a presença de interação em apenas uma das deleções, a chamada A3B1 (Figura 10). Os testes feitos em meio seletivo Leu– Trp– His– Ade– foram todos positivos. A sequência codificadora desta região compreende 332 pares de base, do par 618 ao par 950. A região da proteína tem 109 aminoácidos e compreende os resíduos 207 a 316. O controle positivo utilizado também funcionou em meio seletivo Leu– Trp– His– Ade–. Não foi observada nenhuma interação para os controles negativos utilizados.

Figura 10 – Mapeamento da região de GluRS responsável pela interação com GS através de ensaio de duplo-híbrido. Em todos os casos, GluRS e GS estavam fusionadas a, respectivamente, domínio de ligação (“isca”) e domínio de ativação (“presa”). Todas as interações foram testadas com dois genes repórteres, um para síntese de His e outro para síntese de Ade. Nestes testes, só foram detectadas interações com as duas proteínas inteiras; com a região A1B1 de GluRS, chamada de N-terminal de 316 aminoácidos (resíduos 1 – 316); e com a deleção A3B1 de GluRS de 110 aminoácidos (resíduos 206 – 316)

4.4 Confirmação das interações *in planta* por BiFC

Embora o duplo-híbrido seja uma técnica voltada para detectar interações entre proteínas *in vivo* (GIETZ et al., 1997), ela pode revelar falsas interações, isto é, interações que não poderiam ocorrer em seus organismos originais. Os dados de interação proteína-proteína ganham maior confiabilidade quando são confirmados com o auxílio de outras técnicas, sejam elas *in vitro*, ou *in planta*, como o sistema FRET, Ressonância Fluorescente por Transferência de Energia (CHEN; MILLS; PERIASAMY,, 2003; PERIASAMY, 2000) e o sistema BiFC, Complementação Fluorescente Bimolecular (WALTER et al., 2004).

Neste trabalho, optou-se por confirmar as interações encontradas *in planta* através da técnica de BiFC. Esta abordagem para a visualização de interações baseia-se na formação de um complexo protéico fluorescente, através de dois fragmentos do gene para a *yellow fluorescent protein* (YFP) clonados em dois plasmídios diferentes, sendo que ambos possuem o promotor 35S do vírus do mosaico. Um dos plasmídios codifica a região N-terminal da YFP, que fusiona-se a uma das proteínas estudadas, e o outro plasmídio codifica a região C-terminal da YFP fusionada à outra proteína estudada. Uma vez que as duas proteínas interagem, a YFP é restaurada e emite fluorescência que pode ser visualizada através de microscopia (WALTER et al., 2004).

Os resultados dos ensaios de BiFC mostraram que ocorre interação entre a GluRS (fusionada à região C-terminal da YFP) e a GS (fusionada à região N-terminal da YFP) e que esta interação ocorre no citoplasma das células de planta (Figura 11). Durante as três repetições conduzidas neste ensaio, o resultado alcançado foi o mesmo. Com esta confirmação, a interação entre GluRS e GS ganhou maior confiabilidade.

As plantas transformadas com os controles negativos não apresentaram nenhuma fluorescência. Aquelas transformadas com controles positivos apresentaram fluorescência no núcleo, conforme esperado para a proteína bZIP63, e também difusa pelo citoplasma. Para melhorar a confiabilidade do controle, uma imagem com GFP citosólica foi adicionada, obtida nas mesmas condições que foram obtidas as imagens de BiFC (Figura 11D). É possível observar que o padrão de fluorescência citosólica se repete e é semelhante ao observado nas plantas transformadas com a interação GluRS-GS, corroborando assim os resultados.

O comprimento de onda utilizado para visualizar estes ensaios de BiFC, 488 nm é o comprimento de onda de excitação da GFP. A YFP tem comprimento de onda 512 nm, muito próximo ao comprimento daquela. Na verdade, há uma sobreposição de espectros luminosos entre as duas fluorescências.

Figura 11 – Confirmação da interação entre GluRS e GS através de ensaio de BiFC. (A) Interação entre GluRS e GS; (B) Dict da interação entre GluRS e GS; (C) Controle negativo (YFP N-terminal e YFP C-terminal sem fusão a nenhuma proteína obtidas através de vetores pSPYNE-35S e pSPYCE-35S vazios) ; (D) Controle positivo (proteína bZIP63); (E) Controle citoplasmático positivo para GFP (THIE-THID fusionada a GFP fora de fase, fazendo com que a proteína fique no citoplasma). Visualizações em comprimento de onda de 488 nm

4.5 Análises filogenéticas da região N-terminal de interação da GluRS

Com uma interação proteína-proteína de interesse encontrada entre GluRS e GS e com a região de interação definida, foram feitas análises filogenéticas para determinar a origem e o caminho evolutivo da GluRS citosólica de *Arabidopsis*, buscando entender também em que ponto da história evolutiva esta proteína adquiriu a região N-terminal responsável pelas interações proteína-proteína.

O alinhamento das sequências de GluRSs citosólicas compreendeu aquelas encontradas em espécies de Eubacteria, Archea, fungos e organismos fotossintetizantes, sendo este último grupo compreendendo musgos, algas e plantas vasculares. Além disso, foi incluída a região N-terminal da GluRS de 333 aminoácidos (A1B1 da Figura 4). Quando a região de interação foi definida através dos ensaios de duplo-híbrido com as deleções, revelando a deleção A3B1 como aquela que contém os aminoácidos e possivelmente os domínios essenciais à interação, a sequência também foi alinhada com aquelas previamente encontradas.

Os resultados do alinhamento revelaram que a extensão N-terminal observada na GluRS de *Arabidopsis* está ausente nas bactérias. Por isso, este grupo foi tomado como externo na construção da árvore filogenética. Nas espécies representantes de Archea é possível observar que esta extensão começa a surgir. Além de surgir em Archea, é observada conservação da região quanto à sequência de aminoácidos. Quando alinhada com todos os grupos de fungos, o nível de conservação dentro do grupo é grande, como era esperado, e o grau de conservação é ainda maior, quando comparado àquela encontrada em Archea. A partir do grupo dos eucariotos fotossintetizantes, sobretudo das plantas vasculares, são observadas sequências mais conservadas quando comparadas à GluRS de *Arabidopsis* (ANEXO 1).

Quando ao alinhamento foi adicionada a região de interação correspondente à deleção A3B1, observou-se a presença de dois domínios extremamente conservados entre todos os grupos analisados: o domínio PEAEIG (Pro-Glu-Ala-Glu-Ile-Gly), com padrão de conservação PXAXXG, e o domínio KVKLR (Lys-Val-Lys-Leu-Arg), com padrão de conservação XVXXR (ANEXO 1).

A árvore filogenética resultante revelou suportes variando de 0,66 a 1, sendo que a maioria ficou em torno de 0,9, o que reforça a interpretação das análises (Figura 12).

Ainda de acordo com a árvore, a história evolutiva da GluRS citosólica em questão tem origem muito antiga, há cerca de 3,5 bilhões de anos. Há cerca de 2 a 1,7 bilhão de anos atrás surgiu o ancestral que originou a extensão presente nas sequências observadas. É possível inferir que este ancestral é comum para Archea, fungos e todos os eucariotos fotossintetizantes analisados, incluindo as plantas vasculares e *Arabidopsis*. Todas as bactérias foram agrupadas como grupo externo, uma vez que não apresentam a extensão N-terminal na suas sequências de GluRSs tal qual é observada na de *Arabidopsis*.

(continua)

(página anterior)

Figura 12 – Árvore resultante das análises filogenéticas sobre a origem da GluRS citoplasmática de *Arabidopsis thaliana*. As análises revelaram que o ancestral da GluRS provavelmente apareceu em Archea, entre 2,6 e 1,8 bilhões de anos atrás. O suporte dos dados variou de 0,66 a 1, sendo que a maioria ficou em torno de 0,9. As bactérias, pela ausência da região N-terminal característica de GluRS citoplasmática de *Arabidopsis*, compõem o grupo externo

5 DISCUSSÃO

Os resultados obtidos nos ensaios de duplo-híbrido mostraram interação entre a GluRS citosólica e a GS de DD, ambas de *Arabidopsis thaliana*. Esta interação foi confirmada *in planta* através de experimentos de BiFC. O comprimento de onda utilizado para visualizar estes ensaios de BiFC, 488 nm é o comprimento de onda de excitação da GFP. A YFP tem comprimento de onda 512 nm, muito próximo ao comprimento ótimo da GFP. Na verdade, há uma sobreposição de espectros luminosos entre as duas fluorescências. Muitos trabalhos mostrando ou utilizando a YFP como gene repórter são visualizados com comprimento de onda da GFP com sucesso (OHAD; SHICHRUR; YALOVSKY, 2007; LERMONTOVA et al., 2007).

A região de interação de GluRS responsável pela interação entre as proteínas encontra-se na sua região N-terminal, nos seus primeiros 300 aminoácidos. Os resultados alcançados com experimentos de duplo-híbrido sobre deleções nesta região revelaram a presença de interação em apenas uma das deleções, a chamada A3B1 (Figura 4), de 109 aminoácidos e que vai do resíduo 207 ao 316. Os resultados dos alinhamentos revelaram que a região N-terminal completa da GluRS está ausente nas bactérias. É nas espécies representantes de Archea que é possível observar que esta região começa a surgir com um padrão de conservação. Quando alinhada com todos os grupos de fungos, musgos e plantas vaculares, o nível de conservação é ainda maior. A árvore filogenética mostrou que a GluRS citosólica apareceu em Archea entre 2,6 e 1,8 bilhões de anos. No que diz respeito à região N-terminal da GluRS responsável pelas interações até hoje descritas em *Arabidopsis*, pode-se inferir que o ancestral comum a fungos e plantas que produziu esta extensão surgiu entre 2 a 1,7 bilhões de anos.

Os resultados encontrados no duplo-híbrido utilizando a GluRS como “isca” juntamente com os resultados dos experimentos de confirmação desta interação através do BiFC sugerem que a GluRS pode estar envolvida de alguma maneira na localização subcelular de GS, mais precisamente na sua exportação para as organelas.

O gene que codifica a GS organelar é nuclear e a proteína é sintetizada no citoplasma. Ela atua no metabolismo do nitrogênio, catalisando a conversão de amônia a glutamina (Gln) nas organelas (TAIRA et al., 2004). A Gln gerada neste processo

passa por várias atividades anabólicas de amidotransferases até regenerar ácido glutâmico (Glu). Interessantemente, a GluRS carrega com Glu o tRNA^{Glu}. Sendo assim, é possível especular que a GluRS atuaria como um carreador citoplasmático para a GS, talvez com a participação do Glu, molécula comum à função destas duas proteínas. Essa hipótese pode ser embasada também pelos dados de BiFC obtidos neste trabalho, em que observa-se que a interação entre GluRS e GS ocorre no citoplasma (Figura 11).

Os resultados obtidos neste trabalho aliados ao que já foi observado em outras espécies de eucariotos reforçaria a grande base para a hipótese de que em plantas a GluRS poderia estar associada à localização subcelular de proteínas. Estudos em levedura mostraram que a GluRS interage com outras proteínas no citoplasma e tem um papel associado à localização subcelular das proteínas interagentes (GALANI et al., 2001; GOLINELLI-COHEN; MIRANDE, 2007). Em mamíferos a GluRS também está relacionada à localização subcelular das nove aaRSs com as quais interage (LysRS, LeuRS, IleRS, GluRS, ProRS, MetRS, GlnRS, ArgRS e AspRS) num grande MSC, exportando-as do núcleo (ROBINSON et al., 2000). Entretanto, ainda não foi descrito nenhum complexo multienzimático envolvendo a GluRS em plantas, mas é provável que esta tendência de expansão funcional da GluRs também seja confirmada nas plantas.

As funções não-canônicas relacionadas à GluRS em plantas referem-se à participação indireta na biossíntese de compostos tetrapirrólicos, como clorofila, heme e outros derivados, através da via do C₅, via também encontrada em Archea e na maior parte das bactérias (BEALE et al., 1975; SCHÖN et al., 1986; JAHN et al., 1992). A participação é indireta porque, uma vez que a GluRS carrega o tRNA^{Glu}, passando a Glu-tRNA^{Glu}, este ácido glutâmico é reduzido até dar origem a compostos tetrapirrólicos. Ainda em bactérias, a GluRS participa da regulação da biossíntese de heme (LEVICÁN et al., 2006).

Resultados recentemente obtidos em nosso laboratório também forneceram indícios de que a GluRS também pode estar associada à localização subcelular de outras proteínas destinadas aos cloroplastos em *Arabidopsis*. Neste trabalho, ensaios de duplo-híbrido revelaram a interação entre a extensão N-terminal da GluRS citosólica

e a proteína TH1, que é importada para o cloroplasto e lá participa da via de biossíntese de tiamina (AJJAWI; TSEGAYE; SHINTANE, 2006). Esta interação foi posteriormente confirmada *in planta* através de ensaios de BiFC (MARCELA E. SCARSO, comunicação pessoal).

Ao analisar em conjunto estes dados com aqueles apresentados neste trabalho, em que uma nova interação envolvendo GluRS e uma proteína importada (a GS organelar) foi revelada e confirmada *in planta*, fica mais plausível sugerir que a GluRS citosólica, ou mais precisamente a sua extensão N-terminal, esteja relacionada à importação de proteínas. Os resultados não permitem, porém, especular sobre possíveis complexos multienzimáticos envolvidos.

Em comum com os dados obtidos em levedura, a região de interação da GluRS citosólica localiza-se na porção N-terminal da proteínas. No complexo Arc1p•MetRS•GluRS de levedura, a região envolvendo os resíduos N-terminais 86 – 207 são cruciais para a interação entre a GluRS, MetRS e Arc1p, num total de 121 aminoácidos envolvidos (GALANI et al., 2001). Em *Arabidopsis*, a região N-terminal de aproximadamente 333 aminoácidos foi caracterizada como região de interação entre GluRS e TH1. Esta mesma região serviu como ponto de partida nas triagens em duplo-híbrido em busca de interações entre GluRS e outras proteínas. A região não só foi responsável pela interação, como foi utilizada para encontrar um maior refinamento em busca de algum domínio responsável por estas interações.

Os resultados dos ensaios com as deleções, a região de 109 aminoácidos correspondente aos resíduos 207 – 316 aqui chamada de deleção A3B1, localizada dentro da região N-terminal que envolve os 333 primeiros resíduos da GluRS citosólica, foi aquela crucial para que houvesse interação entre as proteínas testadas. Desta forma, os resultados obtidos nestes experimentos voltados para o mapeamento de interação, aliados aos resultados previamente observados em laboratório e às análises filogenéticas deste trabalho sugerem que a extensão N-terminal presente na GluRS citosólica de *Arabidopsis* representa uma vantagem evolutiva para esta proteína, ao passo que permite interações possivelmente voltadas para a tendência à expansão funcional da GluRS.

Quando foram alinhadas as sequências da GluRS citosólica de *Arabidopsis* com a sequência da GluRSs citosólica de levedura, que é encontrada no complexo Arc1p•MetRS•GluRS, é possível sugerir que os domínios PEAEIG e KVKLR são fundamentais para a interação entre GluRS e outras proteínas. O que se observa é uma grande conservação da região, inclusive de alguns resíduos dos domínios PEAEIG e KVKLR. É interessante observar que, na GluRS de levedura, este domínio está contido na região crucial para a GluRS permanecer no complexo. Tal fato reforça ainda mais a hipótese de que é nesta região que está(ão) localizada(s) o(s) domínio(s) fundamental(is) para a região N-terminal da GluRS promover interações proteína-proteína.

A árvore filogenética gerada pela análise das sequências de GluRSs apresentou suporte em torno de 0,9. Observa-se que a similaridade entre as sequências da GluRS de *Arabidopsis* e bactérias é pequena, notadamente na região N-terminal. Entretanto, é possível observar apenas uma pequena sequência conservada na região N-terminal localizada poucos resíduos antes da região do domínio KVKLR, com padrão de conservação XVXXR, o que sugere que este domínio pode ter origem mais antiga, por volta de 3 bilhões de anos.

Se a GluRS iniciou seu caminho evolutivo há aproximadamente 3,5 bilhões de anos, a região N-terminal da GluRS citosólica de *Arabidopsis* é aparentemente um pouco mais recente. Provavelmente, ela apareceu em *Archea* entre 2,6 e 1,8 bilhões de anos. O outro gene que codifica a GluRS organelar de *Arabidopsis* possivelmente também tem origem antiga, com seu ancestral filogenético sendo derivado de uma proteobactéria (BRANDÃO; SILVA-FILHO, submetido).

É possível inferir que a região N-terminal da GluRS responsável pelas interações proteína-proteína descritas em *Arabidopsis* possui ancestral comum a fungos e plantas. Além disso, estima-se que esta extensão surgiu entre 2 a 1,7 bilhões de anos, quando houve a divisão entre os ramos que deram origem aos eucariotos fotossintetizantes e aos fungos. Este dado complementa as observações dos alinhamentos nas quais a região N-terminal de plantas e fungos encontra-se bem conservada.

Embora evolutivamente distantes, as GluRSs citosólicas de *Arabidopsis* e de levedura, precisamente aquela que tem função na localização subcelular do complexo

Arc1p•MetRS•GluRS, surgiram em sua forma moderna por volta de 290.000 anos, num evento mais recente que o surgimento dos domínios PEAEIG e KVKLR, este último surgido há cerca de 3 bilhões de anos.

6 CONSIDERAÇÕES FINAIS

Os ensaios de duplo-híbrido mostraram uma interação entre a glutamil-tRNA sintetase (GluRS) citosólica e a glutamina sintetase (GS) de duplo direcionamento (DD), ambas de *Arabidopsis thaliana*. Esta interação foi confirmada *in planta* através de experimentos de BiFC. A região de interação de GluRS responsável pela interação entre as proteínas foi mapeada e encontra-se na sua extensão N-terminal, precisamente na região de 109 aminoácidos e que vai do resíduo 207 ao 316. Estes resultados em conjunto sugerem que pode haver também em plantas um complexo formado através da interação de GluRS com outras proteínas. Este complexo estaria voltado para a regulação da localização subcelular protéica através de interações estabelecidas na extensão N-terminal de GluRS, tal qual é observado noutras espécies de eucariotos.

Os resultados obtidos nas análises filogenéticas mostram que a extensão N-terminal da GluRS está ausente nas bactérias. É nas espécies representantes de Archea que é possível observar que esta região surgiu entre 2,6 e 1,8 bilhões de anos. Quando a extensão N-terminal é alinhada com todos os grupos de fungos, musgos e plantas vasculares, o nível de conservação é ainda maior e pode-se dizer que o ancestral comum a fungos e plantas que produziu esta extensão surgiu entre 2 a 1,7 bilhões de anos.

REFERÊNCIAS

ABASCAL, F.; ZARDOYA, R.; POSADA, D. ProtTest: selection of best-fit models of protein evolution. 2005, p. 2104-2105.

AGATEP, R.; KIRKPATRICK, R.D.; PARCHALIUK, D.L.; WOODS, R.A.; GIETZ, R.D. Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single-stranded carrier DNA/polyethylene glycol (LiAc/ss-DNA/PEG) protocol. **Technical Tips Online**, 1998. Disponível em: <http://tto.trends.com/>. Acesso em: 04 abr 2008.

AJJAWI, I.; TSEGAYE, Y.; SHINTANI, D. Determination of the genetic, molecular, and biochemical basis of the *Arabidopsis thaliana* thiamin auxotroph th1. 2007 **Archives of Biochemistry and Biophysics**, New York, v. 459, p. 107-14, 2007.

AMBROGELLY, A.; I. AHEL, C.; POLYCARPO, S.; BUNJUN-SRIHARI, B.; KRETT, C.; JACQUIN- BECKER, B.; RUAN, C.; KOHRER, C.; STATHOPOULOS, U.; BHANDARY, R.; D. SÖLL. *Methanocaldococcus jannaschii* prolyl-tRNA synthetase charges tRNA^{Pro} with cysteine. **The Journal of Biological Chemistry**, Bethesda, v. 277, p. 34749–34754, 2002.

AMBROGGIO, E.E.; AUSTEN, B.; FIDELIO, G.D. Biophysical properties of a synthetic transit peptide from wheat chloroplast ribulose 1,5-bisphosphate carboxylase. **Journal of Peptide Science**, Chichester, v. 13, p. 245–252, 2007.

AN, S.; MUSIER-FORSYTH, K. Trans-editing of Cys-tRNA^{Pro} by *Haemophilus influenzae* YbaK protein. **The Journal of Biological Chemistry**, Bethesda, v. 279, p. 42359–42362, 2004.

ANDRÈS, C.; AGNE, B.; KESSLER, F. The TOC complex: Preprotein gateway to the chloroplast. **Biochimica et Biophysica Acta**, Amsterdam, doi: 10.1016 / j.bbamcr.2010.03.004, 2010.

BAILLY, M.; BLAISE, M.; LORBER, B.; BECKER, H.D.; KERN, D. The transamidosome: a dynamic ribonucleoprotein particle dedicated to prokaryotic tRNA-dependent asparagine biosynthesis. **Molecular Cell**, Cambridge, v. 28, p. 228–239, 2007.

BEALE, S.I.; GOUGH, S.P.; GRANICK, S. Biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 72, p. 2719–2723, 1975.

BEC, G.; KERJAN, P.; WALLER, J.P. Reconstitution in vitro of the valyl-tRNA synthetase-elongation factor (EF) 1 beta gamma delta complex. Essential roles of the NH2-terminal extension of valyl-tRNA synthetase and of the EF-1 delta subunit in complex formation. **The Journal of Biological Chemistry**, Bethesda, v. 269, p. 2086–2092, 1994.

BHUSHAN, S.; KUHN, C.; BERGLUND, A.K.; ROTH, C.; GLASER, E. The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and missorting. **FEBS Letters**, Amsterdam, v. 580, p. 3966–3972, 2006.

BRANDÃO, M.M.; SILVA-FILHO, M.C. Evolutionay history of Arabidopsis thaliana aminoacyl-tRNA synthetase dual-targeting proteins. Submetido.

BROWN, J. R.; DOOLITTLE, W.F. Gene descent, duplication and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. **Journal of Molecular Evolution**, New York, v. 49, p. 485–495, 1999.

BUDDHA, M.R.; CRANE, B.R. Structure and activity of an aminoacyl-tRNA synthetase that charges tRNA with nitrotryptophan. **Nature Structural & Molecular Biology**, New York, v. 12, p. 274–275, 2005.

BUECHTER, D.D.; SCHIMMEL, P. Aminoacylation of RNA minihelices – implications for tRNA synthetase structural design and evolution. **Critical Reviews in Biochemistry and Molecular Biology**, Boca Raton, v. 28, p. 309–322., 1993.

CAHUZAC, B.; BERTHONNEAU, E.; BIRLIRAKIS, N.; GUITTET, E.; MIRANDE, M. A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. **The EMBO Journal**, London, v. 19, p. 445–452, 2000.

CARRIE, C.; GIRAUD, E.; WHELAN, J. Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. **FEBS Journal**, Oxford, v. 276, p. 1187-95, 2009.

CHABREGAS, S.M.; LUCHE, D.D.; FARIAS, L.P.; RIBEIRO, A.F.; VAN SLUYS, M.A.; MENCK, C.F.; SILVA-FILHO, M.C. Dual targeting properties of the N-terminal signal sequence of Arabidopsis thaliana THI1 protein to mitochondria and chloroplasts. **Plant Molecular Biology**, Dordrecht, v. 46, p. 639-650, 2001.

CHABREGAS, S.M.; LUCHE, D.D.; VAN SLUYS, M.A.; MENCK, C.F.; SILVA-FILHO, M.C. Differential usage of two in-frame translational start codons regulates subcellular localization of Arabidopsis thaliana THI1. **Journal of Cell Science**, London, v. 116, p. 285-291, 2003.

CHACINSKA, A.; LIND, M.; FRAZIER, A.E.; DUDEK, J.; MEISINGER, C.; GEISSLER, A.; SICKMANN, A.; MEYER, H.E.; TRUSCOTT, K.N.; GUIARD, B.; PFANNER, N.; REHLING, P. Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. **Cell**, Cambridge, v. 120, p. 817–829, 2005.

CHACINSKA, A.; KOEHLER, C.M.; MILENKOVIC, D.; LITHGOW, T.; PFANNER, N. Importing mitochondrial proteins: machineries and mechanisms. **Cell**, Cambridge, v. 138, n. 4, p. 628-644, 2009.

CHAN, N.C.; LIKIC, V.A.; WALLER, R.F.; MULHERN, T.D.; LITHGOW, T. The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. **Journal of Molecular Biology**, London, v. 358, p. 1010-1022, 2006.

CHEN, Y.; MILLS, J.D.; PERIASAMY, A. Protein localization in living cells and tissues using FRET and FLIM. **Differentiation**, Berlin, v. 71, p. 528–541, 2003.

CHEW, O.; WHELAN, J. Dual targeting ability of targeting signals is dependent on the nature of the mature protein. **Functional Plant Biology**, Victoria, v. 30, p. 805–812, 2003.

CHEW, O.; LISTER, R.; QBADOU, S.; HEAZLEWOOD, J.L.; SOLL, J.; SCHLEIFF, E.; MILLAR, A.H.; WHELAN, J. A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. **FEBS Letters**, Amsterdam, v.557, p. 109-114, 2004.

CLAROS, M.G.; VINCENS, P. Computational method to predict mitochondrially imported proteins and their targeting sequences. **European Journal of Biochemistry**, Berlin, v. 241, p. 779–786, 1997.

CREISSEN, G.; REYNOLDS, H.; XUE, Y.; MULLINEAUX, P. Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. **The Plant Journal**, New York, v. 8, p. 167-175, 1995.

CUSACK, S.; HARTLEIN, M.; LEBERMAN, R. Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. **Nucleic Acids Research**, Oxford, v. 19, p. 3489–3498, 1990.

CUSACK, S. Aminoacyl-tRNA synthetases. **Current Opinion in Structural Biology**, London, v. 7, p. 881–889, 1997.

DAGKESSAMANSKAIA, A.; MARTIN-YKEN, H.; BASMAJI, F.; BRIZA, P.; FRANCOIS, J. Interaction of Knr4 protein, a protein involved in cell wall synthesis, with tyrosine tRNA synthetase encoded by TYS1 in *Saccharomyces cerevisiae*. **FEMS Microbiology Letters**, Amsterdam, v. 200, p. 53–58, 2001.

DAY, I.S.; GOLOVKIN, M.; REDDY, A.S.N. Cloning of the cDNA for glutamyl-tRNA synthetase from *Arabidopsis thaliana*. **Biochimica et Biophysica Acta**, Amsterdam, v. 1399, p. 219–224, 1998.

de CASTRO SILVA FILHO, M.; CHAUMONT, F.; LETERME, S.; BOUTRY, M. Mitochondrial and chloroplast targeting sequences in tandem modify protein import specificity in plant organelles. **Plant Molecular Biology**, Dordrecht, v. 30, p. 769–780, 1996.

DEKKER, P.J.T.; MARTIN, F.; MAARSE, A.C.; BÖMER, U.; MÜLLER, H.; GUIARD, B.; MEIJER, M. RASSOW, J.; PFANNER, N. The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. **The EMBO Journal**, London, v. 16, p. 5408–5419, 1997.

DOOLITTLE, R.F.; FENG, D.F.; TSANG, S.; CHO, G.; LITTLE, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. **Science**, New York, v. 271, p. 470–477, 1996.

DRUMMOND, A.J.; RAMBAUT, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, London, v. 7, p. 214, 2007.

DUCHÊNE, A.M.; GIRITCH, A.; HOFFMANN, B.; COGNAT, V.; LANCELIN, D.; PEETERS, N.M.; ZAEPFEL, M.; MARÉCHAL-DROUARD, L.; SMALL, I.D. Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 102, p. 16484-16489, 2005.

DURFEE, T.; BECHERER, K.; CHEN, P.L.; YEH, S.H.; YANG, Y.; KILBURN, A.E.; LEE, W.H.; ELLEDGE, S.J. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. **Genes & Development**, New York, v. 7, p. 555-69, 1993.

ERIANI, G.; DELARUE, M.; POCH, O.; GANGLOFF, J.; MORAS, D. Partition of tRNA synthetases in two classes based on mutually exclusive sets of sequence motifs. **Nature**, London, v. 374, p. 203-206, 1990.

FRECHIN, M.; KERN, D.; MARTIN, R.B.; BECKER, H.D.; SENGER, B. Arc1p: Anchoring, routing, coordinating. **FEBS Letters**, Amsterdam, v. 584, p. 427-433, 2010.

FREIST, W.; GAUSS, D.H.; SOLL, D.; LAPOINTE, J. Glutamyl-tRNA synthetase. **Biological Chemistry**, Berlin, v. 378, p. 1313-1329, 1997.

GALANI, K.; GROBHANS, H.; DEINERT, K.; HURT, E.C.; SIMOS, G. The intracellular location of two aminoacyl-tRNA synthetases depends on a complex formation with Arc1p. **The EMBO Journal**, London, v. 20, p. 6889-6898, 2001.

GIETZ, R.D.; TRIGGS-RAINE, B.; ROBBINS, A.; GRAHAM, K.C.; WOODS, R.A. Identification of proteins that interact with proteins of interest: Applications of the yeast two-hybrid system. **Molecular and Cellular Biochemistry**, The Hague, v. 172, p. 67-79, 1997.

GIETZ, R.D.; WOODS, R.A. Transformation of yeast by the Liac/SS carrier DNA/PEG method. two-hybrid system. **Methods in Enzymology**, New York, v. 350, p. 87-96, 2002.

GODINIC, V.; MOCIBOB, M.; ROCAK, S.; IBBA, M.; WEYGAND-DURASEVIC, I. Peroxin Pex21p interacts with the C-terminal noncatalytic domain of yeast seryl-tRNA synthetase and forms a specific ternary complex with tRNA^{Ser}. **FEBS Journal**, Oxford, v. 274, p. 2788–2799, 2007.

GOLINELLI-COHEN, M.P.; MIRANDE, M. Arc1p is required for cytoplasmic confinement of synthetases and tRNA. **Molecular and Cellular Biochemistry**, The Hague, v. 300, p. 47-59, 2007.

GROSSHANS, H.; HURT, E.; SIMOS, G. An aminoacylation-dependent nuclear tRNA export pathways in yeast. **Genes & Development**, New York, v. 14, p. 830-840, 2000.

GUIGOU, L.; SHALAK, V.; MIRANDE, M. The tRNA-interacting factor p43 associates with mammalian arginyl-tRNA synthetase but does not modify its tRNA aminoacylation properties. **Biochemistry**, Washington, v. 43, p. 4592–4600, 2004.

GUO, F.; GABOR, J.; CEN, S.; HU, K.; MOULAND, A.J.; KLEIMAN, L. Inhibition of cellular HIV-1 protease activity by lysyl-tRNA synthetase. **The Journal of Biological Chemistry**, Bethesda, v. 280, p. 26018–26023, 2005.

GUZZO, C.M.; YANG, D.C. Lysyl-tRNA synthetase interacts with EF1alpha, aspartyl-tRNA synthetase and p38 in vitro. **Biochemical and Biophysical Research Communications**, Orlando, v. 365, p. 718–723, 2008.

HAMADA, S.; ITO, H.; HIRAGA, S.; INAGAKI, K.; NOZAKI, K.; ISONO, N.; YOSHIMOTO, Y.; TAKEDA, Y.; MATSUI, H. Differential characteristics and subcellular localization of two starch-branching enzyme isoforms encoded by a single gene in *Phaseolus vulgaris*. **The Journal of Biological Chemistry**, Bethesda, v. 277, p. 16538-16546, 2002.

HASHIMOTO, N.N.; CARNEVALLI, L.S.; CASTILHO, B.A. Translation initiation at non-AUG codons mediated by weakened association of eukaryotic initiation factor (eIF) 2 subunits. **Biochemical Journal**, London, v. 367, p. 359-368, 2002.

HAUSMANN, C.D.; PRAETORIUS-IBBA, M.; IBBA, M. An aminoacyl-tRNA synthetase: elongation factor complex for substrate channeling in archaeal translation. **Nucleic Acids Research**, Oxford, v. 35, p. 6094–6102, 2007.

HAUSMANN, C.D.; IBBA, M. Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. **FEMS Microbiology Reviews**, Amsterdam, v.32, p. 705–721, 2008.

HEAZLEWOOD, J.L.; TONTI-FILIPPINI, J.; VERBOOM, R.E.; MILLAR, A.H. Combining experimental and predicted datasets for determination of the subcellular location of proteins in *Arabidopsis*. **Plant Physiology**, Washington, v. 139, p. 598-609, 2005.

HECKMAN, D.S.; GEISER, D.M.; EIDELL, B.R.; STAUFFER, R.L.; KARDOS, N.L.; HEDGES, S.B. Molecular evidence for the early colonization of land by fungi and plants. **Science**, New York, v. 293, p. 1129-1133, 2001.

HEDGES, S.B.; BLAIR, J.E.; VENTURI, M.L.; SHOE, J.L. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. **BMC Evolutionary Biology**, London, v. 4, p. 2, 2004.

HOFFMAN, C.S.; WINSTON, F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. **Gene**, Amsterdam, v. 57, p. 267- 272, 1987.

HÖRMANN, F.; SOLL, J.; BÖLTER, B. The chloroplast protein import machinery: a review. **Methods in Molecular Biology**, Clifton, v. 390, p. 179-194, 2007.

HUELSENBECK, J.P.; RONQUIST, F. MRBAYES: Bayesian inference of phylogenetic trees. **Bioinformatics**, Oxford, v. 17, p. 754-755, 2001.

IBBA, M.; SÖLL, D. Aminoacyl-tRNA synthesis. **Annual Review of Biochemistry**, Palo Alto, v. 69, p. 617-650, 2000.

IBBA, M.; SÖLL, D. Aminoacyl-tRNAs: setting the limits of the genetic code. **Genes & Development**, New York, v. 18, p. 731-738, 2004.

IVAKHNO, S.S.; KORNELYUK, A.I. Bioinformatic analysis of changes in expression level of tyrosyl-tRNA synthetase during sporulation process in *Saccharomyces cerevisiae*. **Mikrobiol Z**, Berlin, v. 67, p. 37–49, 2005.

JAHN, D.; VERKAMP, E.; SÖLL, D. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. **Trends in Biochemical Sciences**, Amsterdam, v. 17, p. 215–218, 1992.

JAVANBAKHT, H.; HALWANI, R.; CEN, S.; SAADATMAND, J.; MUSIER-FORSYTH, K.; GOTTLINGER, H.; KLEIMAN, L. The interaction between HIV-1 Gag and human lysyl-tRNA synthetase during viral assembly. **Journal of Biological Chemistry**, Bethesda, v. 278, p. 27644–27651, 2003.

KARANASIOS, E.; SIMADER, H.; PANAYOTOU, G.; SUCK, D.; SIMOS, G. Molecular determinants of the yeast Arc1p-aminoacyl-tRNA synthetase complex assembly. **Journal of Molecular Biology**, London, v. 374, p. 1077–1090, 2007.

KARNIELY, S.; PINES, O. Single translation–dual destination: mechanisms of dual protein targeting in eukaryotes. **EMBO Reports**, London, v. 6, p. 420–425, 2005.

KIRCHER, S.; GIL, P.; KOZMA-BOGNÁR, L.; FEJES, E.; SPETH, V.; HUSSELSTEIN-MULLER, T.; BAUER, D.; ADÁM, E.; SCHÄFER, E.; NAGY, F. Nucleocytoplasmic partitioning of the plant photoreceptors phytochrome A, B, C, D, and E is regulated differentially by light and exhibits a diurnal rhythm. **The Plant Cell**, Rockville, v. 14, p. 1541-55, 2002.

KISSELEV, L.L.; JUSTESEN, J.; WOLFSON, A.D.; FROLOVA, L.Y. Diadenosine oligophosphates (Ap(n)A), a novel class of signalling molecules? **FEBS Letters**, Amsterdam, v. 427, p. 157–163, 1998.

KOVALESKI, B.J.; KENNEDY, R.; HONG, M.K.; DATTA, S.A.; KLEIMAN, L.; REIN, A.; MUSIER-FORSYTH, K. In vitro characterization of the interaction between HIV-1 Gag and human lysyl-tRNA synthetase. **Journal of Biological Chemistry**, Bethesda, v. 281, p. 19449–19456, 2006.

KOVÁCS-BOGDÁN, E.; SOLL, J.; BÖLTER, B. Protein import into chloroplasts: The Tic complex and its regulation. **Biochimica et Biophysica Acta**, Amsterdam, doi: 10.1016/j.bbamcr.2010.01.015, 2010.

LEE, S.W.; CHO, B.H.; PARK, S.G.; KIM, S. Aminoacyl-tRNA synthetase complexes: beyond translation. **Journal of Cell Science**, London, v. 117, p. 3725–3734, 2004.

LERMONTOVA, I.; SCHUBERT, V.; BÖRNKE, F.; MACAS, J.; SCHUBERT, I. Arabidopsis CBF5 interacts with the H/ACA snoRNP assembly factor NAF1. **Plant Molecular Biology**, Dordrecht, v. 65, p. 615–626, 2007.

LEVITAN, A.; TREBITSH, T.; KISS, V.; PEREG, Y.; DANGOOR, I.; DANON A. Dual targeting of the protein disulfide isomerase RB60 to the chloroplast and the endoplasmic reticulum. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 102, p. 6225–6230, 2005.

LIPMAN, R.S.; CHEN, J.; EVILIA, C.; VITSEVA, O.; HOU, Y.M. Association of an aminoacyl-tRNA synthetase with a putative metabolic protein in archaea. **Biochemistry**, Washington, v. 42, p. 7487–7496, 2003.

LISTER, R.; O'MALLEY, R.C.; TONTI-FILIPPINI, J.; GREGORY, B.D.; BERRY, C.C.; MILLAR, A.H. ECKER, J.R. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. **Cell**, Cambridge, v. 133, p. 523–536, 2008.

LUND, E.; DAHLBERG, S.G. Proofreading and aminoacylation of tRNAs before export from the nucleus. **Science**, New York, v. 282, p. 2082-2085, 1998.

MACASEV, D.; NEWBIGIN, E.; WHELAN, J.; LITHGOW, T. How do plant mitochondria avoid importing chloroplast proteins? Components of the import apparatus Tom20 and Tom22 from Arabidopsis differ from their fungal counterparts. **Plant Physiology**, Washington, v. 123, p. 811-816, 2000.

MACASEV, D.; WHELAN, J.; NEWBIGIN, E.; SILVA-FILHO, M.C.; MULHERN, T.D.; LITHGOW, T. Tom220, an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes, **Molecular Biology and Evolution**, Chicago, v. 21, p. 1557-1564, 2004.

MACKENZIE, J.A.; PAYNE, R.M. Ribosomes specifically bind to mammalian mitochondria via protease-sensitive proteins on the outer membrane. **Journal of Biological Chemistry**, Bethesda, v. 279, p. 9803-9810, 2004.

MACKENZIE, S.A. Plant organellar protein targeting: a traffic plan still under construction. **Trends in Cell Biology**, Cambridge, v. 15, p. 548-554, 2005.

MARC, P.; MARGEOT, A.; DEVAUX, F.; BLUGEON, C.; CORRAL-DEBRINSKI, M.; JACQ, C. Genome-wide analysis of mRNAs targeted to yeast mitochondria. **EMBO Reports**, London, v. 3, p. 159-64, 2002.

MARTIN, T.; SHARMA, R.; SIPPEL, C.; WAEGEMANN, K.; SOLL, J.; VOTHKNECHT, U.C. A protein kinase family in Arabidopsis phosphorylates chloroplast precursor proteins. **Journal of Biological Chemistry**, Bethesda, v. 281, p. 40216–40223, 2006.

MAZUMDER, B.; SESHADRI, V.; FOX, P.L. Translational control by the 30-UTR: the ends specify the means. **Trends in Biochemical Sciences**, Amsterdam, v. 28, p. 91–98, 2003.

MEINECKE, M.; WAGNER, R.; KOVERMANN, P.; GUIARD, B.; MICK, D.U.; HUTU, D.P.; VOOS, W.; TRUSCOTT, K.N.; CHACINSKA, A.; PFANNER, N.; REHLING, P. Tim50 maintains the permeability barrier of the mitochondrial inner membrane. **Science**, New York, v. 312, p. 1523–1526, 2006.

MICHALECKA, A.M.; SVENSSON, A.S.; JOHANSSON, F.I.; AGIUS, S.C.; JOHANSON, U.; BRENNICKE, A.; BINDER, S.; RASMUSSEN, A.G. Arabidopsis genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light. **Plant Physiology**, Washington, v. 133, p. 642–652, 2003.

MILLAR, A.H.; WHELAN, J.; SMALL, I. Recent surprises in protein targeting to mitochondria and plastids. **Current Opinion in Plant Biology**, London, v. 9, p. 610–615, 2006.

MILISAV, I.; MORO, F.; NEUPERT, W.; BRUNNER, M. Modular structure of the TIM23 preprotein translocase of mitochondria. **Journal of Biological Chemistry**, Bethesda, v. 276, p. 25856–25861, 2001.

MIRANDE, M. Aminoacyl-tRNA synthetase family from prokaryotes and eukaryotes: structural domains and their implications. **Progress in Nucleic Acid Research & Molecular Biology**, New York, 40, 95–142, 1991.

MITSCHE, J.; FUSS, J.; BLUM, T.; HÖGLUND, A.; RESKI, R.; KOHLBACHER, O.; RENSING, S.A. Prediction of dual protein targeting to plant organelles. **New Phytologist**, Cambridge, v. 183, :p. 224-235, 2009.

MOBERG, P.; NILSSON, S.; STAHL, A.; ERIKSSON, A.C.; GLASER, E.; MALER, L. NMR solution structure of the mitochondrial F1beta presequence from *Nicotiana plumbaginifolia*. **Journal of Molecular Biology**, London, v. 336, p. 1129–1140, 2004.

MOKRANJAC, D.; SICHTING, M.; POPOV-CELEKETIC, D.; MAPA, K.; GEVORKYAN-AIRAPETOV, L.; ZOHARY, K.; HELL, K.; AZEM, A.; NEUPERT, W. Role of Tim50 in the transfer of precursor proteins from the outer to inner membrane of mitochondria. **Molecular Biology of the Cell**, Bethesda, v. 20, p. 1400–1407, 2009.

MORGANTE, C.V.; RODRIGUES, R.A.O.; MARBACH, P.A.S.; BORGONOV, C.M.; MOURA, D.S.; SILVA-FILHO, M.C. Conservation of dual-targeted proteins in Arabidopsis and rice points to a similar pattern of gene family evolution. **Molecular Genetics and Genomics**, Berlin, v. 281, n. 5, p. 525-38, 2009.

MOTORIN, Y.; WOLFSON, A.D.; ORLOVSKY, A.F.; GLADILIN, K.L. Mammalian valyl-tRNA synthetase forms a complex with the first elongation factor. **FEBS Letters**, Amsterdam, v. 238, p. 262–264, 1988.

NEGRUTSKII, B.S.; SHALAK, V.F.; KERJAN, P.; EL'SKAYA, A.V.; MIRANDE, M. Functional interaction of mammalian valyl-tRNA synthetase with elongation factor EF-1alpha in the complex with EF-1H. **Journal of Biological Chemistry**, Bethesda, v. 274, p. 4545–4550, 1999.

NEUPERT, W.; HERMANN, J.M. Translocation of proteins into mitochondria. **Annual Review of Biochemistry**, Palo Alto, v. 76, p. 723-749, 2007.

O'DONOGHUE, P.; LUTHEY-SCHULTEN, Z. On the Evolution of Structure in Aminoacyl-tRNA synthetases. **Microbiology and Molecular Biology Reviews**, New York, v. 67, p. 550-573, 2003.

OBARA, K.; SUMI, K.; FUKUDA, H. The use of multiple transcription start sites causes the dual targeting of Arabidopsis putative monodehydroascorbate reductase to both mitochondria and chloroplasts. **Plant Cell Physiology**, Oxford, v. 43, p. 697-705, 2002.

OHAD, N.; SHICHRUR, K.; YALOVSKYL, S. The Analysis of Protein-Protein Interactions in Plants by Bimolecular Fluorescence Complementation. **Plant Physiology**, Washington, v. 145, p. 1090–1099, 2007.

PARK, S.G.; EWALT, K.L.; KIM, S. Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives and housekeepers. **Trends in Biochemical Sciences**, Amsterdam, v. 30, p. 569-574, 2005.

PEETERS, N.; SMALL, I. Dual targeting to mitochondria and chloroplasts. **Biochimica et Biophysica Acta**, Amsterdam, v. 1541, p. 54–63, 2001.

PERRY, A.J.; RIMMER, K.A.; MERTENS, H.D.T.; WALLER, R.F.; MULHERN, T.D.; LITHGOW, T.; GOOLEY P.R. Structure, topology and function of the translocase of the outer membrane of mitochondria. **Plant Physiology and Biochemistry**, Paris, v. 46, p. 265-274, 2008.

PERIASAMY, A. Fluorescence resonance energy transfer microscopy: a mini review. **Journal for Biomedical Optics**, Amsterdam, v. 6, p. 287–291, 2000.

PETROVA, V.Y.; DRESCHER, D.; KUJUMDZIEVA, A.V.; SCHMITT, M.J. Dual targeting of yeast catalase A to peroxisomes and mitochondria. **Biochemical Journal**, London, v. 380, p. 393-400, 2004.

PRAETORIUS-IBBA, M.; ROGERS, T.E.; SAMSON, R.; KELMAN, Z.; IBBA, M. Association between archaeal prolyl- and leucyl-tRNA synthetases enhances tRNA^{Pro} aminoacylation. **Journal of Biological Chemistry**, Bethesda, v. 280, p. 26099–26104, 2005.

PRAETORIUS-IBBA, M.; HAUSMANN, C.D.; PARAS, M.; ROGERS, T.E.; IBBA, M. Functional association between three archaeal aminoacyl-tRNA synthetases. **Journal of Biological Chemistry**, Bethesda, v. 282, p. 3680–3687, 2007.

Pujol, C.; Marechal-Drouard, L.; Duchêne, A.M. How can organellar protein N-terminal sequences be dual targeting signals? In silico analysis and mutagenesis approach. **Journal of Molecular Biology**, London, v. 369, p. 356–367, 2007.

RAY, P.S.; ARIF, A.; FOX, P.L. Macromolecular complexes as depots for releasable regulatory proteins. **Trends in Biochemical Sciences**, Amsterdam, v. 32, p. 158–164, 2007.

REED, V.S.; YANG, D.C. Characterization of a novel N-terminal peptide in human aspartyl-tRNA synthetase. Roles in the transfer of aminoacyl-tRNA from aminoacyl-tRNA synthetase to the elongation factor 1 α . **Journal of Biological Chemistry**, Bethesda, v. 269, p. 32937–32941, 1994.

REED, V.S.; WASTNEY, M.E.; YANG, D.C. Mechanisms of the transfer of aminoacyl-tRNA from aminoacyl-tRNA synthetase to the elongation factor 1 α . **Journal of Biological Chemistry**, Bethesda, v. 269, p. 32932–32936, 1994.

REUMANN, S.; BABUJEE, L.; MA, C.; WIENKOOP, S.; SIEMSEN, T.; ANTONICELLI, G.E.; RASCHE, N.; LU, D. F., WECKWERTH, W.; JAHN, O. Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. **The Plant Cell**, Rockville, v. 19, p. 3170–3193, 2007.

ROCAK, S.; LANDEKA, I.; WEYGAND-DURASEVIC, I. Identifying Pex21p as a protein that specifically interacts with yeast seryl-tRNA synthetase. **FEMS Microbiology Letters**, Amsterdam, v. 214, p. 101–106, 2002.

ROBINSON, J.C.; KERJAN, P.; MIRANDE, M. Macromolecular assemblage of aminoacyl-tRNA synthetases: quantitative analysis of protein–protein interactions and mechanism of complex assembly. **Journal of Molecular Biology**, London, v. 304, p. 983–994, 2000.

ROKOV-PLAVEC, J.; DULIC, M.; DUCHÊNE, A.M.; WEYGAND-DURASEVIC, I. Dual targeting of organellar seryl-tRNA synthetase to maize mitochondria and chloroplasts. **Plant Cell Reports**, Berlin, v. 27, p. 1157–1168, 2008.

RUAN, B.; SÖLL, D. The bacterial YbaK protein is a CysRNA Pro and Cys-tRNA^{Cys} deacylase. **Journal of Biological Chemistry**, Bethesda, v. 280, p. 25887–25891, 2005.

RUDHE, C.; CLIFTON, R.; CHEW, O.; ZEMAM, K.; RICHTER, S.; LAMPPA, G.; WHELAN, J.; GLASER, E. Processing of the dual targeted precursor protein of glutathione reductase in mitochondria and chloroplasts. **Journal of Molecular Biology**, London, v. 343, p. 639–647, 2004.

RYCKELYNCK, M.; GIEGE, R.; FRUGIER, M. tRNAs and tRNAs mimics as cornerstones of aminoacyl-tRNA synthetase regulations. **Biochimie**, Paris, v. 87, p. 835–845, 2005.

SAKAMOTO, K; BRIGGS, W.R. Cellular and subcellular localization of phototropin 1. **The Plant Cell**, Rockville, v. 14, p. 1723-1735, 2002.

SAMBROOK, J.; RUSSEL, D.W. **Molecular Cloning: A Laboratory Manual**. 3 ed., New York: Cold Spring Harbor Laboratory Press, 2001, p. 1.31-1.34.

SAMPATH, P.; MAZUMDER, B.; SESHADRI, V.; GERBER, C.A.; CHAVATTE, L.; KINTER, M.; TING, S.M.; DIGNAM, J.D.; KIM, S.; DRISCOLL, D.M.; FOX, P.L. Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. **Cell**, Cambridge, v. 119, p. 195–208, 2004.

SAPIR-MIR, M.; METT, A.; BELAUSOV, E.; TAL-MESHULAM, S.; FRYDMAN, A.; GIDONI, D.; EYAL, Y. Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. **Plant Physiology**, Washington, v. 148, p. 1219–1228, 2008.

SARKAR, S.; AZAD, A.K.; HOPPER, A.K. Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in *Saccharomyces cerevisiae*. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 96, p. 14366-14371, 1999.

SCHIMMEL, P.; GIEGÉ, R.; MORAS, D.; YOKOYAMA, S. An operational RNA code for amino acids and possible relationship to genetic code. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 90, p. 8763-8768, 1993.

SCHÖN, A.; KRUPP, G.; GOUGH, S.; BERRY-LOWE, S.; KANNANGARA, C.G.; SÖLL, D. The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. **Nature**, London, v. 322, p. 281–284, 1986.

SCHWACKE, R.; FISCHER, K.; KETELSEN, B.; KRUPINSKA, K.; KRAUSE, K. Comparative survey of plastid and mitochondrial targeting properties of transcription factors in *Arabidopsis* and rice. **Molecular Genetics and Genomics**, Berlin, v. 277, p. 631–646, 2007.

SILVA-FILHO, M.C. One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. **Current Opinion in Plant Biology**, New York, v. 6, p. 589–595, 2003.

SIMADER, H.; HOTHORN, M.; KOHLER, C.; BASQUIN, J.; SIMOS, G.; SUCK, D. Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. **Nucleic Acids Research**, Oxford, v. 34, p. 3968–3979, 2006.

SIMOS, G.; SEGREF, A.; FASIOLO, F.; HELLMUTH, K.; SHEVCHENKO, A.; MANN, M.; HURT, E.C. The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl and glutamyl-tRNA synthetases. **The EMBO Journal**, London, v. 15, p. 5437-5448, 1996.

SIMOS, G.; SAUER, A.; FASIOLO, F.; HURT, E.C. A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. **Molecular Cell**, Cambridge, v. 1. p. 235–242, 1998.

SMALL, I.; PEETERS, N.; LEGEAI, F.; LURIN, C. Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. **Proteomics**, Weinheim, v. 4, p. 1581–1590, 2004.

TAIRA, M.; VALTERSSON, U.; BURKHARDT, B.; LUDWIG, R.A. Arabidopsis thaliana GLN2-encoded glutamine synthetase is dual targeted to leaf mitochondria and chloroplasts. **The Plant Cell**, Rockville, v. 16, p. 2048-2058, 2004.

TAMURA, Y., HARADA, Y., SHIOTA, T.; YAMANO, K.; WATANABE, K.; YOKOTA, M.; YAMAMOTO, H.; SESAKI, H.; ENDO, T. Tim23-Tim50 pair coordinates functions of translocators and motor proteins in mitochondrial protein import. **The Journal of Cell Biology**, New York, v. 184, p. 129–141, 2009.

TRUSCOTT, K.N., KOVERMANN, P.; GEISLER, A.; MERLIN, A.; MEIJER, M.; DRIESSEN, A.J.; RASSOW, J.; PFANNER, N.; WAGNER, R. A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. **Nature Structural Biology**, New York, v. 8, p. 1074–1082, 2001.

UEDA, M.; NISHIKAWA, T.; FUJIMOTO, M.; TAKANASHI, H.; ARIMURA, S.I.; TSUTSUMI, N.; KADOWAKI, K.I. Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. **Molecular Biology and Evolution**, Chicago, v. 25, p. 1566–1575, 2008.

VENEMA, R.C.; PETERS, H.I.; TRAUGH, J.A. Phosphorylation of valyl-tRNA synthetase and elongation factor 1 in response to phorbol esters is associated with stimulation of both activities. **Journal of Biological Chemistry**, Bethesda, v. 266, p. 11993–11998, 1991.

VOINNET, O.; RIVAS, S.; MESTRE, P.; BAULCOMBE, D. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. **The Plant Journal**, Oxford, v. 33, p. 949-956, 2003.

von HEIJNE, G.; NISHIKAWA, K. Chloroplast transit peptides. The perfect random coil? **FEBS Letters**, Amsterdam, v. 278, p. 1–3, 1989.

WAEGERMANN, K.; SOLL, J. Phosphorylation of the transit sequence of chloroplast precursor proteins. **Journal of Biological Chemistry**, Bethesda, v. 271, p. 6545–6554, 1996.

WALTER, M.; CHABAN, C.; SCHÜTZE, K.; BATISTIC, O.; WECKERMANN, K.; NÄKE, C.; BLAZEVIC, D.; GREFFEN, C.; SCHUMACHER, K.; OECKING, C.; HARTER, K.; KUDLA, J. Visualization of protein interactions in living plant cells using biomolecular fluorescent complementation. **The Plant Journal**, New York, v. 40, p. 428-438, 2004.

WHELAN, S.; GOLDMAN, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. **Molecular Biology and Evolution**, Chicago, v.18, n. 5, p. 691-699, 2001.

WOESE, C.R.; OLSEN, G.J.; IBBA, M.; SÖLL, D. Aminoacyl-tRNA synthetases, the genetic code and evolutionary process. **Microbiology and Molecular Biology Reviews**, New York, v. 64, p. 202-236, 2000.

APÊNDICES

APÊNDICE 1 – Sequência dos iniciadores utilizados nas clonagens e na confirmação das clonagens por sequenciamento das sequências codificadoras de GluRS e de GS. O sítio de restrição das enzimas utilizadas encontra-se destacado em sublinhado

	Iniciador	Iniciador reverso
GluRS (vetores pGADT7 e pGBKT7)	5' GGGC <u>ATCGAT</u> ACAAGATGGAT GGGATGAAGCTTTTCG 3'	5' TACGGGATCCATTCACTTAGC GGCTCTTCCATC 3'
GluRS (vetor pSPYCE-35S)	5' CACCGGATCC <u>AAGAT</u> GGATGG GATGAAGCTTTTCG 3'	5' TACATCGATAGTCTTAGCGGC TCTTCCATCTGG 3'
GS (vetores pGADT7 e pGBKT7)	5' ACT <u>CATATG</u> CCAATGGCTCAG ATCTTAGCAGCT 3'	5' ACTGAATTCATTTTAAACATTC AAAGAAAGCTT 3'
GS (vetor pSYNE-35S)	5' CACC <u>ATCGAT</u> CCAATGGCTCA GATCTTAGCAGCT 3'	5' TACCTCGAGGGTAACATTCAA AGAAAGCTTTTG 3'
GS sequenciamento 2	5' TACCGGTCAAGCACCTGGTGA AGATAGTGA 3'	
GS sequenciamento 3	5' TACAAAGCTTGTTTATATGCTG GAATTAAC 3'	
GS sequenciamento 4	5' TGACCGGAAAGCACGAGACAG CTAGTATTG 3'	
GluRS sequenciamento 1	5' CCTAAGTCTAAAGATAGCCA 3'	
GluRS sequenciamento 2	5' GAGAGAGGATTACAGTGCTG 3'	
GluRS sequenciamento 3	5' TGCCCTAGACACACTGCTGT 3'	

APÊNDICE 2 – Sequência dos iniciadores utilizados nas clonagens e na confirmação das clonagens da região codificadora da extensão N-terminal de GluRS. O sítio de restrição das enzimas utilizadas encontra-se destacado em sublinhado

	Iniciador	Iniciador reverso
A1B1	5' <u>GGGATCC</u> AAAAGATGGATGGG ATGAAGCTTTTCG 3'	5' CAT <u>GTCGACTT</u> CTGCATCTGC TCCCTC 3'
A1B2	5' <u>GGGATCC</u> AAAAGATGGATGGG ATGAAGCTTTTCG 3'	5' CAT <u>GTCGACG</u> AACCAGAAGAG AAGACA 3'
A1B3	5' <u>GGGATCC</u> AAAAGATGGATGGG ATGAAGCTTTTCG 3'	5' CAT <u>GTCGAC</u> AAGTCCACTTCA GGCTTACCT 3'
A2B1	5' <u>GGGATCC</u> CAGAGTTTGAGAAT GCTTGTGGT 3'	5' CAT <u>GTCGACTT</u> CTGCATCTGC TCCCTC 3'
A3B1	5' <u>GGGATCCT</u> GCCGGAAGCGGA GATTGGAAA 3'	5' CAT <u>GTCGACTT</u> CTGCATCTGC TCCCTC 3'

APÊNDICE 3 – Sequências das regiões correspondentes a cada uma das deleções utilizadas para o mapeamento da região de interação de GluRS nos ensaios de duplo-híbrido. O sítio de anelamento dos iniciadores encontra-se destacado em sublinhado

A1B1 – tamanho da deleção: 950 pares de bases (1 a 950)

ATGGATGGGATGAAGCTTTCGTTCCCACCGGAAAGTCCACCACTTTCAGTCATCGT
TGCTCTTTCTCTCTCAGCTTCTCCGGTGACGATTGATTCTTCCGCCGCTGCAACAAC
CGTCCCTTCTTTTGTCTTCTCCGACGGGAGGAAATTGAATGGAGCCACCGTTCTTC
TTCGCTATGTTGGTCGATCAGCGAAAAAGCTTCCTGATTTCTATGGCAACAATGCTT
TTGATTCTTCTCAGATTGATGAGTGGGTAGATTACGCATCTGTCTTCTCTTCTGGTT
CAGAGTTTGAGAATGCTTGTGGTCGTGTTGATAAGTATCTCGAGAGTAGCACGTTT
CTTGTTGGCCATTCTCTTTCCATTGCTGATGTCGCTATTTGGTCAGCTCTTGCTGGA
ACTGGTCAAAGATGGGAAAGTTTGAGGAAATCTAAAAAGTATCAGAGTCTTGTTAGA
TGGTTCAATTCGATATTAGACGAGTACAGTGAGGTGCTTAACAAGGTTCTAGCAACT
TATGTTAAGAAAGGATCAGGGAAGCCTGTTGCTGCACCTAAGTCTAAAGATAGCCA
ACAAGCTGTGAAAGGAGATGGTCAGGATAAAGGTAAGCCTGAAGTGGACTTGCCG
GAAGCGGAGATTGGAAAGGTTAAACTCCGGTTTGCTCCAGAGCCAAGTGGTTATCT
TCACATAGGACATGCTAAGGCTGCGTTGCTGAACAAGTATTTGCTGAGCGTTACC
AAGGGGAAGTGATTGTGCGTTTTGATGATACTAACCTGCTAAAGAAAGCAATGAG
TTTGTGGATAATCTTGTGAAGGATATTGGGACCTTGGGGATCAAGTATGAGAAAGT
GACATACACTTCGGACTATTTTCCTGAATTGATGGATATGGCGGAAAAACTGATGC
GTGAGGGTAAGGCATATGTTGATGACACACCGAGGGAGCAGATGCAGAA

A1B2 – tamanho da deleção: 283 pares de bases (1 a 284)

ATGGATGGGATGAAGCTTTCGTTCCCACCGGAAAGTCCACCACTTTCAGTCATCGT
TGCTCTTTCTCTCTCAGCTTCTCCGGTGACGATTGATTCTTCCGCCGCTGCAACAAC
CGTCCCTTCTTTTGTCTTCTCCGACGGGAGGAAATTGAATGGAGCCACCGTTCTTC
TTCGCTATGTTGGTCGATCAGCGAAAAAGCTTCCTGATTTCTATGGCAACAATGCTT
TTGATTCTTCTCAGATTGATGAGTGGGTAGATTACGCATCTGTCTTCTCTTCTGGTT

C

A1B3 – tamanho da deleção: 616 pares de bases (1 a 617)

ATGGATGGGATGAAGCTTTCGTTCCCACCGGAAAGTCCACCACTTTCAGTCATCGT
TGCTCTTTCTCTCTCAGCTTCTCCGGTGACGATTGATTCTTCCGCCGCTGCAACAAC
CGTCCCTTCTTTTGTCTTCTCCGACGGGAGGAAATTGAATGGAGCCACCGTTCTTC
TTCGCTATGTTGGTCGATCAGCGAAAAAGCTTCCTGATTTCTATGGCAACAATGCTT
TTGATTCTTCTCAGATTGATGAGTGGGTAGATTACGCATCTGTCTTCTCTTCTGGTT
CAGAGTTTGAGAATGCTTGTGGTCGTGTTGATAAGTATCTCGAGAGTAGCACGTTT
CTTGTTGGCCATTCTCTTTCCATTGCTGATGTCGCTATTTGGTCAGCTCTTGCTGGA
ACTGGTCAAAGATGGGAAAGTTTGAGGAAATCTAAAAAGTATCAGAGTCTTGTTAGA
TGGTTCAATTCGATATTAGACGAGTACAGTGAGGTGCTTAACAAGGTTCTAGCAACT
TATGTTAAGAAAGGATCAGGGAAGCCTGTTGCTGCACCTAAGTCTAAAGATAGCCA
ACAAGCTGTGAAAGGAGATGGTCAGGATAAAGGTAAGCCTGAAGTGGACTTG

A2B1 – tamanho da deleção: 667 pares de bases (285 a 950)

GAGTTTGAGAATGCTTGTGGTCGTGTTGATAAGTATCTCGAGAGTAGCACGTTTCTT
GTTGGCCATTCTCTTTCCATTGCTGATGTCGCTATTTGGTCAGCTCTTGCTGGA
GGTCAAAGATGGGAAAGTTTGAGGAAATCTAAAAAGTATCAGAGTCTTGTTAGATG
GTTCAATTCGATATTAGACGAGTACAGTGAGGTGCTTAACAAGGTTCTAGCAACTTA
TGTTAAGAAAGGATCAGGGAAGCCTGTTGCTGCACCTAAGTCTAAAGATAGCCAAC
AAGCTGTGAAAGGAGATGGTCAGGATAAAGGTAAGCCTGAAGTGGACTTGCCGGA
AGCGGAGATTGGAAAGGTTAACTCCGGTTTGCTCCAGAGCCAAGTGGTTATCTTC
ACATAGGACATGCTAAGGCTGCGTTGCTGAACAAGTATTTGCTGAGCGTTACCAA
GGGGAAGTGATTGTGCGTTTTGATGATACTAACCCTGCTAAAGAAAGCAATGAGTT
TGTGGATAATCTTGTGAAGGATATTGGGACCTTGGGGATCAAGTATGAGAAAGTGA
CATACACTTCGGACTATTTTCCTGAATTGATGGATATGGCGGAAAACTGATGCGTG
AGGGTAAGGCATATGTTGATGACACACCGAGGGAGCAGATGCAGAA

A3B1 – tamanho da deleção: 332 pares de bases (618 a 950)

CCGGAAGCGGAGATTGGAAAGGTTAACTCCGGTTTGCTCCAGAGCCAAGTGGTT
ATCTTCACATAGGACATGCTAAGGCTGCGTTGCTGAACAAGTATTTGCTGAGCGT
TACCAAGGGGAAGTGATTGTGCGTTTTGATGATACTAACCCCTGCTAAAGAAAGCAA
TGAGTTTGTGGATAATCTTGTGAAGGATATTGGGACCTTGGGGATCAAGTATGAGA
AAGTGACATACACTTCGGACTATTTTCCTGAATTGATGGATATGGCGGAAAAACTGA
TGCGTGAGGGTAAGGCATATGTTGATGACACACCCGAGGGAGCAGATGCAGAA

ANEXOS

	360	370	380	390	400
18978125 Pyrococcus Euryarchaeota	IKP---	DEIVYASDRLEIYYKYAEELIKMGKAYVCTCPDFRRLDQC			
170290688 Candidatus Korarchaeota	LSFGNDEEIEESYHEEMYSLAKKLEAGKAYVCTCPAAFRVRKRLC				
161527819 Nitrosopumilus Crenarchaeota	IEF---	DKVSTSDCMVEFYEKGLINSKAYVCTCKREDISQNRREI			
Enterocytozoon Eurotiomycetes Microsporidia 169806531	---	IKNYITVRSDDHFDLYNYAQLICLGLAYCNDIQCMRREERTK			
Encephalitozoon Eurotiomycetes Microsporidia 19074040	---	INETYLSHTSDYFDKIIDLCVELIGEGKAYADNTPQVNRDEGRG			
Oryza Liliopsida Streptophyta 115481716	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Sorghum Liliopsida Streptophyta 242036979	---	IKYIVVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Zea Liliopsida Streptophyta 226530698	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Sorghum Liliopsida Streptophyta 242039397	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Oryza Liliopsida Streptophyta 115435896	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Ricinus Eurotiomycetes Streptophyta 255543064	---	VYVYKVIHYSDDYFDELMKKAELIRQKRAYVDDTPRECMRSEMDG			
Populus Eurotiomycetes Streptophyta 224124284	---	IKYETIITHSDYFPQLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Populus Chlorophyceae Streptophyta 224122796		---	MAESLIQKRAYVDDTPRECMRSEMDG		
A3B1	---	IKYEVVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
GluRS	---	IKYEVVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Vitis Eurotiomycetes Streptophyta 225450407	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Vitis Eurotiomycetes Streptophyta 225450409	---	IKYDAVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Physcomitrella Bryopsida Streptophyta 168020667	---	VKGIIVITYSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Physcomitrella Bryopsida Streptophyta 168016970	---	VKGIIVITYSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Pichia 254567253	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Ajellomyces 261201127	---	IYEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Coccidioides Eurotiomycetes Ascomycota 119187457	---	IYEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Uncinocarpus Eurotiomycetes Ascomycota 258563978	---	IYEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 115383992	---	VIPTD-----	QMEACGLISGKAYADNTPRECMRSEMDG		
Aspergillus Eurotiomycetes Ascomycota 169783372	---	IKYEVVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Schizosaccharomyces Ascomycota 213407330	---	IKYEVVIYTSDDYFPKLMEMAEKLIKQKRAYVDDTPRECMRSEMDG			
Schizosaccharomyces Ascomycota 19114395	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Zygosaccharomyces Saccharomycetes Ascomycota 254583225	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Saccharomyces Saccharomycetes Ascomycota 37362649	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Kluyveromyces Saccharomycetes Ascomycota 50302189	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Eremothecium Saccharomycetes Ascomycota 45198782	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Lachancea Saccharomycetes Ascomycota 255717577	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Vanderwaltozyma Saccharomycetes Ascomycota 156840649	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Nakaseomyces Saccharomycetes Ascomycota 50285663	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Pichia Saccharomycetes Ascomycota 146422191	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Debaryomyces Saccharomycetes Ascomycota 50425499	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Lodderomyces Saccharomycetes Ascomycota 149234782	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Candida Saccharomycetes Ascomycota 255731680	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Candida Saccharomycetes Ascomycota 68471145	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Candida Saccharomycetes Ascomycota 241957009	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Pichia Saccharomycetes Ascomycota 126136887	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Saccharomyces Saccharomycetes Ascomycota 260942845	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Yarrowia Saccharomycetes Ascomycota 50554199	---	IKGDKITYSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Magnaporthe Sordariomycetes Ascomycota 39976235	---	IKADKFSYTSDDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Podospota Sordariomycetes Ascomycota 171693025	---	IKCDITLYTSDDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Neurospora Sordariomycetes Ascomycota 85091295	---	IKFDLSYTSDDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Chaetomium Sordariomycetes Ascomycota 116198903	---	IKYHPTVTSDDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Gibberella Sordariomycetes Ascomycota 46117480	---	IKEDCVSYTSDDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Pyrenophora Dothideomycetes Ascomycota 189206994	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Penicillium Eurotiomycetes Ascomycota 212535232	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Talaromyces Eurotiomycetes Ascomycota 242792479	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 169776501	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 238503207	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 115433080	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Penicillium Eurotiomycetes Ascomycota 255955665	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 121718334	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 70984964	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Neosartorya Eurotiomycetes Ascomycota 119498841	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Aspergillus Eurotiomycetes Ascomycota 145242672	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Emericella Eurotiomycetes Ascomycota 67902474	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Ajellomyces Eurotiomycetes Ascomycota 154288154	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Filobasidiella Tremellomycetes Basidiomycota 58268242	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Filobasidiella Tremellomycetes Basidiomycota 134113522	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Coprinopsis Agaricomycetes Basidiomycota 169852668	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Malassezia Saccharomycetes Basidiomycota 164657287	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
Ustilago Ustilaginomycetes Basidiomycota 71021397	---	IKEDKITSFSSDYFQIMYDYAVCMIRKDGKAYCDTPTERMREERMG			
42523768 Bdellovibrio Proteobacteria	---	IKVGPNGPYRQSERLIIYKIEADQLIKGKAYCFMTSEEDIEKQ---			
75911057 Anabaena Cyanobacteria	---	---	GPFFGQRLDLYQKAVKLLDQGLAYCYTSEELEAQRAC		
222530218 Anaerococcus Firmicutes	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		
146295745 Caldicellulosiruptor Firmicutes	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		
52078587 Bacillus Firmicutes	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		
120554808 Marinobacter Proteobacteria	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		
70729429 Pseudomonas Proteobacteria	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		
15598330 Pseudomonas Proteobacteria	---	---	GPYGPYRSTEFVDIYKAVVLEKGYAYCYTSEELEAQRAC		